Ann. Hum. Genet. (2001), **65**, 583–601
DOI: 10.1017/S0003480001008909 *Printed in the United Kingdom*

Power of regression and maximum likelihood methods to map QTL from sib-pair and DZ twin data

P. M. VISSCHER[1,2] AND J. L. HOPPER[1]

[1] *Centre for Genetic Epidemiology, The University of Melbourne, Melbourne, Australia*
[2] *Institute of Cell, Animal and Population Biology, The University of Edinburgh, Edinburgh, Scotland*

(Received 3.7.01. Accepted 19.9.01)

SUMMARY

A common study design to map quantitative trait loci (QTL) is to compare the phenotypes and marker genotypes of two or more siblings in a sample of unrelated sib groups, and to test for linkage between chromosome location and quantitative trait values. The simplest case is sib pairs only, in particular dizygotic twin pairs, and a simple and elegant regression method was proposed by Haseman & Elston in 1972 to test for linkage. Since then, several other methods have been proposed to test for linkage. In this study, we derived the statistical power of linear regression and maximum likelihood methods to map QTL from sib pair data analytically, and determined which methods are superior under which set of population parameters. In particular, we considered four regression-based and three maximum likelihood-based approaches, and derived asymptotic approximations of the mean test statistic and statistical power for each method. It was found, both analytically and by computer simulation, that the revisited or new Haseman–Elston method (based upon the mean-corrected crossproduct of the observations on sib-pairs) is less powerful than a full maximum likelihood approach and is also inferior to the Haseman–Elston method under a realistic range of values for the population parameters. We found that a simple regression method, based upon both the squared difference and the mean-corrected squared sum of the observations on sib-pairs, is as powerful as a full maximum likelihood approach. Our derivations of statistical power for regression and maximum likelihood methods provide a simple way to compare alternative methods and obviate the need to perform elaborate computer simulations. DZ twin pairs are likely to be more powerful for linkage analysis than ordinary siblings because they may share more common environmental effects, thereby increasing the proportion of within-family variance that is explained by a QTL.

INTRODUCTION

Quantitative traits in human populations are generally thought to be influenced by multiple genetic loci, as well as environmental factors. These traits, such as blood pressure, body-mass index and cholesterol levels, are known to be risk factors for common non-Mendelian disorders such as diabetes and cardiovascular diseases. Quantitative trait variation is proving difficult to dissect into contributions from specific genetic loci, because the genotype of an individual cannot be uniquely determined based upon phenotype and because variation from other genetic loci and environmental factors obscures the effects of variation at any single locus. Nevertheless, linkage analysis is a proven method to detect chromosome regions which co-segregate with a disease, and the same principles have been used to map quantitative trait loci (QTL) in animal and human populations. One common

Correspondence: P. M. Visscher, Institute of Cell, Animal and Population Biology, The University of Edinburgh, West Mains Road, Edinburgh EH9 3JT, UK. Tel: +44 131 650 7702; Fax: +44 131 650 6564.
E-mail: peter.visscher@ed.ac.uk

Table 1. *Proposed linkage analyses for quantitative traits on sib-pair data*

Variables[a]	Method	References
Y_1 and Y_2	ML parametric linkage analysis	Amos & de Andrade (2001) and references therein
D^2	Linear regression	Haseman & Elston (1972)
D^2 and S^2	Linear regression	Drigalenko (1998); Xu *et al.* (2000); Forrest (2001)
CP	Linear regression	Drigalenko (1998); Elston *et al.* (2000)
D	ML	Kruglyak & Lander (1995)
D and S	ML	Wright (1997)
Y_1 and Y_2	ML variance components	Goldgar (1990); Amos (1994); Fulker & Cherny (1996); Wright (1997)

[a] Y_1 and Y_2 are observations on the sib-pairs. D, S, D^2, S^2 and CP are the difference, mean-corrected sum, squared difference, squared mean-corrected sum and the mean-corrected cross-product, respectively.

sampling design used in human populations to detect linkage is to obtain phenotypes and DNA markers from multiple families consisting of two or more siblings, and to test the correlations between each phenotype and the marker genotypes of the sibs. The simplest of these sampling schemes, the sib-pair design, is when exactly two sibs are scored for a quantitative phenotype and a set of genetic markers.

Much has been written about linkage analysis for complex quantitative traits under such a sib-pair design, despite the fact that the nature of the data is exceedingly simple, for each of a set of n sib-pairs, we have two phenotypic observations and measures on the number of alleles that the sibs share identical by descent (IBD). In the case of fully informative markers, when the proportion of alleles that two sibs or twins share at a particular locus in the genome is known without error, the data are even simpler because the IBD proportion has only three distinct values, 0, $\frac{1}{2}$ or 1.

Nevertheless, a plethora of methods has been proposed for analysis of linkage using the sib-pair design. The landmark paper by Haseman & Elston (1972) proposed using classic linear regression to regress the squared sib-pair difference on the proportion of alleles shared IBD. As in Elston *et al.* (2000) and Allison *et al.* (2000), we refer to this method as the old Haseman–Elston (HE) method. Following this paper, it was suggested that the difference between the sib-pairs and maximum likelihood (ML) be used to estimate relevant genetic parameters (Kruglyak & Lander, 1995), the individual observations be used in a full ML analysis (Goldgar, 1990; Fulker & Cherny, 1996), the mean-corrected cross-product of the sib-pair phenotypes be used in a regression analysis (Drigalenko, 1998; Elston *et al.* 2000; Allison *et al.* 2000) – called the revisited or new Haseman–Elston method (NHE) or a combination of the mean-corrected squared sum and squared difference be used in a regression analysis (Drigalenko, 1998; Xu *et al.* 2000; Forrest, 2001; Sham & Purcell, 2001; Wang *et al.* 2001). The proposed methods have been tabulated in Table 1. In this study we will ignore standard parametric linkage analysis for quantitative traits, in which the relationship between genotype and phenotype is explicitly modelled in terms of the fixed effects of a finite number of unobserved alleles, because of the well-documented problems with model specification and computation (e.g. Amos & de Andrade, 2001, and references therein).

Currently it is difficult to compare the methods listed in Table 1 and suggested above, because the distribution of the test statistic under the null hypothesis (no linked QTL), and under the alternative hypothesis (linked QTL), has not been documented for all methods under similar models. Comparisons in the literature have been based on theoretical calculations of the expected test statistic under the alternative hypothesis of a linked QTL (Wright, 1997; Drigalenko, 1998; Sham *et al.* 2000; Rijsdijk *et al.* 2001; Sham & Purcell, 2001), and on simulation studies calculating the proportion of times the test statistic under the alternative hypothesis of a linked QTL exceeds a tabulated value from a known distribution (e.g. Amos *et al.* 1997; Elston *et al.* 2000).

A direct comparison has been hampered by:

(1) different assumptions about the genetic model. In simulation studies, a diallelic QTL is usually simulated, even when the assumptions for the method of analysis require bivariate or multivariate normality of QTL (e.g. Kruklyak & Lander, 1995; Fulker & Cherny, 1996; Wright, 1997; Forrest, 2001; Sham & Purcell, 2001).
(2) different parameters of the genetic model. For example, Elston *et al.* (2000), in their recent paper in which they proposed the mean-corrected sib-pair crossproduct, used a single set of parameters (a single diallelic additive QTL explaining 1/3 of the total variance, with no other sources of covariance between sibs) to determine power. Their analysis showed that using the cross-product was more efficient, and that the regression based upon cross-products appeared to be equivalent in power to a full bivariate maximum likelihood method. However, Xu *et al.* (2000) showed by simulation that in most cases the 'old' HE method was more powerful than the new HE method, but this was based on a relatively rare QTL with an allele frequency of 0.10 that explained 10–40% of total variance.
(3) investigating the distribution under the null hypothesis separately from the power calculation. For example, Allison *et al.* (1999, 2000) and Elston *et al.* (2000) addressed the distribution of the test statistic under the null hypothesis of no linked QTL, whereas Fulker & Cherny (1996), Wright (1997), Drigalenko (1998), Xu *et al.* (2000), Sham *et al.* (2000), Forrest (2001), Rijsdijk *et al.* (2001) and Sham & Purcell (2001) assumed that this distribution was known and therefore compared models based upon expected test statistic under the alternative hypothesis of a linked QTL.

In this study, we have compared all the proposed methods, both theoretically and by simulation, and calculated which methods are superior under which sets of population parameters. We have also provided accurate predictions of previously published results, and explained why some methods are superior to others. We have shown that an adaptation of the proposed method by Drigalenko (1998) and Xu *et al.* (2000) provides a simple new test which is as powerful as a full maximum likelihood approach. The proposed method is very similar to the weighted regression analysis recently proposed by Forrest (2001). To derive the predictions, the variances of the squared differences, squared sums and crossproducts of sib-pair phenotypes were derived under a fixed and random QTL model, and the expected test statistics under the null and alternative hypotheses were calculated.

MATERIALS AND METHODS

Assumptions and notation

Throughout, we assume (i) a single additive QTL, (ii) normally distributed polygenic, common environmental and residual effects, and (iii) fully informative markers to determine the proportion of alleles IBD at the QTL. Although we will present a simple derivation in the prediction of the powers of the statistical tests to take account of markers not being fully-informative, it is unlikely that there will be an interaction between the informativeness of markers and the relative statistical powers of the tests. For the additive QTL, we assume either a QTL with a normal distribution of allelic effects in the population, or a diallelic QTL.

At the population level, for sib j ($j = 1, 2$) in sib-pair i, we decompose the phenotype (Y) into causal components,

$$Y_{ij} = \mu_{ij} + Q_{ij} + F_i + R_{ij}, \tag{1}$$

where μ is the fixed effect of covariates, Q is the QTL effect, F is the family effect common to the sibs,

Table 2. *Probabilities of sib-pair genotypes for a diallelic QTL in Hardy–Weinberg equilibrium*[a]

Genotype		Conditional probability, given			Unconditional probability
Sib 1	Sib 2	$\pi = 0$	$\pi = \frac{1}{2}$	$\pi = 1$	
BB	BB	p^4	p^3	p^2	$\frac{1}{4}p^2(1+p)^2$
bb	bb	$(1-p)^4$	$(1-p)^3$	$(1-p)^2$	$\frac{1}{4}(1-p)^2(2-p)^2$
Bb	Bb	$4p^2(1-p)^2$	$p(1-p)$	$2p(1-p)$	$p(1-p)[1+p(1-p)]$
BB	Bb	$2p^3(1-p)$	$p^2(1-p)$	0	$\frac{1}{2}p^2(1-p^2)$
Bb	BB	$2p^3(1-p)$	$p^2(1-p)$	0	$\frac{1}{2}p^2(1-p^2)$
Bb	bb	$2p(1-p)^3$	$p(1-p)^2$	0	$\frac{1}{2}(1-p)^2[1-(1-p)^2]$
bb	Bb	$2p(1-p)^3$	$p(1-p)^2$	0	$\frac{1}{2}(1-p)^2[1-(1-p)^2]$
BB	bb	$p^2(1-p)^2$	0	0	$\frac{1}{4}p^2(1-p)^2$
bb	BB	$p^2(1-p)^2$	0	0	$\frac{1}{4}p^2(1-p)^2$
Sum		1	1	1	1

[a] Based upon Table I of Haseman & Elston (1972).

and R is a residual effect specific to each sibling. It is assumed that $\mathrm{E}(F) = \mathrm{E}(R) = 0$, so $\mu = \mathrm{E}(Y) - \mathrm{E}(Q)$. Family effects can be both genetic (due to variation at other loci) and due to a shared environment ''common'' to sibs in the same pair. Likewise, the residual effect consists of an environmental effect and a measurement error. This simple model leads to the following partition of variation,

$$\mathrm{var}(Y) = \mathrm{var}(Q) + \mathrm{var}(F) + \mathrm{var}(R). \tag{2}$$

Without loss of generality, for a given parameterisation of the mean μ, we scale the phenotypic standard deviation to unity, and define the variance ratios as, $q^2 + f^2 + r^2 = 1$. Hence, q^2 is the proportion of population variance of the trait, adjusted for given fixed effects, explained by the QTL.

QTL models

Two models for the QTL are used in this study, the random additive QTL model and the diallelic additive QTL model. For the random additive QTL model, the two QTL genotype effects of the parents of the two sibs are presumed to have been chosen at random and independently from a normal distribution of QTL genotype effects. Then, conditional on the proportion of marker alleles (and hence QTL alleles) that the sibs share IBD (π), and using standard normal theory, $Q_{i1} \sim \mathrm{N}(0, q^2)$, $Q_{i2}\,|\,\pi = \pi Q_{i1} + e_{i2}\,|\,\pi$, $e_{i2} \sim \mathrm{N}(0, q^2)$ and $\mathrm{var}(e_{i2}\,|\,\pi) = (1-\pi^2)q^2$. Hence, for the sibs sharing 0, 1, or 2 alleles IBD, corresponding to $\pi = 0, \frac{1}{2}, 1$, $Q_{i2}\,|\,\pi = 0 \sim \mathrm{N}(0, q^2)$, $Q_{i2}\,|\,\pi = \frac{1}{2} \sim \mathrm{N}(\frac{1}{2}Q_{i1}, \frac{3}{4}q^2)$, and $Q_{i2}\,|\,\pi = 1 \sim \mathrm{N}(Q_{i1}, 0)$. The joint distribution of the sibs' phenotypes in a population of sib-pairs is a mixture of three bivariate normal distributions, and not a single bivariate distribution as is commonly assumed.

For an additive diallelic QTL, suppose the three genotypes at the QTL are bb, Bb and BB, with means of $-\alpha$, 0, and $+\alpha$, respectively, and the frequency allele B is p. Assuming that genotypes are in Hardy–Weinberg equilibrium, $\mathrm{E}(Q) = (2p-1)\alpha$ and $\mathrm{var}(Q) = q^2 = 2p(1-p)\alpha^2$.

The distributions of the sibs' QTL genotypes, conditional on π, are more complicated than in the random QTL model because they depend on the genotypes of the parental mating type. Following Fisher (1918), they can be derived from the probabilities of the parental genotypes and are shown in Table 2. Except for the additional column of the unconditional probabilities, this table is the same as Table 1 of Haseman & Elston (1972).

Conditional on the proportion of QTL alleles shared IBD, the covariance between the traits on the sib-pairs is $\mathrm{cov}(Y_{i1}, Y_{i2}\,|\,\pi) = \mathrm{cov}_\pi = f^2 + \pi q^2$ (e.g. Fulker & Cherny, 1996). Because we have scaled the

Table 3. *Differences, sums and cross-products of sib genotypes for a diallelic QTL in Hardy–Weinberg equilibrium*

Genotype					
Sib 1	Sib 2	D/α[a]	S/α[b]	CP/α[c]	Probability
BB	BB	0	$-2x+2$	x^2-2x+1	$\frac{1}{4}p^2(1+p)^2$
bb	bb	0	$-2x-2$	x^2+2x+1	$\frac{1}{4}(1-p)^2(2-p)^2$
Bb	Bb	0	$-2x$	x^2	$p(1-p)[1+p(1-p)]$
BB	Bb	1	$-2x+1$	x^2-x	$\frac{1}{2}p^2(1-p^2)$
Bb	BB	-1	$-2x+1$	x^2-x	$\frac{1}{2}p^2(1-p^2)$
Bb	bb	1	$-2x-1$	x^2+x	$\frac{1}{2}(1-p)^2[1-(1-p)^2]$
bb	Bb	-1	$-2x-1$	x^2+x	$\frac{1}{2}(1-p)^2[1-(1-p)^2]$
BB	bb	2	$-2x$	x^2+1	$\frac{1}{4}p^2(1-p)^2$
bb	BB	-2	$-2x$	x^2+1	$\frac{1}{4}p^2(1-p)^2$

[a] $D=(Y_1-\mu)-(Y_2-\mu)$; $\mu=(2p-1)\alpha$; 2α is the difference between the BB and bb genotypes.
[b] $S=(Y_1-\mu)+(Y_2-\mu)$; $x=2p-1$.
[c] $CP=(Y_1-\mu)\,(Y_2-\mu)$.

total phenotypic variance to unity, the conditional covariances correspond to conditional correlations, $\rho_\pi=f^2+\pi q^2$. For an additive QTL model, the sib correlation for sibs sharing 1 allele IBD is equivalent to the population-wide correlation, $\rho_\pi=\rho=f^2+\frac{1}{2}q^2$ (see also Wright, 1997).

It is important to note that, even conditional on the number of alleles shared IBD, the distribution of a pair's phenotypes is *not* bivariate normal. There are nine different combinations of the three genotypic values for the sibs in the population $(-\alpha\ -\alpha;\ldots;\ 0\ 0;\ldots;\ \alpha\ \alpha)$, so the genotypic distribution is multinomial and not continuous, and therefore the distribution of the phenotypes is not (bivariate) normal. For a QTL with a small effect, however, the assumption of bivariate normality for each IBD class may be a good approximation of the actual distribution.

Regression methods

Where possible we will drop the subscript i which denotes the i-th sib-pair. Let Z be the difference, squared difference, mean corrected squared sum, or mean corrected cross-product. These variables are listed in Table 1. All proposed regression methods (Haseman & Elston, 1972; Drigalenko, 1998; Elston *et al.* 2000; Forrest 2001; Xu *et al.* 2000) are of the form

$$Z=\gamma+\beta\Pi+E, \tag{3}$$

where Π is the proportion of alleles shared IBD, and in full generality must be considered to be a random variable. These methods proceed by calculating the usual least squares estimator, $\hat{\beta}$, of β and applying tests for significance, see below. We note that if the joint distribution of Z and Π is bivariate normal, the exact variance of $\hat{\beta}$ is $\mathrm{var}(\hat{\beta})=\mathrm{var}(E)/[(n-3)\mathrm{var}(\Pi)]$ (e.g. Kendall & Stuart, 1977). However, the residual or error term, E, is not presumed to have a normal distribution and in most applications has a skewed distribution and heteroscedastic variance. For sib-pair i, we define a squared difference, a mean corrected squared sum, and and mean corrected cross-product as,

$$D_i^2=(Y_{i1}-Y_{i2})^2 \tag{4a}$$

$$S_i^2=[(Y_{i1}-\mu_{i1})+(Y_{i2}-\mu_{i2})]^2 \tag{4b}$$

$$CP_i=(Y_{i1}-\mu_{i1})(Y_{i2}-\mu_{i2}). \tag{4c}$$

Although we note that the effects of covariates of the sib-pairs (μ_{ij}) can be different, in this study we are concerned only with the variances and subsequently we will assume a single overall fixed effect (μ) in our derivations. In previous regression-based linkage analyses, the choice of Z has been either D^2 (Haseman & Elston, 1972), S^2, CP (Drigalenko, 1998; Elston *et al.* 2000), or a weighted average of D^2 and S^2 (Drigalenko, 1998; Xu *et al.* 2000; Forrest, 2001; Sham & Purcell, 2001). The value of β is a simple function of q^2, the proportion of variance in the population explained by the QTL, being $-2q^2$, $+2q^2$ and q^2 for D^2, S^2 and CP, respectively (Haseman & Elston, 1972; Drigalenko, 1998; Elston *et al.* 2000) The intercept γ is of no further interest.

To make a comparison between these regression methods, the variances of Z and E need to be derived, both under the null hypothesis of no QTL effects and under the alternative hypothesis of a linked QTL. The exact variance of Z is derived in Appendix I under the assumption of normality of common family effects and individual residual effect (i.e. of F_i and R_{ij}). For the full model (including a putative QTL),

$$\mathrm{var}(E) = \mathrm{var}(Z) - \beta^2 \mathrm{var}(\Pi). \tag{5}$$

The variance of Π is the same for all regression models, because it depends only on the observed genetic marker information of the sibs (and their parents). For a fully informative marker, the mean and variance of Π are $\frac{1}{2}$ and $\frac{1}{8}$, respectively (e.g. Rijsdijk *et al.* 2001). The usual way to test the significance of $\hat{\beta}$ is to calculate the ratio of the regression mean square to the residual mean square under the full model (e.g. Allison *et al.* 2000; Elston *et al.* 2000; Haseman & Elston (1972) used the square root of this test). The expectation of this test statistic, T, is

$$\begin{aligned}
\mathrm{E}(T) &\approx \mathrm{E}[\text{Mean Square Regression}]/\mathrm{E}[\text{Mean Square Error}] \\
&= [(n-1)\mathrm{var}(Z) - (n-2)\mathrm{var}(E)]/\mathrm{var}(E) \\
&= [\mathrm{var}(Z) + (n-2)\beta^2/8]/[\mathrm{var}(Z) - \beta^2/8] \\
&\approx 1 + [(n-2)\beta^2/8]/\mathrm{var}(Z).
\end{aligned} \tag{6}$$

If Z were to be normally distributed, then T is distributed as a central or non-central $F_{(1,n-2)}$ distribution, or, if n is large, a central or non-central $\chi^2_{(1)}$. However, in practice the test would be one-sided, because the alternative hypothesis is that β in Equation (3) is either negative (when using D^2) or positive (when using S^2 or CP). Hence, the test statistic is zero if the estimated regression coefficient is of the 'wrong' sign, and the asymptotic distribution of the test statistic under the null hypothesis of no linked QTL is $\frac{1}{2}\chi^2_{(1)}$ and $\frac{1}{2}\chi^2_{(0)}$. This suggests that for the different Z values, the means of the test statistic T pertaining to D^2, S^2 and CP are, respectively,

$$\mathrm{E}(T_{D2}) \approx (1 - \mathrm{Prob}(\hat{\beta}_{D2} > 0)) + \tfrac{1}{2}(n-2)q^4/\mathrm{var}(D^2) \tag{7a}$$

$$\mathrm{E}(T_{S2}) \approx (1 - \mathrm{Prob}(\hat{\beta}_{S2} < 0)) + \tfrac{1}{2}(n-2)q^4/\mathrm{var}(S^2) \tag{7b}$$

$$\mathrm{E}(T_{CP}) \approx (1 - \mathrm{Prob}(\hat{\beta}_{CP} < 0)) + \tfrac{1}{8}(n-2)q^4/\mathrm{var}(CP). \tag{7c}$$

Under the null hypothesis of no linked QTL, the $(1-\mathrm{Prob})$ terms are $\frac{1}{2}$. From these equations for the mean test statistic, and assuming that the distributions of these T statistics under the null hypothesis are approximately the same, it can be concluded that (i) a test based upon S^2 is always less powerful than a test based upon D^2, because the variance of S^2 is always larger than the variance of D^2 (see Appendix I, Forrest, 2001, and Sham & Purcell, 2001, and (ii) a test based upon CP is more powerful than a test based upon D^2 only if the $\mathrm{var}(CP) < \frac{1}{4}\mathrm{var}(D^2)$. In Appendix II we derive a quadratic equation which specifies when the power of a test based upon D^2 and CP is equal.

Several authors have used the classical result of statistics that D and S, and thus D^2 and S^2, are independent and therefore combining the information on the difference and the sum of the sib-pair phenotypes is more efficient than using the difference or sum (or their square) only (Wright, 1997; Drigalenko, 1998; Fulker & Cherny, 1996; Amos *et al.* 1997; Forrest 2001). Because D^2 and S^2 are independent, the expected correlation between the estimated regression coefficients from regression of D^2 and S^2 on π is also zero (Drigalenko, 1998). This suggests combining the two regression coefficients, by weighting them according to the inverse of their variance, to give $\hat{\beta}_{D^2,S^2} = \frac{1}{2}[(1-w_1)\ \hat{\beta}_{S^2} - w_1\hat{\beta}_{D^2}]$, with $w_1 = [1/\mathrm{var}(\hat{\beta}_{D^2})]/[1/\mathrm{var}(\hat{\beta}_{D^2}) + 1/\mathrm{var}(\hat{\beta}_{S^2})]$. The properties of this test are, $\beta_{D^2,S^2} = q^2$, $\mathrm{var}(\hat{\beta}_{D^2,S^2}) = \frac{1}{4}w_1\mathrm{var}(\hat{\beta}_{D^2})$ and, $T_{D^2,S^2} = (\hat{\beta}_{D^2,S^2})^2/\mathrm{var}(\hat{\beta}_{D^2,S^2})$. Using the previous results (Equations (5) to (7)), the expected value of this test statistic is,

$$\mathrm{E}(T_{D^2,S^2}) \approx (1-\mathrm{Prob}(\hat{\beta}_{D^2,S^2} < 0)) + \tfrac{1}{2}(n-2)q^4/\mathrm{var}(D^2) - \tfrac{1}{2}(n-2)q^4/\mathrm{var}(S^2). \quad (8)$$

By comparing equation (8) to equation (7), and noting that the new test is also based upon 1 degree of freedom, it is clear that the expected test statistic is always larger than the expected test statistic from using either D^2 or S^2. This argument is similar to that used by Wright (1997), who showed elegantly that, when using maximum likelihood, the expected LOD score from using the difference and the sum of the sibs' phenotypes is larger than the LOD score from using either, and the derivation of Drigalenko (1998), who showed that both D^2 and S^2 are a simple linear function of the proportion IBD. The test statistic from the regression models can also be expressed as a likelihood ratio test (LRT), $LRT \approx n\ln(1 + T\{df1/df2\})$, where *df1* and *df2* are the degrees of freedom in the numerator and denominator for test statistic T (e.g. Baret *et al.* 1998). For large n, and $df1 = 1$, $LRT \approx n\ln(1 + T/df2) \approx T$, and, $\mathrm{E}(LRT_{D^2,S^2}) \approx \mathrm{E}(LRT_{D^2}) + \mathrm{E}(LRT_{S^2}) - 1$, which is similar to Wright's (1997) derivation for the test statistic under maximum likelihood when using both the difference and sum of observations on sib-pairs. Note that Wright used ELOD to indicate the expected value of the LOD score, but that strictly speaking his ELOD are not expected values of the test statistic, because they ignore the expectation of the test statistic under the null hypothesis; ELOD are always greater than zero if the variance components are tested with a likelihood ratio test. In effect, the terms called ELOD are non-centrality parameters of the χ^2 distribution (Fulker & Cherny, 1996).

The suggested new and simple test based upon the two regression coefficients is very similar to the test proposed by Drigalenko (1998) and Xu *et al.* (2000). The difference is that we have used the fact that the difference and mean-corrected sum are uncorrelated, and that the correlation between the regression coefficients based upon these traits is zero. Drigalenko (1998) showed this independence, but assumed that the variance of the residuals from the regressions of D^2 and S^2 were the same. Xu *et al.* (2000) proposed to estimate the covariance between the regression coefficients from the data, which makes the method inexact, somewhat cumbersome, and not easy to use in standard regression software; this is not necessary. It is shown below that our simple test has, approximately, equivalent power to a full maximum likelihood approach.

Very recently, Forrest (2001) proposed a weighted regression analysis which is almost identical to the one proposed above. The only (slight) difference is that Forrest (2001) proposed to re-estimate the residual variances after estimating a single regression coefficient ($\hat{\beta}_{D^2,S^2}$) from a weighted analysis using the separate regressions of D^2 and S^2. This makes the method iterative, because the residual variances depend on the value of the regression coefficient and *vice versa*. Our method appears identical to that of Forrest (2001) if the iterations are stopped after one round. Forrest (2001) noted that only a few iterations are needed to achieve convergence, which suggests that the difference in inference based upon our proposed method and his iterative weighted analysis is likely to be small.

Maximum likelihood

Two maximum likelihood (ML) approaches have been proposed in the literature, one based upon the sib-pair difference (D) (Haseman & Elston, 1972; Kruglyak & Lander, 1995) and the other a full bivariate analysis of the individual observations on the sib-pairs (Hopper & Mathews, 1982; Goldgar, 1990; Amos, 1994; Fulker & Cherny, 1996; Wright, 1997; Sham *et al.* 2000; Rijsdijk *et al.* 2001). Although there is more information in the full bivariate model, as was shown clearly by Wright (1997) and Fulker & Cherny (1996), we have derived the expected test statistics for both approaches so that they can be compared to the test statistics from the regression analyses.

ML on sib-pair difference

In the case of fully informative markers, the likelihood function for the full model can be expressed as,

$$-2\ln(\mathrm{L_u}\,|\,D) = \Sigma n_\pi \ln(\sigma^2_\pi) + \Sigma SS_\pi/(\sigma^2_\pi), \tag{9}$$

where the summation is over the three IBD groups ($\pi = 0, \frac{1}{2}, 1$), and SS is the sum of squares of the sib-pair differences. If we do not put any (biologically plausible) restrictions on the variance components, the ML estimates from the complete (full) model are $\sigma^2_\pi = SS_\pi/n_\pi$. Under the alternative hypothesis of no QTL effect, i.e. for the reduced model, all three variances are equal, and the estimate of the population variance (of sib differences) is, $\sigma^2 = \Sigma SS_\pi/\Sigma n_\pi = SS/n$. Values of (minus twice) the maximum log-likelihood are obtained by substituting the estimates of the variance for the full and reduced model in Equation (9). To calculate the expected test statistic under the alternative hypothesis of a linked QTL, we assume a large number of sib-pairs, so that $n_\pi = \frac{1}{4}n$, $\frac{1}{2}n$ and $\frac{1}{4}n$, for $\pi = 0$, $\frac{1}{2}$ and 1, and that the expected value of the population variance under the reduced model is $\mathrm{E}(\sigma^2) = 2(r^2 + \frac{1}{2}q^2)$. The likelihood ratio test (LRT) is $T_{\mathrm{ML}(D)} = \{-2\ln(\mathrm{ML_{reduced}}) - 2\ln(\mathrm{ML_{full}})\}$, and its expectation is

$$\begin{aligned}
\mathrm{E}(T_{\mathrm{ML}(D)}) &= df + n\lambda \\
&= 2 + n\ln[(r^2 + \tfrac{1}{2}q^2)^{1/2}]/[(r^2)^{1/4}(r^2+q^2)^{1/4}] \\
&= 2 + n\ln[(1-\rho_{1/2})^{1/2}]/[(1-\rho_0)^{1/4}(1-\rho_1)^{1/4}],
\end{aligned} \tag{10}$$

where λ is the non-centrality parameter of a non-central χ^2 distribution with 2 degrees of freedom (df). If no restrictions are applied to the three variances in the full model, this test is simply a test for heterogeneous variances. The non-centrality parameter is identical to that used by Fulker & Cherny (1996). Wright (1997) used a similar derivation, but termed λ 'ELOD' (expectation of the LOD score), i.e. ignoring the degrees of freedom.

In practice, we would wish to impose restrictions on the three variances estimated under the full model to ensure the fitted model was biologically plausible (Kruglyak & Lander, 1995). For an additive model, the likelihood function for the full model is then $-2\ln(\mathrm{L_r}\,|\,D) = \Sigma n_\pi \ln[2(r^2 + (1-\pi)q^2)] + \Sigma SS_\pi/[2(r^2 + (1-\pi)q^2)]$. The non-centrality parameter is the same as before, but the LRT now has only 1 degree of freedom. In addition, because the alternative hypothesis proposes a positive QTL variance, the test is one-sided and the expected test statistic becomes $\mathrm{E}(T_{\mathrm{ML}(D)}) = (1 - \mathrm{Prob}(\mathrm{LRT} = 0)) + n\lambda$. Under the null hypothesis of no QTL variance, (i.e. $\lambda = 0$) the probability that the LRT is zero is $\frac{1}{2}$, for large n (Self & Liang, 1987; Almasy & Blangero, 1998; Baret *et al.* 1998). For a powerful experiment $\mathrm{Prob}(\mathrm{LRT} = 0) = 0$, the expected test statistic is $1 + n\lambda$, i.e. the same as in the unrestricted model, apart from the difference in degrees of freedom.

If both f^2 and q^2 are small relative to r^2, the expected test statistics from a regression analysis of D^2 (Equations (7)) and an ML analysis of D are identical (i.e. $\lambda = q^4/16$), which suggests that for large studies that aim to detect QTL effects of moderate size (say, $q^2 < 0.20$), there is no difference in power between a regression approach and a maximum likelihood approach based upon the difference of the sib-pair observations.

Bivariate analysis

We write the full likelihood function as

$$-2\ln(\mathrm{L}\,|\,y_{ij}) = \Sigma n_\pi \ln|V_\pi| + \Sigma[SS_\pi(f^2+q^2+r^2) - 2CP_\pi(f^2+\pi q^2)]/[|V_\pi|] \tag{11}$$

with $|V_\pi| = (f^2+q^2+r^2)^2 - (f^2+\pi q^2)^2$. SS_π and CP_π are the sum of the squares, $\Sigma\Sigma(y_{ij}-\mu)^2$ and sum of cross-products $\Sigma(y_{i1}-\mu)(y_{i2}-\mu)$ for the IBD group determined by π. For the reduced model ($q^2 = 0$), two parameters are estimated. Their ML estimates are, $f^2 = \Sigma CP_\pi/n$, and $r^2 = \Sigma[SS_\pi - 2CP_\pi]/(2n)$. Asymptotically, the resulting value of (minus twice) the maximum log-likelihood for the reduced model is $-2\ln(\mathrm{ML}_{\mathrm{reduced}}) = n[\ln(1-(f^2+\frac{1}{2}q^2)^2)+2]$.

For the full model, in which three variances are estimated, $-2\ln(\mathrm{ML}_{\mathrm{full}}) = n[\frac{1}{4}\ln(1-f^4)+\frac{1}{2}\ln(1-(f^2+\frac{1}{2}q^2)^2)+\frac{1}{4}\ln(1-(f^2+q^2)^2)+2]$. The resulting LRT has expectation, $\mathrm{E}(T_{\mathrm{ML}}) = z + n\lambda_{\mathrm{ML}}$, with z the expected value of the test statistic when q^2 is zero, and

$$\begin{aligned}\lambda_{\mathrm{ML}} &= \ln\{[1-(f^2+\tfrac{1}{2}q^2)^2]^{1/2}/[(1-f^4)^{1/4}(1-(f^2+q^2)^2)^{1/4}]\} \\ &= \ln\{(1-\rho_{1/2}^2)^{1/2}/[(1-\rho_0^2)^{1/4}(1-\rho_1^2)^{1/4}]\}.\end{aligned} \tag{12}$$

Wright (1997) and Fulker & Cherny (1996) derived the ratio of the non-centrality parameters from the full ML bivariate model (Equation (12)) and the ML model based upon the differences (Equation (10)). This ratio quantifies the extra information provided by the sum of the sib-pair observations. Assume that ρ_π^2 is small, so that $\ln(1-\rho_\pi^2) \approx -\rho_\pi^2$ gives a value of $\lambda_{\mathrm{ML}} = q^4/8$, i.e. twice the value of the parameter from using the sib-pair difference only (Sham *et al*. 2000). The same approximation is obtained for the new regression test based upon D^2 and S^2. Hence, this shows that, for a realistic range of parameters (say, $f^2+q^2 < 0.3$), the simple regression approach is as powerful as a full bivariate maximum likelihood method.

The distribution of the LRT under the null hypothesis is complicated, because it depends on the population value of f^2. If f^2 is not zero, then the estimate of the QTL variance is expected to be zero with a probability of $\frac{1}{2}$. However, if both f^2 and q^2 are zero, the LRT to test the QTL variance is expected to be zero with a probability $> \frac{1}{2}$, because the estimate of the QTL variance is zero if there is no evidence of a sibling covariance (i.e. estimate of f^2 from the reduced model is zero) or if there is no evidence of a QTL variance in addition to a familial correlation. Asymptotically, the probability of $\hat{f}^2 = 0$ is $\frac{1}{2}$ and the probability of $\hat{q}^2 = 0$ given that $\hat{f}^2 > 0$ is $(\frac{1}{2})(\frac{1}{2}) = \frac{1}{4}$, so that we expect the LRT for $q^2 = 0$ to be zero with a probability of $\frac{3}{4}$ and the expected LRT $= \frac{1}{4}(df)+\frac{3}{4}(0) = \frac{1}{4}$.

Simulation of test statistics and powers, and their predictions

All the theoretical predictions were validated by simulation, performed by programs written by the first author in Fortran and are available upon request. Phenotypes and the proportion of alleles IBD for sib-pairs were simulated according to the random or fixed QTL model from equation (1), using a normal distribution of family and residual effects. Each simulated data set on n sib-pairs was analysed using the four linear regression models (D^2, S^2, CP, D^2 & S^2) and three maximum likelihood

Table 4. *Comparison of simulated mean test statistics with theoretical values and proportions in the 5% and 1% tails, for different methods of linkage analysis and model parameters under the null hypothesis of no linked QTL*

n	q^2	f^2	Method[a]	T[b]	E(T)[c]	$\alpha_{0.05}$[d]	$\alpha_{0.01}$[d]
100	0.0	0.0	D^2	0.50	0.5	0.049	0.010
			S^2	0.50	0.5	0.048	0.009
			CP	0.52	0.5	0.052	0.011
			D^2 & S^2	0.50	0.5	0.048	0.009
			$ML_u(D)$	2.06	2	0.053	0.011
			$ML_r(D)$	0.52	0.5	0.051	0.012
			ML	0.21	0.25[e]	0.033	0.003
1000	0.0	0.0	D^2	0.50	0.5	0.051	0.010
			S^2	0.50	0.5	0.051	0.010
			CP	0.50	0.5	0.051	0.011
			D^2 & S^2	0.50	0.5	0.051	0.011
			$ML_u(D)$	1.98	2	0.049	0.010
			$ML_r(D)$	0.50	0.5	0.051	0.011
			ML	0.19	0.25	0.028	0.003
1000	0.0	0.25	D^2	0.49	0.5	0.051	0.009
			S^2	0.51	0.5	0.052	0.011
			CP	0.50	0.5	0.049	0.011
			D^2 & S^2	0.50	0.5	0.051	0.010
			$ML_u(D)$	1.98	2	0.049	0.011
			$ML_r(D)$	0.49	0.5	0.052	0.010
			ML	0.50	0.5[f]	0.052	0.011

[a] D^2, regression of squared differences on proportion IBD; S^2, regression of mean-corrected squared sum on proportion IBD; CP, regression of mean-corrected cross-product on proportion IBD; D^2 & S^2, weighted regression of D^2 and S^2 on the proportion IBD; $ML_u(D)$, maximum likelihood of sib-pair difference without a restriction of the parameter estimates; $ML_r(D)$, maximum likelihood of sib-pair difference, restricting estimates of the causal components to be postive; ML, full bivariate maximum likelihood.

[b] Mean test statistic from 10,000 simulations.

[c] Theoretical values from the main text.

[d] Empirical 5% and 1% type-I error rates from 10,000 simulations.

[e] Assuming $f^2 = 0$.

[f] Assuming $f^2 > 0$.

models (unconstrained ML on D, constrained ML on D, full bivariate ML). The maximum of the likelihood was found through a hill-climbing routine (Press *et al.* 1992). To investigate the behaviour of the test statistic under the null hypothesis of no linked QTL, and to calculate the mean test statistic and power in the presence of QTL effects, 10,000 replicates were simulated for each set of population parameters.

Calculations of the expected test statistics for each of the four regression models and three ML models were based upon the above derivations. Statistical power was calculated assuming that all tests followed a non-central χ^2 distribution, with the non-centrality parameter as calculated above. The thresholds used in these predictions were taken from the 1% and 5% tails of a central χ^2 distribution with 1 degree of freedom (all regression models, restricted ML using sib-pair differences, bivariate ML) or 2 degrees of freedom (unrestricted ML using sib-pair differences). For those tests based upon 1 degree of freedom, a one-sided test was assumed, so that the threshold corresponded to a test statistic which is a mixture of zero and $\chi^2_{(1)}$.

RESULTS

Data were generated under the null hypothesis of no QTL effects, and the mean test statistics over 10,000 replicates, and their predictions based upon the derived analytical expressions are presented in Table 4. In general, predictions were very close to the simulated values. The empirical 5% and

Table 5. *Comparison of simulated mean test statistics and power with theoretical values for different regression methods of linkage analysis*

n	q^{2}[a]	f^2	Method[b]	T[c]	E(T)[d]	Power(0.01)	E(Power(0.01))[d]
1000	0.2	0.0	D^2	4.0	4.0	0.29	0.28
			CP	5.7	5.7	0.44	0.44
			D^2 & S^2	5.8	5.9	0.45	0.45
		0.2	D^2	5.8	5.9	0.45	0.46
			CP	5.2	5.4	0.40	0.41
			D^2 & S^2	7.1	7.3	0.56	0.57
		0.4	D^2	10.2	10.3	0.78	0.77
			CP	4.8	4.8	0.36	0.36
			D^2 & S^2	11.2	11.4	0.83	0.81
	0.3	0.0	D^2	8.3	8.4	0.66	0.65
			CP	10.5	10.8	0.79	0.79
			D^2 & S^2	11.6	11.9	0.84	0.84
		0.2	D^2	13.0	13.2	0.88	0.88
			CP	10.0	10.0	0.76	0.75
			D^2 & S^2	15.7	15.8	0.94	0.94
		0.4	D^2	24.1	24.2	1.00	0.99
			CP	8.8	8.9	0.68	0.68
			D^2 & S^2	26.1	26.3	1.00	1.00

[a] Additive diallelic QTL with allele frequency of allele B of 0.10.
[b] See Table 4 for definition of methods.
[c] Test statistic, from 10,000 simulations.
[d] Theoretical values from the main text.

1 % type-I error rates, assuming that the tests were distributed proportional to a central χ^2, are also shown in Table 4. The SE of the mean test statistics were in the range of 0.01 to 0.02, and the SE of the proportions of test statistics in the 5 % and 1 % tails of the distribution were, approximately, 0.002 and 0.001, respectively. For all tests, except for the likelihood ratio test from the bivariate model when there was no family correlation ($f^2 = 0$), the observed type-I errors were remarkably close to the theoretical ones, given that the χ^2 distribution is an asymptotic result and requires normality of the residuals in the linear model. In the case of squared differences and sums and the cross-products, the Z values are clearly not normally distributed, but apparently the test is robust to these violations of assumptions, at least at the 5 % and 1 % tails of the distribution. When $f^2 = 0$, the mean of the LRT from the bivariate model is reasonably predicted (0.25 predicted versus ~ 0.2 observed), but the test, assuming $\frac{1}{4}\chi^2_{(1)}$ and $\frac{3}{4}\chi^2_{(0)}$, is too conservative. In the presence of a residual family correlation ($f^2 > 0$), the test statistic appears to be close to its approximation of $\frac{1}{2}\chi^2_{(1)}$ and $\frac{1}{2}\chi^2_{(0)}$.

Table 5 shows a comparison of the power of QTL detection for three regression models, assuming a diallelic QTL with a frequency of the increasing allele of 0.10. The parameters chosen were a subset of those considered by Xu *et al.* (2000), who used the z-statistic (the square root of our test statistic) in their simulations. The SEs of the mean test statistics were in the range of 0.04 to 0.09, and the SEs of the observed powers were in the range of 0 to 0.005. The theoretical values of the expected test statistic, E(T), and power, E(Power), were in excellent agreement with values observed from simulations. Clearly, our new method based upon a regression of both D^2 and S^2 is superior to both the old and new Haseman-Elston methods, especially when $f^2 > 0.2$, and the power of the new HE depends critically on the value of f^2, as predicted theoretically (see Appendix II). In Figure 1, the relative power of the old and new HE method is shown as a function of the population sib correlation and the proportion of variance explained by the QTL effect. Unless the population sib correlation is moderate to low (< 0.27), the old HE method has superior power to the new HE method.

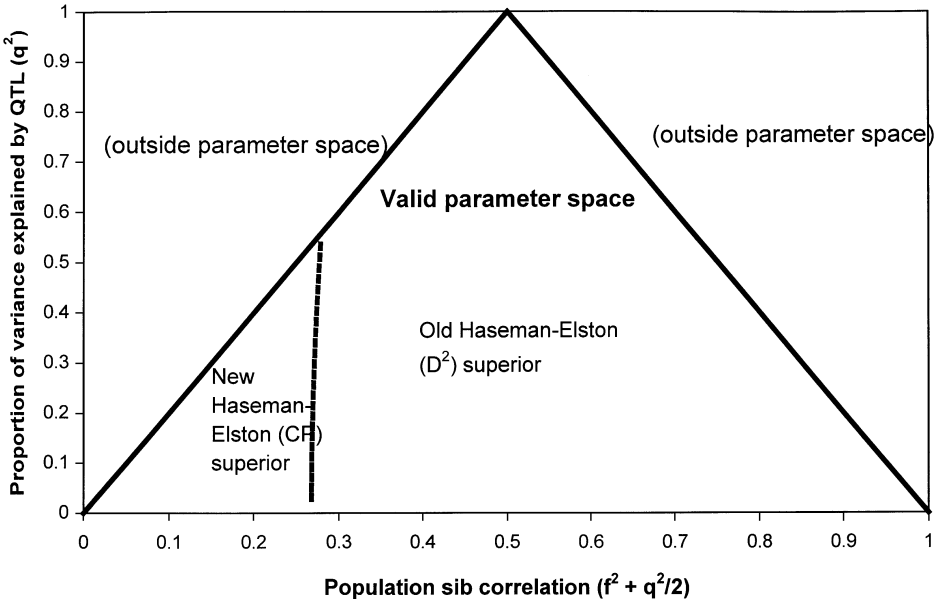

Figure 1. Relative power of the old and new Haseman–Elston method calculated from the theoretical predictions of the test statistics. The curve of equal expected test statistic was calculated in Appendix II for a random QTL.

Table 6. *Comparison of simulated mean test statistics and power with theoretical values for regression and maximum likelihood methods of linkage analysis*

n	q^{2a}	f^2	Method[b]	T[c]	E(T)[d]	Power(0.01)[e]	E(Power(0.01))[e]
100	0.5	0.0	Reg	4.9	4.9	0.36	0.37
			ML	4.3	5.0	0.32	0.37
		0.25	Reg	7.9	8.0	0.63	0.62
			ML	8.6	8.9	0.68	0.69
1000	0.1	0.0	Reg	2.2	2.3	0.06	0.11
			ML	1.7	2.3	0.11	0.11
		0.25	Reg	2.6	2.7	0.14	0.15
			ML	2.6	2.7	0.15	0.15
	0.2	0.0	Reg	6.2	6.2	0.48	0.48
			ML	5.7	6.2	0.45	0.48
		0.25	Reg	8.3	8.3	0.63	0.65
			ML	8.4	8.4	0.68	0.65

[a] Additive diallelic QTL with allele frequency of allele B of 0.50.
[b] Reg = regression method based upon D^2 and S^2 , ML = bivariate maximum likelihood method.
[c] Test statistic, from 10,000 simulations.
[d] Theoretical values, assuming that the probability of observing a zero test statistic was 0.
[e] Thresholds used for the observed and theoretical power calculations were from a $\frac{1}{2}\chi^2$ distribution.

In Table 6, a comparison is made between the regression method based upon D^2 and S^2, and a full bivariate maximum likelihood method. The SEs of the mean test statistics were in the range of 0.03 to 0.06, and the SEs of the powers in the range of 0.002 to 0.005. Theoretical predictions were close to the observed values from simulation, and the methods are, as anticipated, virtually identical in statistical power. The prediction of the test statistic and power was made assuming that the probability of observing a zero test statistic under the null hypothesis was $\frac{1}{2}$, and that the probability of observing a zero test statistic under the alternative hypothesis was 0. The predictions in Table 6 assume that the probability of observing a zero test statistic was negligible. However, for less

powerful designs (e.g. $n = 100$ and $q^2 < 0.03$) this probability can be relatively large (0.10–0.20), so that the average test statistic and power were lower than predicted. These designs were not further explored, because they correspond to low power (< 0.20 for a 1% type-I error rate) and are therefore not of practical importance.

The results from Tables 5 and 6 highlight the relatively low power of the sib-pair method to detect a QTL. For example, even with 1000 sib-pairs, the QTL needs to explain 20% or more of the total variance to have a reasonable chance of being detected.

DISCUSSION

We have shown, through analytical predictions and simulations studies, that for an additive QTL and a fully informative marker, it is relatively straightforward to predict the power of regression and maximum likelihood linkage approaches to map a QTL. Our results can therefore be used to predict the power of study designs without having to perform elaborate simulation studies. We have also shown that a simple combination of the regression coefficients obtained from regressions of the squared difference and of the squared sum of the sib-pair observations gives an expected test statistic with power which is essentially the same as that from a full maximum likelihood analysis.

Very recently, Sham & Purcell (2001) showed the asymptotic equivalence between the regression and maximum likelihood approach, by deriving the non-centrality parameter of a generalised likelihood-ratio test for both methods. Their derivations were based upon the assumption of a bivariate normal distribution of the sibs' trait values in the population, and, implicitly, on normality of D^2, S^2 and CP. Sham & Purcell (2001) confirmed the equivalence through a simulation study of 10000 sib-pairs in which a di-allelic QTL explained 5 or 10% of the total variance. From our study, the simulation results suggest that the assumption of normality of the Z values gives very good approximations of the expected value of the test statistic and statistical power, despite the fact that the assumption is per definition not correct. The assumption of a single bivariate normal distribution of sib trait values is equivalent to assuming that the QTL explains a small (strictly infinitessimal) proportion of total variance. In that case, our results in terms of the variances of D^2, S^2 and CP and the expected test statistic or non-centrality parameter (Equations 7–9 and Appendix I) are equivalent to those of Sham & Purcell (2001), who showed elegantly that the variances of D^2, S^2 and CP become $8(1-\rho)^2$, $8(1+\rho)^2$ and $(1+\rho^2)$, respectively. For a QTL of small effect, such as used in the simulation study of Sham & Purcell (2001), the approximations of the variances and of subsequent power are close to the true values. In the most extreme case where all variance is explained by a random QTL, the true variance of D^2 is 3.5 (Equation [A2]) whereas the approximation assuming bivariate normality gives a value of 2.0.

Assumptions

A number of assumptions have been made in this study, and the possible impact of these will be discussed in turn. At the individual's level, the random family and residual effects were assumed to be normally distributed. This is a fairly standard assumption. In practice, non-normal trait could be transformed to normality, or the statistical analysis could take account of the actual distribution of the data. However, following the results from mapping QTL in line-crosses (e.g., Visscher *et al.* 1996*a*), for the purposes of this study, i.e. QTL detection, it is doubtful if much power is gained by transforming the Y values or by performing a statistical analysis which properly takes account of the

distribution of the random effects. For the QTL, we have assumed either a fixed or a random effects model. In practice, the 'true' model will be somewhere between these extremes; perhaps a few alleles with large effects and many alleles with small effects. Our results show that if the QTL does not explain a large amount of variance, say, less than 20% of total variance, then the normal QTL model is a good approximation of the diallelic model.

For the regression-based analyses, we have calculated and tested for significance the least-squares estimator ($\hat{\beta}$) of β. In practice, other robust methods for estimating could also be used to protect against influential points and outliers, and add credibility to findings.

We have assumed an additive QTL model in our predictions and simulations. It would be relatively straightforward (but tedious) to incorporate dominance effects in the predictions by expanding Table 3. At present, no human QTLs have been cloned, and it is too early to be dogmatic about their gene action. In any case, even when no dominance effects are fitted in the model of analysis, an additive model would absorb some of the dominance variation. Variance component analysis on extended pedigrees usually fit additive QTL effects only (e.g. Almasy & Blangero, 1998).

We have assumed that the IBD status of sib-pairs can be assigned unambiguously, i.e. the markers are fully informative. It is simple to relax this assumption, because the only impact of uncertainty of IBD status is on the variance of Π in the population (Sham *et al.* 2000). The less marker information available, the more sib-pairs will be in or close to the group of $\pi = \frac{1}{2}$, because this is the expected value in the absence of any marker information. If P_{inf} is the proportion of sib-pairs where we can assign IBD status without error, and $(1-P_{inf})$ the proportion of sib-pairs with non-informative markers, then the three IBD groups are in proportion $\frac{1}{4}P_{inf}$, $1-\frac{1}{2}P_{inf}$ and $\frac{1}{4}P_{inf}$, and the resulting variance of the proportion IBD in the population is, $\mathrm{var}(\Pi) = P_{inf}/8$. These results can be used in, for example, Equations (5) and (6), to adjust predictions (e.g. Sham *et al.* 2000). Elston *et al.* (2000) and Fulker & Cherny (1996) considered examples of 10 cM spaced markers with 5 equifrequent alleles each, an additive QTL which explained $\frac{1}{3}$ of the total variance and 1000 sib-pairs. For a single marker, this results in a probability of informativeness of 0.576, and, assuming that at any point of the chromosome we have two closely spaced markers, we can approximate the overall probability of informativeness as $1-(1-0.576)^2 = 0.82$. For $P_{inf} = 0.82$, we predict an average test statistic of 8.5 when using D^2 and 11.4 when using CP. These values are very close to the simulated values derived by Fulker & Cherny (1996) and Elston *et al.* (2000).

When markers are not fully informative, methods have been proposed to take account properly of the uncertainty in estimating π (Kruglyak & Lander, 1995). However, several authors have shown, both in line crosses and for sib-pair designs, that the simple approach which uses the estimates of π as if they were known (the 'expectation approach') is as powerful as the complete 'distributional approach' (e.g. Haley & Knott, 1992; Fulker & Cherny, 1996; Gessler & Xu, 1996).

In our predictions for the regression models, we assume that the test statistic is distributed as a central χ^2 under the null hypothesis of no linked QTL. Our results suggest that the predictions are robust to violations of this assumption, at least at the 5% and 1% levels of significance. Elston *et al.* (2000) and Allison *et al.* (2000) investigated the behaviour of the test statistic under the null hypothesis in more detail, and noted departures in the extreme tails. In practice, it would be straightforward to set one's own null distribution, by using a permutation test in which IBD coefficients are permuted across sib pairs.

In both our predictions and simulations, we have ignored the estimation of the mean. We do not believe that this has influenced any of the results, because for a reasonable sample size ($n > 100$), the mean can be estimated with little error, and pre-correcting the observations for the overall mean only induces a very small correlation between residuals (Elston *et al.* 2000; Wang *et al.* 2001). Very recently, Wang *et al.* (2001) investigated the variance of cross-products of sib-pair values when the

product was uncorrected, corrected for the overall mean, or corrected for the family mean. Power of regression methods was parameterised as the ''residual ratio'', which is the ratio of unexplained to total variance of the dependent variable, i.e. similar to our use of the residual variance in Equation (5). However, Wang *et al.* (2001) derived their results from a one-way ANOVA between and within IBD groups, and their analysis is therefore based upon two degrees of freedom. The authors appear to suggest that a standard F-test is employed to test for linkage. This suggested test is equivalent to a test of heterogeneous variances between the three IBD groups, without taking into account the biological nature of the data, and, as the authors acknowledge, is less powerful than a regression approach (Wang *et al.* 2001). Notwithstanding the above, practitioners need to be aware that essentially different 'traits' are being considered for each set of covariates taken into account in 'adjusting' the mean, and these adjustments could have a major impact on the QTL effect (Hopper, 1993).

Comparison of predictions with published results

There have been a number of published results on sib-pairs from simulation studies which we can compare with our theoretical calculations. For example, in addition to the case of $n = 1000$, $q^2 = \frac{1}{3}$, $f^2 = 0$ (Fulker & Cherny, 1996; Wright, 1997; Elston *et al.* 2000), Xu *et al.* (2000) simulated various scenarios for an additive QTL with allele frequency of 0.10 and fully informative markers, for 1000 sib-pairs. Their results (expressed in the test statistic z, the square root of our T) are well predicted when using our predictions (Equations (7) and (8)). For example, for $q^2 = 0.3$ and $f^2 = 0.5$, we predict z-statistics of 6.0, 2.9 and 6.1, when using D^2, CP, and a combination of D^2 and S^2. From the graphs of Xu *et al.* (2000), the corresponding values from simulations were 5.6, 2.7, and 5.9. Our predicted (and simulated) test statistics are slightly larger, which may be because we have not estimated the sampling correlation between the two regression coefficients from the data, as Xu *et al.* (2000) proposed, but have used the correct value of zero correlation.

Twin studies

Dizygotic (DZ) twin pairs are a subset of the general population of sib-pairs and are particularly useful given that there is no need to adjust mean values for age and for sex if same-sex pairs only are used, so the methodology described in this study is appropriate for analysis of collections of DZ pairs. In addition, DZ twin pairs are likely to share more environmental effect than ordinary sib-pairs because they are matched for age, resulting in a larger between family component of variance, so that the proportion of within-family variance that is explained by the QTL, and hence power, may be larger for twin pairs. Schork & Xu (2000) recently suggested that the general framework of ML variance component analysis, which in the case of pair of sibs reduces to the bivariate analysis described in this study, is appropriate for QTL detection using both DZ and MZ pairs. However, MZ pairs do not add information on linkage, because they are identical at all genetic loci, including the QTL. One reason to include MZ pairs would be to get more accurate estimates of fixed effects when adjusting the mean for measured covariates, and to allow partitioning of family effects into a shared environmental and a (polygenic) genetic effect in a full likelihood analysis (Schork & Xu, 2000).

Conclusions

Which method to use in practice depends on a number of factors, but usually the availability of appropriate software is one consideration. If there are many covariates to be adjusted for in the sibs means, then a full maximum likelihood may be recommended. However, the full ML approach

assumes random samples and this may not be appropriate. All regression models are based upon functions of the individual sib-pair observations adjusted for the sib means. However, when there are few covariates to be fitted, the regression method based upon the regression coefficients obtained from the squared differences and squared mean-corrected sums could be used because of its simplicity and power. This analysis is easy to perform, using standard statistical software, and lends itself to computer intensive resampling schemes such as permutations tests (Churchill & Doerge, 1994) and bootstrapping (Visscher *et al.* 1996*b*).

If the choice is between the old and new Haseman-Elston method, then the preference depends on the (unknown) population parameters f^2 and q^2. Results from twin analyses of risk factors for common diseases suggest that the value of $(f^2+\frac{1}{2}q^2)$, i.e. the correlation between DZ twins, is rarely greater than 0.5 and more typically in the range of 0.2 to 0.4 (e.g. Poulsen *et al.* 1999; Harrap *et al.* 2000). Therefore, there is no strong preference for either method (see Figure 1 and Appendix II). One could estimate the sib correlation from the data (ignoring a putative QTL), and decide on the method to use. However, the simple weighted average of the regression coefficients from using D^2 and S^2 is, on average, more powerful.

Numerous studies, including the present one, have highlighted the need for large numbers of sib-pairs to increase the power of detection. From studies on experimental populations and outbred animal populations (e.g. Lynch & Walsh, 1998), it appears that the largest QTLs segregating in those populations explain up to $\sim 20\%$ of the phenotypic variance. If the effect of QTLs in human populations is similar, large numbers of sib-pairs (> 1000) are needed to map such loci. Power can be increased dramatically by including other relatives of the sibs for whom phenotype and marker information has also been collected (Duggirala *et al.* 1996; Almasy & Blangero, 1998).

This work was supported by the the UK Biotechnology and Biological Sciences Research Council, and by the National Health and Medical Research Council of Australia. We thank Chris Haley and Mike Goddard for helpful discussions, and two referees for useful comments.

APPENDIX I

Variances of functions of observations on sib-pairs

Under the assumption of a bivariate normal distribution of the random family and residual effects, we derive the mean and variance of the sib-pair squared difference, mean-corrected squared sum and crossproduct, for a fixed and random QTL. The results are exact, i.e. no approximations are involved. All derivations were validated by computer simulation.

We have used the following, for two traits Y_1 and Y_2 , with means μ_1 and μ_2, which are jointly distributed as bivariate normal, $\mathrm{E}(Y_i-\mu_i)^2 = \mathrm{var}(Y_i)$; $\mathrm{E}[(Y_1-\mu_1)(Y_2-\mu_2)] = \mathrm{cov}(Y_1, Y_2)$; $\mathrm{E}[(Y_1-\mu_1)(Y_2-\mu_2)^2] = 3\mathrm{cov}(Y_1, Y_2)\mathrm{var}(Y_2)$; $\mathrm{E}[(Y_1-\mu_1)^2(Y_2-\mu_2)] = 3\mathrm{cov}(Y_1, Y_2)\mathrm{var}(Y_1)$; $\mathrm{E}[(Y_1-\mu_1)^2(Y_2-\mu_2)^2] = \mathrm{var}(Y_1)\mathrm{var}(Y_2)+2\{\mathrm{cov}(Y_1, Y_2)\}^2$; $\mathrm{E}(Y_i-\mu_i)^4 = 3\{\mathrm{var}(Y_i)\}^2$ (see Kendall & Stuart, 1977). The general problem in working out the variances of the squared sums, squared differences and crossproducts is that, (i) the existence of a QTL makes the variances in the three IBD groups unequal, so that the population variance is a weighted average of the group variances, and (ii), for a diallelic QTL, even within an IBD group the distribution is not bivariate normal. The first issue was tackled by first deriving higher order terms for each of the three IBD groups, and then weighting these terms by the expected proportion in each class. The second issue was addressed by extending the table I of Haseman & Elston (1972) to higher order terms (see Table 3). For a normally distributed random QTL, we assume that $\mu_1 = \mu_2 = \mu$ and $\mathrm{var}(Y_1) = \mathrm{var}(Y_2) = \mathrm{var}(Y)$.

Variance of D^2

For the variance of the squared difference, we need to calculate,

$$\mathrm{var}(Y_1-Y_2)^2 = \mathrm{E}[(Y_1-\mu_1)-(Y_2-\mu_2)]^4-\{\mathrm{E}[(Y_1-\mu_1)-(Y_2-\mu_2)]^2\}^2. \quad \text{(A1)}$$

The second term has the same expectation for a random or fixed QTL, and is the variance of the difference, $\mathrm{E}[(Y_1-\mu_1)-(Y_2-\mu_2)]^2 = \mathrm{var}(D) = 2(1-(f^2+\frac{1}{2}q^2))$. For a normally distributed random QTL, using the above bivariate moments, it follows that for each of the three IBD groups, $\mathrm{E}((Y_1-Y_2)^4\,|\,\pi) = 12(\mathrm{var}(Y)-\mathrm{cov}_\pi)^2$; $\mathrm{E}((Y_1-Y_2)^2\,|\,\pi) = 2(\mathrm{var}(Y)-\mathrm{cov}_\pi) = \mathrm{var}(D\,|\,\pi)$, with $\mathrm{cov}_\pi = f^2+\pi q^2$. At the population level, the expectations of the unconditional moments are, $\mathrm{E}(Y_1-Y_2)^2 = \Sigma \mathrm{P}_\pi\{2(\mathrm{var}(Y)-\mathrm{cov}_\pi)\} = \mathrm{var}(D)$, and $\mathrm{E}(Y_1-Y_2)^4 = \Sigma \mathrm{P}_\pi\{12(\mathrm{var}(Y)-\mathrm{cov}_\pi)^2\}$, with $\mathrm{P}_\pi = \frac{1}{4}, \frac{1}{2}$ and $\frac{1}{4}$, for $\pi = 0, \frac{1}{2}$ and 1, i.e. the simple Mendelian proportions of sibs sharing 0, 1, or 2 alleles IBD. After some algebra the variance of the squared difference simplifies to,

$$\mathrm{var}(Y_1-Y_2)^2 = \mathrm{var}(D^2) = 8r^2(1-f^2)+(7/2)q^4. \quad \text{(A2)}$$

For a fixed diallelic QTL, the individual terms $\mathrm{E}(Y_1-Y_2)^4$ and $\mathrm{E}(Y_1-Y_2)^2$ were calculated using the results in Table 3. After some tedious algebra, it follows that, $\mathrm{E}(Q_1-Q_2)^4 = q^4(3/2+1/(2p(1-p)))$, $\mathrm{E}(Y_1-Y_2)^4 = \mathrm{E}(Q_1-Q_2)^4+12r^2(q^2+r^2)$, and

$$\mathrm{var}(Y_1-Y_2)^2 = 8r^2(1-f^2)+[\tfrac{1}{2}+1/\{2p(1-p)\}]q^4. \quad \text{(A3)}$$

Unless the proportion of variance explained by the QTL is large, equations (A2) and (A3) indicate that the variance of the squared of the sib-pair difference is similar for a random and fixed QTL.

Variance of S^2

For the variance of the squared mean-corrected sum,

$$\mathrm{var}((Y_1-\mu)+(Y_2-\mu))^2 = \mathrm{E}[(Y_1-\mu)+(Y_2-\mu)]^4-\{\mathrm{E}[\,(Y_1-\mu)+(Y_2-\mu)]^2\}^2. \quad \text{(A4)}$$

As with the squared difference, the second terms in [A4] are the same for random and fixed QTLs, and are, $\mathrm{E}[(Y_1-\mu_1)+(Y_2-\mu_2)]^2 = \mathrm{var}(S) = 2(1+f^2+\frac{1}{2}q^2)$. For the random QTL model, $\mathrm{E}((Y_1-\mu_1)+(Y_2-\mu_2)])^4 = \Sigma \mathrm{P}_\pi\{12(\mathrm{var}(Y)+\mathrm{cov}_\pi)^2\}$, and,

$$\mathrm{var}(S^2) = 8(1+f^2)(1+q^2+f^2)+(7/2)q^4. \quad \text{(A5)}$$

Comparing (A5) to (A2) shows that the variance of the squared sum is, as expected, always larger than the variance of the squared difference. For the fixed QTL, the relevant terms were calculated from Table III and the resulting variance can be expressed as

$$\mathrm{var}(S^2) = 8(1+f^2)(1+q^2+f^2)+\tfrac{1}{2}[9/\{p(1-p)\}-47]q^4. \quad \text{(A6)}$$

The second terms of (A5) and (A6) are equal for an allele frequency of $p = \frac{1}{2}\pm\sqrt{3}/6$, i.e. for $p = 0.2113$ or $p = 0.7887$. The largest difference between (A5) and (A6) is obtained for $p = \frac{1}{2}$. However, unless the QTL has a large effect, the variance of the mean-corrected squared sum of the sib-pair observations is similar for a random and fixed QTL model.

Variance of CP

For the variance of the squared mean-corrected cross-product,

$$\mathrm{var}((Y_1-\mu)(Y_2-\mu)) = \mathrm{E}[(Y_1-\mu)(Y_2-\mu)]^2-\{\mathrm{E}[(Y_1-\mu)(Y_2-\mu)]\}^2. \quad \text{(A7)}$$

As with the squared difference, the second terms in [A7] are the same for random and fixed QTLs, and are, $E[(Y_1-\mu_1)(Y_2-\mu_2)] = \mathrm{cov}(Y_1, Y_2) = f^2+\frac{1}{2}q^2$. For the random QTL, we use from the bivariate normal distribution that, $E[(Y_1-\mu)^2(Y_2-\mu)^2 \mid \pi)] = (f^2+q^2+r^2)^2+2(\mathrm{cov}_\pi)^2$, which, after taking the weighted average over the three IBD groups and subtracting the square of the covariance term, results in,

$$\mathrm{var}(CP) = 1+f^2(1-r^2)+\tfrac{1}{2}q^4. \tag{A8}$$

For the fixed QTL, the entries in Table 3 were used, giving,

$$\mathrm{var}(CP) = 1+f^2(1-r^2)+\tfrac{1}{2}[(2p-1)^2/(2p(1-p))]q^4. \tag{A9}$$

Contrasting (A8) and (A9) results in the same values for the allele frequency to obtain equal variances of the cross-products ($p = \frac{1}{2}\pm\sqrt{3}/6$), and the same value ($p = \frac{1}{2}$) to obtain their largest difference.

APPENDIX II

Old versus New Haseman–Elston method

From the results in the main text (Equations (5) to (7)), and the variances of squared sib-pair differences and mean-corrected crossproducts (Appendix I, Equations (A2) and (A8)), we can derive when the test statistic under the new HE approach (i.e. regression of CP on proportion IBD) is larger than the test statistic under the old HE method (regression of D^2 on proportion IBD),

$$E(T_{D^2}) = E(T_{CP}) \Leftrightarrow [8r^2(1-f^2)+(7/2)q^4] = 4[1+f^2(1-r^2)+\tfrac{1}{2}q^4].$$

When parameterised in q^2 and f^2, this results in

$$q^2 = (4/3)(2-f^2)-(2/3)[4f^2(4-f^2)+20]^{1/2}. \tag{B1}$$

Figure 1 shows this equality, and highlights the areas for which using the trait CP is superior to using D^2. For $q^2 \to 0, f^2 = 2-\sqrt{3} = 0.27$ (see also Sham & Purcell, 2001). For ease of comparison, the values on the x-axis are in $(f^2+\frac{1}{2}q^2)$, which is the population sib correlation under an additive QTL model, and also the correlation between dizygotic (r_{DZ}) twin pairs.

REFERENCES

Allison, D. B., Neale, M. C., Zannolli, R., Schork, N. J., Amos, C. I. & Blangero, J. (1999). Testing the robustness of the likelihood-ratio test in a variance-component quantitative-trait loci-mapping procedure. *Am. J. Hum. Genet.* **65**, 531–544.

Allison, D. B., Fermandex, J. R., Heo, M. & Beasley, T. M. (2000). Testing the robustness of the new Haseman–Elston quantitative-trait loci-mapping procedure. *Am. J. Hum. Genet.* **67**, 249–252.

Almasy, L. & Blangero, J. (1998). Multipoint quantitative-trait linkage analysis in general pedigrees. *Am. J. Hum. Genet.* **62**, 1198–1211.

Amos, C. I. (1994). Robust variance-components approach for assessing genetic linkage in pedigrees. *Am. J. Hum. Genet.* **54**, 535–543.

Amos, C. I., Krushkal, J., Thiel, T. J., Young, A., Zhu, D. K., Boerwinkle, E. & de Andrade M. (1997). Comparison of model-free linkage mapping strategies for the study of a complex trait. *Genet. Epidemiol.* **14**, 743–748.

Amos, C. I. & de Andrade M. (2001). Genetic linkage methods for quantitative traits. *Stat. Meth. Med. Res.* **10**, 3–25.

Baret, P. V., Knott, S. A., & Visscher, P. M. (1998). On the use of regression and maximum likelihood for QTL mapping in halfsib designs. *Genet. Res.* **72**, 149–158.

Churchill, G. A. & Doerge, R. W. (1994). Empirical threshold values for quantitative trait mapping. *Genetics* **138**, 963–971.

Drigalenko, E. (1998). How sib pairs reveal linkage. *Am. J. Hum. Genet.* **63**, 1242–1245.
Duggirala, R., Stern, M. P., Mitchell, B. D., Reinhart, L. J., Shipman, P. A., Uresandi, O. C., Chung, W. K., Leibel, R. L., Hales, C. N., O'Connell, P. & Blangero, J. (1996). Quantitative variation in obesity-related traits and insulin precursors linked to the OB gene region on human chromosome 7. *Am. J. Hum. Genet.* **59**, 694–703.
Elston, R. C., Buxbaum, S., Jacobs, K. B. & Olson, J. M. (2000). Haseman & Elston revisited. *Genet. Epidemiol.* **19**, 1–17.
Fisher, R. A. (1918). The correlation between relatives on the supposition of Mendelian inheritance. *Trans. Royal Soc. Edin.* **52**, 399–433.
Forrest, W. F. (2001). Weighting improves the 'New Haseman-Elston' method. *Hum. Hered.* **52**, 47–54.
Fulker, D. W. & Cherny, S. S. (1996). An improved multipoint sib-pair analysis of quantitative trait loci. *Behav. Genet.* **26**, 527–532.
Gessler, D. G. & Xu, S. (1996). Using the expectation or the distribution of the identity by descent for mapping quantitative trait loci under the random model. *Am. J. Hum. Genet.* **59**, 1382–1390.
Goldgar, D. E. (1990). Multipoint analysis of human quantitative variation. *Am. J. Hum. Genet.* **47**, 957–967.
Harrap, S. B., Stebbing, M., Hopper, J. L., Hoang, H. N. & Giles, G. G. (2000). Familial patterns of covariation for cardiovascular risk factors in adults. *Am. J. Epidemiol.* **152**, 704–715.
Haseman, J. K. & Elston, R. C. (1972). The investigation of linkage between a quantitative trait and a marker locus. *Behav. Genet.* **2**, 3–19.
Hopper, J. L. (1993). Variance components for statistical genetics, applications in medical research to characteristics related to human diseases and health. *Stat. Meth. Med. Res.* **2**, 199–223.
Hopper, J. L. & Mathews, J. D. (1982). Extensions to multivariate normal models for pedigree analysis. *Ann. Hum. Genet.* **46**, 373–383.
Kendall, M. & Stuart, A. (1977). *The advanced theory of statistics. Vol. 1. Distribution theory.* 4th edition. New York: Macmillan.
Kruglyak, L. & Lander, E. S. (1995). Complete multipoint sib-pair analysis of qualitative and quantitative traits. *Am. J. Hum. Genet.* **57**, 439–454.
Lynch, M. & Walsh, B. (1998). *Genetic and analysis of quantitative traits.* Sunderland, MA: Sinauer Associates, Inc.
Sham, P. C. & Purcell, S. (2001). Equivalence between Haseman-Elston and variance-components linkage analyses for sib pairs. *Am. J. Hum. Genet.* **68**, 1527–1532.
Schork, N. J. & Xu, X. (2000). The use of twins in quantitative trait locus mapping. In *Advances in twin and sib-pair analysis* (eds. J. T. Spector, H. Snieder, A. J. MacGregor), pp. 190–202. London: Greenwich Medical Media Ltd.
Self, S. G. & Liang, K.-Y. (1987). Asymptotic properties of maximum likelihood estimators and likelihood ratio test under non-standard conditions. *J. Am. Stat. Assoc.* **82**, 605–610.
Sham, P. C., Cherny, S. S., Purcell, S. & Hewitt, J. K. (2000). Power of linkage versus association analysis of quantitative traits, by use of variance-components models, for sibship data. *Am. J. Hum. Genet.* **66**, 1616–1630. .
Poulsen, P., Ohm Kyvik, K., Vaag, A. & Beck-Nielsen, H. (1999). Heritability of Type II. non-insulin-dependent, diabetes mellitus and abnormal glucose tolerance – a population-based twin study. *Diabetologia* **42**, 139–145.
Press, W. H., Teukolsky, S. A., Vetterling, W. T. & Flannery, B. P. (1992). *Numerical recipes in Fortran 77, the art of scientific computing.* 2nd edition. Cambridge: Cambridge University Press.
Rijsdijk, F. V., Hewitt, J. K. & Sham, P. C. (2001). Analytical power calculation for QTL linkage analysis of small pedigrees. *Eur. J. Hum. Genet.* **9**, 335–340.
Visscher, P. M., Haley, C. S. & Knott, S. A. (1996*a*). Mapping QTLs for binary traits in backcross and F_2 populations. *Genet. Res.* **68**, 55–63.
Visscher, P. M., Thompson, R. & Haley, C. S. (1996*b*). Confidence intervals for QTL locations using bootstrapping. *Genetics* **143**, 1013–1020.
Wang, D., Lin, S., Cheng, R., Gao, X. & Wright, F. A. (2001). Transformation of sib-pair values for the Haseman–Elston method. *Am. J. Hum. Genet.* **68**, 1238–1249.
Wright, F. A. (1997). The phenotypic difference discards sib-pair QTL linkage information. *Am. J. Hum. Genet.* **60**, 740–742.
Xu, X., Weiss, S., Xu, X.-P. & Wei, L. J. (2000). A unified Haseman–Elston method for testing linkage with quantitative traits. *Am. J. Hum. Genet.* **67**, 1025–1028.