Multiple Regression Analysis of Sib-Pair Data on Reading to Detect Quantitative Trait Loci

D. W. FULKER,[1] L. R. CARDON,[1] J. C. DEFRIES,[1]
W. J. KIMBERLING,[3] B. F. PENNINGTON [2] and S. D. SMITH[3]

[1] *Institute for Behavioral Genetics, University of Colorado, Boulder, CO 80309* [2] *Department of Psychology, University of Denver, Denver, CO 80208* [3] *Boys Town National Institute for Communication Disorders in Children, Omaha, NE 68010*

ABSTRACT: A simple extension of the DeFries and Fulker multiple regression model for twin analysis is applied to the problem of detecting linkage in a quantitative trait. The method, employing sib pairs, is based on that of Haseman and Elston. Reading data from 19 extended pedigrees were analyzed employing RLFPs as markers on chromosome 15 and using the widely available statistical applications software package, SAS. A number of possible linkages were detected, indicating that this approach is both powerful and effective, especially in the case of selected samples. Detecting genotype–environment interaction and the issue of power are briefly discussed. The programs used are available upon request.

KEYWORDS: multiple regression, quantitative trait loci, sib-pair data, reading disability, linkage analysis.

INTRODUCTION

This paper describes a simple application of the DeFries and Fulker (1985, 1988) multiple regression analysis of twin data to the problem of detecting linkage in a quantitative trait. It combines the regression approach with that of Haseman and Elston (1972), which uses sibling data on the trait together with information on identity by descent (ibd) for marker loci to which the quantitative trait loci (QTLs) may be genetically linked. While the approach we suggest is not a radical departure from that of Haseman and Elston, we believe it offers a number of advantages over their approach.

Firstly, our approach is conceptually very straightforward. Secondly, it is simple to apply, only requiring one of the widely used statistical packages such as SAS or SPSS. Thirdly, it is equally applicable to selected samples as well as unselected samples, providing a simple unified approach to sibling linkage analysis. Fourthly, it is very flexible, permitting the evaluation of variables that may interact with QTL genotypic effects such as sex or age. And fifthly, it appears to be more statistically powerful than the standard Haseman and Elston approach, particularly when applied to selected samples.

Reading and Writing: An Interdisciplinary Journal **3**: 299–313, 1991.

While other approaches may be more appropriate in some specific situations we believe the features listed above make the regression approach extremely useful for those who need a simple straightforward method for detecting polygenes or QTLs in order to undertake exploratory data analysis or rapid screening of genetic markers.

In this paper the method is outlined and then illustrated using sibling data on reading performance as a quantitative trait and restriction fragment length polymorphisms (RFLPs) on chromosome 15 as genetic markers.

REGRESSION MODEL

The regression approach to the analysis of twin data exists in two forms, one employing the basic model and the other the augmented model. The basic model is intended for use with selected samples and involves the idea that differential regression to the population mean of MZ and DZ cotwins of selected probands indicates the presence of heritable variation. The cotwin score on a quantitative trait (C) is entered into the analysis as the dependent variable in a regression equation with the proband score (P) and the coefficient of relationship (R), which takes values 1.0 for MZ pairs and 1/2 for DZ pairs, entered as independent variables. The regression equation, including the constant term A, is given below.

$$C = B_1P + B_2R + A \tag{1}$$

The B_1 regression weight adjusts the cotwin score for average twin resemblance while the B_2 term measures the extent of differential regression of cotwins' means back towards the mean of the population. A significant B_2 term indicates the presence of a heritable component in the proband mean. The method has the advantage of increasing statistical power as a function of the degree of selection imposed on the probands, thus requiring fewer and fewer twin pairs the more intense the selection becomes.

The underlying principle of the method may be seen in Figure 1, in which the distribution at the top of the figure represents that of the base population with a selected group of low scoring individuals at the tail of the distribution labeled probands. The distributions of the MZ and DZ cotwins are shown below and can be seen to have means regressing back towards that of the population, μ, but more so for DZ than for MZ pairs. It can be shown (DeFries and Fulker, 1988) that B_1 is an estimate of the average twin correlation and that $B_2 = 2[(\overline{C}_{MZ} - \overline{C}_{DZ}) - B_1 (\overline{P}_M - \overline{P}_{DZ})]$ or just $2(\overline{C}_{MZ} - \overline{C}_{DZ})$ when the two kinds of probands have the same mean, which is usually expected to be the case. This regression coefficient, when divided by the selection differential, $\overline{P} - \mu$, is an estimate of the herit-

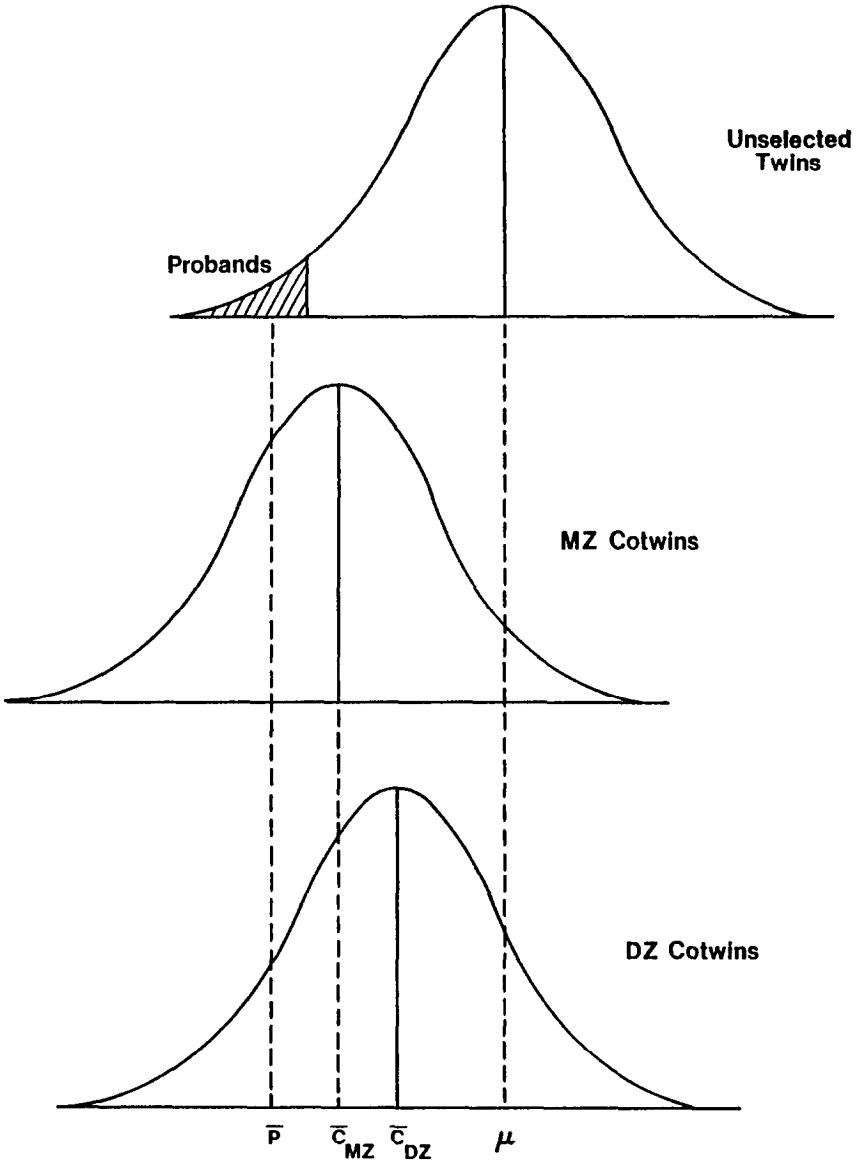

Fig. 1. Hypothetical distributions for reading performance of an unselected sample of twins and of the identical (MZ) and fraternal (DZ) cotwins of probands with a reading disability. The differential regression of the MZ and DZ cotwin means toward the mean of the unselected population (μ) provides a test of genetic etiology. [From DeFries, Fulker, and LaBuda (1987). Reprinted by permission from *Nature*, Vol. 329, p. 537. Copyright © 1987, Macmillan Magazines Ltd.]

ability of the proband deficit, h_g^2. If this deficit is due to the same factors that cause individual differences in the general population, then it is also an estimate of heritability (h^2) in the population as a whole.

In the augmented model a fourth term, PR, the product of proband score and the coefficient of relationship, is added to equation (1) to give (2).

$$C = B_3P + B_4R + B_5PR + A \tag{2}$$

In this form the coefficient B_5 is a direct estimate of h^2 and B_3 is a direct estimate of shared environmental variance c^2. When scores are expressed as deviations from the population mean (μ) and divided by $\overline{P} - \mu$, the coefficient B_4 estimates $h_g^2 - h^2$. The addition of the product PR assesses differential twin resemblances as a function of zygosity, which is the basis for inferring h^2 from the twin design and is no more than $2(B_{MZ} - B_{DZ})$, or $2(R_{MZ} - R_{DZ})$, where B and R are simple twin regressions and correlations calculated from a random sample of twins, which is a standard way to estimate h^2.

The advantages of the regression method over the evaluation of correlations is its convenience for those not familiar with model fitting procedures, its flexibility for testing for interactions with other variable such as gender (Cyphers et al., 1990), and the fact that by using regression rather than correlations a correction for selection on the probands is applied. Thus, the augmented model is applicable to the analysis of continuous variation in both unselected and selected samples.

HASEMAN AND ELSTON LINKAGE MODEL

The Haseman and Elston (1972) approach to linkage uses information on marker loci in siblings and their parents to determine the proportion of alleles two siblings share ibd. This number, which they call π, can only take values zero, one half, or one, corresponding to ibd status zero, one, and two, and indicates how closely the pair resemble each other genetically at this locus. With a value of zero siblings are no more alike than totally unrelated individuals, with a value of one half they are as alike as ordinary siblings are on average for any locus, and with a value of one they are identical just like identical twins.

Thus, if a QTL is at that locus or closely linked to it, the three kinds of siblings should show differential resemblance, those with π equal to one being more alike than those with π equal to one half, who in turn should be more alike than those pairs with π equal to zero. Haseman and Elston use the sib pair difference squared (Y) as a measure of resemblance. Since this is technically twice the within pair variance for each pair of sibs, the

three values of π should relate inversely to this measure of resemblance if there is any linkage. Consequently the regression of Y on π is expected to be negative if the marker is near a QTL influencing the phenotype. They show that the regression coefficient will be equal to the additive genetic variance of the QTL multiplied by $-2(1 - 2\theta)^2$, where θ is the recombination fraction between the marker locus and the QTL ($0 \leqslant \theta \leqslant 0.5$). Therefore, when θ is zero the regression will detect all the genetic variance due to the QTL; however, when θ is as much as 0.12, or approximately twelve centimorgans away from the marker, only half the genetic variance of the QTL will be detected.

The ideal situation for the application of their approach is when the marker locus is completely informative regarding the ibd status of the sibs. For two allele markers, which are the most common, this is seldom the case and the method was initially criticized on these grounds (Robertson, 1973). However, with the advent of more recent molecular markers, which are increasingly polymorphic, the method has become much more promising. Nance and Neale (1989), who recently modified the Haseman and Elston approach for use with twin data using LISREL, illustrate this fact with a table of parental genotypes involving a four allele system in which the parents are both heterozygous for different alleles. That is, one parent is A_1A_2 and the other A_3A_4. Under these conditions ibd status is clearly unambiguous since it is obvious from which parent the alleles came and whether or not they are the same. Table 1, modified from their paper, illustrates the point.

However, not all markers are so informative; indeed, some are totally uninformative and Haseman and Elston introduced an ingenious refinement into their method to account for such markers. When they cannot

Table 1. Number of Alleles ibd for Marker Locus

		Parents' Genotypes $A_1A_2 \times A_3A_4$			
		Sib 1			
		A_1A_3	A_2A_3	A_1A_4	A_2A_4
Sib 2	A_1A_3	2	1	1	0
	A_2A_3	1	2	0	1
	A_1A_4	1	0	2	1
	A_2A_4	0	1	1	2

# *IBD*	π
0	0
1	1/2
2	1

determine π unambiguously from the markers, they estimate it ($\hat{\pi}$) instead. The method is shown in Table 2 taken from their paper and it involves forming a weighted average of the probabilities that 0, 1, or 2 sibling alleles are identical by descent given both sibling and the parental genotypes. These probabilities they denote as f_{j0}, f_{j1}, and f_{j2}, respectively, for the j^{th} locus. The estimate of $\hat{\pi}$ is then simply $1/2f_{j1} + f_{j2}$. It can be seen from the table that the estimates ($\hat{\pi}$) are the same as π when sufficiently informative markers are available, as in those cases where only one value of f_j appears in any given row. From this table any value of $\hat{\pi}$ can be obtained by simply locating the appropriate combination of parental and sibling genotypes from among the 34 possibilities. The original paper proves that $\hat{\pi}$ provides an unbiased estimate of π. The computer program SIBPAL carries out the Haseman and Elston analysis and is available commercially.

COMBINED MODEL

The essence of the approach we suggest is simply to replace the coefficient of relationship in the DeFries and Fulker regression approach with the

Table 2. $\hat{\pi}_j$ When Both Parental and Sib Genotypes Are Known

Mating type	Sib-pair type	f_{j0}	f_{j1}	f_{j2}	$\hat{\pi}_j$
I: $A_iA_i \times A_iA_i$	I: $A_iA_i - A_iA_i$	1/4	1/2	1/4	1/2
II: $A_iA_i \times A_jA_j$	V: $A_iA_j - A_iA_j$	1/4	1/2	1/4	1/2
III: $A_iA_i \times A_iA_j$	I: $A_iA_i - A_iA_i$	0	1/2	1/2	3/4
	III: $A_iA_i - A_iA_j$	1/2	1/2	0	1/4
	V: $A_iA_j - A_iA_j$	0	1/2	1/2	3/4
IV: $A_iA_i \times A_jA_k$	V: (2)	0	1/2	1/2	3/4
	VI: $A_iA_j - A_iA_k$	1/2	1/2	0	1/4
V: $A_iA_j \times A_iA_j$	I: (2)	0	0	1	1
	II: $A_iA_i - A_jA_j$	1	0	0	0
	III: (2)	0	1	0	1/2
	V: $A_iA_j - A_iA_j$	1/2	0	1/2	1/2
VI: $A_iA_j \times A_iA_k$	I: $A_iA_i - A_iA_i$	0	0	1	1
	III: (2)	0	1	0	1/2
	IV: $A_iA_i - A_jA_k$	1	0	0	0
	V: (3)	0	0	1	1
	VI: $A_iA_j - A_iA_k$	1	0	0	0
	VI: $A_iA_j - A_jA_k$	0	1	0	1/2
	VI: $A_iA_k - A_jA_k$	0	1	0	1/2
VII: $A_iA_j \times A_kA_l$	V: (4)	0	0	1	1
	VI: (4)	0	1	0	1/2
	VII: (2)	1	0	0	0

value of $\hat{\pi}$, and employ data on pairs of sibs in the place of those on twins. Thus, the basic model becomes

$$C = B_1P + B_2\hat{\pi} + A \tag{3}$$

and the augmented model becomes

$$C = B_3P + B_4\hat{\pi} + B_5P\hat{\pi} + A \tag{4}$$

Since $\hat{\pi}$ performs the same function as R in modelling sib resemblance in terms of additive genetic variance — not of the whole genome but for the QTL associated with specific marker locus in question — the regression coefficients B_2 and B_5 test for h_g^2 and h^2 of the QTL when the linkage is complete. When the QTL is θ centimorgans from the marker, these heritabilities are reduced by a factor $(1 - 2\theta)^2$, as shown in Haseman and Elston's paper.

The coefficient B_1 in the basic model provides a measure of average sib resemblance. However, B_3 in the augmented model assesses the average sib resemblance due to all sources of variation, both genetic and environmental, other than that due to the QTL. Since this resemblance is often substantial, the control for this source of variation in the regression approach should add power to that of Haseman and Elston.

Although the sample size in the present analysis is not sufficient to allow tests of gender X genotype interaction, it is of interest to show how simple it is to incorporate interactions into the model. If the main effect of gender (or any other main effect such as remediation) is designated S and introduced into the basic model, then three more regression coefficients are required.

$$C = B_1P + B_2\hat{\pi} + B_{10}S + B_{11}SP + B_{12}S\hat{\pi} + A \tag{5}$$

In the augmented model four more regression coefficients are required.

$$C = B_3P + B_4\hat{\pi} + B_5P\hat{\pi} + B_6S + B_7SP + B_8S\hat{\pi} + B_9SP\hat{\pi} + A \tag{6}$$

The terms B_{12} in the basic model and B_9 in the augmented model test for genotype X gender interaction.

METHODS

Subjects

The subjects used to illustrate the method are members of 19 three-

generation families with a history of reading disability that have been the subject of a series of linkage studies which started with nine families and a chromosomal marker on 15 (Smith et al., 1983). Since that time families have been added to the study and 9 RFLP markers have been typed on chromosome 15. A recent update on the Colorado Reading Study provides a summary of research methods and findings (DeFries et al., 1991). For the present purpose it is sufficient to note that a variety of analyses have indicated linkage to the disorder on chromosome 15 and familial heterogeneity. The subjects are those used by Smith et al. in the present volume.

These 19 families yield 161 sib pairs for analysis. Although the pairs are not all independent of each other when formed in this way, it appears that the assumption of independence is not important (Blackwelder and Elston, 1985). In order to label one sib a proband and the other a cosib, we employed two procedures depending on whether selection was used or not. With no selection all pairs were double entered and standard errors of the regression weights adjusted by a factor of the square root of two, a procedure routinely applied in the regression procedure when used with pairs of twins. In the case of subsets of the sibs selected for extreme scores all pairs were again double entered before selection in order to allow either sib to be a proband if he or she met the selection criterion. Then standard errors were rescaled, this time by the square root of the ratio of the total number of pairs entered into the analysis to that number minus the number of double entered pairs. These procedures take account of the method of ascertainment and the statistical problems associated with double entry (DeFries et al., 1991).

Test Scores

The subjects have been evaluated in a variety of ways. However, for the present analysis a discriminant score based on the Peabody Individual Achievement Test (PIAT; Dunn and Markwardt, 1970) Reading Recognition, Reading Comprehension and Spelling scores of the sibs was used. Details of how the discriminant score was constructed are given by DeFries (1985). Typical z scores range from about plus two to minus five with those in the present sample being low due to the initial identification of families with reading problems.

Markers

The 9 markers on chromosome 15 used in the present analysis are all RFLPs typed in the laboratory of Drs. Smith and Kimberling. They are shown in Table 3 in the results section. A more detailed description is given in Smith et al. in the present volume.

Analysis

The analysis was conducted using the models described above and a SAS program (SAS Institute, 1988) that read in data in the standard pedigree format employed by the program SIBPAL, sorted subjects into sib pairs, carried out the required degree of phenotypic selection for the purpose of comparisons of power and input the data into the regression routine of SAS. Where marker information was not sufficient to estimate $\hat{\pi}$ using Table 2, we omitted the sib pair from the analysis. Haseman and Elston provide a more elegant solution to this problem, but we chose to take the present more conservative approach. These programs were written by the second author (LRC) to run on mainframe or personal computers and are available on one diskette with explanatory notes, free of charge, upon request.

It should be noted that the sorting of the data into pairs and calculation of the $\hat{\pi}$s is the major task performed by SAS. Once these tasks have been performed the regression analysis can be undertaken using any simple statistical package on a PC or indeed using a pocket calculator.[1]

RESULTS

Univariate Analyses

Five analyses were carried out for each of the 9 markers. The first was a direct application of the Haseman and Elston approach, but using our own program, and was employed as a comparison with our own approach. The second was an application of the DeFries and Fulker augmented model, which differs from the Haseman and Elston approach only insofar that the average sib resemblance is controlled for in the analysis, presumably increasing power. The remaining three analyses involved the basic model applied to probands selected for a phenotypic score of less than 0, -1, and -2, respectively. The results of these analyses are given in Table 3, in the form of t-tests (adjusted for double entry) of the significance of β for the Haseman and Elston analysis, B_5 in the case of the augmented DeFries and Fulker model and B_2 in the case of the basic model.

The Haseman and Elston analysis given in the first row of Table 3 detects a significant effect for the marker ju201 at the end of the long arm. The same result was obtained using SIBPAL (Smith et al., this volume). The augmented DeFries and Fulker model detects the same effect. In addition, however, one other locus is detected, ynz90. When selection is imposed and the basic model is employed, an additional effect due to th114 becomes statistically significant and ynz90 and ju201 tend to become more so.

Table 3. *t*-Values for RFLP Markers on Chromosome 15 (Ns in parentheses)

Model	Marker								
	c15	cmw1	th114	dp151	ms114	efd49	ynz90	thh55	ju201
(1) Haseman—Elston									−1.86* (111)
(2) Augmented D-F			1.58 (40)				1.79* (54)		1.98* (111)
(3) D-F Basic: Selection (<0)			−1.58 (12)				−1.19 (30)		−1.91* (44)
(4) D-F Basic: Selection (<-1)			−2.22* (11)				−2.82** (19)		−2.93*** (50)
(5) D-F Basic: Selection (<-2)			−2.26* (9)				−3.16*** (14)		−2.68*** (29)

* $p < 0.05$.
** $p < 0.01$.
*** $p < 0.005$.

The most noteworthy feature of the result of selection, however, is the marked tendency of the *t*-values to increase with selection. In these cases there is clearly a great increase in power with selection and this increase continues since, in spite of the decrease in sample size, the values of *t* either remain almost constant or increase. The analysis of the selected sample reaches high levels of significance for very modest sample sizes suggesting that the pessimism associated with the sib-pair method (Robertson, 1973) may be misplaced when selected samples are employed.

The fit of the basic model to data for marker ju201, which is the most informative marker available in this data set, is presented in more detail in Table 4. The expectation is that the basic model will detect an inverse linear relationship with the means of cosibs for probands of increasing values of $\hat{\pi}$. On the other hand probands' scores should show no relationship to $\hat{\pi}$. A fortuitous relationship is corrected for by the regression analysis. Table 4 shows precisely such a result with a degree of regularity unusual in real data. Probands show a reasonable uniformity of values but cosibs regress progressively back towards the population mean with smaller values of $\hat{\pi}$, with the single possible exception of $\hat{\pi}$ equal to 1.00, which is only based on a single sib pair. The table is instructive in illustrating the simple nature of the selected sample procedure. In essence it just involves a statistical comparison of the means in column five which have been corrected for discrepancies in the values $\overline{P}$ and with the average sib resemblance removed from the estimate of error.

Multivariate Analyses

A series of univariate analyses of the same phenotypic scores is less convincing than a multivariate one. Therefore, we repeated the analyses of

Table 4. Regression Selection Model (<0) Observed and Expected Co-sibling Means for Marker ju201

$$E(\overline{C}_s) = B_1\overline{P} + B_2\hat{\pi}_j + A$$

$B_1 = -0.12$ $B_2 = -1.75$ $A = 0.31$ $\overline{P} = -1.54$

N	$\hat{\pi}$	$\overline{P}$	$\overline{C}$	$E(\overline{C}_s)$
1	0.00	−0.82	2.60	0.41
44	0.25	−1.21	−0.25	−0.02
35	0.50	−1.40	−0.94	−0.40
59	0.75	−1.20	−1.08	−0.86
1	1.00	−1.53	1.73	−1.26

Note: These Ns include the double entries required for the correct calculation of the means.

the selected samples entering the $\hat{\pi}$ simultaneously in a multiple regression equation. The results are shown in Table 5. Markers th114 and thh55 were excluded from this analysis because of insufficient sample sizes. The same two loci (ynz90 and ju201) manifest linkage with QTLs, but with somewhat increased significance.

DISCUSSION

It seems clear that the linkage analysis employing the DeFries and Fulker model is successful in terms of ease of application, substantial agreement with results of previous analyses, and a marked increase in power when used with selected samples. At least one QTL for reading disability is detected on chromosome 15, where we expected to find one, and with a very high level of significance even with quite small samples of selected probands. The method appears both reliable and powerful.

We have not presented systematic derivations of the models as these are discussed in detail in the literature cited. Neither have we presented systematic power calculations since the results are self evident in this respect. These calculations have been performed and it is hoped they will form the basis of a subsequent report (Carey and Williamson, in preparation) which considers other complications that we have ignored such as dominance and the effect of a single locus compared to that of several QTLs.

An approximate formula for calculating power for the analysis of selected samples and completely informative markers is

$$N = \frac{8(\text{non-central } \chi^2)\,(1 - R^2)}{[(\bar{P} - \mu)h^2]^2} \tag{7}$$

where $\bar{P}$ is the proband mean, μ is the population mean, h^2 is the heritability of the QTL, R is the overall sib correlation, and the non-central χ^2 is obtained from Pearson and Hartley (1972). N will be very sensitive to the degree of selection, $\bar{P} - \mu$, which is squared in the denominator of (7). Thus, small increases in selection have relatively large effects on power. The use of this formula gave the expected sample sizes shown in Table 6. These Ns seem roughly compatible with the levels of significance we obtained with the Ns in the present analyses.

In part, the effectiveness of these analyses could be due to the fact that the families from which the sibs pairs are drawn are few in number and selected for a high incidence of reading disability. The effect of such selection could lead to a few heterogeneous major locus systems, each segregating in only a few families or possibly only one family. In that case

Table 5. t-Values for RFLP Markers on Chromosome 15 for Multivariate Regression Selection Models (Ns in parentheses)

Model	Marker								
	c15	cmw1	th114	dp151	ms114	efd49	ynz90	thh55	ju201
(1) Multivariate Regression (<0)							−1.27 (20)		−1.18 (44)
(2) Multivariate Regression (<-1)							−4.41*** (19)		−4.68** (50)

Note: Markers th114 and thh55 omitted from analysis due to insufficient sample sizes.
** $p < 0.01$.
*** $p < 0.005$.

Table 6. Sib Pairs Needed to Reach Power of 0.50 and 0.90 for α levels 0.05 and 0.01 With Two Standard Deviations Selection Cutoff for Probands and Average Sib Correlations of 0.50

	$h^2 = 0.1$		$h^2 = 0.2$		$h^2 = 0.25$		$h^2 = 0.50$	
$1 - \beta$	0.50	0.90	0.50	0.90	0.50	0.90	0.50	0.90
α								
0.05	417	1148	104	287	51	137	16	46
0.01	717	1617	179	404	86	194	29	65

Note: For 3 SD cut off, divide Ns by 1.92.

the model would still provide a test for linkage, but the h^2 estimates would be higher than those typically found in the general population.

However, both our approximate power calculations and our substantive findings suggest that the regression approach applied to selected samples may have considerable utility in the detection of QTLs in behavioral and other human phenotypes.

ACKNOWLEDGEMENTS

This work was supported in part by program project and center grants from NICHD (HD-11681 and HD-27802) to J. C. DeFries.

NOTE

[1] Requests for the SAS files and accompanying documentation should be sent to Lon R. Cardon, Institute for Behavioral Genetics, University of Colorado, Boulder, Colorado 80309–0447, or by electronic mail to cardon@abacus.colorado.edu. No knowledge of SAS is required to run these programs.

REFERENCES

Blackwelder, W. C. and Elston, R. C. 1985. A comparison of sib-pair linkage tests for disease susceptibility loci. *Genetic Epidemiology, 2*, 85–97.

Cyphers, L. H., Phillips, K., Fulker, D. W., and Mrazek, D. A. 1990. Twin temperament during the transition from infancy to early childhood. *American Academy of Child and Adolescent Psychiatry, 29*, 392–397.

DeFries, J. C. 1985. Colorado reading project. In D. B. Gray and J. F. Kavanagh (eds.), *Biobehavioral Measures of Dyslexia* (pp. 123–135). Parkton, MD: York Press.

DeFries, J. C. and Fulker, D. W. 1985. Multiple regression analysis of twin data. *Behavior Genetics, 15*, 467–473.

DeFries, J. C. and Fulker, D. W. 1988. Multiple regression analysis of twin data: Etiology of deviant scores versus individual differences. *Acta Geneticae Medicae et Gemellologiae, 37,* 205–216.

DeFries, J. C., Olson, R. K., Pennington, B. F., and Smith, S. D. 1991. Colorado Reading Project: An update. In D. B. Gray and D. Duane (eds.), *The Reading Brain: The Biological Basis of Dyslexia.* Parkton, MD: York Press.

Dunn, L. M. and Markwardt, F. C. 1970. *Examiner's Manual: Peabody Individual Achievement Test.* Circle Pines, MN: American Guidance Service.

Haseman, J. K. and Elston, R. C. 1972. The investigation of linkage between a quantitative trait and a marker locus. *Behavior Genetics, 2,* 3–19.

Nance, W. E. and Neale, M. C. 1989. Partitioned twin analysis: A power study. *Behavior Genetics, 19,* 143–150.

Pearson, E. S. and Hartley, H. O. (eds.). 1972. *Biometrika Tables for Statisticians.* Cambridge: Cambridge University Press.

Robertson, A. 1973. Linkage between marker loci and those affecting a quantitative trait. *Behavior Genetics, 2,* 389–391.

SAS Institute, Inc. 1988. *SAS/STAT Users's Guide: Release 6.03 Edition.* Cary, NC: SAS Institute.

Smith, S. D., Kimberling, W. J., Pennington, B. F., and Lubs, H. A. 1983. Specific reading disability: Identification of an inherited form through linkage analysis. *Science, 219,* 1345–1347.