Am. J. Hum. Genet. 63:1190–1201, 1998

Multiple Phenotype Modeling in Gene-Mapping Studies of Quantitative Traits: Power Advantages

David B. Allison,[1] Bonnie Thiel,[2] Pamela St. Jean,[2] Robert C. Elston,[2] Ming C. Infante,[1] and Nicholas J. Schork[2,3,4,5]

[1]Obesity Research Center, St. Luke's/Roosevelt Hospital, Columbia University College of Physicians and Surgeons, New York; [2]Department of Epidemiology and Statistics and [3]Department of Genetics, Case Western Reserve University, Cleveland; [4]Department of Biostatistics and Program for Population Genetics, Harvard University School of Public Health, Boston; and [5]Jackson Laboratory, Bar Harbor, ME

Summary

Genomewide searches for loci influencing complex human traits and diseases such as diabetes, hypertension, and obesity are often plagued by low power and interpretive difficulties. Attempts to remedy these difficulties have typically relied on, and have promoted the use of, novel subject-ascertainment schemes, larger sample sizes, a greater density of DNA markers, and more-sophisticated statistical modeling and analysis strategies. Many of these remedies can be costly to implement. We investigate the utility of a simple statistical model for the mapping of quantitative-trait loci that incorporates multiple phenotypic or diagnostic endpoints into a gene-mapping analysis. The approach considers finding a linear combination of multiple phenotypic values that maximizes the evidence for linkage to a locus. Our results suggest that substantial increases in the power to map loci can be obtained with the proposed technique, although the increase in power obtained is a function of the size and direction of the residual correlation among the phenotypes used in the analysis. Extensive simulation studies are described that justify these claims, for cases in which two phenotypic measures are analyzed. This approach can be easily extended to cover more-complex situations and may provide a basis for more insightful genetic-analysis paradigms.

Introduction

Modern geneticists have accepted the challenge of localization of genes that influence traits and diseases of all types. The tools and basic approaches for this task are numerous and varied, but one of the most widely used involves a total genome scan with anonymous DNA markers. With this approach, a large number of related individuals are sampled who are thought to be segregating for genes that influence a particular disease or trait. These families are then genotyped for numerous DNA markers that are known to be associated with various landmark positions (i.e., loci) along the genome. This marker genotype information is then evaluated for statistical linkages or associations with putative disease-susceptibility or trait-influencing loci, in an effort to "map" (i.e., locate the rough genomic position of) those loci. Subsequent genotyping and analysis can be pursued not only to refine the position of those loci but also ultimately to determine gene sequences and mutations that causally influence the trait or disease of interest. Although this description is an oversimplification, this basic approach has been used with great success to map loci influencing simple (Mendelian) diseases, such as cystic fibrosis and neurofibromatosis, that are determined by a single locus whose relationship with the disease or trait is largely unequivocal.

Extensions and implementations of genome-scan technologies to more-complex traits and diseases have been plagued by numerous problems. "Complex" traits and diseases, such as hypertension, diabetes, and obesity, are influenced by numerous genetic and nongenetic factors, each of which may contribute to a trait or disease in only a small way. As such, the detection or characterization of any one of the relevant genetic factors might be obscured or confounded by the influence of others. Thus, the genetic dissection of complex traits and diseases may require study designs and research protocols that are more sophisticated than those used in the analysis of simple Mendelian genetic traits and diseases (Lander and Schork 1994). One set of traits that are particularly difficult to deal with are those that exhibit continuous or metrical variation in the population at large. Such traits are often complex in nature, in that multiple genetic and nongenetic factors contribute to their population-level variation.

Perhaps the greatest challenge in the mapping of loci for quantitative traits is combating the necessarily large sample size that one must use to detect locus effects on

Received December 5, 1997; accepted for publication August 4, 1998; electronically published October 2, 1998.

Address for correspondence and reprints: Dr. Nicholas J. Schork, Department of Epidemiology and Biostatistics, Case Western Reserve University, R215 Rammelkamp Building, MetroHealth Medical Center, 2500 MetroHealth Drive, Cleveland, Ohio 44109-1998. E-mail: njs2@po.cwru.edu

0002-9297/98/6304-0034$02.00

the trait that are small or moderate in size (Blackwelder and Elston 1982; Eaves 1994; Allison and Schork 1997). Many investigators have shown that power can often be substantially increased by adopting different analytic strategies (e.g., see Fulker and Cherny 1996), sampling large sibships (Todorov et al. 1997), sampling phenotypically extreme sibling pairs (Eaves and Meyer 1994; Risch and Zhang 1995, 1996; Allison 1996), making use of multiple marker data (Fulker et al. 1995; Olson 1995), and a variety of other methods (Allison and Schork 1997). However, one method whose power has been only minimally explored is the use of multivariate phenotypes in mapping strategies.

In this paper we evaluate a method for multivariate linkage analysis originally proposed by Amos et al. (1990). We show that, under some circumstances, such multivariate analysis can substantially increase the power of quantitative-trait locus (QTL)–mapping studies. We begin, however, with a *brief* review of approaches to the use of multivariate data in human linkage analysis. Work of relevance to QTL-mapping studies involving model organisms includes that of Jiang and Zeng (1995) and Korol et al. (1995).

Table 1

Observed Type I Error Rates (with 1 df *t*-Test) for Standard Univariate Haseman-Elston Test, Based on 1,000 Simulations

Sample Size and Type of Value[a]	Type I Error Rate for Nominal α = .200	.100	.050	.010	.001
25:					
TCV	−.86	−1.32	−1.71	−2.50	−3.48
ER	.206	.106	.0535	.0075	.0005
ECV	−.90065	−1.338	−1.733	−2.458	−3.591
50:					
TCV	−.85	−1.30	−1.68	−2.40	−3.26
ER	.202	.103	.053	.012	.001
ECV	−.858	−1.310	−1.720	−2.461	−3.352
100:					
TCV	−.85	−1.29	−1.66	−2.36	−3.17
ER	.212	.109	.054	.010	.002
ECV	−.8790	−1.323	−1.701	−2.293	−3.499
250:					
TCV	−.84	−1.28	−1.64	−2.33	−3.09
ER	.212	.111	.065	.014	.002
ECV	−.8895	−1.339	−1.763	−2.442	−3.647
500:					
TCV	−.84	−1.28	−1.64	−2.33	−3.09
ER	.189	.094	.049	.012	.0005
ECV	−.799	−1.252	−1.685	−2.355	−3.180

[a] Sample size is no. of sib pairs; TCV = standard one-sided *t*-test critical value, ER = estimated error rate for critical value from simulations, and ECV = estimated critical value from empirical distribution of test statistics gathered from simulations.

Some Approaches to Multivariate Phenotypic Data in Linkage Analysis

Multiple phenotypes are frequently included in gene-mapping studies. Several authors have pointed out that effective use of multivariate phenotypic measures can potentially enhance the power of linkage studies (Amos et al. 1986, 1990; Elston 1991; Amos and Liang 1993; Blangero et al. 1993; Schork 1993; Markel and Corley 1994; Dupuis et al. 1995; Jiang and Zeng 1995; Korol et al. 1995; Boomsma 1996), and several approaches have been taken to handle such multivariate data.

Use of Several Univariate Analyses

One approach is to run a separate univariate linkage analysis on each phenotype. Examples can be found in articles by Reed et al. (1995), Duggirala et al. (1996), and Norman et al. (1995). The obvious advantage to this approach is its simplicity of execution and interpretation. Unfortunately, there are two major disadvantages. First, it does not make use of the multivariate structure of the data and capitalize on its potential power advantages. Second, the use of multiple phenotypes can increase the studywise type I error rate if not accounted for properly (Lander and Schork 1994). Although this inflated type I error rate could be managed with Bonferroni-type corrections, such corrections are likely to be overly stringent and to result in an increased type II error rate, because the phenotypes are very unlikely to be independent (Hochberg and Tamhane 1987; Allison and Beasley 1998).

Methods Based on a Priori Composites

A second approach is to combine the multiple phenotypes into a single composite score. Although this may initially seem to be a radical proposal, testing for linkage to dichotomous phenotypes (i.e., affected vs. unaffected) is quite familiar when the dichotomous phenotype is a "syndrome" (e.g., schizophrenia); and a syndrome is nothing but a composite of several phenotypic measurements (albeit a potentially nonlinear composite) (Bailey 1973).

Using a composite in linkage analysis requires that one have some reasonable method for determining a good way to combine multiple variables. There are several ways in which one could select a composite prior to conducting a linkage analysis. For example, Jones (1971) and Grove (1994) developed methods for extracting the linear composite of several variables that maximizes the broad-sense heritability when twin data are available. Analogously, in the dichotomous situation (i.e., affected vs. unaffected), McGuffin et al. (1993) described the use of twin data to derive the syndrome definition that yields the maximum heritability. Similarly,

Table 2

Observed Type I Error Rates for Bivariate (Linear-Combination) Haseman-Elston Test, Based on 1,000 Simulations and Assumption of .00 Residual Correlation between Traits

Sample Size and Type of Value[a]	Type I Error Rate at Nominal α = .200	.100	.050	.010	.001
	Residual Correlation between Traits = .00				
25:					
TCV	−.86	−1.32	−1.71	−2.50	−3.48
ER	.443	.263	.153	.030	.002
ECV	−1.534	−1.986	−2.337	−2.978	−5.121
50:					
TCV	−.85	−1.30	−1.68	−2.40	−3.26
ER	.436	.259	.148	.036	.006
ECV	−1.480	−1.936	−2.278	−2.969	−3.435
100:					
TCV	−.85	−1.29	− 1.66	−2.36	−3.17
ER	.467	.278	.156	.037	.004
ECV	−1.504	−1.895	−2.265	−2.815	−3.722
250					
TCV	−.84	−1.28	−1.64	−2.33	−3.09
ER	.466	.285	.171	.042	.006
ECV	−1.541	−1.928	−2.284	−2.945	−4.324
500:					
TCV	−.84	−1.28	−1.64	−2.33	−3.09
ER	.414	.235	.127	.036	.006
ECV	−1.376	−1.814	−2.202	−2.821	−4.003
	Residual Correlation between Traits = .50				
25:					
TCV	−.86	−1.32	−1.71	−2.50	−3.48
ER	.364	.186	.100	.023	.001
ECV	−1.271	−1.716	−2.092	−2.888	−4.674
50:					
TCV	−.85	−1.30	−1.68	−2.40	−3.26
ER	.336	.180	.105	.019	.001
ECV	−1.232	−1.710	−2.029	−2.587	−3.289
100:					
TCV	−.85	−1.29	−1.66	−2.36	−3.17
ER	.391	.218	.111	.031	.005
ECV	−1.339	−1.714	−2.109	−2.716	−3.787
250:					
TCV	−.84	−1.28	−1.64	−2.33	−3.09
ER	.370	.207	.106	.019	.001
ECV	−1.300	−1.679	−1.987	−2.650	−3.206
500:					
TCV	−.84	−1.28	−1.64	−2.33	−3.09
ER	.349	.205	.112	.025	.002
ECV	−1.306	−1.674	−2.003	−2.689	−3.149
	Residual Correlation between Traits = −.50				
25:					
TCV	−.86	−1.32	−1.71	−2.50	−3.48
ER	.51	.31	.18	.035	.005
ECV	−1.66	−2.04	−2.36	−3.13	−4.18
50:					
TCV	−.85	−1.30	−1.68	−2.40	−3.26
ER	.51	.33	.19	.04	.005
ECV	−1.65	−2.02	−2.33	−2.98	−3.79
100:					
TCV	−.85	−1.29	−1.66	−2.36	−3.17
ER	.51	.29	.17	.035	.003
ECV	−1.54	−1.94	−2.21	−2.88	−3.59
250:					
TCV	−.84	−1.28	−1.64	−2.33	−3.09
ER	.51	.31	.17	.039	.005
ECV	−1.56	−1.95	−2.88	−2.81	−3.51
500:					
TCV	−.84	−1.28	−1.64	−2.33	−3.09
ER	.52	.32	.19	.046	.005
ECV	−1.60	−1.95	−2.32	−2.87	−3.32

[a] Definitions are as in table 1.

Boomsma (1996) showed that, using twin data, one can compute genetic-factor scores to represent an underlying genetic predisposition toward increasing values on several variables. These factor scores can then be used as the phenotype in subsequent linkage analysis. Boomsma (1996) showed that, under certain circumstances, the use of such factor scores resulted in substantially greater power than the use of the individual phenotypic scores.

One limitation of the approaches above is that the composite that maximizes an overall heritability, or that best correlates with a latent genetic factor, may not be the optimal composite for mapping a particular locus (Eaves et al. 1996). Pedigree discriminant analysis (Goldin et al. 1980; Zlotnik et al. 1983; Amos et al. 1986) determines the composite that best fits major-gene transmission. Similarly, Blangero and Konigsberg (1991) offered a method of multivariate segregation analysis that maximizes the strength of a hypothesized major-gene effect and suggested that the resulting composite might be useful as a dependent variable for linkage analysis. Although these methods may provide an optimal linear composite for the largest "major gene," they may not provide the optimal composite for other QTLs of interest. This is because the covariances induced among a set of phenotypes by one locus may be different from the covariances induced among the same set of phenotypes by a different locus (Eaves et al. 1996).

Methods Based on Simultaneous Linkage to Multiple Variables

Another set of approaches simultaneously includes several separate phenotypes in a single linkage analysis. Some of these are based on the generation of a composite score, whereas others are not.

Table 3

Power (i.e., 1 − Type II Error Rate) of Univariate Test versus Bivariate Test

Model and σ^2_M [a]	$p(a)$ [b]	μ_{AA} [c]	μ_{Aa} [c]	μ_{aa} [c]	Residual Correlation	Power of Bivariate Test/Power of Univariate Test, When Type I Error = .200	.100	.050	.010
Dominant:									
.100	.1	0	.85	.85	0	.295/.279	.157/.148	.081/.082	.019/.012
					.5	.271/.284	.161/.150	.066/.080	.015/.014
					−.5	.395/.284	.242/.156	.161/.088	.036/.019
	.2	0	.69	.69	0	.276/.273	.161/.148	.082/.077	.026/.017
					.5	.250/.284	.147/.136	.068/.069	.023/.018
					−.5	.381/.274	.244/.145	.148/.081	.043/.017
.250	.1	0	1.47	1.47	0	.519/.444	.353/.270	.185/.171	.064/.040
					.5	.428/.422	.272/.255	.148/.147	.045/.036
					−.5	.642/.396	.479/.246	.61/.140	.117/.035
	.2	0	1.20	1.20	0	.484/.405	.302/.252	.175/.140	.059/.033
					.5	.416/.416	.266/.244	.146/.144	.051/.041
					−.5	.713/.417	.550/.260	.444/.149	.164/.038
Additive:									
.100	.1	0	.79	1.58	0	.290/.280	.156/.152	.082/.080	.026/.017
					.5	.292/.295	.167/.155	.089/.082	.022/.019
					−.5	.367/.271	.222/.151	.147/.077	.039/.020
.100	.2	0	.59	1.18	0	.279/.277	.130/.144	.072/.072	.023/.016
					.5	.267/.262	.145/.145	.074/.079	.024/.021
					−.5	.432/.292	.272/.169	.177/.092	.048/.019
.250	.1	0	1.36	2.73	0	.469/.429	.319/.250	.192/.154	.068/.040
					.5	.428/.418	.280/.261	.146/.152	.046/.036
					−.5	.652/.418	.484/.247	.364/.144	.133/.041
	.2	0	1.02	2.04	0	.528/.425	.342/.275	.215/.167	.088/.049
					.5	.434/.415	.279/.245	.143/.149	.052/.039
					−.5	.697/.416	.558/.255	.422/.150	.161/.041
Recessive:									
.100	.1	0	0	3.35	0	.234/.255	.126/.134	.064/.071	.014/.014
					.5	.259/.266	.149/.154	.074/.082	.013/.018
					−.5	.226/.262	.125/.128	.068/.063	.015/.014
.101	.2	0	0	1.71	0	.264/.257	.129/.137	.068/.066	.017/.13
					.5	.247/.259	.134/.129	.051/.064	.014/.011
					−.5	.294/.259	.167/.138	.103/.067	.028/.014
.250	.1	0	0	5.81	0	.227/.310	.112/.163	.044/.079	.006/.013
					.5	.246/.316	.128/.164	.046/.078	.010/.012
					−.5	.221/.317	.111/.169	.046/.079	.002/.009
	.2	0	0	2.95	0	.375/.405	.216/.232	.119/.141	.031/.032
					.5	.376/.385	.227/.228	.113/.130	.028/.031
					−.5	.380/.373	.240/.221	.145/.122	.030/.025

[a] Variation explained by biallelic locus for *both* traits.
[b] Allele frequency for biallelic locus; $p(A) = 1 - p(a)$.
[c] Mean for both traits, for genotype.

Schork (1993)

Schork (1993) extended Goldgar's (1990) variance-components approach in several ways, including the use of multiple phenotypes, initially put forth by Lange and Boehnke (1983), and showed that, under some circumstances, power could be increased substantially by consideration of multivariate phenotypes (also see Schork et al. 1994). Schork also showed that, within the variance-components approach, the degree of residual correlation among the phenotypes can also dramatically affect power.

Eaves et al. (1996)

Eaves et al. (1996) proposed a multivariate linkage approach that simultaneously incorporates both the phenotypic and genetic-marker information into a single structural equation model (SEM). Eaves et al.'s SEM approach essentially models the cross-sib covariances as a function of the estimated degree to which alleles are shared identical by descent at points along the genome. Through simulation, they show that the method is capable of detecting multiple QTLs with pleiotropic ef-

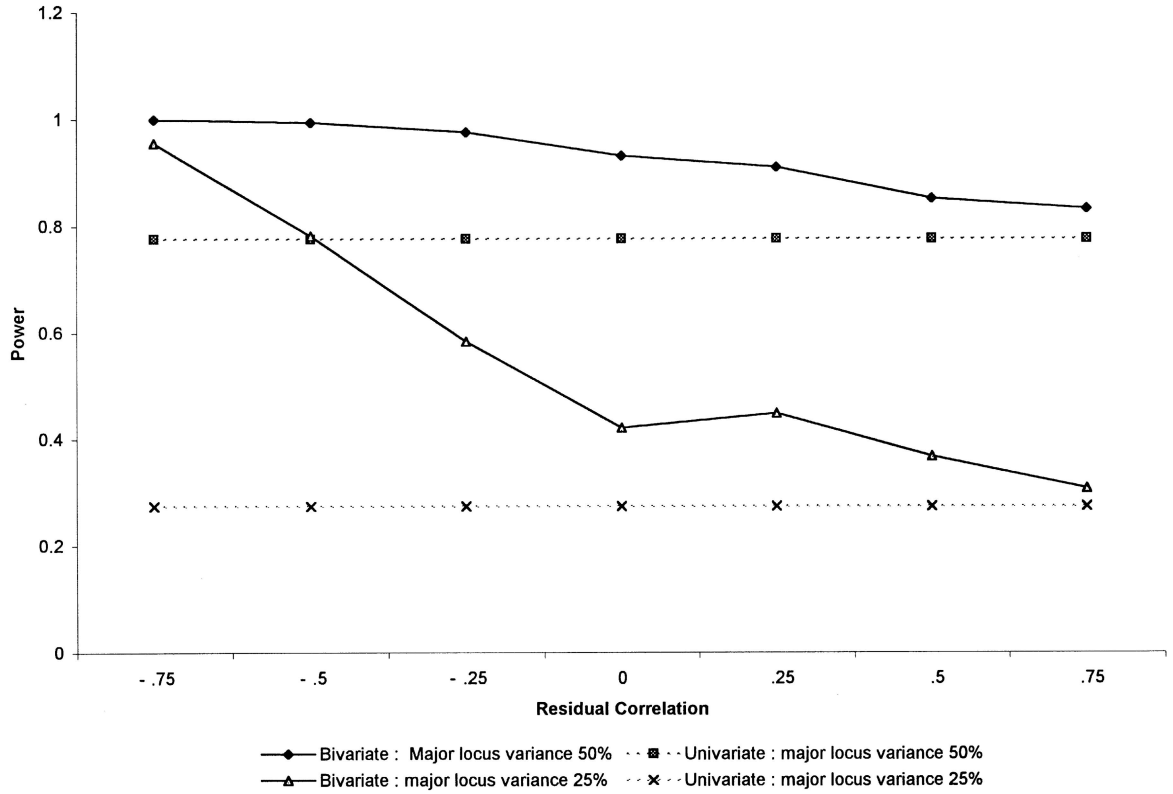

Figure 1 Simulation results examining the effect of the residual correlation between two traits on the power to detect a major locus (allele frequency = .20) with pleiotropic effects on two traits. A fully informative marker at distance of 1 cM from the trait locus was assumed. Figure 1 depicts the power of the proposed bivariate mapping method as a function of the residual correlation with the locus accounting for 50% or 25% of the variance of each trait.

fects. The relative power of this multivariate approach was not explicitly evaluated.

Moldin and Van Eerdewegh (1995) and Bonney et al. (1988)

Moldin and Van Eerdewegh (1995) and Bonney et al. (1988) used a regressive model to conduct joint segregation and linkage analysis with both a continuous and dichotomous trait simultaneously. The method appeared to work well in settings involving oligogenes. As presented, it had several limitations, including the fact that it (*a*) could accommodate only two phenotypic variables at a time and (*b*) assumed that, conditional on the QTL, one variable had no causal influence on the other. In addition, its power was not explicitly evaluated in comparison with other techniques.

Amos et al. (1990) and Amos and Liang 1993)

Amos et al. (1990) and Amos and Liang (1993) extended the Haseman and Elston (1972) linkage model for sib pairs to multiple traits, as explained below, illustrating the method by applying it to apolipoprotein and cholesterol levels in sib pairs from a large family that had many members diagnosed with heart disease.

The Amos et al. (1990) Model

The Amos et al. (1990) model to be studied herein is described within the context of a sib-pair study (Schork and Xu 1997). The procedure involves the estimation of a linear combination of the phenotypes that maximizes the linkage information (Amos et al. 1990). It is a simple extension of the well-known Haseman-Elston model (Haseman and Elston 1972). Let $y_{i,j}$ be the ith standardized phenotype ($i = 1,\ldots v$ for the jth sibling ($j = 1$ or 2) within a sib pair. Then the model can be written as

$$[\alpha_1(y_{1,2} - y_{1,2}) + \ldots\alpha_v(y_{v,1} - y_{v,2})]^2 = \beta_0 + \beta_1\hat{\pi} + e \ , \quad (1)$$

where $\hat{\pi}$ is an estimate of the fraction of alleles shared, at a locus, between sibs, β_0 is an intercept term, β_1 is a regression parameter that essentially quantifies the de-

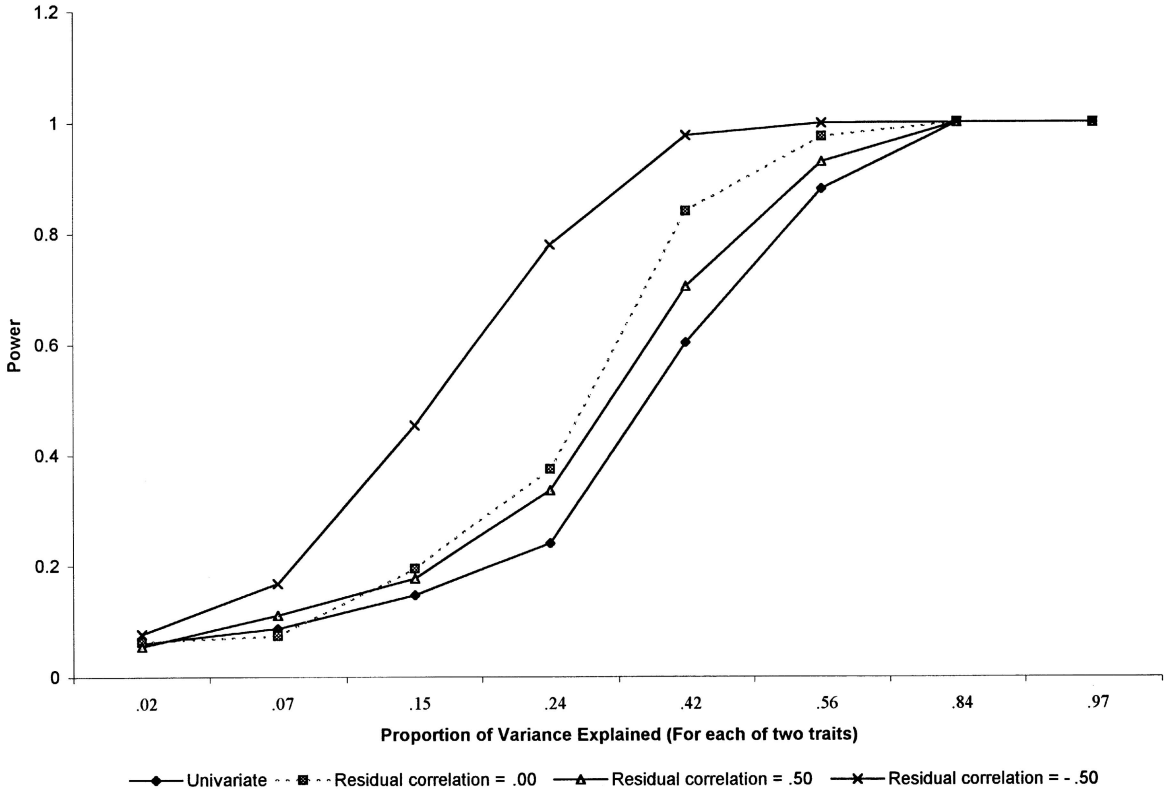

Figure 2 depicts power of the proposed bivariate mapping method as a function of the effect of a locus on two traits (i.e., the locus was assumed to influence the two traits in an equivalent manner). Three scenarios were investigated, each assuming a different residual correlation between the two traits.

gree to which variation at the locus in question explains variation in the phenotypes, and e is an error term. The α terms are estimable quantities and represent "loadings" on the phenotypic variables, loadings whose relative magnitudes could, conceivably, be interpreted as an indication as to how strongly each variable is influenced by the locus in question. Estimates of the α terms can be obtained in a variety of ways but should be estimated simultaneously with both the intercept, β_0, and the allele-sharing–effect parameter, β_1. We chose to estimate the α_1, β_0, and β_1 terms by determining the values that maximize the evidence for linkage. Since $\beta_1 < 0$ under linkage, we find the α_i that minimizes β_1. This can be achieved by assuming different values for the α's, fitting the regression in equation (1) via least squares, to obtain β_0 and β_1, and finding that α parameterization that results in the most "negative" β_1 value. It is necessary to place some constraint on the sum of the α_1, so that the model is identifiable. For convenience, we choose the arbitrary constraint that $\Sigma_{i=1}^{v} = 1$. We also assume that the phenotype values y_{ij} have been standardized to avoid scaling and interpretive difficulties.

The one-tailed test that $\beta_1 < 0$ is a test of linkage. Ordinarily, $\hat{\beta}_1/\hat{\sigma}_{\beta_1} \sim t[\text{df} = (n - 2)]$. However, in the multivariate case, this is not true, because the α_i values involve the estimation of more than two parameters. Amos et al. (1990) showed that, at least in large samples, a test statistic can be formulated that conservatively follows an F distribution with 2 and $(n - 3)$ df. The approach that we took to this problem was to estimate empirical thresholds corresponding to the 100α percentile of the distribution of $\hat{\beta}_1/\hat{\sigma}_{\beta_1}$, via simulation under the null hypothesis. A result is considered significant when the observed value of $\hat{\beta}_1/\hat{\sigma}_{\beta_1}$ is less than T_α, where T_α is the value corresponding to the empirical estimate of the 100α percentile of the distribution of $\hat{\beta}_1/\hat{\sigma}_{\beta_1}$.

Evaluation of the Model's Performance

To assess the utility of this approach, we conducted extensive simulation studies. We began by simply determining whether our simulation software was performing adequately, by simulating, via the Haseman-Elston regression methods for correlated traits the type I error rates for univariate linkage analyses. Simulations assuming no locus effect but different residual correlations

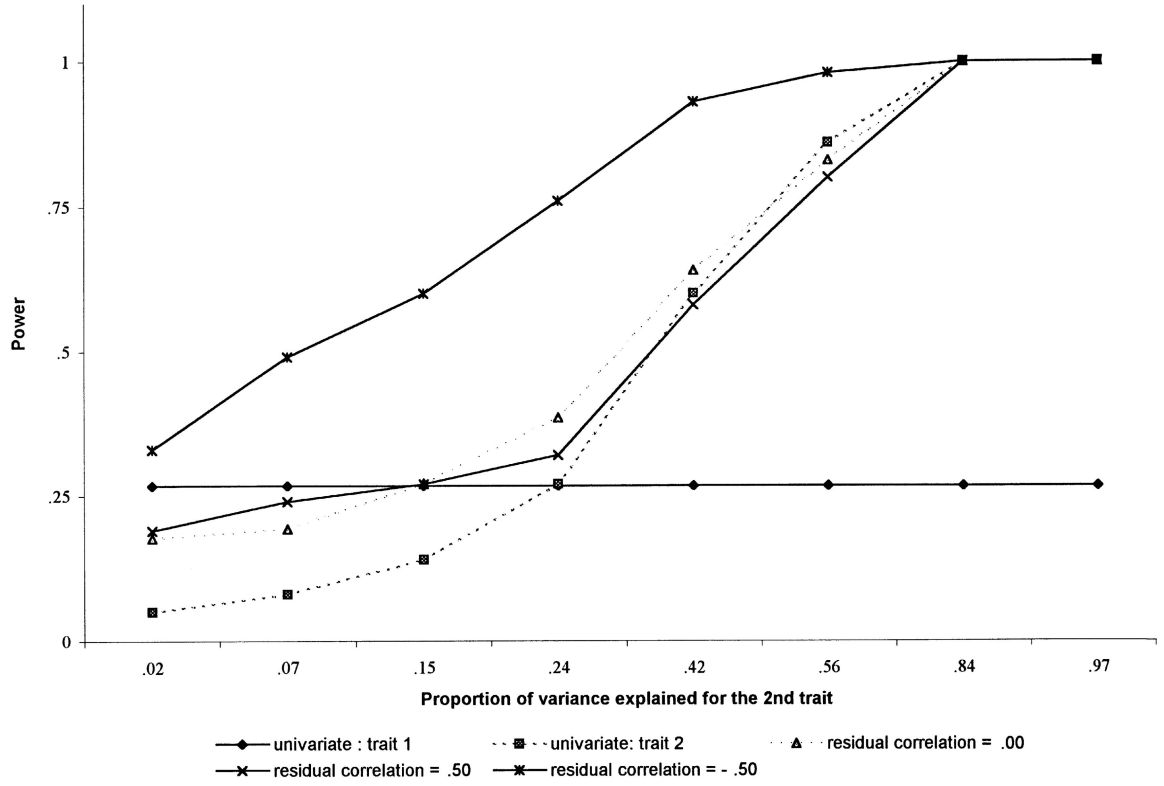

Figure 3 The effect of a second trait and marker distance on the power to detect linkage. The diagram depicts the power of proposed bivariate mapping method to detect linkage as a function of the effect of a relevant locus on the second of two traits (the locus was assumed to explain 25% of the variation of the first trait for all simulations). An allele frequency of 0.2 was assumed for all simulations. The residual correlation between the traits was assumed to be 0.0, 0.5, or −0.5. A fully informative marker at distance of 1 cM from the trait locus was assumed.

were performed, and empirical type I error rates were estimated. Results are based on 1,000 simulations, with sample sizes of 25–500 sib pairs and with one-tailed α levels of .200–.01. These results are given in table 1. Given 1,000 simulations per parameter set, the standard error of the estimated type I and type II error rates (below) are always $\leqslant$.0158. As can be seen, the empirical critical values correspond almost exactly to the theoretical critical values, and the observed error rates correspond almost exactly to the theoretical error rates. Thus, our software appears to be functioning properly.

We then simulated data under the null hypothesis of no linkage and conducted a bivariate linkage analysis (eq. [1]). We conducted these simulations with residual correlations, among the two traits, of −.5, .0, and .5. As can be seen in tables 2, the critical value estimated in the simulations is uniformly higher than the critical value needed in a univariate analysis. This is to be expected, given the extra parameter being estimated. It is apparent that, if bivariate linkage analyses were conducted and if one naively used the critical values for a univariate test, then one's actual type I error rate would be approximately two to six times greater than the nominal α level.

Next we simulated data under the alternative hypothesis for dominant, additive, and recessive models. In each case, we assumed perfectly informative markers and a recombination fraction of 0. We simulated with cross-phenotype residual correlations of .5, .0, and −.5; cross-sibling residual correlations of .0; sample sizes of 100 sibling pairs; QTL effects that explained 10%–25% of the phenotypic variance; allele frequencies (for the phenotype-increasing alleles) of .1 or .2; a residual variance of 1.0 for each trait; and one-tailed α levels of .2, .1, .05, and .01. The results are described in table 3. In every situation, we generated locus effects that induced a *positive* correlation between the traits. Thus, in many instances, the locus-induced correlations were opposite in sign to the residual correlation between the traits. This is an extremely important aspect of the simulations, as will be discussed below. As can be seen in table 3, in every situation the power of the bivariate test is at least equal to the power of univariate situation, apart from very small differences due to sampling variability. Again,

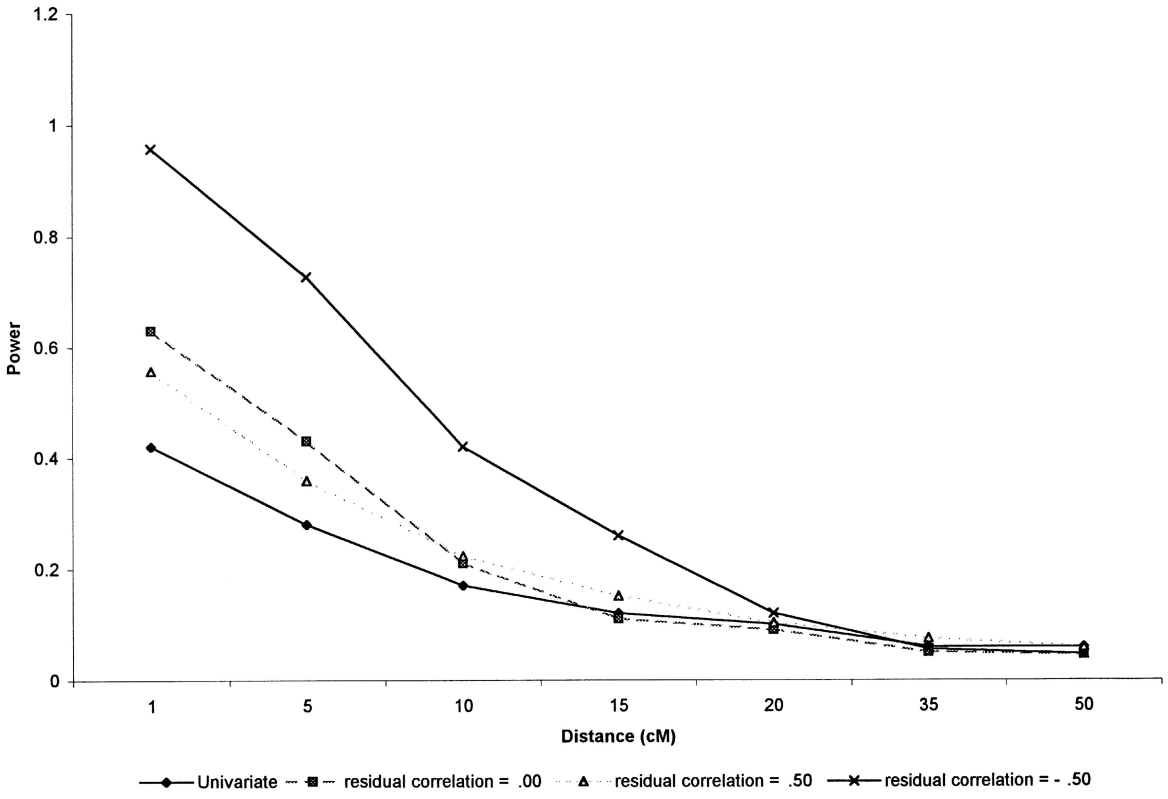

Figure 4 depicts the power of the proposed bivariate mapping method as a function of the distance (in cMs) between the marker locus and the trait locus. The trait locus was assumed to explain 33% of each of the two traits' variance. An allele frequency of 0.5 was assumed for all simulations. Fully informative markers were also assumed.

this is an expected result. However, what is clear and more interesting is that the power of the bivariate model is, in many situations, far greater than the power of the univariate model. To illustrate this with just one strong example, when a QTL explains 25% of the variance in each of the two traits and acts additively with an allele frequency (for the increasing allele) of .2, 100 sib pairs yield only 42% power to detect each of the univariate traits, even at the .20 α level, but yield 70% power in the bivariate model.

Finally, to explore the performance of the model with a more powerful (i.e., larger) sample and other circumstances, we conducted simulations with a sample size of 250 sibling pairs and an α level of .05. Again, 1,000 simulations for each situation were conducted. The residual correlation among the siblings was assumed to be .0.

Figure 1 displays results examining the effect that the residual correlation between two traits has on the power to detect a major locus with pleiotropic effects on two traits. For scenarios in which critical values were unknown, 1,000 simulations assuming no locus effect but an appropriate residual correlation were performed, and critical values were estimated from the empirical distribution of the test statistic obtained from these simulations. A fully informative marker at a distance of 1 cM from the trait locus was assumed. All simulations assumed an allele frequency (for the increasing allele) of .2. Two scenarios were investigated. One assumed that the locus accounted for 50% of the variance of each trait, and the other assumed that the locus accounted for 25% of the variance. Figure 1 shows that the increase in power with use of the bivariate model, relative to the univariate model, depends strongly on the residual correlation between the two traits. The lower (i.e., more negative) the residual correlation, the greater the power. In our simulations, the greatest power increase in the multivariate model accrued when the residual correlation was lowest. To again illustrate with a strong example, when a QTL explains 25% of the variance in each of the two traits with an allele frequency of .20 and when the residual correlation is $-.75$, 250 sib pairs yield barely 30% power to detect each of the univariate traits but yield >90% power in the bivariate model.

Figure 2 depicts the power of the proposed bivariate mapping method, as a function of the effect that a locus

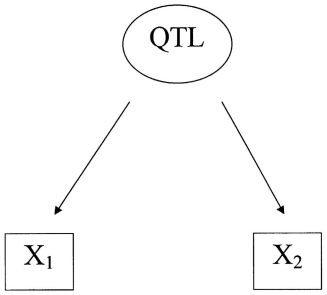

Figure 5 depicts a "Mosaic" pleiotropy relationship among a QTL and phenotypes X_1 and X_2.

has on two traits (i.e., the locus was assumed to influence the two traits in an equivalent manner). Three scenarios were investigated, each assuming a different residual correlation between the two traits. Figure 2 shows the expected result–that is, that power increases as the proportion of variance in each trait explained by the QTL increases.

In figure 3, the effect that a second trait and marker distance have on the power to detect linkage is portrayed. The figure depicts the power of the proposed bivariate mapping method to detect linkage, as a function of the effect of a relevant locus on the second of two traits (the locus was assumed to explain 25% of the variation of the first trait, for all simulations). An allele frequency of .2 was assumed for all simulations. The residual correlation between the traits was assumed to be .0, .5, or −.5. A fully informative marker at a distance of 1 cM from the trait locus was assumed. Figure 3 shows the interesting result that the bivariate analysis does not appear to increase power unless the QTL explains as much variance in the second trait as it does in the first *or* the residual correlation between the two traits is <0. Whether this result will hold up as a generality or is specific to the circumstances that we simulated is unclear.

Figure 4 depicts the power of the proposed bivariate mapping method, as a function of the distance (in cM) between the marker locus and the trait locus. The trait locus was assumed to explain 33% of each of the two traits' variance. An allele frequency of .5 was assumed for all simulations. The residual correlation between the two traits was assumed to be .0, .5, or −.5. Fully informative markers were also assumed. Figure 4 shows the expected result—that is, that power drops off dramatically as the distance between the marker and QTL increases but that the general pattern of results, in terms of the power advantages of the bivariate approach, remains unchanged.

Discussion

The model proposed has a number of advantages. It is quite simple and flexible and can easily accommodate covariates (as in the program SIBPAL; SAGE 1997) and multiple linked loci (Elston 1995; Tiwari and Elston 1997). Similarly, other multivariate approaches (e.g., see Eaves et al. 1996) could also accommodate these additional features and, when considering the full covariance structure among the siblings, may be even more powerful (Fulker and Cherny 1996).

In this paper, we have considered only two phenotypes. However, the model as formulated in equation (1) is obviously expandable to a theoretically unlimited number of phenotypes. In practice, the grid-search approach to derivation of the α coefficients for more than two phenotypes would involve a greater computational demand as the number of phenotypes increases. For example, for three phenotypes, one would fix the level of α_1 and then would conduct a grid search in which α_2 varied from 0 to $1 - \alpha_1$ while α_3 is set to $1 - \alpha_2 - \alpha_1$. This would then be repeated over the entire range of potential α_1 values (i.e., for $\alpha_1 = 0$ to $\alpha_1 = 1$). Although this is conceptually simple, it could become computationally intense. Therefore, as one moves to an increasing number of phenotypes, a minimization procedure that is more efficient than the grid search might be employed. In this context, it is noteworthy that many variance-component models are now becoming available that can accommodate multiple phenotypes and that may be desirable in this context (Almasy and Blangero 1988).

We have shown that the required critical values of the test statistic depend on the residual correlation. In practice, the residual correlations among the traits are unknown. This problem in easily solved by pointing out that, under the null hypothesis, the phenotypic correlations are the residual correlations. Therefore, the phenotypic correlations can be used as estimates of the residual correlations. The investigators must then simulate their own critical values, given the presumed residual

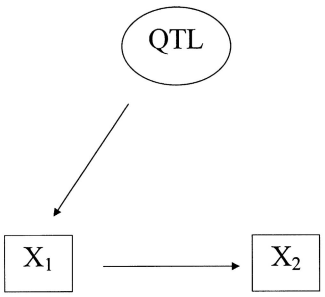

Figure 6 depicts a "Relational" pleiotropy relationship among a QTL and phenotypes X_1 and X_2.

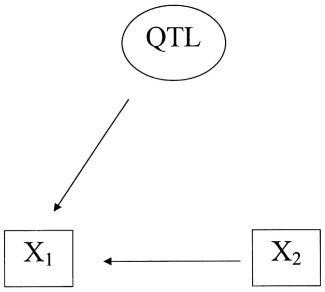

Figure 7 depicts a relationship among a QTL and phenotypes X_1 and X_2 with exogenous phenotype.

correlations. This method of estimating the null distribution of a test statistic—that is, on the basis of simulation in a population with parameters set equal to the corresponding observed sample statistics—is widely used in simulation-based inference (Ott 1989; Schork et al. 1990; Churchill and Deorge 1994; Deorge and Churchill 1996).

Correlations among phenotypes can arise from several different causal processes, and these different causal processes may have different implications for the power and conduct of multivariate linkage analysis. Figures 5–79 graphically display five different models involving a QTL and two phenotypes of interest, labeled "X_1" and "X_2." In model 1, X_1 and X_2 are both functions of the QTL; this is both the situation that the multivariate simulations conducted herein represent and the situation that Boomsma (1996) simulated. In keeping with Hadorn (as cited in Rieger et al. 1991), we refer to this situation as "'mosaic' pleiotropy." In this case, substantial power can be gained by conducting a multivariate linkage analysis, as we and Boomsma (1996) have shown.

Model 2, in Hadorn's (as cited in Rieger et al. 1991) terminology, is an example of "relational" pleiotropy. Here, QTL directly impacts X_1, and X_1 in turn directly impacts X_2. Including X_1 and X_2, as depicted in figure 6, simultaneously within the multivariate linkage analysis that we have constructed may not add additional power above and beyond that associated with X_1 alone, unless the residual correlation between X_1 and X_2 is strong and opposite to the linked-locus–induced correlation. However, if X_2 is caused by a "true" X_1 that is only observable with error and if that error is large (relative to total variance) in comparison with the error in the observed X_2 (relative to the total variance in X_2), then inclusion of X_2 within in a bivariate analysis may increase power.

Model 3 depicts the situation in which X_2 might be termed an "exogenous" variable (Neale and Cardon 1992). Here, both the QTL and X_2 exert a causal influence on X_1. However, the QTL does not influence X_2. In this situation, addition of X_2 to the linkage analysis as an additional variable might also increase power, as might inclusion of X_2 as a covariate on the right side of the equations.

Distinguishing among these models will allow one to decide whether it is best to include X_2 as an additional phenotype in the linkage analysis or as a covariate. One may have strong theoretical reasons for including an additional variable in the analysis—for example, that it is known to be physiologically related to the primary phenotype. Alternatively, an empirical approach may be superior. In other words, one might wish to try a model that has X_2 as a covariate and then try a model that has X_2 as an additional dependent variable. A decision rule might then be constructed to decide which model is best. It might be that the better-fitting model (i.e., that with the smaller P value) would indicate the more probable mode of action of the covariate. However, such multiple model fitting might warrant additional adjustments to the per-test α level chosen.

Model 4 depicts the situation in which X_1 and X_2 are correlated because each is influenced by an additional variable, Z. However, Z is not observed. In this case, as indicated above, including X_2 as an additional variable will probably not increase power, unless there is more error of measurement (relative to total variance) in X_1 than in X_2. However, using X_2 as a covariate may well be better. This is because X_2 may act as a poor proxy for Z and, therefore, will be a useful covariate. However, it seems to be unlikely to be as good a covariate as X_2 was in model 3.

Finally, in model 5, X_2 is an intermediary phenotype between the QTL and X_1. Here, including X_1 as a covariate would probably work against one's purposes and markedly reduce power. However, including X_2 within the model as an additional phenotype should markedly

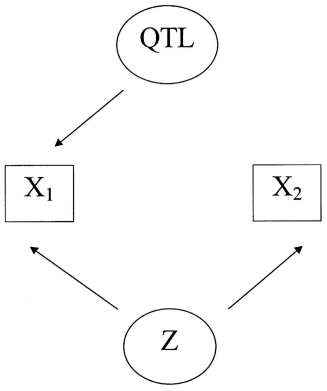

Figure 8 depicts a relationship among a QTL and phenotypes X_1 and X_2 with correlated phenotype.

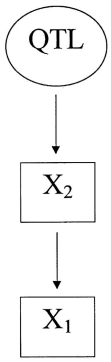

Figure 9 depicts a relationship among a QTL and phenotypes X_1 and X_2. With intermediate phenotype.

increase power. Obviously, further studies that specifically model these five situations are called for.

Two important results of our analyses deserve further emphasis. First, the greatest power increases occur when the locus-induced correlation is opposite in sign to the correlation induced by "residual" factors. This finding is consistent with the analytic results of Jiang and Zeng (1995). This result has important implications for mapping studies, since one will never know a priori whether a locus will induce a correlation between variables that is opposite to the residual correlation. This fact provides further motivation for an "empirical" approach, in which variables are tested to see whether they increase evidence for linkage. Second, the power of our approach in the bivariate setting is virtually always greater than or equal to the power in the univariate setting. This is due to the fact that, when α_1 or α_2 is 0.0, the model reduces to the univariate Haseman-Elston procedure. As such, our procedure can be used as a general screening tool (as long as relevant critical values are determined), without a tremendous cost in terms of type I and type II error rates.

Another issue that demands emphasis is the distinction between using multivariate data to increase mapping power (due to, e.g., pleiotropy of the mapped locus) and *testing* for pleiotropy. This article has considered methods for increasing the mapping power. Further work on tests for pleiotropy is needed.

In conclusion, this article has shown that, under many circumstances, modeling multiple phenotypes in a single linkage analysis simultaneously can markedly increase power, compared with modeling of each phenotype separately. This same strategy should be extendible to other existing linkage procedures based on variance components (e.g., see Schork 1993; Fulker and Cherny 1996) and may yield even more power.

Acknowledgments

This paper was supported in part by National Institutes of Health grants P30DK26687, O1DK51716, R29DK47526, HL94-011, HL54998-01, GM28356, and RR03655.

References

Allison DB (1996) The use of discordant sibling pairs for finding genetic loci linked to obesity: practical considerations. Int J Obes Relat Metab Disord 20:553–560

Allison DB, Beasley M (1998) A method and computer program for controlling the family-wise alpha rate in gene association studies involving multiple phenotypes. Genet Epidemiol 15:87–101

Allison DB, Schork NJ (1997) Selected methodological issues in meiotic mapping of obesity genes in humans: issues of power and efficiency. Behav Genet 27:401–421

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

Amos CI, Elston RC, Bonney GE, Keats BJB, Berenson GS (1990) A multivariate approach for detecting linkage, with application to a pedigree with an adverse lipoprotein phenotype. Am J Hum Genet 47:247–254

Amos CI, Liang AE (1993) A comparison of univariate and multivariate tests for genetic linkage. Genet Epidemiol 10: 671–676

Amos CI, Wilson AF, Rosenbaum PA, Srinivasan SR, Weber LS, Elston RC, Berenson GS (1986) An approach to the multivariate analysis of high-density-lipoprotein cholesterol in a large kindred: the Bogalusa Heart Study. Genet Epidemiol 3:255–268

Bailey KD (1973) Monothetic and polythetic typologies and their relation to conceptualization, measurement and scaling. Am Sociol Rev 38:18–33

Blackwelder WC, Elston RC (1982) Power and robustness of sib-pair linkage tests and extensions to large sibships. Commun Stat (Theory Meth) 11:449–484

Blangero J, Konigsberg LW (1991) Multivariate segregation analysis using the mixed model. Genet Epidemiol 8:299–316

Blangero J, Williams-Blangero S, Mahaney MC (1993) Multivariate genetic analysis of APO A1 concentration and HDL subtractions: evidence for major locus pleiotropy. Genet Epidemiol 10:617–622

Bonney GE, Lathrop GM, Lalouel J-M (1988) Combined linkage and segregation analysis using regressive models. Am J Hum Genet 43:29–37

Boomsma DI (1996) Using multivariate genetic modeling to detect pleiotropic quantitative trait loci. Behav Genet 26: 161–166

Churchill GA, Doerge RW (1994) Empirical threshold values for quantitative trait mapping. Genetics 138(3): 963–971

Doerge RW, Churchill GA (1996) Permutation tests for multiple loci affecting a quantitative character. Genetics 142(1): 285–294

Duggirala R, Stern MP, Mitchell BD, Reinhart LJ, Shipman PA, Uresandi OC, Chung WK, et al (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
Dupuis J, Brown PE, Siegmund D (1995) Statistical methods for linkage analysis of complex traits from high-resolution maps of identity by descent. Genetics 140:843–856
Eaves LJ (1994) Effect of genetic architecture on the power of human linkage studies to resolve the contribution of quantitative trait loci. Heredity 72:175–195
Eaves L, Meyer J (1994) Locating human quantitative trait loci: guidelines for the selection of sibling pairs for genotyping. Behav Genet 24:443–455
Eaves LJ, Neale MC, Maes H (1996) Multivariate multipoint linkage analysis of quantitative trait loci. Behav Genet 26: 519–525
Elston RC (1991) Genetic analysis of multivariate traits. In: Anderson VE, Hauser IE, Leppik JL, Noebels JL, Rich SS (eds) Genetic strategies in epilepsy research. Epilepsy Res Suppl 4:161–171
Elston RC (1995) The genetic dissection of multifactorial traits. Clin Exp Allergy 25, Suppl 2:103–106
Fulker DW, Cherny SS (1996) An improved multipoint sib-pair analysis of quantitative traits. Behav Genet 26:527–532
Fulker DW, Cherny SS, Cardon LR (1995) Multipoint interval mapping of quantitative trait loci, using sib pairs. Am J Hum Genet 56:1224–1233
Goldgar DE (1990) Multipoint analysis of human quantitative genetic variation. Am J Hum Genet 47:957–967
Goldin LR, Elston RC, Graham JB, Miller CH (1980) Genetic analysis of von Willebrand's disease in two large pedigrees: a multivariate approach. Am J Med Genet 6:279–293
Grove WM (1994) Maximizing the heritability of a linear combination of traits. Psychol Rep 75:467–476
Haseman JK, Elston RC (1972) The investigation of linkage between a quantitative trait and a marker locus. Behav Genet 2:3–19
Hochberg Y, Tamhane AC (1987) Multiple comparison procedures. In: Wiley series in probability and mathematical statistics: applied probability and statistics. John Wiley & Sons, New York
Jiang C, Zeng Z (1995) Multiple trait analysis of genetic mapping for quantitative trait loci. Genetics 140:1111–1127
Jones MB (1971) Heritability in the construction of psychological tests. Psychol Bull 75:92–96
Korol AB, Ronin YI, Kirzhner VM (1995) Interval mapping of quantitative trait loci employing correlated trait complexes. Genetics 140:1137–1147
Lander ES, Schork NJ (1994) Genetic dissection of complex traits. Science 265:2037–2048
Lange K, Boehnke M (1983) Extensions to pedigree analysis. IV. Covariance components models for multivariate traits. Am J Med Genet 14:513–524
Markel PD, Corley RP (1994) A multivariate analysis of repeated measures: linkage of the albinism gene (Tyr) to a QTL influencing ethanol-induced anesthesia in laboratory mice. Psychiatr Genet 4:205–210
McGuffin P, Katz R, Rutherford J, Watkins S, Farmer AE, Gottesman II (1993) Twin studies as vital indicators of phenotypes in molecular genetic research. In: Bouchard TJ Jr, Propping P (eds) Twins as a tool for behavioral genetics. Life Sci res rep 53. John Wiley & Sons, New York, pp 243–256
Moldin SO, Van Eerdewegh P (1995) Multivariate genetic analysis of oligogenic disease. Genet Epidemiol 12:801–806
Neale MC, Cardon LR (1992) Methodology for genetic studies of twins and families. Kluwer Academic, Boston
Norman RA, Bogardus C, Ravussin E (1995) Linkage between obesity and a marker near the tumor necrosis factor-alpha locus in Pima Indians. J Clin Invest 96:158–168
Olson JM (1995) Robust multipoint linkage analysis: an extension of the Haseman-Elston method. Genet Epidemiol 12: 177–193
Ott J (1989) Computer simulation methods in human linkage analysis. Proc Natl Acad Sci USA 86:4175–4178
Reed DR, Ding Y, Xu W, Cather C, Price RA (1995) Human obesity does not segregate with the chromosomal regions of Prader-Willi, Bardet-Biedl, Cohen, Borjeson, or Wilson-Turner syndromes. Int J Obes Relat Metab Disord 19: 599–603
Rieger R, Michaelis A, Green MM (1991) Glossary of genetics: classical and molecular, 5th ed. Springer-Verlag, New York, pp 384-385
Risch N, Zhang H (1995) Extreme discordant sib pairs for mapping quantitative trait loci in humans. Science 268: 1584–1589
Risch NJ, Zhang H (1996) Mapping quantitative trait loci with extreme discordant sib pairs: sampling considerations. Am J Hum Genet 58:836–843
SAGE (1997) Statistical analysis for genetic epidemiology, release 3.0. Computer program package available from the Department of Epidemiology and Biostatistics, Rammelkamp Center for Education and Research, Case Western Reserve University, Cleveland
Schork NJ (1993) Extended multipoint identity-by-descent analysis of human quantitative traits: efficiency, power, and modeling considerations. Am J Hum Genet 53:1306-1319
Schork NJ, Boehnke M, Terwilliger JD, Ott J (1993) Two-trait-locus linkage analysis: a powerful strategy for mapping complex genetic traits. Am J Hum Genet 55:1127–1136
Schork NJ, Weder AB, Schork MA (1990) On the asymmetry of biological frequency distributions. Genet Epidemiol 7(6): 427–446
Schork NJ, Weder AB, Trevisan M, Laurenzi M (1994) The contribution of pleiotropy to blood pressure and body-mass index variation: the Gubbio Study. Am J Hum Genet 54: 361–373
Schork NJ, Xu X (1997) Sibpairs versus pedigrees: what are the advantages? Diabetes Rev 5:1–7
Tiwari HK, Elston RC (1997) Linkage of multilocus components of variance to polymorphic markers. Ann Hum Genet 61:253–261
Todorov AA, Province MA, Borecki IB, Rao DC (1997) Trade-off between sibship size and sampling scheme for detecting quantitative trait loci. Hum Hered 47:1–5
Zlotnik LH, Elston RC, Namboordiri KK (1983) Pedigree discriminant analysis: a method to identify monogenic segregation. Am J Med Genet 15:307–313