Am. J. Hum. Genet. 60:1211–1221, 1997

Cost-Effective Sib-Pair Designs in the Mapping of Quantitative-Trait Loci

Hongyu Zhao,[1] Heping Zhang,[1] and Jerome I. Rotter[2]

[1]Department of Epidemiology and Public Health, Yale University School of Medicine, New Haven; and [2]Division of Medical Genetics, Departments of Medicine and Pediatrics, Cedars-Sinai Research Institute, and UCLA School of Medicine, Los Angeles

Summary

The extreme discordant-sib-pair design has been found to be the most powerful, across most genetic models. In this paper, we address two of the most frequently asked questions related to this design. First, under the extreme discordant-sib-pair design, a large number of people have to be screened for the phenotype of interest, before the desired number of discordant sibs can be collected for genotyping and linkage analysis. When the phenotyping cost is not negligible compared with the genotyping cost, such methods might not be cost effective. The second question is how sensitive the cost is to the genetic model and allele frequency. In this paper, we compare the cost under different sampling strategies, different genetic models, and different phenotyping:genotyping cost ratios. Because our knowledge of the underlying genetic model for a trait is limited, the discordant-sib-pair design proves to be the most robust. When the cost for screening probands is not included, the design that genotypes sibs with one sib in the top 10% and the other sib in the bottom 30% of the population with respect to the trait of interest is, across most models studied, the optimum among the designs considered in this paper. The cost under this design, across different genetic models, appears to be relatively robust to allele frequency and model type, whether additive or dominant. If probands initially must be screened as well, then 25% appears to be the optimal portions of the upper and lower distributions to be studied.

Introduction

Sib-pair methods are gaining more and more popularity in the linkage study of complex traits, because of the relative ease of data collection and the simplicity of data analysis. In the context of mapping genes for quantitative traits, linkage analysis often has been applied to sib pairs constituting a random sample from the population (Haseman and Elston 1972; Fulker and Cardon 1994; Kruglyak and Lander 1995). Carey and Williamson (1991) noted that, by selecting sib pairs on the bais of probands with extreme values, the sample size required for detection of linkage can be reduced. The logic of this latter approach is further supported when one realizes that, if one is ascertaining such pairs through disease, one is more likely to obtain individuals and pairs with extreme values of traits correlated with that disease. Recently, Eaves and Meyer (1994) and Risch and Zhang (1995, 1996) proposed the use of extreme discordant sib pairs for mapping quantitative-trait loci in humans, on the basis of their observation that the extreme discordant-sib-pair design has the greatest power across most genetic models that they studied. They demonstrated that the number of extreme discordant sib pairs needed for detection of linkage is much smaller than the number of sib pairs selected at random from the population. Furthermore, Zhang and Risch (1996) noted that the sample size can be further reduced by using the phenotypic information available from parents.

The main motivation of Risch and Zhang's work was to reduce the effort and cost of genotyping, because, in general, genotyping is expensive and time consuming, especially as regards a whole-genome scan. However, with the advances in biotechnology, the cost of genotyping has decreased during the past several years, and the task of genotyping will not be as formidable as it used to be. On the other hand, phenotyping can be quite costly for the study of some traits—for example, formal measures of insulin resistance (the glucose clamp, insulin suppression, or intravenous glucose challenge) used in studying the risk for diabetes, measures of central or abdominal obesity (computed-tomography or magnetic-resonance-imaging scan), and imaging methods used in the study of arteriosclerosis and hypertension (either ultrasound assessment of coronary artery thickness or ventricular mass or assessment of coronary calcification) (King et al. 1992). Under the extreme discordant-sib-pair design paradigm, a large number of individuals have to be screened before the desired number of extreme discordant sib pairs can be selected for genotyping.

Received October 23, 1996; accepted for publication February 28, 1997.

Address for correspondence and reprints: Dr. Hongyu Zhao, Department of Epidemiology and Public Health, Yale University School of Medicine, 60 College Street, New Haven, CT 06520. E-mail: hongyu.zhao@yale.edu

0002-9297/97/6005-0023$02.00

When the phenotyping cost is not inexpensive compared with the genotyping cost, both kinds of cost need to be taken into consideration in the planning of a genetic-linkage study. Although the more extreme the discordant sib pairs are, the more powerful the design is, in the sense that the number of sib pairs needed to detect linkage is smaller, it might not be cost effective to use such very extreme sib pairs. For example, under one additive model studied by Risch and Zhang (1996), an estimated 21,441 sibs need to be phenotyped before 342 discordant sib pairs can be identified for detection of linkage under the top-10%/bottom-10% design, whereas only 7,481 sibs need to be phenotyped to collect the desired 673 discordant sib pairs for linkage analysis under the top-10%/bottom-30% design (see Risch and Zhang 1996, tables 1 and 4). When phenotyping costs as much as genotyping, it is apparent that the top-10%/bottom-30% design apparently costs much less than the top-10%/bottom-10% design, to achieve the same power to detect linkage. Although the effect that the cost of phenotyping, relative to the cost of genotyping, has on sampling strategy was mentioned by Risch and Zhang (1995, 1996), no detailed analysis of these considerations was reported.

In the planning of a genetic-linkage study, a certain power is desired, and the costs of genotyping and phenotyping often can be estimated reliably. A design that can balance the number of people to be phenotyped and the number of people to be genotyped, in order to minimize the cost while attaining the desired power, is always preferable. To address this practical issue, we systematically studied the effect of the cost of genotyping, relative to the cost of phenotyping, on the sampling strategy. Some general recommendations are made after a variety of genetic models and a range of cost ratios have been examined. A related question addressed in this paper is the sensitivity of the cost to the genetic model.

Methods

The genetic model used in the following analysis has been described by Risch and Zhang (1995). Assume a locus A with two alleles, A_1 and A_2. The allele frequency of A_1 is p, and that of A_2 is $q = 1 - p$. Let x_{i1} and x_{i2} be the observed trait values for the two sibs in the ith sib pair. The model is

$$x_{i1} = \mu + g_{i1} + e_{i1} ,$$
$$x_{i2} = \mu + g_{i2} + e_{i2} ,$$

where μ is the overall mean, g_{ij} is the genetic effect of locus A, and e_{ij} is the sum of other genetic and environmental contributions. The g_{ij} is a, d, and $-a$ for individuals having genotype A_1A_1, A_1A_2, and A_2A_2, respectively. Without loss of generality, the variance of e_{ij} is assumed to be 1. The (e_{i1}, e_{i2}) are assumed to follow a bivariate normal distribution, with the correlation between e_{i1} and e_{i2} being ρ. The additive and dominance variances from locus A under this model are $\sigma_a^2 = 2pq[a - d(p - q)]^2$ and $\sigma_d^2 = (2pqd)^2$, respectively. The heritability in the broad sense is $h = \sigma_g^2/(\sigma_g^2 + 1)$, where $\sigma_g^2 = \sigma_a^2 + \sigma_d^2$.

We first compare 15 symmetric sampling strategies in this paper, which fall into T*x*T*x*, T*x*B*x*, and B*x*B*x* classes; that is, *T* is the "top" and *B* is the "bottom" of the distribution of the quantitative trait of interest. When the trait is studied because it is associated with a certain disease, we assume that, without loss of generality, people with higher trait values are more likely to be affected by the disease associated with the trait. When the trait is associated with covariates, the top or bottom portion is defined on the basis of covariate-adjusted trait values—that is, covariate effects have been removed or reduced by regression methods. The T*x*T*x* represents the sample design in which both sibs to be genotyped are in the top *x*% of the population, with respect to the trait of interest. The T*x*B*x* design is the one with one sib in the top *x*% and with the other in the bottom *x*%. For the B*x*B*x* design, we genotype both those sibs in the bottom *x*%. For symmetric designs, the *x* values considered are 5, 10, 15, 20, and 25. For example, T05B05 represents the sampling strategy with one sib in the top 5% of the population and with the other sib in the bottom 5% of the population. These designs are called "symmetric" because the same value of *x*% is used for both sibs. To generalize further, we also considered T*x*T*y*, T*x*B*y*, and B*x*B*y* designs, where *x* and *y* may differ. We will discuss the symmetric designs first, and then we will consider the general T*x*B*y* designs. For asymmetric discordant designs T*x*B*y*, the *x* values considered are 10 and 20, and the *y* values considered are 10, 20, 30, 40, and 50. There are several ways to obtain concordant (T*x*T*x* or B*x*B*x*) and discordant (T*x*B*x*) sib pairs. In this paper, the probands are ascertained on the basis of either (1) clinical records with trait values either in the top *x*% (T*x*T*x* and T*x*B*x*) or in the bottom *x*% (B*x*B*x*) of the population, in which case there is no phenotyping cost for obtaining probands, or (2) screening of a specified population, in which case the identification of probands might be costly. The sibs of the probands are phenotyped, to collect concordant or discordant sib pairs. These concordant or discordant sib pairs then are genotyped for linkage analysis. In the case of ascertainment on the basis of clinical records, the phenotyping cost is due to the phenotyping of the probands' sibs, because we assume that the probands already have been phenotyped. For example, in the case of the T05B05 design with ascertainment on the basis of clinical records, the only phenotyping is done in sibs of those probands who are in the top 5%, and only sib pairs with the other sib in the bottom 5% are genotyped for

linkage analysis. On the other hand, when probands are ascertained through population screening, the phenotyping cost at the screening stage must be added into the total cost.

For a given sampling strategy, TxTx, TxBy, or BxBx, let Z_i be the probability that two sibs in the sample share i alleles by descent, where $i = 0$, 1, or 2. The Z_i can be calculated as described by Risch and Zhang (1995). Define X_i as the number of alleles shared (identical) by descent for the ith sib pair and let $\bar{X} = \Sigma_{i=1}^{n} X_i/2n$. On the basis of the central limit theorem, $\bar{X}$ is approximately normally distributed, with mean τ and variance $[\tau(1 - 2\tau) + Z_2]/2n$, where $\tau = Z_2 + Z_1/2$. When there is no linkage, $Z_0 = 1/4$, $Z_1 = 1/2$, and $Z_2 = 1/4$. We assume that there is no recombination between the marker and trait locus and that the marker is fully polymorphic. Although one often-cited advantage of sib-pair methods is that they make no assumption with regard to the mode of inheritance, it has been shown that different test statistics in sib-pair analysis do correspond to "optimal" test statistics under different models (Whittemore 1996). In this paper, we employ the test used by Risch and Zhang (1995) to test for linkage. This test is designed to test the null hypothesis: $\tau = 1/2$. The alternative hypothesis is either $\tau < 1/2$ (for the TxBx design) or $\tau > 1/2$ (for the TxTx or BxBx design). Therefore the power of a one-sided test against $\tau = 1/2$ is

$$\Phi\left(\frac{z_\alpha/2 + |\tau - 1/2|\sqrt{2n}}{\sqrt{\tau(1 - 2\tau) + Z_2}}\right),$$

where Φ is the cumulative standard normal distribution function and z_a is the upper αth percentile of the standard normal distribution. The sample size required for achieving $1 - \beta$ power is

$$\frac{1}{2}\left(\frac{z_{1-\beta}\sqrt{\tau(1 - 2\tau) + Z_2} - z_\alpha/2}{\tau - 1/2}\right)^2.$$

In this paper, the significance level for the test is set at $\alpha = .001$, and the power is set at $1 - \beta = .8$. The level of $\alpha = .001$ corresponds approximately to a LOD score of 2, which, under a whole-genome scan, would give suggestive evidence for linkage (Lander and Kruglyak 1995).

As mentioned above, it is assumed either that clinical records are readily available for ascertainment of probands or that the probands are obtained on the basis of screening a specified population. The sibs of the probands then are phenotyped, to yield a selected sample of sib pairs for genotyping. In the following discussion we assume that an arbitrarily large number of probands in the upper or lower tail of the distribution can be obtained.

When clinical records are readily available for ascertainment of probands, the cost of the study consists of two parts: phenotyping of the probands' sibs and genotyping of the selected sib pairs. Without loss of generality, assume that genotyping each person costs 1 unit and that phenotyping each person costs r units, where r is the phenotyping:genotyping cost ratio. Depending on the specific trait to be phenotyped, the phenotyping cost can be relatively inexpensive (e.g., height and weight or simple blood tests), moderately expensive (multiple blood tests), expensive (provocative testing), or very expensive (detailed physiological tests and imaging studies). In this paper, we consider r values of 0.02, 0.05, 0.2, 1, 5, 20, and 50. For example, if genotyping each marker costs \$1, a whole-genome scan using 300 markers will cost \$300 for each individual. Thus, the r's considered in such a case correspond to phenotyping costs of \$6, \$15, \$60, \$300, \$1,500, \$6,000, and \$15,000.

Use s to denote the conditional probability that the sib of the proband satisfies the selection criterion (either falling in the top x% or bottom x% of the population), given that the proband is in the top x% or bottom x% of the population. If the required sample size is n, the average number of sibs to be screened to get n desired sib pairs is n/s. As above, use x_{i1} and x_{i2} to denote the trait values of the proband and of his or her sib, respectively. Conditional probability s can be calculated easily if we assume that the conditional distribution of x_{i2} given x_{i1}, denoted as $f(x_{i2}|x_{i1})$, is the same as the conditional distribution of y_{i2} given y_{i1}, denoted as $f(y_{i2}|y_{i1})$, where y_{i2} and y_{i1} are the trait values of two sibs selected at random from the general population. The assumption that $f(x_{i2}|x_{i1})$ is the same as $f(y_{i2}|y_{i1})$ means that, if both a proband identified on the basis of clinical records and a person in the general population have the same trait value, then their sibs should have the same conditional trait distribution. In practice, this can be a reasonable assumption. For example, HDL levels have been measured in unselected nuclear families, and correlations have been sought between family members. The sib-sib HDL-cholesterol levels correlated, with an r of .2–.3, and parent-child levels also correlated, with an r of .1–.3 (Namboodiri et al 1983). On the basis of the assumption that the HDL levels for relative pairs have a bivariate normal distribution, the conditional probability that one individual is in the top decile (or bottom decile) of the distribution, given that his or her relative is in the top decile (or bottom decile), is 14%–23%, for $r = .1$–.3. In the Princeton School District Family Study (Laskarzewski et al. 1982), for 23 probands with bottom-decile HDL levels, 16% of their relatives had the same phenotype. For 23 probands with top-decile HDL levels, 22% of their relatives had the same phenotype. These observed proportions agree well with the proportions predicted on the basis of a bivariate normal distribution.

To calculate s for the TxTx design, the probability that both sibs are in the top x percent of the population, P(TxTx), can be calculated as illustrated by Risch and Zhang (1995); therefore, $s = P(\text{T}x|\text{T}x) = P(\text{T}x\text{T}x)/(x\%)$. The calculation of s for the BxBx design is similar. For the TxBx design, let P(TxBx) denote the probability that one sib is in the top x% and the other sib in the bottom x% of the population. Then the conditional probability that the proband's sib is in the bottom x%, given that the proband is in the top x%—$P(\text{B}x|\text{T}x)$—is $P(\text{T}x\text{B}x)/(2x\%) = P(\text{T}x\text{B}x)/(2x) \times 100$. There is a factor 2 because the order of the two sibs is irrelevant in the calculation of P(TxBx). Given s, the total cost T, which consists of the genotyping cost for n sib pairs ($2n$) and of the phenotyping cost for screening the probands' sibs (rn/s), is $n(2 + r/s)$.

When probands are obtained on the basis of screening a specified population, the phenotyping cost for the first-round screening must be included, in addition to the cost of genotyping selected sib pairs and phenotyping the probands' sibs. To have n sib pairs for genotyping under the TxTx, BxBx, or TxBx design, $m = n/s$ individuals are needed for the second-round screening; therefore, $m/(x\%) = 100m/x$ people must be screened in the first round. The total cost is $T = 2n + rn/s + rm/(x\%) = n[2 + r/s + r/(sx\%)] = n[2 + r/s + 100r/(sx)]$.

Results

Symmetric Designs

For each ascertainment scenario, based on either clinical records or population screening, the total cost under each symmetric sampling strategy was evaluated, under both additive and dominant models, with heritability h of .05, .10, and .20, gene frequency p of .1, .4, and .8, and residual correlation ρ (due to the effects of environment and other genes) set at 0 or .4. Therefore, within each model type, additive or dominant, $3 \times 3 \times 2 = 18$ possible combinations of h, p, and ρ were considered, resulting in a total of 36 genetic models. The cost ratio r varied from 0.02, 0.05, 0.2, 1, 5, 20, and 50. In the following, we consider, separately, ascertainment on the basis of clinical records and ascertainment on the basis of population screening.

Ascertainment on the basis of clinical records.—For a cost ratio $r = 1$, table 1 (18 additive models) and table 2 (18 dominant models) present both the best design among five x values and its associated cost within each design type: TxTx, TxBx, or BxBx, for each model. It can be seen from tables 1 and 2 that, under the TxTx and the BxBx designs, total cost varies considerably. Sometimes the cost of the TxTx design is an order of magnitude higher than that of the BxBx design, and sometimes the cost of the BxBx design is similarly higher than that of the TxTx design. For the same model type (additive or dominant), the same h, and the same residual correlation ρ, the cost within the same design can differ by an order of magnitude, for a different allele frequency p.

In contrast, the minimal cost under the TxBx design exhibits very good robustness to the model type, allele frequency p, and even residual correlation ρ, with the exception of the dominant model with a high allele frequency, $p = .8$. For example, when the heritability from the locus studied is .1, the cost is ~15,000 units across all p and ρ values, under both additive and dominant models, except under the dominant model with $p = .8$. The robustness in terms of the cost of the TxBx design, with respect to model type and p, can be appreciated by noting that the sample size and the number of sibs needed to be screened for a different model type and p are very similar when h and ρ are fixed, as demonstrated in the tables published by Risch and Zhang (1996). It is interesting that, although the sample size differs by as much as fivefold and although the x giving the minimal cost differs for different ρ, the minimal cost is similar for $\rho = 0$ and $\rho = .4$, although the cost when $\rho = .4$ is always slightly less than the cost when $\rho = 0$. The x that gives the lowest cost within the TxBx class, under models with $\rho = 0$, is 5 or 10, whereas the best x under models with $\rho = .4$ is 20. The exception, the dominant model with high allele frequency p, corresponds to the case in which the population is a mixture of two normal distributions, with 96% of the population from one normal distribution and 4% from the other distribution. This special case also was noted by Risch and Zhang (1996), and it was argued that this case can be identified a priori by evaluation of the role of dominance variance for the trait.

Tables 1 and 2 show that, when $r = 1$, the cost under the TxBx design is always between that of the TxTx design and that of the BxBx design. This is in contrast to the results seen when only the genotyping cost is of concern, in which the sample size of extreme discordant pairs is the key factor in the determination of the design. In that case, the TxBx design often results in the smallest number of sib pairs necessary for genotyping. When both phenotyping and genotyping costs are considered and when phenotyping costs are as much as genotyping, under a given model, the TxBx design is not the most cost effective. But overall the TxBx design is still the best design among the three sampling classes. This is so because, although it is not the one with the lowest cost compared with the other two alternatives under a given model, the increase in cost can be justified by the fact that (1) our knowledge regarding the true genetic model for the trait of interest is often extremely limited and (2) as illustrated in tables 1 and 2, the TxBx design has good robustness properties whereas the other two designs exhibit great variability.

Table 1

Optimal Design within Each Sampling Class, T*x*T*x*, T*x*B*x*, and B*x*B*x*, under Additive Models

	Optimal Design[a]					
	ρ = 0			ρ = .4		
p	h = .05	h = .1	h = .2	h = .05	h = .1	h = .2
.1	T15T15 (193)	T20T20 (73)	T25T25 (31)	T05T05 (146)	T05T05 (59)	T15T15 (27)
	T05B05 (63)	T05B05 (16)	T10B10 (4)	T20B20 (54)	T20B20 (14)	T15B15 (4)
	B05B05 (20)	B05B05 (4)	B05B05 (1)	B05B05 (22)	B05B05 (4)	B05B05 (1)
.4	T10T10 (78)	T05T05 (22)	T10T10 (6)	T05T05 (61)	T05T05 (17)	T05T05 (5)
	T05B05 (61)	T10B10 (15)	T10B10 (4)	T20B20 (52)	T20B20 (13)	T20B20 (3)
	B05B05 (54)	B05B05 (14)	B05B05 (4)	B05B05 (45)	B05B05 (11)	B05B05 (3)
.8	T05T05 (33)	T05T05 (7)	T05T05 (2)	T05T05 (31)	T05T05 (7)	T05T05 (2)
	T05B05 (61)	T10B10 (15)	T10B10 (4)	T20B20 (52)	T20B20 (13)	T20B20 (3)
	B10B10 (122)	B15B15 (40)	B15B15 (14)	B05B05 (92)	B05B05 (31)	B05B05 (12)

NOTE.—The phenotyping:genotyping ratio r is 1.

[a] Data are x values resulting in the least cost. Data in parentheses are costs (in $1,000s); the cost for screening of probands is not included.

We also looked into other cost ratios r and found similar patterns—namely, the following: (1) the cost under the T*x*T*x* and the B*x*B*x* designs exhibited considerable variation even when only p varies while the other model parameters remain fixed; and (2) for the same h, the cost under the T*x*B*x* design is very robust under different models and, on average, is much lower than the cost under the T*x*T*x* and the B*x*B*x* designs. As expected, when r = 5 the optimal x for each design class is the same as or higher than that when r = 1. This is because there is a larger fiscal penalty for having too large a population to screen. The cost under the T*x*B*x* design is still always between the cost under the T*x*T*x* design and that under the B*x*B*x* design. When r = .2, the optimal x is smaller, since phenotyping costs less. For the dominant model with ρ = .4 and p = .4, the cost under the T*x*B*x* design is the smallest among three design classes, whereas for all other models it is still between the costs under the other two design classes. Similar patterns also hold for different significance levels α and different power β.

Since the T*x*B*x* design is more cost effective overall, the issue of choosing x to minimize the cost while achieving the desired power was examined. We calculated the cost for each of the five possible x values, for each of the 36 models, and for each of the seven phenotyping:genotyping cost ratios r. The results are summarized below.

Table 2

Optimal Design within Each Sampling Class, T*x*T*x*, T*x*B*x*, and B*x*B*x*, under Dominant Models

	Optimal Design[a]					
	ρ = 0			ρ = .4		
p	h = .05	h = .1	h = .2	h = .05	h = .1	h = .2
.1	T15T15 (174)	T20T20 (65)	T25T25 (28)	T05T05 (132)	T05T05 (53)	T15T15 (24)
	T05B05 (62)	T10B10 (16)	T10B10 (4)	T20B20 (53)	T20B20 (13)	T20B20 (3)
	B05B05 (26)	B05B05 (6)	B05B05 (1)	B05B05 (25)	B05B05 (5)	B05B05 (1)
.4	T05T05 (50)	T10T10 (13)	T10T10 (4)	T05T05 (41)	T05T05 (10)	T05T05 (3)
	T10B10 (61)	T10B10 (15)	T10B10 (4)	T20B20 (52)	T20B20 (13)	T20B20 (3)
	B10B10 (99)	B15B15 (32)	B20B20 (11)	B05B05 (75)	B05B05 (24)	B10B10 (9)
.8	T05T05 (5)	T05T05 (1)	T05T05 (.2)	T05T05 (7)	T05T05 (1)	T05T05 (.3)
	T05B05 (75)	T05B05 (22)	T05B05 (7)	T15B15 (60)	T10B10 (16)	T10B10 (5)
	B20B20 (787)	B25B25 (509)	B25B25 (441)	B15B15 (668)	B25B25 (401)	B25B25 (326)

NOTE.—See footnotes to table 1.

Under the additive models with residual correlation $\rho = 0$, in general the T05B05 design is the best one when phenotyping costs are less than genotyping costs. When phenotyping is more expensive, the best design is T10B10. For the higher heritability value $h = .2$, the T10B10 design outperforms the T05B05 design even when phenotyping is somewhat less expensive.

For the additive models with correlation $\rho = .4$, the T20B20 design is the best for all nine models studied, when phenotyping is more expensive than genotyping. The T15B15 design is the best when the phenotyping:genotyping cost ratio r is .05–1. When phenotyping is much cheaper than genotyping ($r = .02$–.05), the T10B10 design is the best.

For the dominant models with residual correlation $\rho = 0$, the T05B05 design dominates other designs, for all cost ratios r considered, when the allele frequency p is .8. For other allele frequencies, in general the T05B05 design is optimal when phenotyping costs less than genotyping, and the T10B10 design is the best when phenotyping costs more.

For the dominant models with correlation $\rho = .4$, when $p = .1$ and .4 the T20B20 design is the best when phenotyping is more expensive, the T15B15 design is the best when the phenotyping:genotyping cost ratio is moderate, and the T10B10 design is the best when phenotyping is much cheaper than genotyping; when $p = .8$, the best design is T10B10, for most cost ratios, except when phenotyping is much cheaper than genotyping, in which case the T05B05 design is the best.

For the models examined, residual correlation ρ plays a more important role than heritability h, allele frequency p, and model type (additive or dominant), in determining the optimal sampling strategy, except in the case of the dominant model with high allele frequency. When the cost of phenotyping is approximately the same as or more than genotyping, the best design overall is T10B10 when ρ is 0 and T20B20 when ρ is .4. Because, for most quantitative traits under linkage study, the overall heritability h (from all genes) is .1–.5, and because no major genes are expected to contribute more than half of the total variation, models with $\rho = .4$ might be closer to reality than are models with $\rho = 0$. Therefore, results for $\rho = .4$ are of more practical relevance. In these cases, the T15B15 or T20B20 design is most cost effective when the phenotyping cost is $\geq 20\%$ of the genotyping cost; and the T15B15 design is most cost effective otherwise, unless phenotyping is much cheaper than genotyping, in which case the best design is T10B10. For the models considered, the cost varies only a little for the same design, for models with the same h and ρ values.

Ascertainment on the basis of population screening.—When clinical records are not readily accessible, a specified population has to be phenotyped for identification of probands; for example, for the T10T10 design, only the top 10% of the individuals in the initial population will enter into the next stage of the genetic-linkage study. This will add considerably more cost to the study, especially when the phenotyping cost is expensive.

For $r = 1$, as expected, the optimal x for each design is larger when the phenotyping cost for the initial population is taken into account than when such phenotyping cost is not included. Most optimal x's are 20 or 25. The TxTx and BxBx designs still exhibit considerable variation for different allele frequencies, when other model parameters are fixed. The minimal cost under the TxBx design is robust with respect to allele frequency and model type, but it does differ for different ρ values. When $\rho = 0$, the cost under the TxBx design still falls between the cost under the TxTx design and that under the BxBx design. When $\rho = .4$, the cost under the TxBx design is the minimal one among three types of designs, for more than half of the models examined. The general patterns for other cost ratios are the same as those for $r = 1$.

As for the previous ascertainment scheme, we studied the optimal x under the TxBx design, for each genetic model and each phenotyping:genotyping cost ratio. Compared with the optimal x when clinical records are readily accessible, the optimal x in this case is more consistent for different cost ratios. The residual correlation ρ also plays an important role in determining the best x. When $\rho = .4$, the T25B25 design is the optimal one unless phenotyping costs much less than genotyping. When $\rho = 0$, under most models the T25B25 design is the optimal one when phenotyping costs more, and the T20B20 design is the optimal one when the phenotyping:genotyping cost ratio is .1–1. As mentioned above, because models with $\rho = .4$ might be closer to reality than are models with $\rho = 0$, the T25B25 design for x varying from 5, 10, 15, 20, to 25 seems to be the optimal one when probands have to be identified through screening of a defined population. For illustration, the relative costs for nine additive models with $\rho = .4$ are plotted in figure 1. In figure 1, the cost under the T25B25 design is used as the baseline cost, and the costs under other designs—$x = 5$, 10, 15, and 20—are shown as the ratio between the cost of the TxBx design and that of the T25B25 design. The costs for the same design are similar for additive and dominant models with the same h and ρ values, regardless of the allele frequency p.

Asymmetric Designs

As noted in the discussion of symmetric designs, concordant designs (TxTx or BxBx) exhibit considerable variation. In this section, we consider asymmetric discordant designs (TxBy). The x value was chosen as either 10 or 20, and the y value was varied from 10, 20, 30, 40, to 50, to cover a wide range of possibilities. The

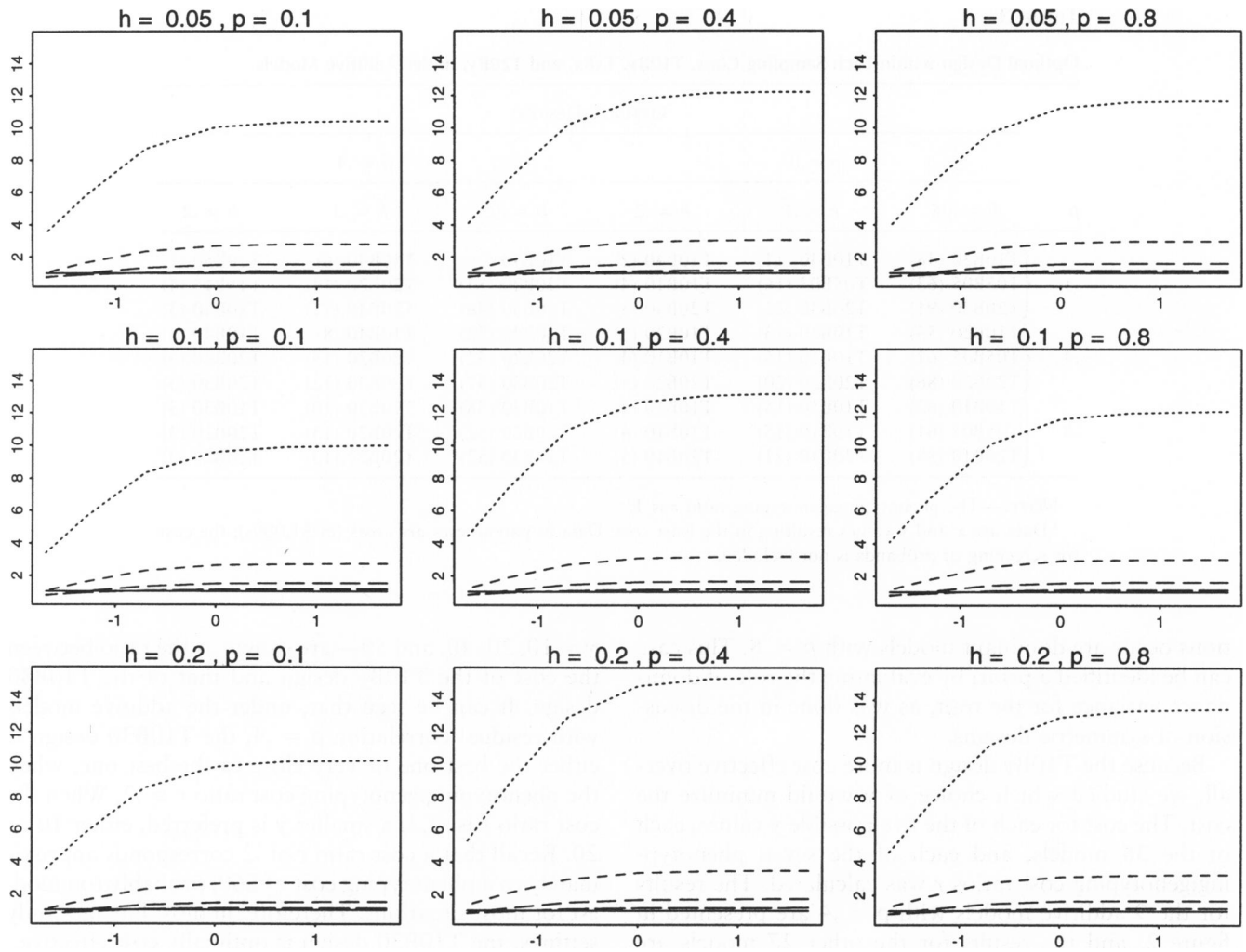

Figure 1 Relative costs among different TxBx designs, for nine additive models with residual correlation ρ = .4, when the phenotyping cost for screening of probands is included—that is, probands are ascertained on the basis of population screening. The nine (3 × 3) panels corresponding to nine models are arranged as follows: three rows, from top to bottom, correspond to three heritabilities—h = .05, .1, and .2, respectively—three columns, from left to right, correspond to three allele frequencies—p = .1, .4, and .8, respectively. For example, the model for the panel in the middle (row 2, column 2) has h = .1 and p = .4. For each plot, the X-axis is the $\log_{10}$ of the phenotyping:genotyping cost ratio. The Y-axis is the relative cost under the TxBx design, compared with that under the T25B25 design. The five curves correspond to x = 5, 10, 15, 20, and 25: T05B05 (---), T10B10 (– – –), T15B15 (— — —), T20B20 (—— —— ——), and T25B25 (——————).

same 18 additive models and 18 dominant models were studied. As before, we separately considered ascertainment on the basis of clinical records and ascertainment on the basis of population screening.

Ascertainment on the basis of clinical records.—For a cost ratio r = 1, table 3 (additive models) and table 4 (dominant models) present the optimal design and its associated cost for symmetric designs (TxBx) and for asymmetric designs of different x values—that is, the T10By and the T20By designs. When p = .1 or .4, for the same h, p, and ρ, the optimal y's for the T10By and T20By designs are similar. The minimal cost under the T10By design is always smaller than the minimal cost under the T20By design, for all 36 models considered. Therefore, the T10By design is more cost effective than the T20By design, when r = 1. Comparison between the T10By design with the symmetric TxBx design, in tables 3 and 4, shows that the T10By design is more cost effective than the TxBx design, except in the dominant models with high p values. Unlike the case for the TxBx designs, where the minimal costs are similar between models with the same h values and different ρ values, the minimal cost under the model with ρ can be half of the minimal cost under the model with ρ = 0.

Examination of other cost ratios shows that, for most models, the minimal cost under the T10By designs is less than the minimal costs under the TxBx and T20By designs, when r > .2. The models under which excep-

Table 3

Optimal Design within Each Sampling Class, T10B*y*, T*x*B*x*, and T20B*y*, under Additive Models

	Optimal Design[a]					
	ρ = 0			ρ = .4		
p	*h* = .05	*h* = .1	*h* = .2	*h* = .05	*h* = .1	*h* = .2
.1	T10B30 (48)	T10B30 (11)	T10B40 (2)	T10B40 (26)	T10B40 (6)	T10B50 (1)
	T05B05 (63)	T05B05 (16)	T10B10 (4)	T20B20 (54)	T20B20 (14)	T15B15 (4)
	T20B30 (91)	T20B30 (22)	T20B30 (5)	T20B30 (46)	T20B40 (11)	T20B40 (3)
.4	T10B20 (53)	T10B20 (13)	T10B20 (3)	T10B40 (32)	T10B40 (8)	T10B40 (2)
	T05B05 (61)	T10B10 (15)	T10B10 (4)	T20B20 (52)	T20B20 (13)	T20B20 (3)
	T20B20 (86)	T20B20 (20)	T20B20 (5)	T20B30 (47)	T20B30 (12)	T20B30 (3)
.8	T10B10 (62)	T10B10 (15)	T10B10 (4)	T10B30 (38)	T10B30 (10)	T10B30 (3)
	T05B05 (61)	T10B10 (15)	T10B10 (4)	T20B20 (52)	T20B20 (13)	T20B20 (3)
	T20B10 (88)	T20B10 (21)	T20B10 (5)	T20B30 (52)	T20B20 (13)	T20B20 (3)

NOTE.—The phenotyping:genotyping ratio *r* is 1.

[a] Data are *x* and *y* values resulting in the least cost. Data in parentheses are costs (in $1,000s); the cost for screening of probands is not included.

tions occur are dominant models with $p = .8$. This case can be identified a priori by evaluating the role of dominance variance for the trait, as was done in the discussion of symmetric designs.

Because the T10B*y* design is more cost effective overall, we studied which choice of *y* would minimize the cost. The cost for each of the five possible *y* values, each of the 36 models, and each of the seven phenotyping:genotyping cost ratios *r* was calculated. The results for the 9 additive models with $\rho = .4$ are presented in figure 2, and the results for the other 27 models are summarized below.

In figure 2, the cost under the T10B30 design is used as the baseline cost, and the costs under other designs—*y* = 10, 20, 40, and 50—are shown as the ratio between the cost of the T10B*y* design and that of the T10B30 design. It can be seen that, under the additive models with residual correlation $\rho = .4$, the T10B30 design is either the best one or very close to the best one, when the phenotyping:genotyping cost ratio $r \geq .2$. When the cost ratio *r* is <.2, a smaller *y* is preferred, either 10 or 20. Recall that a cost ratio *r* of .2 corresponds approximately to a phenotyping cost of $60, probably too modest for most situations. Therefore, in most linkage-study settings, the T10B30 design is optimally cost effective.

For the additive models with correlation $\rho = 0$, the T10B30 design is either the best or close to the best, for all nine models studied, when phenotyping is more

Table 4

Optimal Design within Each Sampling Class, T10B*y*, T*x*B*x*, and T20B*y*, under Dominant Models

	Optimal Design[a]					
	ρ = 0			ρ = .4		
p	*h* = .05	*h* = .1	*h* = .2	*h* = .05	*h* = .1	*h* = .2
.1	T10B30 (48)	T10B30 (11)	T10B40 (2)	T10B40 (26)	T10B40 (6)	T10B50 (1)
	T05B05 (62)	T10B10 (16)	T10B10 (4)	T20B20 (53)	T20B20 (13)	T20B20 (3)
	T20B30 (88)	T20B30 (21)	T20B40 (5)	T20B30 (46)	T20B40 (11)	T20B40 (3)
.4	T10B20 (58)	T10B20 (14)	T10B20 (3)	T10B30 (36)	T10B30 (9)	T10B30 (2)
	T10B10 (61)	T10B10 (15)	T10B10 (4)	T20B20 (52)	T20B20 (13)	T20B20 (3)
	T20B20 (84)	T20B20 (20)	T20B20 (4)	T20B30 (50)	T20B30 (12)	T20B30 (3)
.8	T10B10 (90)	T10B10 (31)	T10B10 (13)	T10B20 (59)	T10B10 (16)	T10B10 (5)
	T05B05 (75)	T05B05 (22)	T05B05 (7)	T15B15 (60)	T10B10 (16)	T10B10 (5)
	T20B10 (99)	T20B10 (31)	T20B10 (13)	T20B10 (62)	T20B10 (16)	T20B10 (5)

NOTE.—See footnotes to table 3.

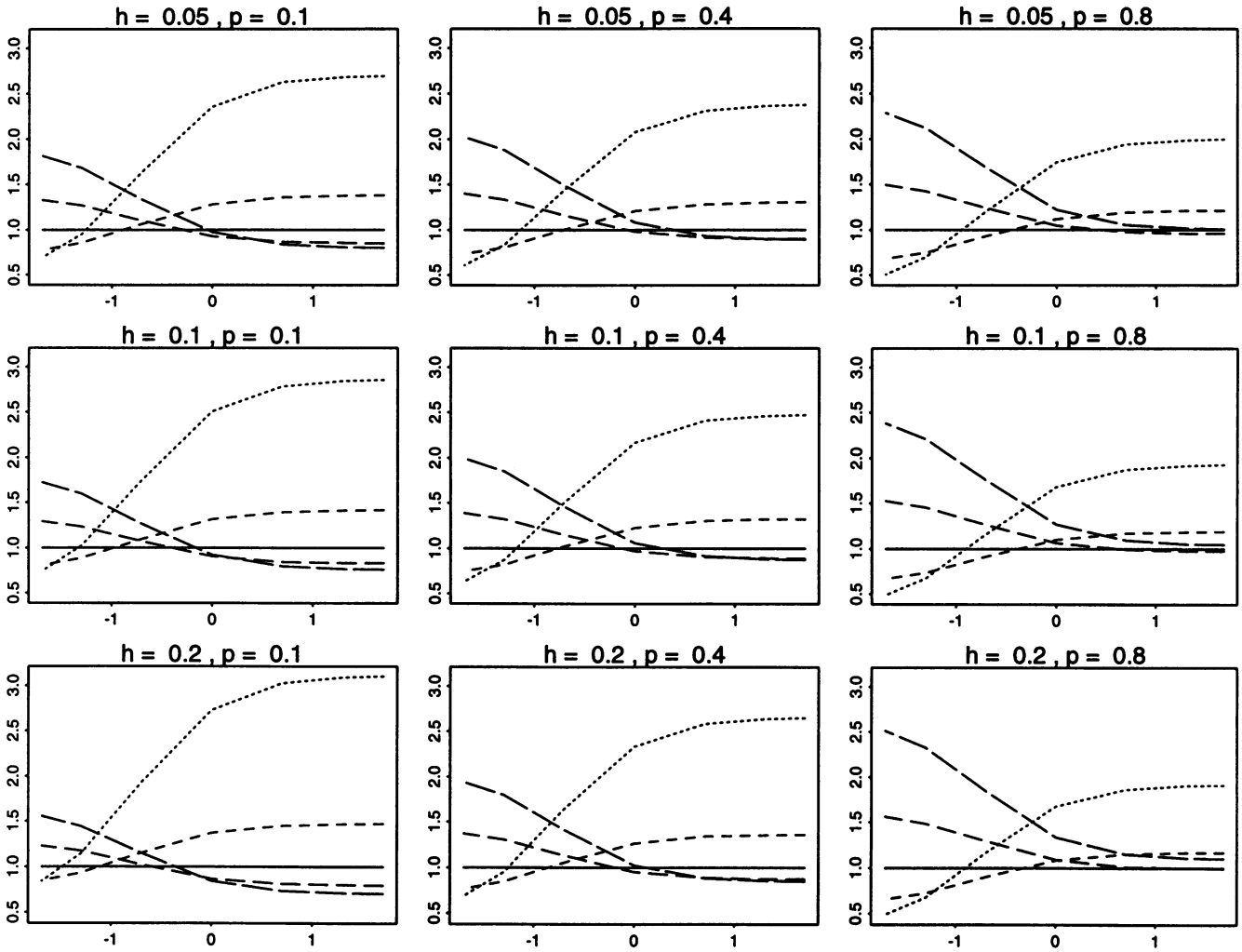

Figure 2 Relative costs among different T10B*y* designs, for nine additive models with residual correlation ρ = .4, when the phenotyping cost for screening of probands is not included—that is, probands are ascertained on the basis of clinical records. The nine (3 × 3) panels corresponding to nine models are arranged as are the panels in figure 1. For each plot, the *X*-axis is the $\log_{10}$ of the phenotyping:genotyping cost ratio. The *Y*-axis is the relative cost under the T10B*y* design, compared with that under the T10B30 design. The five curves correspond to *y* = 10, 20, 30, 40, and 50: T10B10 (· · ·), T10B20 (– – –), T10B30 (— — —), T10B40 (— — — —), and T10B50 (————).

expensive than genotyping. The T10B10 design is the best overall when the phenotyping:genotyping cost ratio *r* is <1.

For the dominant models with residual correlation ρ = 0, the T10B10 design dominates other designs, for all cost ratios considered, when the allele frequency *p* is .8. For other allele frequencies, in general the T10B20 design is either the optimal one or close to the optimal one when phenotyping costs ≥.2 of the cost of genotyping, and the T10B10 design is the best otherwise. For these nine models, although overall the T10B30 design does not perform as well as the T10B20 design, the costs under these two designs are rather similar.

For the dominant models with correlation ρ = .4 and *p* = .1, the T10B40 design is the best overall when phenotyping is more expensive, and the T10B30 design is the best when phenotyping costs less than genotyping. When ρ = .4 and *p* = .4, in most cases the T10B30 design is best when the phenotyping:genotyping cost ratio *r* is ≥.2, and the T10B20 is the best otherwise. When *p* = .8, the best design is T10B10, for most cost ratios.

In summary, among the T10B*y*, T20B*y*, and T*x*B*x* designs, when clinical records are readily available, the T10B*y* designs outperform the T20B*y* and T*x*B*x* designs. Overall, when the phenotyping:genotyping cost ratio *r* is ≥.2—that is, phenotyping costs ~≥$60—if there is a single design that we must choose, it should be the T10B30 design. This design and the T10B10 design were discussed in detail by Risch and Zhang (1996). The costs for the T10B30 design is not as robust as the cost for symmetric designs. The higher the *p*, the more the design will cost. For the same *h*, *p*, and ρ, the costs are very

similar for additive and dominant models, except when p is .8.

Ascertainment on the basis of population screening.—As for symmetric designs, the total cost is considerably more when probands have to be identified on the basis of population screening. When all 36 models are examined, with seven cost ratios, among the T10B*y*, T20B*y*, and T*x*B*x* designs, the best design in the T10B*y* class is the most cost effective when $p = .1$, the best design in the T*x*B*x* class is the most cost effective when $p = .4$, and the best design in the T20B*y* class is the most cost effective when $p = .8$, for all cost ratios. For most models, the minimal costs under the T*x*B*x* designs are similar to the minimal costs under the T20B*y* designs, when $p = .8$. Because it is not known a priori that the allele frequency p is low, symmetric designs, T*x*B*x*, are better overall than the other two (T10B*y* and T20B*y*) designs. Therefore, we come to the same general conclusions as were reached above, in the discussion of the symmetric designs for ascertainment on the basis of population screening—that is, the T25B25 design is the optimal one under most situations.

Discussion

In this paper, we have studied the effect of the phenotyping:genotyping cost ratio on the sampling strategy when sib pairs are used to map quantitative-trait loci. Risch and Zhang (1995, 1996) showed that extreme discordant-sib-pair design is the most powerful design for most genetic models examined. But a large number of individuals might have to be screened in order to obtain a given number of sib pairs for linkage analysis under this study design. When the phenotyping cost is not inexpensive compared with the genotyping cost, resources have to be properly allocated for genotyping and phenotyping, to make the overall design maximally cost effective.

Thirty-six genetic models have been studied in this paper. For both ascertainment through clinical records and ascertainment through population screening, the cost was calculated for 15 symmetric sampling strategies and for phenotyping:genotyping cost ratios, under each genetic model. It was found that both the T*x*T*x* and B*x*B*x* designs, in which, with respect to the trait of interest, both sibs to be phenotyped are either in the top x% (T*x*T*x*) or the bottom x% (B*x*B*x*) of the population, have great variability in cost when only the allele frequency p is varied while other model parameters remain fixed. This makes the planning of a genetic-linkage study difficult, because allele frequency is often unknown before linkage analysis is performed. On the other hand, the minimal cost under the T*x*B*x* design, in which one sib is in the top x% and the other sib is in the bottom x% of the population, shows very good robustness against both model type (additive or dominant) and allele frequency. This cost is also robust to residual correlation for the same heritability, when ascertainment is performed on the basis of clinical records. This robustness comes as a nice and unexpected property of the T*x*B*x* design under this ascertainment scheme, because, although, for different residual correlations, the sample size necessary for detection of linkage does vary by several fold, after the phenotyping cost is taken into account, the total cost remains approximately the same, although the optimal x often differs. The only exception is when the genetic model is dominant with high allele frequency, in which case the population consists of two normal distributions, with >95% of the population from one distribution. Although the T*x*B*x* design is not always the one with the lowest cost under each model, the cost is not very far from that of the best design and, on average, is much smaller than the cost under either the T*x*T*x* design or the B*x*B*x* design. Therefore the T*x*B*x* design stands out as the best choice among the T*x*T*x*, T*x*B*x*, and B*x*B*x* designs, when the total linkage-study cost is considered.

General T*x*B*y* designs, where x and y may differ, also were considered in order to find the discordant design with the minimal cost. (1) For ascertainment on the basis of clinical records, under the models studied, symmetric designs (T*x*B*x*) and the T20B*y* designs are inferior to the T10B*y* designs. When the phenotyping:genotyping cost ratio is $\geq .2$, the T10B30 design appears to be the most cost effective for most models. (2) For ascertainment on the basis of phenotyping of a specified population, to obtain the probands, symmetric designs, in general, perform better than the T10B*y* and T20B*y* designs. When the residual correlation $\rho = 0.4$, the optimal design is T25B25. When $\rho = 0$, the best design is T25B25 when phenotyping costs more and is T20B20 when the phenotyping:genotyping cost ratio is .1–1. Because, for most models, the cost ratio under the T20B20 design and the T25B25 design is very close to 1, in practice both designs can be considered.

The statistical test used here is for the null hypothesis that the average allele sharing by descent is 1/2. There is evidence that other tests may be more powerful (H. Zhao and H. Zhang, unpublished data). When a more powerful test is employed in the linkage analysis, the required sample size can be reduced. For simplicity, these more powerful tests have not been considered in this paper. Therefore, the cost figures in this paper are somewhat conservative, in that the sample size may be able to be reduced further, thus saving costs by employing more powerful tests.

For symmetric designs, the largest x studied in this paper is 25. For the T*x*B*x* design, when the cost under the T25B25 design is smaller than that under the

T20B20 design, the ratio is close to 1, for most cases; that is, the two designs will cost approximately the same. Therefore, a larger x was not considered in this paper, because a larger x will not lead to large cost savings for a genetic-linkage study.

We have assumed that there is no recombination between the marker and trait locus and that the marker is fully polymorphic. There are easy ways to adjust the sample size when the recombination fraction is not 0 and when markers are not fully polymorphic (Risch and Zhang 1996). This will increase the cost but will not change our general conclusions regarding the sampling strategy. The significance level considered in this paper is .001, which approximately corresponds to suggestive linkage in a whole-genome scan. Other significance levels, $\alpha = .01$ and $\alpha = .0001$, also were studied, and they yielded essentially the same results.

It is very common for several traits to be analyzed simultaneously in a single genetic study. If we apply the above-recommended sampling strategy to each trait, some sib pairs might be genotyped in the study of one trait, whereas other sib pairs in the sample might be used for the study of another trait. If, for each trait, the TxBy design is used, then, as the number of traits increases, more and more sib pairs would have to be genotyped. Apparently the best strategy will no longer be the same, and this will depend on how these traits are related. We will be addressing this issue in a future study.

Acknowledgments

We thank two referees for their constructive comments, which greatly helped to improve the presentation of the manuscript, and we thank Rita Cantor for her insightful and stimulating discussions. This work was supported in part by grants HG01093 (to H. Zhao) and HD30712 (to H. Zhang) from the National Institutes of Health and by the Cedars-Sinai Board of Governors' Chair in Medical Genetics (J.I.R.).

References

Carey G, Williamson JA (1991) Linkage analysis of quantitative traits: increased power by using selected samples. Am J Hum Genet 49:786–796

Eaves L, Meyer J (1994) Locating human quantitative trait loci: guidelines for the selection of sibling pairs for genotyping. Behav Genet 24:443–455

Fulker DW, Cardon LR (1994) A sib-pair approach to interval mapping of quantitative trait loci. Am J Hum Genet 54: 1092–1103

Haseman JK, Elston RC (1972) The investigation of linkage between a quantitative trait and a marker locus. Behav Genet 2:3–19

King RA, Rotter JI, Motulsky AG (1992) The genetic basis of common diseases. Oxford University Press, New York

Kruglyak L, Lander ES (1995) Complete multipoint sib-pair analysis of qualitative and quantitative traits. Am J Hum Genet 57:439–454

Lander E, Kruglyak L (1995) Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. Nat Genet 11:241–247

Laskarzewski PM, Khoury P, Morrison JA, Kelly K, Mellies MJ, Glueck CJ (1982) Prevalence of familial hyper- and hypolipoproteinemias: the Princeton school district family study. Metabolism 31:558–577

Namboodiri KK, Green PP, Kaplan EB, Tyroler HA, Morrison JA, Chase GA, Elston RC, et al (1983) Family aggregation of high density lipoprotein cholesterol: Collaborative Lipid Research Clinics Program family study. Arteriosclerosis 3: 616–626

Risch N, Zhang H (1995) Extreme discordant sib pairs for mapping quantitative trait loci in humans. Science 268: 1584–1589

——— (1996) Mapping quantitative trait loci with extreme discordant sib pairs: sampling considerations. Am J Hum Genet 58:836–842

Whittemore AS (1996) Genome scanning for linkage: an overview. Am J Hum Genet 59:704–716

Zhang H, Risch N (1996) Mapping quantitative-trait loci in humans by use of extreme concordant sib pairs: selected sampling by parental phenotypes. Am J Hum Genet 59: 951–957