Am. J. Hum. Genet. 63:880–888, 1998

Optimal Ascertainment Strategies to Detect Linkage to Common Disease Alleles

Judith A. Badner, Elliot S. Gershon,* and Lynn R. Goldin

Clinical Neurogenetics Branch, National Institute of Mental Health, Bethesda

Summary

Traditionally, extended pedigrees with many affected individuals have been studied for the purpose of detection of linkage. For traits caused by a rare susceptibility allele, this is a productive strategy. However, this sampling strategy may not work well for traits determined by multiple loci in which one or more have common susceptibility alleles. We simulated three single-additive-locus models of inheritance and two-locus models with additive or multiplicative interactions, all with rare or common susceptibility alleles. A trait locus was linked, with no recombination, to a marker locus with four equally frequent alleles. Family structure varied, but the total number of affected individuals was held constant. Two generations of individuals were genotyped. We used three nonparametric affected-sib-pair programs and two nonparametric pedigree-analysis programs to perform linkage analysis. For single-locus, additive, and multiplicative models, we found that, when the susceptibility allele was rare, (frequency .0025), extended pedigrees with first or second cousins had the most power for detection of linkage. However, when the susceptibility allele was common in the single-locus, additive, and multiplicative two-locus models (frequency .25), extended pedigrees were no more powerful than nuclear families. There was also a decrease in power when the pedigrees had a greater number of affected individuals, more so for the single-locus and multiplicative models than for the additive model. We conclude that for single-locus, additive, and multiplicative models of qualitative traits with common alleles, there is no benefit to the collection of extended pedigrees, and there may be a loss of power in the collection of pedigrees with many affected individuals.

Received January 30, 1998; accepted for publication July 2, 1998; electronically published August 14, 1998.

Address for correspondence and reprints: Dr. Judith A. Badner, National Institute of Mental Health Intramural Research Program, Building 10, Room 3N218, 10 Center Drive, MSC 1274, Bethesda, MD 20892-1274. E-mail: jbadner@helix.nih.gov

* Present affiliation: Department of Psychiatry, University of Chicago, Chicago.

0002-9297/98/6303-0030$02.00

Introduction

For linkage analysis of rare Mendelian traits, large pedigrees have been analyzed by parametric methods, when possible, because of the power that they give for detection of linkage and localization of the susceptibility genes. However, even for common or non-Mendelian traits, large pedigrees often have been collected for linkage analysis (Egeland et al. 1987; Baron et al. 1994) These may not be the best family structures for detection of linkage for a complex trait especially when parametric methods are used. Affected relative pairs that are more distantly related (i.e., second cousins) will contain more information on linkage than will affected sib pairs (ASPs), if the susceptibility allele is rare. This is because they are less likely to share an allele at a marker identical by descent (IBD) by chance alone but are very likely to be IBD if there is linkage to the susceptibility allele. However, if the susceptibility allele is common, then distantly related affected relative pairs are not as likely to be IBD at the susceptibility locus. This is because of the increased probability that the susceptibility allele will be coming from the individuals marrying into the pedigree. Also, when a susceptibility allele is common, densely affected pedigrees (pedigrees with many affected individuals) are more likely to have parents homozygous for the susceptibility allele than are less densely affected pedigrees. This will cause a reduction in power to detect linkage.

Risch (1990*b*) demonstrated that, for the detection of genes with small effect, ASPs are the most powerful relative pairs for linkage analysis for a single-locus or multiplicative model. However, for additive models, there is a greater advantage in distant relationships. We hypothesized that, as the susceptibility-allele frequency becomes more common, the power of pedigrees with distantly related relative pairs will not add additional linkage information, compared with nuclear families with ASPs. Thus, in pedigrees with distantly related ASPs, linkage algorithms that analyze only ASPs will have power comparable to that of those algorithms that analyze pedigrees. We also hypothesize that densely affected pedigrees will provide less power as the susceptibility-allele frequency increases. In this article, these

hypotheses are tested by simulation of genetic models with varying susceptibility-allele frequencies and different pedigree structures and by linkage algorithms that analyze either ASPs or pedigrees.

Material and Methods

Data

Data for the single-locus models were simulated by PSLINK, a parallel version of SLINK (Ott 1989; Weeks et al. 1990; Cottingham et al. 1997) and by the high-performance computational capabilities of the IBM RISC/6000 SP system at the Division of Computer Research and Technology, National Institutes of Health (Bethesda). For the two-locus models, susceptibility-locus genotypes were simulated conditional on phenotypes, by the Pedigree Analysis Package (PAP) (Hasstedt 1994), and the marker-locus genotypes were simulated conditional on the first susceptibility-locus genotype and the recombination value, by PSLINK. Six types of family structure were simulated, which are shown in figure 1: I, nuclear families with an ASP; II, nuclear families with four affected sibs only two of whom were genotyped; III, pedigrees with two ASPs that were first cousins to each other (first-cousin ASPs); IV, pedigrees with two ASPs that were second cousins to each other (second-cousin ASPs); V, pedigrees with an affected–first-cousin pair (first-cousin pairs); and VI, pedigrees with an affected–second-cousin pair (second-cousin pairs). In all cases, only the last two generations were assumed to be genotyped, and the phenotype of individuals who were not affected was assumed to be unknown. The family structures consisting of four sibs with only two genotyped sibs were simulated to analyze the effects of sampling from a densely affected sibship but to keep the number of ASPs the same as that for family structure I.

For each replicate, the number of families required for 100 genotyped affected individuals was simulated. Thus, for family structures I, II, V, and VI, 50 families were in each replicate. For family structures III, and IV, 25 families were in each replicate. Each family structure was simulated with the seven genetic models described below. Each analysis was done with 1,000 replicates.

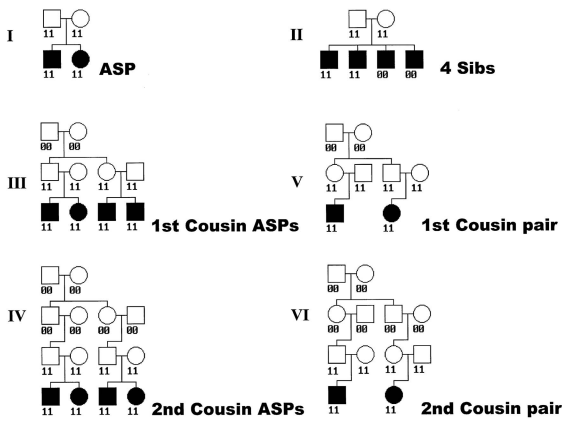

Figure 1 Family structures simulated in linkage analyses

Genetic Models

Seven genetic models were simulated. Table 1 shows the parameters of each single-locus model, and table 2 shows the parameters of the two-locus models. Penetrance was held constant among the single-locus models and the allele frequency was varied. For the two-locus models, an attempt was made to keep the marginal penetrances similar to those for the single-locus models, but this was not possible for the rare multiplicative model. The two additive models had a nonzero penetrance for genotype dd, unlike the single-locus models. For each two-locus model, the two susceptibility loci were unlinked and had equal allele frequencies and penetrances. Risch's (1987) λ refers to the relative risk that sibs will develop the trait, compared with the risk for the general population, and was calculated by use of the formula of James (1971). The values for total λ for the two-locus models were calculated according to the formulas of Risch (1987). A marker with four equally frequent alleles that either was linked to the susceptibility locus (recombination fraction [θ] = 0) or was not linked to it (θ = .50) was simulated. For the multipoint analysis, two flanking markers, each linked to the susceptibility locus, with θ = .05, were simulated.

Linkage-Analysis Methods

Programs used to analyze the data are SIBPAL, version 2.7.2 (SAGE 1994); ASPEX, version 1.62 (sib_ibd) (Hau-

Table 1

Prameters of Single-Locus Genetic Models Used in Simulations

Model	*P*(D)	*P*(aff|DD)	*P*(aff|Dd)	*P*(aff|dd)	Trait Frequency	Risch's λ
Rare	.0025	.10	.05	.00	.00025	100
Intermediate	.025	.10	.05	.00	.0025	11
Common	.25	.10	.05	.00	.025	1.8

ser et al. 1996); MAPMAKER/SIBS, version 2.0 (Kruglyak and Lander 1995); GENEHUNTER, version 1.1 (Kruglyak et al. 1996); and GENEHUNTER-PLUS, version 1.1 (Kong and Cox 1997). SIBPAL, ASPEX, and MAPMAKER/SIBS analyze nuclear families (or pedigrees broken down into nuclear families). GENEHUNTER and GENEHUNTER-PLUS analyze moderate-size pedigrees in their entirety. None of the programs require specification of a genetic model (although GENEHUNTER can perform parametric analyses in addition to nonparametric analyses). All of the programs except SIBPAL can perform multipoint analyses. SIBPAL estimates the observed alleles shared IBD for ASPs and calculates a *t*-statistic for the probability when there is no linkage. ASPEX and MAPMAKER/SIBS estimate a LOD score–based Risch λ for the susceptibility locus. MAPMAKER/SIBS restricts the maximization of LOD scores to the so-called possible triangle (Holmans 1993). Both the nonparametric linkage (NPL) statistic calculated by GENEHUNTER and the (Kong-and-Cox (KAC) statistic calculated by GENEHUNTER-PLUS measure allele sharing among all affected relatives, using the "ALL" score function. When ambiguous IBD sharing is encountered, the NPL statistic averages over all possible IBD-sharing configurations (weighted by likelihood), which is referred to as the "perfect-data approximation" (Kruglyak et al. 1996). The *P* values for the statistics are then based on the respective distributions formed from all possible IBD-sharing scenarios for a given set of pedigrees. *P* values determined by use of the perfect-data approximation are expected to be conservative when the data are not fully informative. The KAC statistic calculates a semiparametric LOD score by use of a single parameter that is a measure of the inheritance vector in the pedigrees and of allele sharing among the affected individuals and that, when data are less than perfectly informative, is less conservative than the NPL statistic.

Table 2

Parameters of Two-Locus Genetic Models Used in Simulations

	Multiplicative		Additive	
	Rare	Common	Rare	Common
$P(A) = P(B)$	.0025	.25	.0025	.25
Penetrances:				
P(Affected|AABB)	1	.4	.4	.26
P(Affected|AaBB)	1	.2	.2	.13
P(Affected|AABb)	1	.2	.2	.13
P(Affected|AaBb)	.5	.1	.1	.064
P(Affected|aaBB)	0	0	.1	.064
P(Affected|AAbb)	0	0	.05	.064
P(Affected|aaBb)	0	0	.05	.032
P(Affected|Aabb)	0	0	0	0
P(Affected|aabb)	0	0	0	0
Penetrance/locus:				
P(Affected|AA)	.005	.1	.1	.1
P(Affected|BB)	.005	.1	.1	.1
P(Affected|Aa)	.0025	.05	.05	.05
P(Affected|Bb)	.0025	.05	.05	.05
P(Affected|aa)	0	0	.00025	.016
P(Affected|bb)	0	0	.00025	.016
Locus λ	100	1.8	26	1.2
Total λ	10,000	3.1	51	1.4

For each program, the type I error rate and the power to detect linkage was determined for each pedigree structure and genetic model. Pedigrees with first-cousin pairs and second-cousin pairs were analyzed with only the NPL statistic and the KAC statistic, since there would be no linkage information in these pedigrees after they had been broken down into nuclear families. The type I error rate and power were

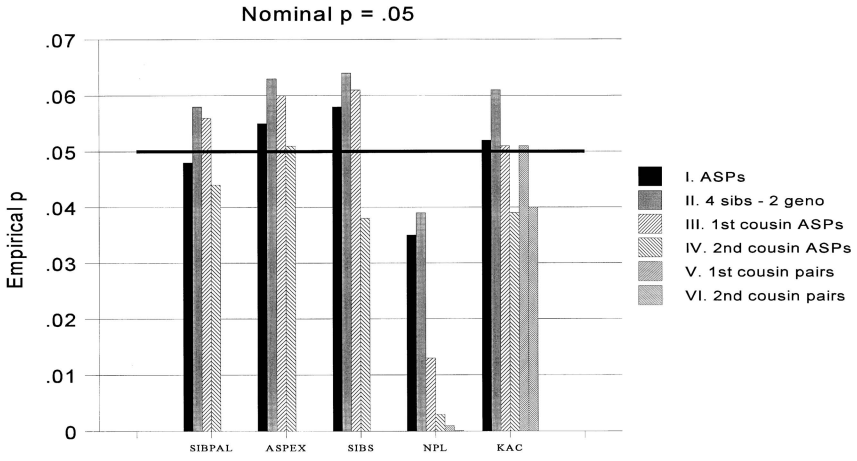

Figure 2 Type I error rate for different test statistics and family structures.

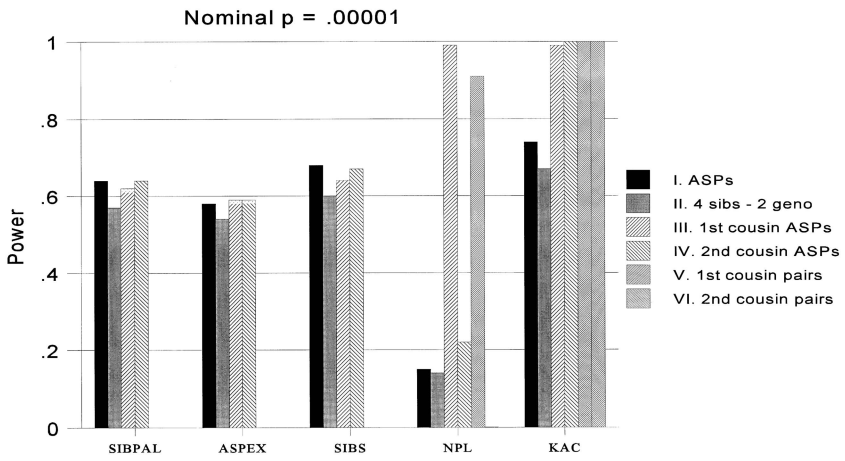

Figure 3 Power for a single locus with susceptibility-allele frequency .0025

analyzed either at nominal probability values of .05, .01, .001, .0001, and .00001 or, for programs that did not output probability values, at their LOD-score equivalents. For ASPEX and the KAC statistic, the LOD score was converted to a one-sided χ^2 with 1 df, by being multiplied by 2ln10. The LOD scores corresponding to the aforementioned nominal probability values were 0.58, 1.17, 2.09, 3.00, and 3.95, respectively. For MAPMAKER/SIBS, in which the LOD scores are maximized by the possible-triangle method, the LOD-score cutoffs were taken from Holmans (1993) and were set to 0.73, 1.36, 2.31, 3.27, and 4.23 for *P* values of .05, .01, .001, .0001, and .00001, respectively.

Results

The type I error rate for the different statistics and pedigree structures was estimated by simulation of a marker locus unlinked to the susceptibility locus and by observation of how often a nominal *P* value of .05 was observed (fig. 2). As expected, the genetic model for the susceptibility locus made little difference. SIBPAL, ASPEX, MAPMAKER/SIBS, and the KAC statistic had a slight increase in false positives for some family structures. The NPL statistic was very conservative.

The power analyses were performed by simulation of linkage at $\theta = 0$ and by observation of the proportion of results that were significant at a particular nominal

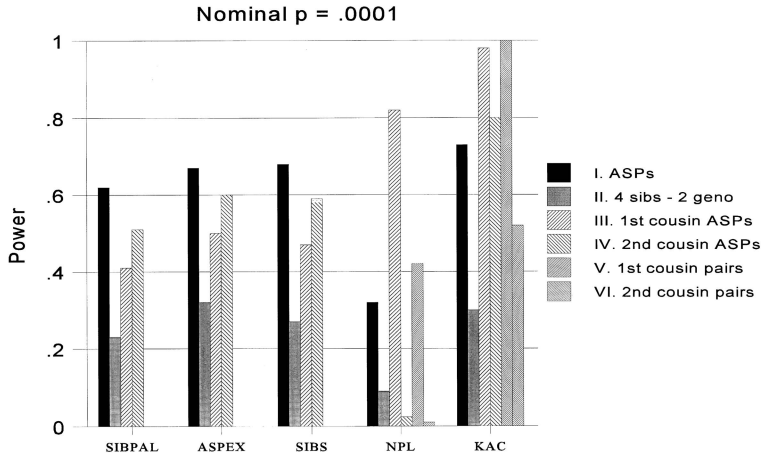

Figure 4 Power for a single locus with susceptibility-allele frequency .025.

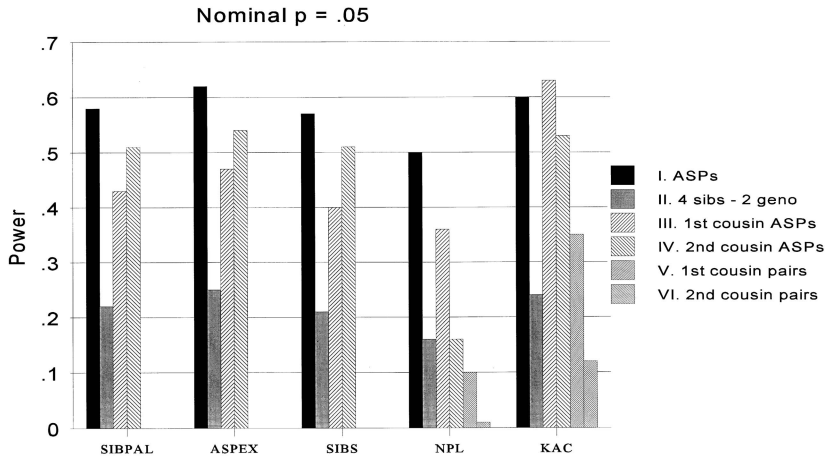

Figure 5 Power for a single locus with susceptibility-allele frequency .25

P value. As expected, the power to detect linkage was significantly higher for the rarer susceptibility loci with the higher λ. Thus, a different nominal *P* value is displayed for each genetic model, to keep the power comparable between the different genetic models.

For the rare susceptibility allele (fig. 3), the KAC statistic had very high power for most of the family structures. Further analysis showed that second-cousin pairs gave the highest power for the KAC statistic. The NPL statistic was very conservative for second-cousin ASPs and second-cousin pairs in which there were many ungenotyped individuals. Thus, in the more powerful pedigree-analysis statistic, the most distant relatives contributed the most linkage information in the case of a rare single-locus susceptibility allele. The multipoint analysis showed very similar results, except that there was substantially greater power for the NPL statistic to detect linkage in second-cousin ASPs and second-cousin pairs, an improvement which is likely to be secondary to the increased information from multipoint analysis (data not shown).

For the intermediate-frequency susceptibility allele (fig. 4), there was a substantial decrease in power in ASPs coming from densely affected sibships, compared with the power in ASPs coming from families with only two affected sibs. For the three-sib-pair statistics, there is a decrease in power for first-cousin ASPs relative to ASPs. These findings may be explained by the increased frequency of parents homozygous for the susceptibility locus. However, for the KAC statistic, the opposite is true. This is because, as with the rare alleles, there is in the cousin pairs some additional information that the KAC statistic is able to use. This is also demonstrated in the high power for the first-cousin pairs when the KAC sta-

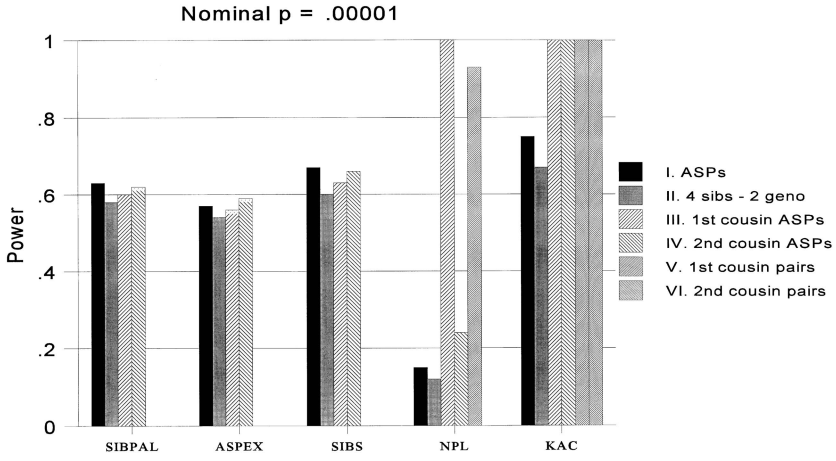

Figure 6 Power for a two-locus multiplicative model with susceptibility-allele frequencies .0025

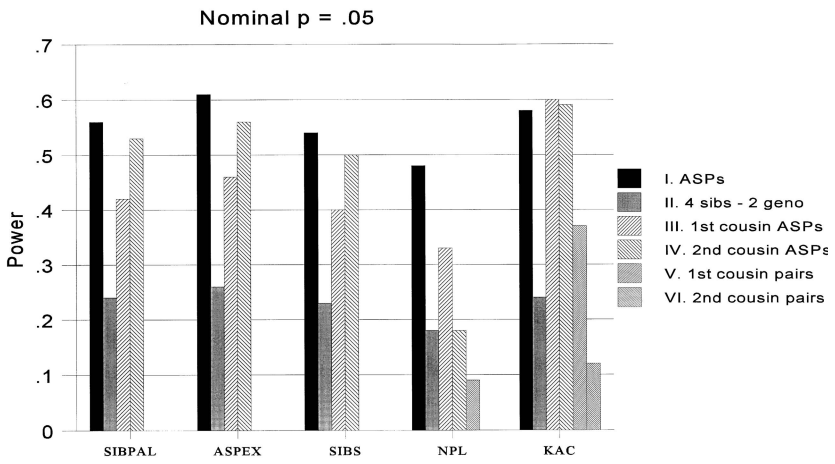

Figure 7 Power for a two-locus multiplicative model with susceptibility-allele frequencies .25

tistic is used. The NPL statistic has the lowest power, for all the samples. For the pedigree-analysis statistics, second-cousin ASPs and second-cousin pairs had decreased power, relative to the first-cousin counterparts. This is likely due to the effects of allelic heterogeneity at the susceptibility locus within the pedigree. The multipoint analysis showed substantially similar results, with the exception of increased power for the NPL statistic to detect linkage in the extended pedigrees (data not shown).

For the common susceptibility allele (fig. 5), there is a substantial loss of power in the analysis of ASPs from densely affected sibships, compared with that for ASPs from sibships with only two affected sibs. There is also a decrease in power for first-cousin ASPs, for the sib-pair statistics. These findings are likely to be due to increased homozygosity at the susceptibility locus of the parents. The KAC statistic is most powerful for families that include ASPs but does not show high power for families with only first-cousin or second-cousin pairs. For this genetic model, the power of the KAC statistic was substantially similar to the power of the sib-pair statistics. The NPL statistic is less powerful than the other statistics, for all family structures. For this model, it is evident that first-cousin and second-cousin pairs contribute much less power to detect linkage than do sib pairs. The results of the multipoint analysis are very similar to those of the single-point analysis (data not shown).

For the two-locus multiplicative models (figs. 6 and 7), the results are substantially similar to those of the single-locus model when susceptibility-allele frequencies are similar. Extended pedigrees have greater power for detection of linkage when the susceptibility allele is rare and have less power for detection of linkage when the susceptibility allele is common. Densely affected pedigrees also have less power for detection of linkage if the susceptibility allele is common.

For the two-locus additive-locus model with rare susceptibility alleles (fig. 8), the KAC statistic shows that extended pedigrees have substantially greater power for detection of linkage than do nuclear families. However, second-cousin ASP pedigrees have lower power for detection of linkage than do first-cousin ASP pedigrees. When the susceptibility alleles are common (fig. 9), extended pedigrees have no greater power for detection of linkage than do nuclear families. Densely affected pedigrees were less likely to allow for detection of linkage, although this result was less pronounced than it was for the single-locus and multiplicative models with common alleles.

Discussion

As has been hypothesized, when the disease allele is common, relative pairs that are more distantly related contribute less power for detection of linkage than do ASPs. Thus, there is also little power in pedigrees with only first-cousin or second-cousin pairs and no ASPs. Our results also show that, as the susceptibility-allele frequency increases, there is a loss of power in more densely affected sibships, for all the models that we simulated. These results are counterintuitive to traditional genetic disease strategies, in which one or a few large pedigrees would suffice for detection of a linkage, and represent yet another instance in which investigative strategies must be reconsidered for common diseases and traits.

The conclusions of our study have some limits, which are due to the type of trait and genetic models simulated, the amount of genotype information, and the analytical methods used. The conclusions might be different if ge-

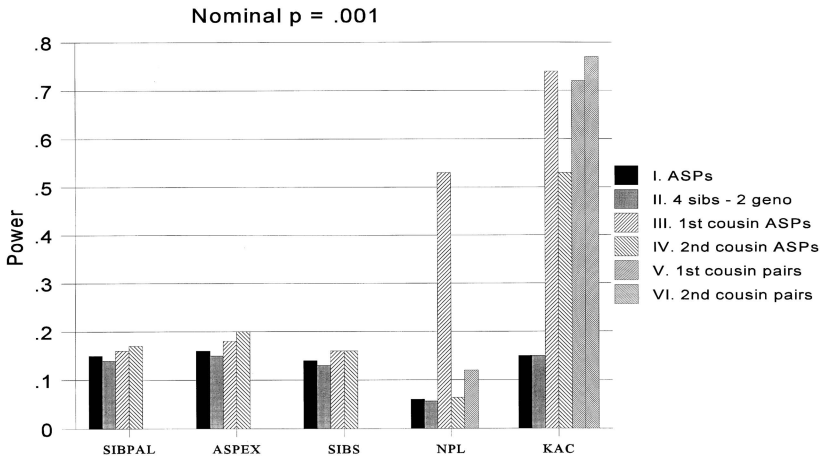

Figure 8 Power for a two-locus additive model with susceptibility-allele frequencies .0025

notype information were available for three generations and if parametric methods were used, since more phase information would be available, which would increase the power for detection of linkage. However, for many qualitative traits with adult onset, it will be possible to genotype only the last two generations of individuals within a pedigree.

Different results also might be found if the trait were continuous. For example, in Genetic Analysis Workshop 10, the power for detection of genes for quantitative traits simulated under the assumption of a model of several loci with common alleles and combinations of additive and multiplicative interactions was greater in extended pedigrees (probands with at least three offspring and three full sibs, their spouses, and all first-degree, second-degree, and third-degree relatives of proband and spouse) than in nuclear families (probands, spouses, and at least two living offspring), when genotyping was complete (Wijsman and Amos 1997). This increase in power was interpreted as arising from the additional meioses available in large pedigrees, coupled with increased ability for determination of marker phase.

In our simulations, the power to detect a single locus was similar to the power to detect one of two multiplicative loci with similar genetic parameters. This is consistent with previous observations that the power to detect multiplicative loci will depend on the effects of the individual loci rather than on either the total number of multiplicative loci or the effects of the other loci (Greenberg and Hodge 1989; Risch 1990*a*). This suggests that our conclusions may apply to multiplicative models with more than two loci and in which the other loci have effects that are different from those of the linked loci. Since the power to detect one of several additive loci will depend on the effects of the other loci (Risch 1990*a*), the results of these simulations of additive models cannot be generalized to all possible additive models.

A caveat in our two-locus simulations is that we assumed that the individual susceptibility loci would have allele frequencies similar for both the multiplicative and the additive models; this assumption was to control for the effect of the individual locus. However, for a given complex genetic disease of known incidence and relative risk to relatives, the locus-specific prevalence rates will be lower for an additive model than for a multiplicative model (Risch 1990*a*). Thus, for the same total incidence and relative risk to relatives, it is likely that the susceptibility-allele frequency will be lower for additive loci than for multiplicative loci. For example, a two-locus additive model with the same total λ as that in our two-locus multiplicative model with common susceptibility alleles would have a susceptibility-allele frequency of $\sim$.05. The simulations for this model were very similar to the simulations for the single-locus model with an intermediate-allele frequency, in that first-cousin ASPs and first-cousin pairs were the family structures with the most power for detection of linkage. Neuman and Rice (1992) have demonstrated that, for two-locus models that are consistent with a specific relative risk to sibs and population incidence, the allele frequencies for the heterogeneity models (which are similar to the additive model) tend to be lower than the allele frequencies for the multiplicative models. For this reason, as Risch (1990*b*) demonstrated, the advantage in having closer relatives available is greater for a multiplicative model than for an additive model, if the total incidence and Risch's λ are held constant.

The effect that family structure has on power when the susceptibility-allele frequency is common is more pronounced for the NPL statistic than for the KAC statistic. Part of the difference between the two algorithms

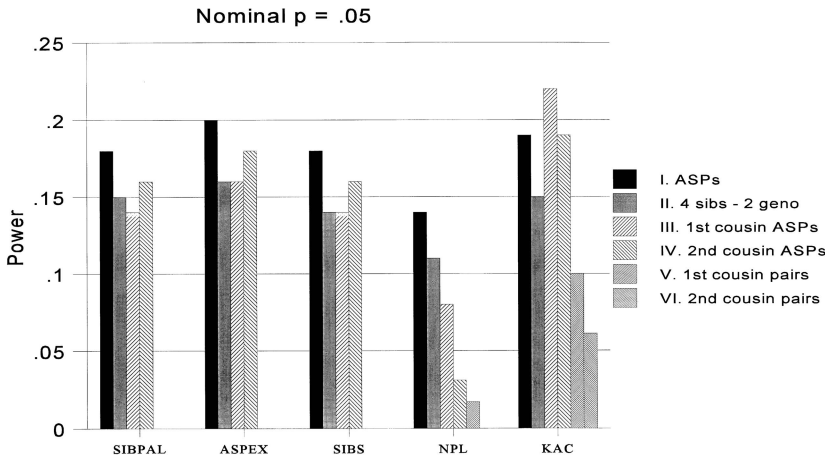

Figure 9 Power for a two-locus additive model with susceptibility-allele frequencies .25

is that the KAC statistic is less conservative when there is incomplete information (Kong and Cox 1997), as was the case for all the non-nuclear pedigrees. This may explain why there was a smaller decrease in power for the KAC statistic as the relatives became more distantly related. The differences between the KAC statistic and the NPL statistic were less in the multipoint analysis, although they still were significant. In simulations of common susceptibility alleles, the statistics that analyzed only ASPs had power similar to that of the KAC statistic. In these analyses, the statistics that analyzed ASPs (i.e., SIBPAL, ASPEX, and MAPMAKER/SIBS) were comparable in power and type I error. The comparability of the ASP statistics is consistent with the findings by Davis and Weeks (1997).

We analyzed only one ASP per sibship in the families with four affected sibs. This was done to keep the number of analyzed relative pairs comparable to that in the other family structures. However, this does not reflect the actual power afforded by such families. We have found, in all simulations, that, when 25 families each with four affected sibs (i.e., the same total no. of affected individuals as was used in the other simulations) are analyzed, the power is much higher than it is for 50 families each with four affected sibs but only two genotyped sibs, although there was in increase in false positives, to almost 7%, for a nominal probability of .05. These sibships were analyzed by the all-possible-pairs option of ASPEX and MAPMAKER/SIBS (this is the only option for SIBPAL), which has been shown, by Davis and Weeks (1997), to have more power and fewer false positives than are seen in other methods of analysis of families with three or four affected sibs.

Although our multipoint analyses gave results substantially similar to those of the single-point analyses, these conclusions need to be tempered by the fact that only two linked loci were simulated. The results might have been different if a more extensive map had been simulated, but this would have required significantly more resources. It is possible that, with a more extensive map—and, hence, more information—the results for the NPL statistic would be more similar to the results for the KAC statistic. It is also possible the extended pedigrees may give more phase information for multipoint analysis, which would make them more powerful, relative to nuclear families, than they are in the analyses presented here.

For complex genetic conditions that are relatively common (~1% population incidence), such as bipolar disorder and schizophrenia, where the observed recurrence risks imply oligogenic inheritance (Risch 1990*a*), the frequency of the susceptibility alleles would depend on the interaction between the loci. The alleles would be common if the interaction is multiplicative and would be rare if the interaction is additive and if many loci are involved. Risch demonstrated that the presence of multiplicative interaction can be determined by the drop-off in relative risk with increasing degree of relationship (Risch 1990*a*). Also, if the concordance of MZ twins is significantly more than twice the concordance of DZ twins, that also would suggest multiplicative interaction. Our results suggest that, if multiplicative interaction is involved, the best strategy would be to collect small families with ASPs. If additive interaction is involved, then lower susceptibility-allele frequencies would be implied and more-extended pedigrees would be best. If both types of interactions are involved or if it is difficult to determine which is involved, then it would be difficult to hypothesize about the susceptibility-allele frequencies of the individual loci. In that case, a strategy of collecting both nuclear families and families with first-cousin pairs and sibships may be the most powerful.

Acknowledgments

We would like to thank Edward Suh, Khiem Tran, and James Malley at the Division of Computer Research and Technology, National Institutes of Health, for the use of PSLINK for our simulations. Portions of the results from this study were obtained by use of the program package SAGE, which is supported by U.S. Public Health Service Resource Grant 1 P41 RR03655 from the Division of Research Resources.

References

Baron M, Endicott J, Lerer B, Loth JE, Alexander JR, Simon R, Sharpe L, et al (1994) A pedigree series for mapping disease genes in bipolar affective disorder: sampling, assessment, and analytic considerations. Psychiatr Genet 4: 43–55

Cottingham RW Jr, Idury RM, Schäffer AA (1993) Faster sequential genetic linkage computations. Am J Hum Genet 53: 252–263

Davis S, Weeks DE (1997) Comparison of nonparametric statistics for detection of linkage in nuclear families: single-marker evaluation. Am J Hum Genet 61:1431–1444

Egeland JA, Gerhard DS, Pauls DL, Sussex JN, Kidd KK, Allen CR, Hostetter AM, et al (1987) Bipolar affective disorders linked to DNA markers on chromosome 11. Nature 325: 783–787

Greenberg DA, Hodge SE (1989) Linkage analysis under "random" and "genetic" reduced penetrance. Genet Epidemiol 6:259–264

Hasstedt SJ (1994) Pedigree analysis package, version 4.02. Department of Human Genetics, University of Utah, Salt Lake City

Hauser ER, Boehnke M, Guo SW, Risch N (1996) Affected sib-pair interval mapping and exclusion for complex genetic traits: sampling considerations. Genet Epidemiol 13: 117–137

Holmans P (1993) Asymptotic properties of affected-sib-pair linkage analysis. Am J Hum Genet 52:362–374

James JW (1971) Frequency in relatives for an all-or-none trait. Ann Hum Genet 35:47–48

Kong A, Cox NJ (1997) Allele-sharing models: LOD scores and accurate linkage tests. Am J Hum Genet 61:1179–1188

Kruglyak L, Daly MJ, Reeve-Daly MP, Lander ES (1996) Parametric and nonparametric linkage analysis: a unified approach. Am J Hum Genet 58:1347–1363

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

Neuman RJ, Rice JP (1992) Two-locus models of disease. Genet Epidemiol 9:347–365

Ott J (1989) Computer-simulation methods in human linkage analysis. Proc Natl Acad Sci USA 86:4175–4178

Risch N (1987) Assessing the role of HLA-linked and unlinked determinants of disease. Am J Hum Genet 40:1–14

——— (1990*a*) Linkage strategies for genetically complex traits. I. Multilocus models [see comments]. Am J Hum Genet 46:222-228

——— (1990*b*) Linkage strategies for genetically complex traits. II. The power of affected relative pairs [see comments]. Am J Hum Genet 46:229-241

SAGE (1994) Statistical analysis for genetic epidemiology, version 2.2. Department of Biometry and Biostatistics, Rammelkamp Center for Education and Research, Metro Health Campus, Case Western Reserve University, Cleveland

Weeks DE, Ott J, Lathrop GM (1990) SLINK: a general simulation program for linkage analysis. Am J Hum Genet Suppl 47:A204

Wijsman EM, Amos CI (1997) Genetic analysis of simulated oligogenic traits in nuclear families and extended pedigrees: summary of GAW10 contributions. Genet Epidemiol 14: 719–735