Letters to the Editor

Am. J. Hum. Genet. 64:1243, 1999

Optimal Ascertainment Strategies to Detect Linkage to Common Disease Alleles

To the Editor:

The genetic dissection of complex diseases is of great current interest. The complexity of the task has led to serious discussion regarding competing strategies for data collection and analysis. In a previous issue of the *Journal,* Badner et al. (1998) contended that extended densely affected pedigrees (multiplex pedigrees with many affected individuals) are of little benefit for detection of linkage to complex traits such as bipolar disorder (a common psychiatric disorder of complex etiology). They state that such pedigrees are no more powerful than nuclear families when the susceptibility allele is common, and there may be loss of power in the collection of pedigrees with many affected individuals. Hence, they voice concern over pedigrees collected by others for linkage analysis of bipolar disorder (Egeland et al. 1987; Baron et al. 1994). However, there is merit to a broader perspective on this important problem.

Badner et al. (1998) simulate single-locus, additive, and multiplicative models with six types of pedigree structures, including nuclear families with affected sibs, pedigrees with first- or second-cousin affected sibs, and pedigrees with an affected first- or second-cousin pair. However, the family structures in the study by Badner et al. (1998) bear little resemblance to the pedigree series they find objectionable (i.e., Egeland et al. 1987; Baron et al. 1994). In particular, unlike the studies they cite, the families used by Badner et al. (1998) in their simulation study are not particularly large, with no direct evidence of vertical transmission (there are only a few affected individuals in these pedigrees, and only in the bottom generations). In addition, only the last two generations are assumed to be genotyped. Also, it is far from clear whether their simulations apply equally well to pedigree data with disparate ascertainment and population structure. For example, the data of Egeland et al. (1987) are based on very large interrelated pedigree structures obtained from a population isolate with a small number of founders (the Old Order Amish). In contrast, the Baron et al. (1994) pedigrees are smaller and derive from a general outbred population.

Second, Badner et al. (1998) determined a priori who were affected (that is, the pedigree structures were fixed) and then selected the genetic models and analyzed the pedigree data under a specified model, which may well have been incorrect. In so doing, they may have reached a foregone conclusion. A more appropriate approach would be to select a model and simulate a population to decide what pedigree structures appear and with what frequency. This would allow a reasonable correspondence between simulated mode of inheritance and the pedigree structures ascertained.

Third, Badner et al. (1998) doubt the utility of parametric methods for complex traits: "These [extended large pedigrees] may not be the best family structures for detection of linkage for a complex trait especially when parametric methods are used" (p. 880). However, the basis for this assertion is unclear, because the investigators confined their simulations to *nonparametric* sib pair and pedigree-analysis methods, to the exclusion of parametric analysis. Moreover, most of the putative linkages of current interest for bipolar disorder were detected in extended, densely affected pedigrees with parametric methods in inbred (Pekkarinen et al. 1995; Freimer et al. 1996; Barden et al. 1998) as well as outbred (Straub et al. 1994; Blackwood et al. 1996; Kelsoe et al. 1998; Aita et al. 1999) populations. Although further work is needed to evaluate these findings, these preliminary results attest to the potential utility of the extended pedigree approach in complex disorders.

Fourth, these researchers' own linkage studies of bipolar disorder (Berrettini et al. 1991) rest with extended pedigrees. Many of their pedigrees are similar, in size and illness density, to pedigrees described in the studies they criticize (e.g., Baron et al. 1994). With an average of 17 informative persons per pedigree (Berrettini et al. 1991), these pedigrees are substantially larger than nuclear families. Curiously, they make no mention of their own pedigree series while voicing concern about studies reported by others. Also, there is an apparent inconsistency between the conclusion drawn by Badner et al. (1998) from their simulations and their treatment of their own (real, not simulated) data—in particular, their recent claim of linkage between bipolar disorder and

chromosome 18 pericentromeric markers, which was based on nonparametric analysis (sib pair and affected–pedigree-member methods) of extended densely affected pedigrees (Berrettini et al. 1997). Although this finding is not generally accepted (Baron 1997; Rice 1997; Knowles et al. 1998), the investigators tout it a confirmed finding (Berrettini et al. 1997), an apparent contradiction of their doubts about the utility of the extended pedigree strategy. Moreover, with nonparametric analysis of their extended pedigrees, Detera-Wadleigh et al. (1996) replicated our linkage finding for bipolar disorder and chromosome 21q22.3; both the original report of this linkage (Straub et al. 1994) and a subsequent supportive analysis (Aita et al. 1999) were based on parametric analysis of the Baron et al. (1994) pedigrees. This illustrates the potential utility of confluent analytic approaches for complex traits in extended multiplex pedigrees.

There is an ongoing debate as to the optimal study design and methods of analysis for complex traits (Vieland et al. 1992; Baron 1997; Greenberg et al. 1997, 1998*b;* Goldgar and Easton 1997; Kruglyak 1997; Terwilliger 1998). The debate often pits analysis of nuclear families with affected sib pairs (ASPs) against analysis of extended high-density pedigrees, and so-called "model-free" methods (e.g., sib-pair analysis) versus "model-based" analysis (LOD scores in pedigrees). The main points can be summarized as follows:

Detractors of the extended pedigree approach argue that such pedigrees (1) incur more opportunities for introducing "extraneous" genes by way of bilineal transmission, increasing intrafamilial heterogeneity and leading to reduced power to detect linkage; (2) represent a particular form of a highly familial disease with a dominant-like effect, to the exclusion of other, more representative genetic mechanisms for complex traits, such as oligogenic transmission; (3) are best suited to detect genes of a relatively large effect and less likely to uncover minor genes that are likely present in a majority of cases; and (4) are hard to come by.

There are counterarguments, however: (1) Bilineal transmission can be screened out as part of the ascertainment scheme; (2) Because "hidden" bilineality can escape detection, two-trait–locus models allowing for more than one disease locus in the pedigree can be applied to bilineal pedigrees, with sufficient linkage information to warrant their inclusion (Schork et al. 1993); also, there are methods to analyze extended pedigrees subdivided into all component nuclear families, to account for intrafamilial heterogeneity (e.g., J. D. Terwilliger's ANALYZE computer program); (3) Small families and sib pairs are not impervious to heterogeneity: phenocopies may be common because of low illness density; (4) Extended pedigrees can contain more genetic information than smaller families and can have higher statistical power, especially when heterogeneity is accounted for; (5) The dominant, "single-gene" appearance in many extended pedigrees may, indeed, favor the detection of genes of a relatively large effect; this, however, is not necessarily a drawback, because such genes can be more easily tractable and may have greater biological importance than minor genes, at least in some cases; (6) As mentioned above, many of the putative linkages of current interest for bipolar disorder were detected in extended pedigrees; and (7) Undoubtedly, ASPs are more readily available than extended pedigrees, but advocates of the extended pedigree strategy argue that "rigorous science" is preferable to "convenient science."

Champions of "model-free" methods contend that these methods (1) are more suitable for complex disorders for which the mode of inheritance is uncertain, because, unlike model-based methods, they are not dependent on particular genetic parameters; (2) are less susceptible to multiple test effects leading to type I error, unlike model-based methods that tend to use several models; and (3) might also be preferable for analysis of bipolar disorder, because their utility has already been demonstrated in several complex traits (e.g., diabetes mellitus type I).

But proponents of model-based methods argue that (1) LOD score analysis in pedigrees generally has greater power and is reasonably robust to model misspecification, provided more than one plausible model is tested; (2) The critical factor in LOD score analysis is the mode of inheritance at the linked locus, not that of the complex trait per se (Greenberg et al. 1998*a*); (3) "Model-free analysis" is not truly model-free and is sometimes statistically equivalent to parametric analysis (Whittemore 1996); (4) With LOD score analysis, there is the option of using several different genetic models, thus covering a range of inheritance patterns with adequate power and little danger of missing a true linkage (such an option—i.e., a range of models—is not available for ASP analysis); and (5) There are no systematic studies supporting the assertion that model-free methods could detect linkage that LOD score analysis would miss.

As aptly put by Suarez et al. (1994), who conducted their own simulations for linkage detection in complex traits, "a simulation could so oversimplify a complex reality as to be misleading" (p. 36). Although some simulations can furnish useful guidelines, the computationally intensive nature of such studies and the complexities of the disorders being considered are inherent limitations. Clearly, there is no one correct strategy for linkage detection in complex traits such as bipolar disorder. Complementary approaches must be considered, including nuclear families with ASPs, extended pedigrees, and model-based and model-free methods of analysis. When genotypic information is available for several genera-

tions, extended pedigrees with vertical transmission may prove propitious for detection of linkage to complex traits.

MIRON BARON

Department of Psychiatry
and Columbia Genome Center
Columbia University, New York

References

Aita VM, Liu J, Knowles JA, Terwilliger JD, Baltazar R, Grunn A, Loth JE, et al (1999) A comprehensive linkage analysis of chromosome 21q22 supports prior evidence for a putative bipolar affective disorder locus. Am J Hum Genet 64: 210–217

Badner JA, Gershon ES, Goldin LR (1998) Optimal ascertainment strategies to detect linkage to common disease alleles. Am J Hum Genet 63:880–888

Barden N, Morissette J, Shink E, Rochette D, Gagne B, Bordeleau L, Villeneuve A, et al (1998) Confirmation of bipolar affective disorder susceptibility locus on chromosome 12 in the region of the Darier disease gene. Am J Med Genet (Neuropsych Genet) 81:475

Baron M (1997) Genetic linkage and bipolar affective disorder: progress and pitfalls. Mol Psychiatry 2:200–210

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

Berrettini WH, Ferraro TN, Goldin LR, Detera-Wadleigh SD, Choi H, Muniec D, Guroff JJ, et al (1997) A linkage study of bipolar disorder. Arch Gen Psychiatry 54:27–35

Berrettini WH, Goldin LR, Martinez MM, Maxell T, Smith AL, Guroff JJ, Kazuba DM, et al (1991) A bipolar pedigree series for genomic mapping of disease genes: diagnostic and analytic considerations. Psychiatr Genet 2:125–160

Blackwood DHR, He L, Morris SW, McLean A, Whitton C, Thomson M, Walker MT, et al (1996) A locus for bipolar affective disorder on chromosome 4p. Nat Genet 12: 427–430

Detera-Wadleigh SD, Badner JA, Goldin LR, Berrettini WH, Sanders AR, Rollins DY, Turner G, et al (1996) Affected-sib-pair analyses reveal support of prior evidence for a susceptibility locus for bipolar disorder, on 21q. Am J Hum Genet 58:1279–1285

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

Freimer NB, Reus VI, Escamilla MA, McInnis LA, Spesny M, Leon P, Service SK, et al (1996) Genetic mapping using haplotype association and linkage methods suggest a locus for severe bipolar disorder (BP) at 18q22–q23. Nat Genet 12: 436–441

Goldgar DE, Easton DF (1997) Optimal strategies for mapping complex diseases in the presence of multiple loci. Am J Hum Genet 60:1222–1232

Greenberg DA, Abreu P, Hodge SE (1998*a*) The power to detect linkage in complex disease by means of simple LOD-score analysis. Am J Hum Genet 63:870–879

Greenberg DA, Hodge SE, Vieland VJ, Spence MA (1997) Reply to Farrall. Am J Hum Genet 61:254–255

——— (1998*b*) Reply to Kruglyak. Am J Hum Genet 62: 202–204

Kelsoe JR, Loetscher E, Spence MA, Foguet M, Sadovnick AD, Remick RA, Flodman P, et al (1998) A genome survey of bipolar disorder indicates a susceptibility locus on chromosome 22. Am J Med Genet (Neuropsych Genet) 81: 461–462

Knowles JA, Rao PA, Cox-Matise T, Loth JE, DeJesus G, Levine L, Alexander JA, et al (1998) No evidence for significant linkage between bipolar affective disorder and chromosome 18 pericentromeric markers in a large series of multiplex extended pedigrees. Am J Hum Genet 62:916–924

Kruglyak L (1997) Nonparametric linkage tests are model free. Am J Hum Genet 61:254–255

Pekkarinen P, Terwilliger J, Bredbacka P-E, Lonnqvist J, Peltonen L (1995) Evidence of a predisposing locus to bipolar disorder on Xq24–q27.1 in an extended Finnish pedigree. Genome Res 5:105–115

Rice J (1997) Genetic analysis of bipolar disorder: summary of GAW10. Genet Epidemiol 14:549–562

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 53:1127–1136

Straub RE, Lehner T, Luo Y, Loth JE, Shao W, Sharpe L, Alexander JR, et al (1994) A possible vulnerability locus for bipolar affective disorder on chromosome 21q22.3. Nat Genet 8:291–296

Suarez BK, Hampe CL, Van Eerdewegh P (1994) Problems of replicating linkage claims in psychiatry. In: Gershon ES, Cloninger CR (eds) Genetic approaches to mental disorders. American Psychiatric Press, Washington, DC, pp 23–46

Terwilliger JD (1998) Linkage analysis B model-based. In: Armitage P, Colton T (eds) Encyclopedia of biostatistics. Vol 3. Wiley, England, pp 2279–2291

Vieland VJ, Hodge SE, Greenberg DA (1992) Adequacy of single-locus approximations for linkage analyses of oligogenic traits. Genet Epidemiol 9:45–59

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

Address for correspondence and reprints: Dr. Miron Baron, Department of Psychiatry and Columbia Genome Center, Columbia University, 722 West 168th Street, Unit 6, New York, NY 10032. E-mail: mb17@columbia.edu

0002-9297/99/6404-0000$02.00