A manic depressive history

Neil Risch & David Botstein

Department of Genetics, Standford University School of Medicine, Stanford, California 94305, USA

The mapping of inherited disease genes onto the human genome by linkage analysis has proceeded at an astonishing rate since the first such achievements in the early 1980's. Many hundreds of traits have been convincingly mapped, and positional cloning of many scores of genes followed, often hard upon the heels of the linkage mapping. The same cannot be said, unfortunately, for traits that are inherited in a more complex fashion. Here the record is one of repeated claims for a variety of different loci followed by counterclaims and even retractions.

In no field has the difficulty been more frustrating than in the field of psychiatric genetics. Manic depression (bipolar illness) provides a typical case in point. Indeed, one might argue that the recent history of genetic linkage studies for this disease is rivaled only by the course of the illness itself. The euphoria of linkage findings being replaced by the dysphoria of non-replication has become a regular pattern, creating a roller coaster-type existence for many psychiatric genetics practitioners as well as their interested observers.

The recent history of linkage studies in manic depression is summarized in Table 1, which lists only studies showing positive evidence for linkage with defined chromosomal locations. Lod scores for the pedigrees in references[1–5] are taken from[6]. Numerous studies (not shown) have also been published that fail to confirm many of these chromosomal regions. Three papers in this issue of *Nature Genetics*, Blackwood *et al.*[7], Ginns *et al.*[8], and Freimer *et al.*[9], have added five additional loci to this list. None of the 14 listed regions has been convincingly, or at least consistently replicated, although some regions have been implicated more than once.

Even the most popular regions have their problems. The region with the longest history is Xq28, starting in 1969 (ref. 10), and several studies have produced sizeable lod scores linking bipolar disease with the phenotypic markers colourblindness and G6PD deficiency[1–5,11,12] The most recent (and impressive) of these studies is that of Baron *et al.*[12], who studied 5 Israeli families and obtained a lod score of 7.5. Unfortunately, additional subsequent analysis of the Israeli families with DNA markers in Xq28 also led to a diminution of the linkage evidence in this region[13]. A similar history plagues the region 11p15 (loci *INS* and *HRAS*) in the Amish, where an initial exciting lod score of 4.9 (ref. 14) was subsequently diminshed to non-significance by additional analysis in the same families[15]. The study by Ginns *et al.*[8] in this issue is yet another study of this same population.

Chromosome 18 has been implicated twice before[16,17], and here for a third time[9]. Lest the reader form the impression of consistency, however, it should be noted that the chromosomal regions covered by these three linkage reports (markers *D18S62* - *D18S70*) appear to span a total distance of about 136 cM, including most of both chromosomal arms.

The distress engendered by the numerous reversals and non-replications has led many to rethink the paradigms being employed. Premier among these have been calls for an increase in the lod score threshold for declaring linkage, to protect against further false positive claims. For example, in a *News and Views* item in *Nature*, Miranda Robertson wrote[18] "readers of *Nature* are entitled to be sceptical about reports of psychiatric disease linkages at lod scores of less than 6." No theoretical basis was given for this number, but perhaps an odds ratio of a million to one seems comforting. Some have argued that the problem is the large number of markers being tested in

Table 1 Reported chromosomal locations for manic-depression genes by linkage analysis

Location	Lod Score	Year	Reference
Xq28	—	1969	Reich *et al.*[10]
	13.4	1972-80	Mendlewicz *et al.*[1–4]
	2.1	1977	Baron[5]
	1.5	1984	Del Zompo *et al.*[11]
	7.5	1987	Baron *et al.*[12]
11p15	4.9	1987	Egeland *et al.*[14]
Xq27	3.1	1987	Mendlewicz *et al.*[26]
	3.9	1992	Lucotte *et al*[27]
	2.2	1993	Jeffries *et al.*[28]
Xq24-26	3.5	1995	Pekkarinen *et al.*[29]
5q35	1.4	1993	Coon *et al.*[30]
21q22	3.4	1994	Straub *et al.*[31]
12q23	2.1	1994	Craddock *et al.*[32]
18p	—	1994	Berrettini *et al.*[16]
18q	1.7–3.1	1995	Stine *et al.*[17]
16p13	2.7	1995	Ewald *et al.*[33]
18q	—	1996	Freimer *et al.*[9]
4p16	4.8	1996	Blackwood *et al.*[7]
6p24, 13q13, 15q11	2.5, 1.4, 1.1	1996	Ginns *et al.*[8]

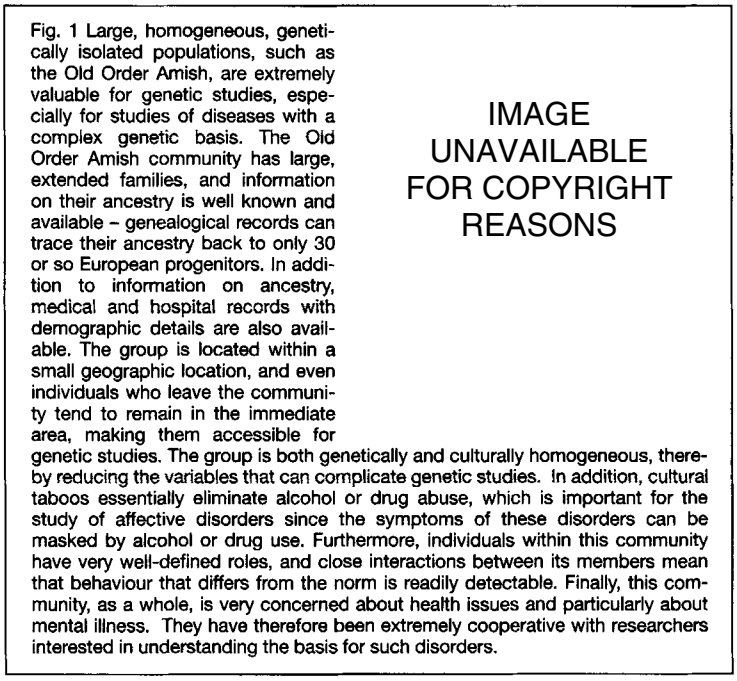

Fig. 1 Large, homogeneous, genetically isolated populations, such as the Old Order Amish, are extremely valuable for genetic studies, especially for studies of diseases with a complex genetic basis. The Old Order Amish community has large, extended families, and information on their ancestry is well known and available – genealogical records can trace their ancestry back to only 30 or so European progenitors. In addition to information on ancestry, medical and hospital records with demographic details are also available. The group is located within a small geographic location, and even individuals who leave the community tend to remain in the immediate area, making them accessible for genetic studies. The group is both genetically and culturally homogeneous, thereby reducing the variables that can complicate genetic studies. In addition, cultural taboos essentially eliminate alcohol or drug abuse, which is important for the study of affective disorders since the symptoms of these disorders can be masked by alcohol or drug use. Furthermore, individuals within this community have very well-defined roles, and close interactions between its members mean that behaviour that differs from the norm is readily detectable. Finally, this community, as a whole, is very concerned about health issues and particularly about mental illness. They have therefore been extremely cooperative with researchers interested in understanding the basis for such disorders.

any given study, which could be thought to lead to an inflated false positive rate[19], and thus the lod threshold needs to be raised. A recent, elegant analysis, firmly based in theory, of the issues raised by false positives arising from analysis of large numbers of markers spanning the genome has been offered by Lander and Kruglyak[20]. On this basis they proposed an increase in the lod score thresholds for significant linkage. However, this increase turned out to be quite modest (3.3 for parametric lod score analysis, 3.6 for sib pair analysis).

It seems most unlikely that testing a dense marker map is the problem here. Dense marker maps have been continuously used to map the loci for classic mendelian disorders and traits, and no such high rate of false positives has been observed; indeed, the lod score threshold of 3 originally proposed by Morton[21] has held up quite well, producing only a small false positive rate, as was also predicted by earlier empirical studies[22].

So then, what is the problem? We believe it is likely to be the complex basis for the diseases being studied. Sometimes, for a common, complex disease, there is a mendelian, autosomal dominant or X-linked subgroup that can be discerned and analysed as such. Parametric lod score analysis has been successfully applied to identify such loci for many familial cancers, including breast cancer, colon cancer, melanoma, as well as Alzheimer's disease. In all these cases, linkage findings were convincingly replicated and the mutant genes have been cloned. Does the same model apply to manic depression? Certainly, multigeneration pedigrees that appear consistent with autosomal dominant inheritance (with reduced penetrance) exist, for example those described in the three reports in this issue[7–9]. Although high lod scores have been generated in single pedigrees, they generally have not been replicated in others, either by the same investigators in a similar study population (as is the case in the study of Blackwood *et al.*[7]) or in the studies of other investigators. Some have argued this is due to genetic heterogeneity – that is, the existence of multiple loci that when mutated can individually lead to disease. But how much heterogeneity? Genetic heterogeneity is a real, and often tractable phenomenon. It has been observed in breast cancer, where there are two predominant loci, in colon cancer, where there are five loci, in early-onset Alzheimer's disease, where there are three loci. While a dozen or more loci could cause bipolar illness, and that is the origin of the non-replications, we believe the explanation lies elsewhere, namely that the genetic mechanism underlying the disease in these families is more complicated than postulated, leading to a reduction in power. When power is low, it is difficult to distinguish a true from a false positive[23,24].

Some of these difficulties are exemplified in the analysis of Ginns *et al.*[8]. These investigators accomplished the prodigious task of genotyping over 551 markers in 207 individuals from the Amish population. It has been repeatedly argued that the Amish are a useful population for studying bipolar illness because of the very limited number of founders. Indeed, as Ginns *et al.*[8] point out, all of the bipolar pedigrees examined trace back to a single common progenitor (see Fig. 1). Furthermore, segregation analysis in this population suggested dominant inheritance[25]. The parametric linkage analysis was performed using multiple mendelian-type models (dominant and recessive). Yet no single large lod score was produced. The authors report quite modest lod scores for three chromosomal regions not previously identified (chromosomes 6, 13 and 15) after testing multiple genetic and diagnostic models. The modest nature of these lod scores, despite their being the maximum values after testing numerous models, and the lack of reports of linkage for these chromosomal regions by others, should engender less than complete confidence that these regions will survive subsequent scrutiny.

Ginns *et al.*[8] conclude that bipolar illness in the Amish is inherited as a compex trait. We agree, but not for the same reason. The authors' conclusion is based on their obtaining modest lod scores for 3 chromosomal regions. We would base our conclusion on the fact that no large lod score was obtained for any single chromosome region. This argument would be more compelling had the authors performed a formal exclusion analysis with their hypothesized autosomal dominant model. Then we could conclude with greater certainty that no such dominant gene accounts for the disease in this population.

The study of Freimer *et al.*[9] also entailed a significant effort, the typing of 473 markers in two Costa Rican pedigrees, which, like the Amish, derive from a limited number of founders. Again, only mod-

est lod scores for linkage were obtained; in this case, the genetic evidence derived from association analyses of several of the markers in the 18q23 region. However, the two most significant associations were with markers 15 cM apart, while other markers in between provided evidence against an association, making it difficult to derive a consistent conclusion.

Given the long and unproductive history of linkage studies in bipolar illness, what is the value in publishing these results, which can only be interpreted in the most tentative way, in *Nature Genetics*? We see two justifications for publication of results of this kind. One derives from the idea that, in the face of complexity in the mode of inheritance, statistically significant linkage in a very small subset of families may indeed represent reality. For example, it is entirely possible that the linkage to markers on 4p16 observed by Blackwood *et al.*[7] could be confirmed in the future in additional, necessarily rare, families and followed up by cloning of a predisposing gene whose function could shed light on the biology of this disease.

The other justification is that a genome-scanning data set from properly defined and executed studies might contribute to the eventual definition of predisposing genes even if the data set by itself does not yield a statistically credible result. One could envision, in this scenario, that a combined analysis of several such data sets together might yield significant results that could then be confirmed, refined and even followed up.

It seems, however, that both scenarios fail if the publication is limited to summary statistics, with only a few loci reported, as is now the common practice. If the justification for publication depends on future work, then the basic underlying data should be made freely available to all investigators.

It is now nearly a decade since modern molecular methods were first applied in the study of psychiatric disorders, and in particular bipolar illness. While numerous studies have been performed and results presented (Table 1), has the field really advanced in that time? Compared to other fields to which genetics has been applied, one would have to argue not, probably because the underlying genetics is now certain to be much more complicated than one might have hoped. To help us deal with this more difficult reality, we make the following suggestions:

1. It is important to have a realistic view about what all these myriad linkage results indicate. Given all the genotypes that have been produced on large samples, it is undoudtedly possible to exclude many simple hypotheses, and it is important to do so. For example, it would appear that the hypothesis that there is a single major locus accounting for the majority of inherited bipolar illness has by now been rejected.

2. At a minimum, lod score thresholds set for the mendelian case (lod = 3) should be adhered to for complex diseases or even increased, as suggested by Lander and Kruglyak[20]. However, it should be clear that this alone will not solve the problem. Applying these thresholds historically (for example, to the results given in Table 1) would not have produced a lower false positive rate. Thus, we need additional criteria beyond thresholds for significance.

3. For complex diseases, where there is no prior evidence for a single major locus, replication of significant linkage findings is essential. Before a finding is accepted formally, replication needs to be obtained in independent studies and point to the same chromosomal region. Even in cases where genetic heterogeneity is postulated, one should consistently find positive, if not convincing evidence at a proposed locus.

4. It is time to start constucting exclusion maps of the genome. This can be done by parametric lod score analysis as well as with more robust methods such as affected sib pair analysis.

5. Investigators should state the power of their study sample for the model they are testing as well as the significance level, so that readers can assess the likelihood of a false positive result. Also, it will help the reader to determine whether the model being tested is realistic.

6. Making primary data available should be a condition of publication of these kinds of linkage studies. That way others can combine it with their own, or a more general, meta-type analysis can be performed.

The lack of consistent replication of any of the observed linkages in Table 1 reinforces the notion that bipolar illness has a genetically complex aetiology. Thus, it is unlikely that any individual study will have sufficient power to produce a definitive result. A simultaneous evaluation of all available data with unified methods of analysis may help eliminate some genetic models for this disease, as well as define the most promising chromosomal regions for follow-up.

1. Mendlewicz, J. *et al. JAMA* **222**, 1624–1627 (1972).
2. Mendlewicz, J. & Fleiss, J.L. *Archs. Gen. Psychiat.* **24**, 721–727 (1974).
3. Mendlewicz, J. *et al. Archs. Gen. Psychiat.* **36**, 1442–1447 (1979).
4. Mendlewicz, J. *et al. Br. J. Psychiatry* **137**, 337–342 (1980).
5. Baron, M. *Archs. Gen. Psychiat.* **24**, 721–727 (1977).
6. Risch, N. & Baron, M. *Ann. Hum. Genet.* **46**, 153–166 (1982).
7. Blackwood, D.H.R. *et al. Nature Genet.* **12**, 427–430 (1996)
8. Ginns, E.I. *et al. Nature Genet.* **12**, 431–435 (1996).
9. Freimer, N.B. *et al. Nature Genet.* **12**, 436–441 (1996)
10. Reich, T. *et al. Am. J. Psychiat.* **125**, 1358–1369 (1969).
11. Del Zompo, M. *et al. Acta Psychiatr. Scand.* **70**, 282–287 (1984).
12. Baron, M. *et al. Nature* **326**, 289–292 (1987).
13. Baron, M. *et al. Nature Genet.* **3**, 49–55 (1993).
14. Egeland, J.A. *et al. Nature* **325**, 783–787 (1987).
15. Kelsoe, J.R. *et al. Nature* **342**, 238–243 (1989).
16. Berrettini, W.H. *et al. Proc. Natl. Acad. Sci. USA* **91**, 5918–5921 (1994).
17. Stine, O.C. *et al. Am J. Hum. Genet.* **56**, 1384–1394 (1995).
18. Robertson, M. *Nature* **342**, 222 (1989).
19. Edward, J.H. & Watt, D.C. *Psychol. Med.* **19**, 273–275 (1989).
20. Lander, E. & Kruglyak, L. *Nature Genet.* **11**, 241–247 (1995).
21. Morton, N.E. *Am. J. Hum. Genet.* **7**, 277–318 (1955).
22. Rao, D.C. *et al. Am. J. Hum.Genet.* **30**, 516–529 (1978).
23. Risch, N. *Am. J. Hum. Genet.* **48**, 1058–1064 (1991).
24. Risch, N. *Science* **255**, 803–804 (1992).
25. Pauls, D.L. *et al. Am. J. Med. Genet.* **60**, 783–787 (1987).
26. Mendlewicz, J. *et al. Lancet* **1**, 1230–1232 (1987).
27. Lucotte, G. *et al. Ann. Genet.* **35**, 93–95 (1992).
28. Jeffries, F.M. *et al. Biol. Psychiatry* **160**, 7–11 (1993).
29. Pekkarinen, P. *et al. Genome Res.* **5**, 105–113 (1995).
30. Coon, H. *et al. Am. J. Hum. Genet.* **52**, 11234–1249 (1993).
31. Straub, R.E. *et al. Nature Genet.* **8**, 291–296 (1994).
32. Craddock, N. *et al. Brit J. Psychiatry* **164**, 355–358 (1994).
33. Ewald, H. *et al. Psych. Gen.* **5**, 71–81 (1995).