EDITORIAL

Caution in locating the gene(s) for affective disorder[1]

Among the common psychiatric illnesses a genetic predisposition is most apparent, and evidence for dominant inheritance strongest, in manic-depressive disease. The protean manifestations encourage the speculation that it is a group of conditions with differing aetiology. Bipolar affective illness in which episodes of illness of both depressive and manic character, frequently unrelated to situational difficulties at onset, are separated by periods of normality and unimpaired functional capacity is the most distinct and delimited form of the disorder and therefore the most promising for a linkage study.

The familial aggregation of cases has long been held to indicate a major genetic factor in manic-depressive disorders but this information has not been sufficient to be of practical use clinically. Considerable attention has therefore been attracted by the report (Egeland *et al.* 1987) that linkage analysis has both defined and located a dominant gene responsible for bipolar affective illness in Amish families.

LINKAGE BETWEEN A DISEASE GENE LOCUS AND A MARKER

Linkage occurs when loci are on the same chromosome and sufficiently close to be transmitted together, from parent to child within a pedigree, more frequently than would occur by random assortment. In attempting to detect linkage a marker is employed. This is any inherited characteristic (e.g. a blood group) which is clearly identifiable, persistent and unaffected by age, illness or accident, and transmitted by a simple mode of inheritance. The scope of linkage has recently been greatly extended by a different kind of marker; the restriction fragment length polymorphism (RFLP), which is defined by an identifiable portion of DNA, a probe, which can be labelled and made to combine with the same sequence of the DNA in the genome after this has been cut into fragments by restriction enzymes. Probes are selected so that they attach to one site only.

If linkage can be established by a consistent tendency to cosegregation between a marker gene and a disease gene it permits certain conclusions. It confirms the Mendelian nature of the disease and identifies the chromosome on which a major predisposing locus lies. In principle it provides the basis for locating a short segment of chromosome, thereby reducing the potential genes to a mere hundred or so, one of which must be the offender. It also provides, in principle, the means of detecting unaffected gene carriers.

LINKAGE INVESTIGATION IN AMISH FAMILIES

The Amish are the descendants of early eighteenth-century immigrants from the Swiss–German border area to the United States who have preserved their identity to such a degree that they constitute a genetic isolate. Extensive pedigrees of this group have been documented and one, which showed transmission of affective disorder, was the subject of the linkage studies carried out by Egeland and her colleagues (1987). These began with 20 markers (Kidd *et al.* 1984), as a result of which some hopeful candidate markers were defined. To confirm this the appropriate procedure would be to use the candidate markers, or markers closely linked to them, on *other* families. One substantial attempt failed (Kidd *et al.* 1984). This is not surprising, as such highly informative

families showing segregation consistent with dominant inheritance are rare; so rare, in fact, that these families with apparent Mendelian inheritance in manic-depressive illness could reasonably be explained by the joint effects of chance and a generalized predisposition based on allelic variation at many loci.

In a further attempt at confirmation of the suggested linkage they had observed (Egeland *et al.* 1987) the authors used the same families with some additional family members and employed additional markers known to be closely linked to their candidate locus.

The results of this procedure hardly support the authors' conclusion that it 'provided convincing evidence for the localisation of a gene that is implicated in the aetiology of a common clinical disorder' for two reasons. The first reason for caution is that no allowance was made in the calculation in the first study for the use of a large number of markers (20), as opposed to a single marker, which increases the possibility that cosegregation of a particular marker and the gene in question may occur by chance.

The second is that the first round of linkage tests defined a set of cosegregating segments of DNA which may or may not include the putative dominant locus. Consequently, further tests using markers known to be on these segments can give only limited information on whether or not these segments contain this putative locus. As there were some additional family members in this second study (Egeland *et al.* 1987) and their candidate locus was uninformative in some individuals, some supporting evidence was provided but it is not clear if it was sufficient to justify the strong claims of linkage which were made.

THE EFFECT OF MULTIPLE MARKERS

If we had two dice, one normal and the other bearing sixes on all faces, which we will term the 'linked' die, and one die had been selected at random and tossed, revealing a six on its upper face, which alone was visible, then the 'odds ratio' favouring the linked die would be $1:\frac{1}{6}$ or $6:1$. If a single die were tossed twice to yield two sixes it would be $36:1$ and if four times $1296:1$. These are odds ratios, i.e. the odds that the die being tossed was the 'linked' die compared with the odds that it was the normal die. Since addition is easier than multiplication it is usual to take logarithms, and to term the logarithm of the odds ratio a 'lod', an acronym for log-odds. As $\log (6)$ is 0.78 these odds ratios give lods of 0.78, 1.56 and 3.12. In human linkage we are testing for identity over 22 chromosomes, excluding the sex chromosomes, and these not only differ in size but usually undergo a crossing-over during the meiotic processes which produce gametes. If we ignore these problems initially, and consider 22 chromosomes of similar size within which recombination is rare, then we can model this by a set of 22 dice, one all sixes; defining a linkage can then be compared to detecting the 'linked' die by tossing one of the 22 dice a number of times. The odds against will be $21:1$ before we have made a throw. Four throws of sixes will favour a 'linkage' with odds of $1296:1$. But the initial odds against are $21:1$ giving total odds of $1296/21:1$ or about $60:1$ favouring the 'linked' die. In human linkage it is usually assumed reasonable to infer linkage if the odds ratio on a test is over $1000:1$ (i.e. the lod is over 3.0). These are not probabilities, in the usual statistical sense, which relate to a range of possibilities, but likelihoods, and merely refer to the relative likelihood of two possibilities, in this case tight linkage and non-linkage.

The $1000:1$ convention stems from Morton (1955), who advanced it in relation to tests of a single marker; it has since become a conventional guideline, well tested in practice and found a useful compromise between error and indecision.

Where more than one marker is used, several opportunities for the marker to show in an affected individual are provided and the likelihood of this happening by chance is therefore increased. If a number of marksmen, instead of a single one, fire at the same target the chance of a hit is increased irrespective of whether any additional skill is provided by the increase in participants. A bet spread over a number of runners will attract lower odds than one backing only a winner. In the case of linkage similar allowance must be made for multiple test. With testing of many markers the criterion of credibility based on the number of opportunities provided for a marker to show in an affected

individual must be increased if the chance of inferring a false linkage is to be kept within acceptable limits. A simple solution often adopted is to divide the odds ratio by the number of markers, which is satisfactory for small numbers of markers, but it over-compensates or, in statistical terms, is 'conservative', so that at least it errs on the side of caution. This procedure is equivalent to adding the logarithm of the number of tests to the acceptance criterion. For example with 10 tests, since log (10) is 1, we would need a criterion of a lod of 4, rather than 3. This conclusion can also be derived by a more complicated argument (Kidd & Ott, 1984; Ott, 1986). An exact analysis gives similar results (Thompson, 1984).

MULTIPLE MARKERS IN THE AMISH STUDIES

In neither of the Amish studies was allowance made for the effect of multiple markers. The genuine strength of evidence is impossible to infer from the published data, due in part to the publication of a pedigree falsified to protect potential depressives from identifying themselves and, by concluding that they carry a pathogenic gene, compounding their innate tendency to depression. If the key information, the phenotypes of all affected members, had been published then an independent assessment of the evidence might have been possible by the reader using more robust even though less efficient methods. These include direct counts of cosegregating pairs of alleles and, since this was one large family, of consistent allelic association. As it is, the joint problems of whether a form of depression consequent upon a single gene exists among the Amish and, if so, whether the position of the gene on the short arm of chromosome 11 has been inferred with reasonable precision, remain unclear. In view of the limited number of Amish clarification may not be possible until a further generation has had time to develop psychiatric symptoms.

MULTIPLE MARKERS IN HLA ASSOCIATION STUDIES

Failure to appreciate the consequences of using multiple markers in one family is common. It is, for example, one of the reasons for the falsely optimistic predictions of the cost of linkage studies in Botstein *et al.* (1980). A similar situation arose in studies of the association of disease with alleles at the HLA locus. Whereas in linkage studies the problem is the number of loci at which markers are available or informative, in HLA association it arises from the large number of alleles. In one study of association of HLA with affective disorder a positive association was reported (Govaerts *et al.* 1977). No correction for the number of alleles had been made; when this was done the apparent association declined to well within the likely vagaries of chance (Mendelwicz *et al.* 1981).

It is clear, therefore, from both theory and practical demonstration that linkage will be mistakenly inferred from random assortment when no allowance has been made for multiple tests.

J. H. EDWARDS AND DAVID C. WATT

References

Botstein, D., White, R. L., Skolnick, M. & Davis, R. W. (1980). Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *American Journal of Human Genetics* 32, 314 331.

Egeland, J. A., Gerhard, D. S., Pauls, D. L., Sussex, J. N., Kidd, K. K., Allen, C. R., Hostetter, A. M. & Housman, D. E. (1987). Bipolar affective disorders linked to DNA markers on chromosome 11. *Nature* 325, 783 787.

Govaerts, A., Mendelwicz, J. & Verbanck, P. (1977). Manic-depressive illness and HLA. *Tissue Antigens* 10, 60 62.

Kidd, K. K., Gerhard, D. S., Kidd, J R , Housman, D. & Egeland, J. A. (1984). Recombinant DNA methods in genetic studies of affective disorders. *Clinical Neuropharmacology* 7, Supp. 1, S104.

Kidd, K. K. & Ott, J. (1984). Power and sample size in linkage studies. *Human Gene Mapping* 7, 510 511.

Mendelwicz, J., Verbanck, P , Linkowski, P. & Govaerts, A. (1981). HLA antigens in affective disorders and schizophrenia. *Journal of Affective Disorders* 3, 17–24.

Morton, N. E. (1955). Sequential tests for the detection of linkage. *American Journal of Human Genetics* 7, 277 318.

Ott, J. (1986). Linkage probability and its approximate confidence interval under possible heterogeneity. *Genetic Epidemiology* Supp. 1, 251 257.

Thompson, E. A. (1984). Interpretation of LOD scores with a set of marker loci. *Genetic Epidemiology* 1, 357 362.

11-2