J Med Genet 1993; 30: 638–639

COMMENT

Association *and* linkage: complementary strategies for complex disorders

Nöthen *et al* argue persuasively for the use of association studies to identify the genetic defects responsible for the functional psychoses (schizophrenia and bipolar affective disorder). We too have advocated this approach[1–3] and agree that advances in DNA technology[4] and analytical methods[5 6] have made association studies even more attractive. However, we believe that there are other, complementary approaches which also take into account the probable complexity of the genetics of psychosis.

Nöthen *et al* follow Sobell *et al*[4] in pointing out the advantages of investigating directly sequence variations which are likely to affect protein structure or expression (VAPSEs) in large samples of unrelated patients and controls since this does not depend upon the existence of linkage disequilibrium between marker and disease locus. Although this approach is attractive it might not be so straightforward as is supposed to prove that any variation found in association with a disease is itself responsible for increasing susceptibility rather than another locus close by and in linkage disequilibrium with it. Moreover, we feel that it may be mistaken to dismiss association studies using anonymous markers and relying upon linkage disequilibrium. A number of recent studies have found allelic association between complex diseases and anonymous DNA markers close to candidate genes, for example, the TGF alpha gene in cleft lip and palate,[7] the glucokinase and glycogen synthase genes in non-insulin dependent diabetes (NIDDM),[8 9] the insulin gene and insulin dependent diabetes,[10] and the myelin basic protein gene in multiple sclerosis.[11] Several of these associations were based upon the use of microsatellite markers, the mutation rate of which must therefore be low enough to permit the maintenance of linkage disequilibrium.[12] Linkage disequilibrium mapping by association studies might therefore prove useful in other complex disorders assuming low mutation rates of both marker and susceptibility loci. Under such circumstances, tight linkage can result in allelic association persisting for many generations. For example, at a recombination fraction of 0.01 the 'half life' of an association would be 69 generations or about 2000 years.[13] This means that it might be feasible to undertake a systematic search of the entire genome for association with susceptibility loci once a dense marker map is available with highly informative polymorphisms roughly evenly spaced at about 2 cM intervals. So far such a map is not available, but current progress with automated methods and microsatellite markers is rapid,[14] and the goal is foreseeable. The main disadvantage of a systematic, whole genome search, quite apart from the volume of work, will be the familiar one of multiple testing and how to set significance levels so as to reduce type 1 errors without overlooking true positives. This can, at least in part, be overcome by collecting a large sample and splitting it into two halves, the first being the test sub-sample in which a not too stringent α level (say 0.01) is set, and the second being a replication subsample. The main advantage of a systematic search is that, as in linkage-positional cloning strategies, no previous knowledge of the pathogenesis of the disorder is required.

As Nöthen *et al* point out, there are also methodological pitfalls in association studies arising from population stratification and these are exemplified in publications on classical markers and schizophrenia.[15] The Haplotype Relative Risk (HRR) method of Falk and Rubinstein,[5] and its recent modification by Terwilliger and Ott,[6] is an apparently elegant solution to this difficulty. However, Falk and Rubinstein trios are significantly more difficult to find than single cases, particularly in the psychoses where onset is in adulthood and family relationships often disrupted. Given these difficulties, it may well be better to extend collection slightly and to identify families where there are two or more sibs affected with the disorder as well as living parents. These will then be useful both for association studies using the HRR statistic and for linkage studies using both parametric and non-parametric methods,[16] though there will be some loss of power to the HRR if cases are ascertained upon the basis of familial loading.

Nöthen *et al* are sceptical of the use of linkage analysis in the psychoses, and this is understandable given the recent 'advances' and subsequent retreats in this field.[17] The approach being adopted by the majority of workers is to study the segregation of markers in large multigenerational families using either a candidate gene or a positional cloning approach.[3] As Nöthen *et al* point out this will only detect linkage if a gene or genes of major effect are segregating in at least a proportion of such families. This approach has been successful in some complex disorders, such as Alzheimer's disease,[18 19] NIDDM,[20] and breast cancer.[21] However, in these conditions there was already some evidence before linkage studies that major gene effects were operating in cases characterised by high familial loading and early age of onset.

Whether or not genes of major effect exist in schizophrenia or other psychoses it seems likely that the commonest mode of transmission is either oligogenic or polygenic. This would accord with evidence from animal breeding experiments, as well as human twin and adoption evidence, that nearly all genetically influenced behaviours, normal or abnormal, are likely to reflect the additive or epistatic effects of several, perhaps many, genes at different loci.[22] In the case of schizophrenia, study of the recurrence risks in various classes of relative[23] shows that they best fit an epistatic model of two or three major loci but not a great many loci or a single locus. Similar conclusions have been reached by McGue and Gottesman.[24]

Given these complexities, there are good arguments for conducting linkage analysis in the psychoses in large samples of small nuclear families. In particular, they are less likely than large, 'high density' families to be segregating more than one disease gene and are less sensitive to alterations in phenotypic or genotypic status.[25] Moreover, as we have seen, this approach can readily be combined with association studies using the Falk and Rubinstein method.[5] Analytical methods for studying linkage in complex diseases are developing rapidly[16,26] and data from studies of linkage disequilibrium can be used to increase the power of linkage studies.[27,28] Furthermore, the power of sib pair methods can be increased by adopting a quantitative trait locus (QTL) approach where a continuous measure of susceptibility to disease, for example, 'Schizotypy' scores,[29] is used rather than a simple dichotomy (affected/unaffected). The usefulness of a method, based on regression analysis but requiring no prior assumptions about the mode of transmission of the disorder, has been shown in reading disability.[30] Finally it is worth reminding ourselves that families containing two affected sibs are much more common than multigenerational, multiplex pedigrees. Not only does this mean that they are easier to collect but also that the cases they contain are likely to be more representative of typical forms of psychosis. Thus, even if linkages are detected in studies of large pedigrees, it will be necessary to test their generality in studies of small families.

M J OWEN*†
P McGUFFIN*
*Department of Psychological Medicine, and
†Department of Medical Genetics,
University of Wales College of Medicine,
Heath Park,
Cardiff CF4 4XN, UK.

1 Sturt E, McGuffin P. Can linkage and marker association resolve the genetic aetiology of psychiatric disorders: review and argument. *Psychol Med* 1985;**15**:455–62.
2 Owen MJ. Will schizophrenia become a graveyard for molecular geneticists? *Psychol Med* 1992;**22**:289–93.
3 Owen MJ, McGuffin P. The molecular genetics of schizophrenia: blind alleys, acts of faith, and difficult science. *BMJ* 1992;**305**:664–5.
4 Sobell JL, Heston LI, Sommer SS. Delineation of genetic predisposition to multifactorial disease: a general approach on the threshold of feasibility. *Genomics* 1992;**12**:1–6.
5 Falk CT, Rubinstein P. Haplotype relative risks: an easy reliable way to construct a proper control sample for risk calculations. *Ann Hum Genet* 1987;**51**:227-33.
6 Terwilliger JD, Ott J. A haplotype-based "haplotype relative risk" approach to detecting allelic association. *Hum Hered* 1992;**42**:337–46.
7 Holder SE, Vintiner GM, Farren B, Malcolm S, Winter RM. Confirmation of an association between RFLPs at the transforming growth factor – alpha locus and non-syndromic cleft lip and palate. *J Med Genet* 1992;**29**:393–7.
8 Chiu KC, Province MA, Permutt A. Glucokinase gene is genetic marker for NIDDM in American blacks. *Diabetes* 1992;**41**:843–9.
9 Groop LC, Kankuri M, Schallin C, *et al.* Association between polymorphism of the glycogen synthase gene and non-insulin dependent diabetes mellitus. *N Engl J Med* 1993;**328**:10–14.
10 Bell GI, Horita S, Karam JH, *et al.* A polymorphic locus near the insulin gene is associated with insulin-dependent diabetes mellitus. *Diabetes* 1984;**33**:176–83.
11 Tienari PJ, Wikstrom J, Sajantila A, *et al.* Genetic susceptibility to multiple sclerosis linked to myelin basic protein gene. *Lancet* 1992;**340**:987–91.
12 Todd JA. La carte des microsatellites est arrivee! *Hum Mol Genet* 1992;**1**:663–6.
13 Morton NE. *Outline of genetic epidemiology*. Basel: Karger, 1982.
14 Weissenbach J, Gyapay G, Dib C, *et al.* A second generation linkage map of the human genome. *Nature* 1992;**359**:794–801.
15 McGuffin P, Sturt E. Genetic markers in schizophrenia. *Hum Hered* 1986;**36**:65–88.
16 Hyer RN, Julier C, Buckley JD, *et al.* High resolution linkage mapping for susceptibility genes in human polygenic disease: insulin-dependent diabetes mellitus and chromosome 11q. *Am J Hum Genet* 1992;**48**:243–57.
17 O'Donovan MC, Owen MJ. Advances and retreats in the molecular genetics of major mental illness. *Ann Med* 1992;**24**:171–7.
18 Goate A, Chartier-Harlin M-C, Mullan M, *et al.* Segregation of a missense mutation in the amyloid precursor protein gene with familial Alzheimer's disease. *Nature* 1991;**349**:704–6.
19 Schellenberg GD, Bird TD, Wijsman EM, *et al.* Genetic linkage evidence for a familial Alzheimer's disease locus on chromosome 14. *Science* 1992;**258**:668–71
20 Vionnet N, Stoffel M, Taheda J, *et al.* Nonsense mutation in the glucokinase gene causes early onset non-insulin dependent diabetes. *Nature* 1992;**356**:721–2.
21 Hall JM, Lee MK, Newman B, *et al.* Linkage of early onset breast cancer to chromosome 17q21. *Science* 1990;**250**:1684–9.
22 Plomin R. The role of inheritance in behaviour. *Science* 1990;**248**:183–8.
23 Risch N. Linkage strategies for genetically complex traits. I. Multilocus models. *Am J Hum Genet* 1990;**46**:222–8.
24 McGue M, Gottesman II. Genetic linkage in schizophrenia; perspectives from genetic epidemiology. *Schizophrenia Bull* 1989;**15**:453–64.
25 Greenberg DA. There is more than one way to collect data for linkage analysis. *Arch Gen Psychiatr* 1992;**49**:745–50.
26 Lathrop GM, Ott J. Analysis of complex diseases under oligogenic models and infra familial heterogeneity for the LINKAGE programs. *Am J Hum Genet* 1990;**47** (suppl):A188.
27 Clerget-Darpoux F, Babron MC, Prum B, *et al.* A new method to test genetic models in HLA associated diseases: the MASC method. *Ann Hum Genet* 1988;**52**:247–58.
28 Bain SC, Prins JB, Hearne CM, *et al.* Insulin gene region-encoded susceptibility to type I diabetes is not restricted to HLA-DR4-positive individuals. *Nature Genet* 1992;**2**:212–5.
29 Claridge G. Schizotypy and schizophrenia In: Bebbington P, McGuffin P, eds. *Schizophrenia, the major issues*. London, Heinemann, 1988: 187–201.
30 Fulker DW, Cardon LR, De Fries JC, *et al.* Multiple regression analysis of sib-pair data on reading to detect quantitative trait loci. *Reading and Writing An Interdisciplinary Journal* 1991;**3**:299–313.