Psychological Medicine, 1992, **22**, 289–293
Printed in Great Britain

EDITORIAL

Will schizophrenia become a graveyard for molecular geneticists?[1]

We have little clear understanding of the structural or functional pathology of schizophrenia despite many years of research. The only aetiological factor established beyond reasonable doubt is inheritance; with firm evidence from family, twin and adoption studies (Gottesman & Shields, 1982; McGuffin, 1988). In contrast, advance in human molecular genetics has been prodigious. The genetic defects causing several major diseases such as cystic fibrosis and Duchenne muscular dystrophy have been identified, and it seems certain that successes such as these will soon be followed by similar achievements in a host of other simply inherited diseases. The great power of the new techniques lies in their ability to determine the chromosomal location of disease genes without *a priori* knowledge of pathological processes. The number of known DNA polymorphisms is increasing rapidly and is already sufficient to allow systematic searches for genetic linkage to be made. Alternatively, clues from cytogenetic abnormalities can focus attention upon a particular chromosomal region that is implicated in the pathogenesis of a disease. Once linkage has been established a number of techniques are available that allow the disease gene to be identified and the pathogenic mutation(s) characterized (Whatley & Owen, 1989). The possibility of making a fresh start in this way is made more compelling by the confusing mass of conflicting evidence concerning the biology of schizophrenia.

Initial optimism was tempered by the expectation that there were likely to be difficulties arising not only from the well-known uncertainties in diagnosis and definition of schizophrenia but also from the apparent complexities of its genetic transmission (Sturt & McGuffin, 1985; Kendler, 1987). Successful applications of molecular genetics so far have been to disorders that show simple Mendelian patterns of inheritance. Schizophrenia shows a pattern that is complex and non-Mendelian in common with other prevalent disorders such as bipolar affective illness, Alzheimer's disease, diabetes mellitus and coronary artery disease. Factors such as diagnostic uncertainty, incomplete penetrance, assortative mating, the existence of non-genetic forms of the disease (phenocopies) and the possibility of genetic heterogeneity would, it was argued, all militate against the finding of linkage even where this exists.

In spite of these reservations, a spectacular start was made when Sherrington and colleagues (1988) presented evidence for genetic linkage between schizophrenia and two DNA polymorphisms on the long arm of chromosome 5. They had been encouraged to study this chromosomal region by the report of a family in which a schizophrenic illness was segregating with an unbalanced translocation of chromosome 5, so that the portion of the chromosome was trisomic in the affected individuals (Bassett *et al.* 1988). However, euphoria rapidly turned to disquiet as several failures to replicate this linkage were reported (Kennedy *et al.* 1988; St Clair *et al.* 1989; Detera-Wadleigh *et al.* 1989; Kaufman *et al.* 1989; McGuffin *et al.* 1990; Aschauer *et al.* 1990). Initially it was proposed that this might reflect 'non-allelic genetic heterogeneity' with a defect on chromosome 5 accounting for only a minority of familial schizophrenia and the majority resulting from defects at another locus or loci (Lander, 1988). However, such an explanation has become increasingly untenable as the number of negative reports has increased, and it seems unlikely that the findings of Sherrington *et al.* are correct. Disappointment has been compounded by the failure of linkage between bipolar affective disorder and markers on chromosome 11 in an Amish pedigree to hold up to further scrutiny (Kelsoe *et al.* 1989).

These findings and the reasons for them have been discussed at length elsewhere (Owen & Mullan,

[1] Address for correspondence: Dr M. J. Owen, Department of Psychological Medicine, University of Wales College of Medicine, Heath Park, Cardiff CF4 4XN.

1990; Owen *et al.* 1990; McGuffin *et al.* 1990). Briefly, it is likely that the false positives reflected spurious inflation of the evidence for linkage by multiple testing with several different phenotypic definitions and estimates of linkage parameters, though systematic error cannot be ruled out. We should also not forget that there is a high risk of false negative findings for reasons already described. A number of methodological refinements have been suggested to reduce the risk of both type 1 and type 2 errors (Baron *et al.* 1990). These include scrupulous 'blindness' between diagnostic procedures and the determination of genotypes, giving greater weight to narrowly defined diagnostic categories and conservative definitions of the phenotype, systematic evaluations of pedigrees for the possibility that disease may be entering from more than one source (bilineal transmission) and more stringent statistical requirements. The conventional criterion for acceptance of linkage is a lod score of 3 (Morton, 1955). This corresponds to odds on linkage of 1000 to 1 but since the prior probability of detecting linkage between two loci is low (it is much lower than the probability of approximately 0·054 that two genes taken at random are on the same chromosome) a lod score of 3 corresponds to a posterior probability or 'reliability' of 0·95 (Morton, 1955, 1982). This only pertains to two-point analysis where both the main trait and the marker locus have simple and well-established modes of transmission. When data from several markers are analysed simultaneously (multi-point analysis) and especially where the mode of transmission of the main trait is unknown, the situation becomes much less clear. One solution is that if there is to be multiple testing with different markers combined with the effects of exploring different models of transmission and different definitions of the phenotype, linkage studies of diseases such as schizophrenia should only be considered definitely positive if a very high lod score is achieved. The danger then is that we become over-conservative and the risk of type 2 errors becomes great. A pragmatic alternative would be to retain a conventional score of 3 as indicating a possible positive which requires independent replication on a second sample before being regarded as a probable positive (McGuffin *et al.* 1990).

Failure to replicate the chromosome 5 linkage has led to disenchantment as well as dysphoria. Indeed, some are beginning to remark that as well as being the 'graveyard of neuropathologists' (Plum, 1972) schizophrenia may also be the 'graveyard of molecular geneticists'. Although the occurrence of false positive findings suggests that refinements in methodology are required, it tells us little about the chances of ultimate success in applying molecular genetics to schizophrenia. The wisdom or otherwise of this endeavour can only be judged by examining the premises upon which it is based and in consideration of these we must turn to what is known of the mode of transmission of the disorder.

Although there is good evidence from recent quantitative genetic analyses that genes account for most of the variance in liability to schizophrenia (McGue *et al.* 1985), the mode or modes of transmission remain obscure. It seems unlikely that a single major locus which is the sole source of resemblance between relatives can account for the transmission of schizophrenia as a whole. It is also unlikely that schizophrenia represents a heterogeneous mixture of single gene disorders only (McGue & Gottesman, 1989). Two major classes of explanation remain and it is probably beyond the power of currently available quantitative genetic techniques to distinguish between them. The first hypothesis is that schizophrenia results entirely from multifactorial polygenic inheritance. According to this view, schizophrenia reflects variation at a number of different genetic loci in combination with a number of different environmental contributors along the lines of say height, blood pressure and intelligence (Gottesman & Shields, 1982). The second kind of model, which is also compatible with the data, recognizes the importance of both polygenic, and major gene, effects. In the 'mixed' model it is proposed that there is a gene of major effect or several different such genes which together with polygenic or multifactorial environmental factors contribute to the familiality of schizophrenia. Of course this begs the question 'how big is major?' In practical terms the best answer is probably 'an effect that is large enough to be detected by linkage analysis'. Finally, it can be postulated that schizophrenia is aetiologically heterogeneous with some cases caused by defects in a single gene, others caused by multifactorial polygenic inheritance with or without major gene effects and finally a group of primarily environmental 'phenocopies'.

If the polygenic theory is correct, then the application of molecular genetics through linkage studies of multiplex families is unlikely to be successful. However, it will be possible to detect genes of major effect using this approach, although, as we have seen, this will be made more difficult by such factors as diagnostic difficulties, incomplete penetrance and the possibility that defects in several genes may result in a schizophrenic phenotype (Sturt & McGuffin, 1985; Owen & Mullan, 1990). However, even if there is quite marked genetic heterogeneity, it should be possible to detect linkage, although it will be necessary to use large sample sizes (Martinez & Goldin, 1989) which will be beyond the resources of a single centre. For this reason a multicentre collaborative programme is being established under the auspices of the European Science Foundation and a similar collaboration is commencing in the US supported by NIMH. Both will take up to five years to complete but should be successful in detecting and locating a major gene (or major genes) if such exist. If defects in a single gene account for the great majority of cases of familial schizophrenia then positive results can be expected from individual centres within the next one or two years given the recent availability of highly polymorphic marker systems (e.g. Weber & May, 1989). Alternatively, groups with access to very large multiply affected families might be successful in demonstrating linkage in the face of heterogeneity, assuming that such extended pedigrees are not the result of undetected bilineal transmission.

The truth of the matter therefore is that, since we do not know the true mode of transmission of schizophrenia, we cannot predict how successful linkage approaches will be. However, recent findings in other common familial diseases showing complex, non-Mendelian transmission are grounds for optimism.

The transmission of both coronary artery disease and diabetes mellitus reflects major gene effects. In an individual heterozygous for a defect in the low density lipoprotein (LDL) receptor gene, elevated serum LDL-cholesterol causes symptomatic arterial disease by age 40 or so. In this case knowledge of the biochemical defect led to the disease gene without the need for linkage analysis.

Familial combined hyperlipidaemia (FCHL) leads to premature coronary artery disease in about one-fifth of cases and appears to have a dominant mode of transmission with reduced penetrance. Recent work has employed both association and linkage analysis to demonstrate tight linkage to the apolipoprotein A1-CIII-AIV gene cluster on chromosome 11q23–q24 in a proportion of families with FCHL (Wojciechowski *et al.* 1991).

Over half the genetically determined susceptibility to insulin-dependent diabetes mellitus (IDDM) maps to the HLA region of chromosome 6 (Todd *et al.* 1988). Diabetes mellitus also displays aetiological heterogeneity with at least one variety of the disease; maturity–onset diabetes of the young (MODY) apparently showing Mendelian inheritance (Tattersall & Fajans, 1975).

In both coronary artery disease and diabetes phenotypic and endophenotypic heterogeneity have helped in substantiating the existence of aetiological heterogeneity. We have so far been unsuccessful in 'splitting' schizophrenia in this way (McGuffin *et al.* 1987) and may have to proceed without the luxury of phenotypic pointers to aetiological heterogeneity apart from familial aggregation. Recent linkage studies of Alzheimer's disease have been carried out against a similar conceptual background. These have been successful in demonstrating that a proportion of familial cases result from a genetic defect on chromosome 21 (St. George-Hyslop *et al.* 1987; Goate *et al.* 1989) in spite of quite considerable heterogeneity (St. George-Hyslop *et al.* 1990). Similar findings with respect to linkage with markers on chromosome 17 have recently been reported for familial breast cancer (Hall *et al.* 1990) although these await replication.

If we are unable to demonstrate linkage in schizophrenia despite having studied a large number of markers spread throughout the genome in a sufficiently large number of families then we will have provided strong evidence against the involvement of major genes and in favour of polygenic causation. It may well be that this is the only way of resolving current uncertainties concerning the mode of transmission . One may ask how we shall know when sufficient markers and families have been studied. Fortunately the development of computer simulation techniques (e.g. Ott, 1989) now allows estimation of the power of different combinations of markers and pedigrees to detect linkage under different assumptions of mode of transmission and degrees of heterogeneity. Simulation

studies are likely to assume increasing importance by allowing the accurate interpretation of negative linkage findings.

The absence of major gene effects need not preclude the application of molecular genetics. It is possible to detect genes that contribute to polygenic inheritance using association strategies (Sturt & McGuffin, 1985; Plomin, 1990). In these studies, frequencies of different alleles of a genetic marker are studied in a sample of patients with the disorder and either a sample of controls without the disorder or a sample drawn from the general population. This is in contrast to linkage studies in which the co-segregation of a marker and the main (disease) trait is examined in families containing several members with the disease (Sturt & McGuffin, 1985; Owen & Whatley, 1988). Association studies can detect genes that contribute as little as 1% to the overall liability of developing a disorder. However, such studies will only reveal susceptibility loci when the marker phenotype is itself of pathological significance or when the marker locus is so close to the susceptibility locus as to be in linkage disequilibrium with it. Moreover, the presence of linkage disequilibrium depends upon low rates of mutation at both susceptibility and marker loci. Examples of the fruitful application of this approach come from HLA associations observed in a number of autoimmune diseases.

There have been a number of association studies in schizophrenia using the so-called 'classical' markers that were available before the advent of recombinant DNA technology such as the blood groups ABO and Rh and the human leucocyte antigen (HLA) system. These have produced inconsistent and contradictory findings (McGuffin & Sturt, 1986; Owen & McGuffin, 1991). To some extent this may have been due to the lack, in early investigations, of strict operational diagnostic criteria. However, there are other methodological problems such as stratification and multiple testing which have also probably operated to produce type 1 errors. The only association to have been independently replicated is that between HLA A9 and the paranoid subform of the illness. The strength of this association is low, with a calculated combined relative risk of paranoid schizophrenia in A9 positive *versus* A9 negative individuals of 1·6 ($\pm$0·1) (McGuffin & Sturt, 1986). There are also several problems with this finding (McGuffin & Sturt, 1986; Owen & McGuffin, 1991). However, it does suggest that in principle at least, association studies may prove fruitful in schizophrenia.

There have been few association studies of schizophrenia using DNA markers. The genetic map is not yet dense enough to allow a systematic search for association using anonymous markers: at least 1500 evenly spaced markers will be required for this. Sufficient markers should be available by the end of the decade given the rate of current advance and the resources available for mapping the human genome. In addition, the development of automated methods for marker typing will render a study of this magnitude increasingly possible. As with linkage studies, it is likely that type 1 errors will be generated by multiple testing and, once again, we should consider positive findings as provisional until they are replicated in an independent sample (McGuffin & Sturt, 1986). In the shorter term association studies should probably focus upon 'candidate' genes encoding proteins of structural or functional relevance to the nervous system. Over 150 such genes have been identified and this number is growing rapidly. It is worth noting that a positive association with a DNA marker at a candidate gene, tyrosine hydroxylase, has been reported for bipolar affective disorder (Leboyer *et al.* 1990). This merits further attention and attempts at replication, as well as suggesting that association with candidate genes is already worth exploring in schizophrenia.

In conclusion, the application of molecular genetics to the study of schizophrenia is still in its infancy, and talk of graveyards is premature. Advances are being made in understanding other common diseases showing complex inheritance and these are grounds for cautious optimism which should be tempered by awareness of the additional difficulties inherent in studying psychiatric diseases. However, there is a strong tradition of scepticism in British academic psychiatry and some, perhaps many, will remain unconvinced until positive findings emerge and are replicated.

MICHAEL J. OWEN

Peter McGuffin and Michael O'Donovan kindly commented upon an earlier version of this paper. Andrea Seymour typed the manuscript.

REFERENCES

Aschauer, H. N., Aschauer-Trieber, G., Isenberg, K. E., Todd, R. D., Knesevich, M. A., Garver, D. L., Reich, T. & Cloninger, C. R. (1990). No evidence for linkage between chromosome 5 markers and schizophrenia. *Human Heredity* **40**, 109–115.

Baron, M., Endicott, J. & Ott, J. (1990). Genetic linkage in mental illness: limitations and prospects. *British Journal of Psychiatry* **157**, 645–655.

Basset, A., McGillvray, B. C., Jones, B. D. & Pantzar, J. T. (1988). Partial trisomy chromosome 5 co-segregating with schizophrenia. *Lancet* i, 799–801.

Detera-Wadleigh, S. D., Goldin, L. R., Sherrington, R., Encio, I., de Miguel, C., Berrettini, W., Gurling, H. & Gershon, E. S. (1989). Exclusion of linkage to 5q11–13 in families with schizophrenia and other psychiatric disorders. *Nature* **340**, 391–392.

Goate, A. M., Haynes, A. R., Owen, M. J., Farrall, M., James, L. A., Lai, L. Y. C., Mullan, M. J., Roques, P., Rossor, M. N., Williamson, R. & Hardy, J. A. (1989). Predisposing locus for Alzheimer's disease on chromosome 21. *Lancet* i, 355–359.

Gottesman, I. I. & Shields, J. (1982). *Schizophrenia: The Epigenetic Puzzle*. Cambridge University Press: Cambridge.

Hall, J. M., Lee, M. K., Newman, B., Morrow, J. E., Anderson, L. A., Hyey, B. & King, M.-C. (1990). Linkage of early-onset familial breast cancer to chromosome 17q21. *Science* **250**, 1684–1689.

Kaufman, C. A., DeLisi, E., Lehner, T. & Gilliam, T. C. (1989). Physical mapping, linkage analysis of the putative schizophrenia locus on chromosome 5q. *Schizophrenia Bulletin* **15**, 441–452.

Kelsoe, J. R., Ginns, E. I., Egeland, J. A., Gerhard, D. S., Goldstein, A. M., Bale, S. J., Pauls, D. L., Long, R. T., Kidd, K. K., Conte, G., Housman, D. E. & Paul, S. M. (1989). Re-evaluation of the linkage relationship between chromosome 11p loci and the gene for bipolar affective disorder in the Old Order Amish. *Nature* **342**, 238–243.

Kendler, K. S. (1987). The feasibility of linkage studies in schizophrenia. In *Biological Perspectives of Schizophrenia* (ed. H. Helmchen and F. A. Henn). John Wiley and Sons: Chichester.

Kennedy, J. L., Giuffra, L. A., Moises, H. W., Cavalli-Sforza, L. L., Pakstis, H. J., Kidd, J. R., Castiglione, C. M., Wetterberg, L. & Kidd, K. K. (1988). Evidence against linkage of schizophrenia to markers on chromosome 5 in a northern Swedish pedigree. *Nature* **336**, 167–170.

Lander, E. S. (1988). Splitting schizophrenia. *Nature* **336**, 105–106.

Leboyer, M., Malafosse, A., Boularand, S., Campion, D., Gheysen, F., Somolyk, D., Henriksson, B., Denise, E., des Lauriers, A., Lepine, J. P., Zarfian, E., Clerget-Darpoux, F. & Mallet, J. (1990). Tyrosine hydroxylase polymorphisms associated with manic-depressive illness. *Lancet*, **335**, 1219.

McGue, M. & Gottesman, I. I. (1989). Genetic linkage in schizophrenia: perspectives from genetic epidemiology. *Schizophrenia Bulletin* **15**, 453–464.

McGue, M., Gottesman, I. I. & Rao, D. C. (1985). Resolving genetic models for the transmission of schizophrenia. *Genetic Epidemiology* **2**, 99–110.

McGuffin, P. (1988). Genetics of schizophrenia. In *Schizophrenia: The Major Issues* (ed. P. Bebbington and P. McGuffin), pp. 107–126. Heinemann Medical: London.

McGuffin, P. & Sturt, E. (1986). Genetic markers in schizophrenia. *Human Heredity* **36**, 65–88.

McGuffin, P., Farmer, A. & Gottesman, I. I. (1987). Is there really a split in schizophrenia? The genetic evidence. *British Journal of Psychiatry* **150**, 581–592.

McGuffin, P., Sargeant, M., Hett, G., Tidmarsh, S., Whatley, S. & Marchbanks, R. M. (1990). Exclusion of a schizophrenia susceptibility gene from the chromosome 5q11–q13 region. New data and reanalysis of previous reports. *American Journal of Human Genetics* **47**, 524–535.

Martinez, M. M. & Goldin, L. R. (1989). The detection of linkage and heterogeneity in nuclear families for complex disorders; 1 *v.* 2 marker loci. *American Journal of Human Genetics* **44**, 552–559.

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

Morton, N. E. (1982). *Outline of Genetic Epidemiology*. Karger: Basel.

Ott, J. (1989). Computer-simulation methods in human linkage analysis. *Proceedings of the National Academy of Sciences, USA*, **86**, 4175–4178.

Owen, M. J. & Mullan, M. J. (1990). Molecular genetic studies of manic-depression and schizophrenia. *Trends in Neuroscience* **13**, 29–31.

Owen, M. J. & Whatley, S. A. (1988). Polymorphic DNA markers and mental disease. *Psychological Medicine* **18**, 529–533.

Owen, M. J. & McGuffin, P. (1991). DNA and classical genetic markers in schizophrenia. *European Archives of Psychiatry and Clinical Neuroscience* **240**, 197–203.

Owen, M. J., Crawford, D. & St. Clair, D. (1990). Localization of a susceptibility locus for schizophrenia on chromosome 5. *British Journal of Psychiatry* **157**, 123–127.

Plomin, R. (1990). The role of inheritance in behaviour. *Science*, **248**, 183–188.

Plum, F. (1972). Prospects for research on schizophrenia. 3. Neurophysiology, neuropathological findings. *Neurosciences Research Program Bulletin* **10**, 384–388.

St. Clair, D., Blackwood, D., Muir, W., Baillie, D., Hubbard, S., Wright, A. & Evans, H. J. (1989). No linkage of chromosome 5q11–q13 markers to schizophrenia in Scottish families. *Nature* **339**, 305–309.

St. George-Hyslop, P. H., Tanzi, R. E., Polinsky, R. J., Haines, J. L., Drachman, D., Growdon, J., Bruni, A., Foncin, J.-F., Salmon, D., Frommelt, P., Amaducci, L., Sorbi, S., Piacentini, S., Stewart, G. D., Hobbs, W. J., Conneally, M. & Gusella, J. F. (1987). The genetic defect causing familial Alzheimer's disease maps on chromosome 21. *Science* **235**, 885–890.

St. George-Hyslop, P. H., Haines, J. L., Farrer, L. A., Polinsky, R., Van Broeckhoven, C., Goate, A., Crapper-McLachlan, D. R., Orr, H., Monesi, N. P., Sorbi, S., Rainero, I., Foncin, J.-F., Pollen, D., Cantu, J.-M., Vartanjan, M., Mayeux, R., Nee, L., Backhovens, H., Martin, J.-J., Rossor, M. N., Owen, M. J., Mullan, M., Percy, M. E., Karlinsky, H., Rich, S., Growdon, J., Heston, L., Gusella, J. F. & Hardy, J. A., (1990). Genetic linkage studies suggest that Alzheimer's disease is not a single homogenous disorder. *Nature* **347**, 194–197.

Sherrington, R., Brynjolfsson, J., Petursson, H., Potter, M., Dudleston, K., Barraclough, B., Wasmuth, J., Dobbs, A. & Gurling, H. (1988). Localization of susceptibility locus for schizophrenia on chromosome 5. *Nature* **336**, 164–167.

Sturt, E. & McGuffin, P. (1985). Can linkage and marker association resolve the genetic aetiology of psychiatric disorders: review and argument. *Psychological Medicine* **15**, 455–462.

Tattersall, R. B. & Fajans, S. (1975). A difference between the inheritance of classical juvenile-onset and maturity-onset type diabetes of young people. *Diabetes* **24**, 44–53.

Todd, J. A., Bell, J. I. & McDevitt, H. O. (1988). A molecular basis for genetic susceptibility to insulin-dependent diabetes mellitus. *Trends in Genetics* **4**, 129–134.

Weber, J. L. & May, P. E. (1989). Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. *American Journal of Human Genetics* **44**, 388–396.

Whatley, S. A. & Owen, M. J. (1989). The application of molecular genetics to psychiatric disorder. *International Review of Psychiatry* **1**, 219–230.

Wojciechowski, A. P., Farrall, M., Cullen, P., Wilson, T. M. E., Bayliss, J. D., Farren, B., Griffin, B. A., Caslake, M. J., Packard, C. J., Shepherd, J., Thakker, R. & Scott, J. (1991). Familial combined hyperlipidaemia linked to the apolipoprotein AI-CIII-AIV gene cluster on chromosome 11q23–q24. *Nature* **349**, 161–164.