Eur Arch Psychiatry Clin Neurosci (1991) 240:169–173

European Archives of Psychiatry and Clinical Neuroscience

The Molecular Genetics of Schizophrenia: An Overview and Forward View

Peter McGuffin[1] **and Michael Owen**[1,2]

[1]Department of Psychological Medicine, University of Wales College of Medicine, Heath Park, Cardiff, UK
[2]Institute of Medical Genetics, University of Wales College of Medicine, Heath Park, Cardiff, UK

The Genetic Basis

The familial nature of schizophrenia is not in doubt, and clustering of the disorder in families has been one of the most constant observations since schizophrenia was first described (Bleuler 1911). Occasionally the clinician interested in familial clustering will come across a pedigree like the one illustrated in Fig. 1, which is heavily "loaded" and features the disorder in several consecutive generations. However, even in families of this type, schizophrenia does not usually comply with the laws of simple Mendelian segregation. For example, the pedigree illustrated, "family 15" from a study by McGuffin et al. (1990), at first glance appears to show a dominant pattern. However, it contains one individual (No. 3) in generation III who is quite healthy but whose mother, individual No. 1 in generation II and daughter, individual No. 5 in generation IV, are both affected. Skipping of generations and other irregularities are the rule rather than the exception in families multiply affected by schizophrenia, and indeed the majority of schizophrenic patients, about two-thirds even in intensively studied series (Bleuler 1978), appear to have no family history of the disorder. Therefore, schizophrenia, along with such conditions as Alzheimer's disease, coronary heart disease and diabetes, falls into the category of common and complex disorders for which straightforward single gene hypotheses do not offer a satisfactory explanation and which can be contrasted with those disorders that have regular Mendelian patterns of transmission; Huntington's disease and Wilson's disease being two examples of psychiatric interest, which are generally much rarer.

In the face of irregular patterns of transmission, the inferences that can be drawn from clinical data alone are severely limited. In particular, it needs to be borne in mind that certain traits can sometimes simulate Mendelism when in fact there is no monogenic transmission or even where there is no genetic transmission at all (Edwards 1960). For example, we have recently shown that a "trait" long known to be familial, i.e. attending medical school, is 60 times more frequent in the first-degree relatives of Cardiff medical students than in the popula-

Offprint requests to: P. McGuffin

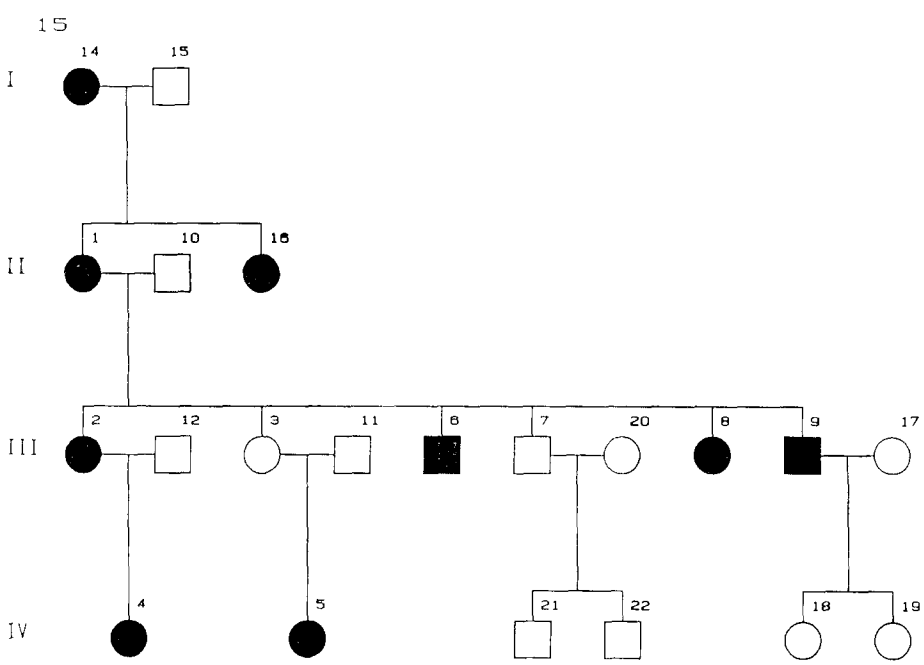

Fig. 1. A multiplex pedigree containing individuals affected by schizophrenia or other non affective psychosis (*filled squares* and *circles*) in four generations. (From McGuffin et al. 1990)

tion at large and furthermore tends to follow a recessive-like pattern of transmission within families which satisfies most of the statistical tests conventionally carried out when modern sophisticated methods of complex segregation analysis are applied (McGuffin and Huckle 1990). Despite this, the authors hope that no-one plans to go out and study the molecular genetic basis of the trait!

However, why then should one feel on firmer ground with schizophrenia, and why should a special issue of *European Archives* be devoted to considering the molecular genetics of this condition? The principal answer is that twin and adoption studies provide convincing evidence that schizophrenia aggregates in families, not because of "cultural" transmission or because of the effects of common family environment but because of genetic factors. It is not within our scope to summarise the genetic evidence here and expert reviews of the classical studies and of modern investigations using twin and adoption designs are provided elsewhere, for example by Gottesman and Shields (1982). Nevertheless, it is pertinent to consider the clinical issues of phenotype definition, strategies for sampling in linkage studies and exploration of the plausible modes of transmission of schizophrenia. Although none of these can strictly be described as belonging to the subject matter of molecular genetics, all are absolutely essential in providing underpinnings for the application of molecular genetic techniques in schizophrenia research and are hence included in this special issue.

Definition of the Phenotype

We have already described a striking but, in the statistical sense, atypical pedigree containing affected individuals in four generations. This issue of *European Archives* contains a description by Wetterberg and Farmer of an even more striking and unusual very large kindred multiply affected by psychosis. Such pedigrees are of great interest and high potential value, since on their own they can provide a sample of sufficient size to give firm evidence of linkage. On the other hand, a disadvantage of such pedigrees, as discussed by Goldin et al. in this issue, is that they may have many affected members because of the segregation of an atypical mutant gene which cannot be found elsewhere. Moreover, when we are dealing with a comparatively common disorder, there may even be two different mutant genes segregating within the pedigree, thus presenting problems for linkage analysis interpretation if one gene is linked and the other unlinked to a certain DNA marker. The other major difficulty is whether at a purely phenotypic level such families contain the same clinical disorder as the more usual, smaller families. Here it becomes absolutely essential to have some standard method of rating signs and symptoms and of arriving at reliable operational diagnostic criteria. The paper by Wetterberg and Farmer suggests that the large Swedish pedigree in which molecular genetic studies have already begun (Kennedy et al. 1988) contains affected individuals who, when a variety of operational definitions are applied, turn out to have a typical schizophrenic pattern of symptoms.

The comparatively recent introduction of operational diagnostic criteria for schizophrenia raises questions not just concerning large atypical pedigrees but has led some to question the conventional acceptance of a genetic contribution to schizophrenia. However, it is worth noting that when strict and reliable modern criteria have been applied to data from an adoptees' family study (Kendler et al. 1981) and to twin study material (McGuffin et al. 1984), the familiality and heritability of the disorder were confirmed. Nevertheless, it was found that not all definitions of schizophrenia provide the same level of heritability and therefore the problem remains of which definition should best be selected for genetic research. At the present time this question is unanswered and therefore we tend to favour a "polydiagnostic" approach where the clinical data are collected in such a way that a variety of competing definitions of the phenotype can be applied. One interesting example of the exploration of phenotypic definitions of varying breadth was provided by the study of Farmer et al. (1987): in a series of monozygotic (MZ) and dizygotic (DZ) twins, a range of DSM-III (American Psychiatric Association 1980) categories was tried in combination with the DSM-III category of schizophrenia to see whether a "more genetic" phenotype resulted. These authors used a somewhat crude but nevertheless useful index, the MZ/DZ concordance ratio, as their measure and showed that for schizophrenia alone this was 5.0 but could be "improved" to 7.7 when the phenotype was broadened to include schizotypal personality, atypical psychosis and affective disorder with mood incongruent delusions. However, when the phenotype was broadened further to include other types of affective disorder, the MZ/DZ concordance ratio began to fall and dropped down to 2.4 when any axis-1 diagnosis was included. It seems probable that attempts to "optimise" the phenotype such as this will become even more feasible if a well-replicated linkage marker is discovered and when the optimisation criterion can then be the magnitude of the lod score (see below) rather than MZ/DZ concordance ratio.

Models and Assumptions

The principal strategies employed in using DNA polymorphisms or indeed any other type of genetic marker in studies of schizophrenia are summarised in this issue by Owen and McGuffin. *Association* studies require no major assumption other than the probable existence of a genetic contribution to the disorder. The idea, then, is simply to detect whether a particular marker allele is more common in a sample of schizophrenics than in healthy controls. By contrast, in *linkage* studies, the focus is on families where the cosegregation of the disease and a marker gene is studied. Here the aim is to detect departure from Mendelian-independent assortment. If we consider a marker locus with two alleles each will, on average, be as common in the affected members of a family as in the unaffected members when the marker and the disease locus are on different chromosomes or are widely separated on the same chromosome, i.e. when there is no linkage and independent assortment occurs. If, on

the other hand, the marker and disease loci are in close proximity on the same chromosome, i.e. they are linked, there will be a tendency for one allele to be more commonly found in diseased members of the family and the other more common in the healthy members. In order to infer that linkage is present, it is not necessary to observe that all the affecteds carry one allele and all of the unaffecteds another, since the association of alleles at linked loci is affected by the phenomenon of recombination or crossing over during meiotic cell division in gametes when homologous pairs of chromosomes lying alongside each other exchange pieces of material.

Within certain limits, the frequency of recombination between two loci is roughly proportional to the distance between them. Thus, for a disease unlinked to a marker locus, when half the time the affecteds carry one marker allele and half the time the other, the recombination fraction is a half.

However, where the disease and the marker are linked, the number of non-recombinant individuals will outnumber the recombinants; the recombination fraction is less than one-half. The aim of linkage studies is therefore to *detect* recombination fractions of less than a half and *estimate* the size of the recombination fraction. The usual statistical approach is to apply a likelihood method and plot a curve of the log of the odds (lod) that the recombination fraction has a certain value, θ, as against a value of one half (Morton 1955). Where the lod score reaches its maximum value provides the maximum likelihood estimate of θ, the recombination fraction. By convention, a lod score of 3 is accepted as firm evidence in favour of linkage, while a lod score of -2 means that linkage can be rejected. However, these values originally proposed by Morton (1955) strictly only apply when one is dealing with simple Mendelian traits where the mode of inheritance is known with certainty. Clearly, as we have already discussed, this is not the case with schizophrenia, where the mode of transmission is unknown. (Uncertainty about the interpretation of lod scores in studies of schizophrenia and how this was resulted in some controversy and confusion is discussed further in this issue by Owen and McGuffin.)

Also in this issue, McGue and Gottesman discuss the possible modes of transmission of schizophrenia and conclude that a pure single locus model, that is a model where a single gene is the sole source of resemblance between relatives, cannot explain the transmission of schizophrenia satisfactorily. They also argue that a simple heterogeneity hypothesis of multiple different single loci can be mathematically rejected. What cannot be rejected is the possibility of a mixed model of transmission (Morton and Maclean 1974), i.e. a major locus plus multifactorial effects contributed either by polygenes or common family environment. If such a major locus exists then obviously linkage strategies become viable, and indeed it is no longer technically difficult to analyse linkage data for disorders that do not show regular Mendelian segregation using computer-based packages (Lathrop and Lalouel 1984). Such programs allow corrections for incomplete penetrance (i.e. where the probability of manifesting the disorder given a certain genotype is less than 1) and variable age of onset. It is also somewhat reassuring that errors in diagnostic classification or misspecification of the mode of transmission should not usually lead to a spurious detection of linkage where in fact no linkage is present (Clerget-Darpoux et al. 1986). However, misspecification and misclassification can result in failure to detect linkage even when true linkage is present.

In the face of these difficulties most researchers attempting linkage studies in schizophrenia have taken a pragmatic approach and have concentrated on "loaded" or multiplex families in which two or more individuals are affected by schizophrenia. The assumptions then are: first that a major gene is segregating in at least some families, second that genetic homogeneity can be assumed within families and third, that the mode of transmission can be inferred approximately (or at least that a plausible and limited range of models of transmission can be explored). To what extent these assumptions are safe is dealt with in part in the paper of McGue and Gottesman in this issue. They point out that on the basis of their simulation studies, sampling multiplex families does not necessarily ensure that all individuals who are affected carry a major gene if that gene happens to have a comparatively low frequency and high penetrance. However, if the gene frequency is high and the penetrance low, affected individuals in multiplex families will nearly always carry the major gene.

Sampling Strategies

The whole issue of appropriate and efficient sampling strategies is obviously of vital importance to the design of linkage studies and this is discussed in greater detail in this issue by Goldin et al. The results of their simulation studies give general encouragement to would-be investigators of linkage in schizophrenia and provide a partial antidote for some of the pessimism which might be engendered by McGue and Gottesman's modelling. Goldin et al. examined the three alternative sampling strategies that can be employed, which range from concentrating on very large multiplex pedigrees, such as the ones described in this issue by Wetterberg and Farmer, to a focus on affected sibling pairs and their parents. In the middle are studies where nuclear families or pedigrees of moderate size are studied. The affected sib-pair strategy has proven to be attractive to some investigators because of the simplicity of the analysis compared with the lod score method and because such analysis are "model-free" and should be robust to the misspecification of mode of transmission. However, sib-pair methods lack power (Sturt and McGuffin 1985), and Goldin et al. concluded that the sib-pair approach is in general less satisfactory than the analysis of families and large pedigrees.

Other interesting and informative results of the studies of Goldin and colleagues include the fact that even in the presence of heterogeneity linkage can be detected in clinically realistic sample sizes. However, there are limits on the detection of linkage so that, for example, heterogeneity presents a serious problem if the gene for schizophrenia is linked to the markers locus in less than 25% of families. Below this limit, only close linkage

can be detected, which would mean that a far denser marker map would be required than is now currently available.

The Findings to Date

The results of linkage and association studies with genetic markers in schizophrenia are reviewed in this issue by Owen and McGuffin. Their survey includes not just DNA markers which now provide a virtually complete, if still low-resolution, map of the human genome (Donis-Keller et al. 1987), but also characters now referred to as "classical" markers, including red cell groups, the HLA system and various protein polymorphisms detectable in the blood. Studies with classical markers have provided a useful test ground for the application of association and linkage strategies in schizophrenia, but the problem has always been that even with good collaboration between laboratories it has seldom been possible to study more than about 25 loci. This provides grossly inadequate coverage of the human gene map and about ten times that number of marker loci are probably required for an adequate systematic search for schizophrenia genes using a linkage approach. Owen and McGuffin briefly describe how restriction fragment length polymorphisms (RFLPs) can be used as random markers throughout the genome or can be applied as potential "candidate genes" where an RFLP is identified using a DNA probe containing a sequence encoding for a protein of neurobiological interest such as a neurotransmitter or an enzyme involved in neurotransmitter metabolism.

At present, there are no well-replicated positive linkage findings in schizophrenia, but there has been much recent interest concerning the report of linkage between a putative schizophrenia susceptibility gene and DNA markers in the q11–q13 region of chromosome 5 (Sherrington et al. 1988). This study, the earlier report of an uncle-nephew pair with schizophrenia and partial trisomy of long arm of chromosome 5 (Basset et al. 1988), together with subsequent failures to detect linkage with markers on chromosome 5q are reviewed by Owen and McGuffin. In addition, one of the first groups to report exclusion of chromosome 5q linkage, Blackwood and coworkers, re-examine their findings in this issue of *European Archives*. This group was unable to detect any linkage and were able to exclude the schizophrenia gene from the chromosome 5q11–q13 region in an impressive collection of moderately large pedigrees from Scotland multiply affected with schizophrenia. Their results provoked some controversy because the application of operational criteria resulted in some individuals in certain families being classified as suffering from bipolar affective disorder. This called into question whether the families studied by the Scottish researchers could appropriately be compared with the clinical material in the positive linkage study by Sherrington et al. (1988). However, in this issue, Blackwood and co-workers argue against reliance purely on clinical signs and symptoms when attempting a classification of affecteds and unaffecteds in linkage studies. Instead they suggest that biological vulnerability markers can provide an important aid to definition of the phenotype and, in particular, they favour measurement of smooth pursuit eye-tracking and an invoked electroencephalographic measure, the auditory P-300. The approach they suggest thus echoes to the theme of earlier authors such as Gottesman and Shields (1972) that reliance on "exophenotypes", clinical signs and symptoms, should be replaced eventually by more dependable "endophenotypes" which might lie a step nearer to the genotype itself.

The future

Although this issue of *European Archives* is devoted to a potentially broad subject, the molecular genetics of schizophrenia, there has been a deliberate emphasis on studies using DNA polymorphisms as linkage markers and much discussion of the relevant clinical and genetic epidemiological issues. Inevitably, in attempting to focus on areas of current interest, some topics have been left out and in particular there has been no coverage of the control of gene expression in the brain and its relevance to the study of mental disease (Harrison and Pearson, 1989) or of the genetic control of brain development (Bloom 1990). However, we believe that a focus on linkage strategies is justifiable, as for the present these appear to offer the most promising and tangible ways of proceeding. The potential benefits of discovering well replicated linkage markers for schizophrenia are obvious. If a major gene can be detected, the aim then becomes one of generating more closely linked markers, followed by flanking markers (i.e. markers on either side of the disease locus), and then applying the techniques of chromosome walking and jumping (Davies and Read 1988) to isolate the gene itself. From there it is theoretically possible and in most cases technically feasible to study the gene products and their effects so that ultimately a rational therapeutic approach to schizophrenia can be devised on the basis of sound knowledge of the biochemistry of the condition. This application of "reverse genetics", moving from linked markers to the gene itself, to gene products is already being applied in Mendelian diseases. Although there are many more technical problems with irregular phenotypes such as schizophrenia, including difficulties with the precision of original linkage study estimates, the strategies an nevertheless still be applied here.

So far, as we have seen, the results of linkage studies in schizophrenia are confusing, contradictory and for some quite disappointing. However, in our view it is not too surprising that clear-cut answers are still lacking, and the difficulties need not be regarded as overwhelming. So far, most attempts at family linkage studies have simply taken "at random" whatever polymorphisms were available in the attempt to detect linkage. Indeed, this was virtually all that could be done in the days when only classical markers were available. Nevertheless, the random search strategy has often paid off with other disorders, and the example of most relevance to psychiatry is Huntington's disease where Gusella et al. (1983) chanced upon linkage with a marker on chromosome 4 having studied eight random DNA polymorphisms.

The prior odds against a marker selected at random being linked to a disease locus are high, and a variety of methods can be used to shorten the odds. The first and most obvious is to follow up on a previous positive result even if this presents only a hint of linkage. The second is to have some prior knowledge of the biochemistry of the disorder and to focus on loci which may represent "candidate genes". As we have discussed earlier, in studies of schizophrenia this usually means genes encoding proteins involved in neurotransmission, particularly dopaminergic systems. The third approach is to gain a clue from chromosomal anomalies, and this has been the rationale for studies of chromosome 5 and, in part, for studies of the proposed pseudoautosomal locus on the sex chromosomes (Crow 1989). However, so far all of these means of focusing on favoured loci have yielded no definite advances and it seems likely that a painstaking and systematic search throughout the entire human genome will need to be undertaken. Effectively, this means screening a large sample of multiplex families with several hundred markers spread throughout the genome at intervals of roughly 5–10 cM. (The whole human genome is roughly 3,000 cM long.) The feasibility of undertaking such a project is increasing all the time, as the number of markers increases and especially as techniques are developed to identify highly polymorphic, multiallelic marker systems (Litt and Luty 1989; Weber and May 1989; Nakamura et al. 1988; Economou et al. 1990). However, it will still be beyond the resources of a single clinical centre or a single laboratory, and so multi-centre collaboration becomes essential. At the time of writing, a European collaborative programme is being established under the auspices of the European Science Foundation and a similar collaborative programme is commencing in the United States supported by the National Institute of Mental Health. Both programme are destined to take about 5 years to complete and, meanwhile, there are also a number of other active groups involved in similar efforts. Most workers are now conserving material from their families in the form of lymphoblastoid cell lines, which effectively provide a infinitely renewable source of DNA. There have also been attempts to standardise the collection and recording of clinical data and the methods of pedigree data storage and linkage analysis. The end result of all this activity must surely be that if a major gene (or major genes) involved in the transmission of schizophrenia exists, detection and localisation will take place by mid-way through this decade.

References

American Psychiatric Association (1980) Diagnostic and Statistical Manual of Mental Disorders, 3rd edn (DSM-III). American Psychiatric Assocation, Washington, DC

Basset AS, McGillivray BC, Jones BD, Pantzar JT (1988) Partial trisomy chromosome 5 cosegregating with schizophrenia. Lancet I:799–801

Bleuler E (1911) Dementia praecox or the group of schizophrenias. Translated from the German by J Zinkin (1950). International Universities Press, New York

Bleuler M (1978) The schizophrenic disorders. Translated from the German by S Clemens. Yale University Press, New Haven

Bloom FE (1990) Strategies for understanding the role of gene defects in the pathogenesis of mental disorders. In: Bulyshenkov V, Christen Y, Prilipko L (eds) Genetic approaches in the prevention of mental disorders. Springer, Paris Berlin Heidelberg New York

Clerget-Darpoux F, Bonaiti-Pellie C, Hochez J (1986) Effects of misspecifying genetic parameters in lod score analysis. Biometrics 42:393–399

Crow TJ (1989) A pseudo-autosomal locus for the cerebral dominance gene. Lancet II:339–340

Davies KA, Read AP (1988) In: Rickwood D, Male D (eds) Molecular basis of inherited disease. IRL Press, Oxford

Donis-Keller H, Green P, Helms C, Cartinhour S, Weiffenbach B, Stephens K, Keith TP, Bowden DW, et al (1987) A genetic linkage map of the human genome. Cell 51:319–337

Economou EF, Bergen AW, Warren AC, Antonarakis SE (1990) The polydeoxyadenylate tract of Alu repetitive elements in polymorphic in the human genome. Proc Natl Acad Sci USA 87:2951–2954

Edwards JH (1960) The simulation of Mendelism. Acta Genet 10:63

Farmer AE, McGuffin P, Gottesman II (1987) Twin concordance for DSM-III schizophrenia: scrutinising the validity of the definition. Arch Gen Psychiatry 44:634–641

Gottesman II, Shields J (1982) Schizophrenia: the epigenetic puzzle. Cambridge University Press, Cambridge

Gusella JF, Wexler NS, Conneally PM, Naylor SL, Anderson MA, et al (1983) A polymorphic marker genetically linked to Huntington's disease. Nature 306:234–238

Harrison PJ, Pearson RCA (1989) Gene expression and mental disease. Psychol Med 19:813–820

Kendler KS, Gruenberg AM, Strauss JS (1981) An independent analysis of the Copenhagen sample of the Danish adoption study of schizophrenia: II. The relationship between schizotypal personality disorder and schizophrenia. Arch Gen Psychiatry 38:982–984

Kennedy JL, Guiffra LA, Moises HW, Cavalli-Sforza LL, Pakstis AJ, Kidd JR, Castiglione CM, Sjogren B, Wetterberg L, Kidd KK (1988) Evidence against linkage of schizophrenia to markers on chromosome 5 in a northern Swedish pedigree. Nature 336:167–170

Lathrop GM, Lalouel JM (1984) Easy calculations of lod scores and genetic risks on small computers. Am J Hum Genet 36:460–465

Litt M, Luty JA (1989) A hypervariable microsatellite revealed by in vitro amplification of a dinucleotide repeat within the cardiac muscle actin gene. Am J Hum Genet 44:397–401

McGuffin P, Farmer AE, Gottesman II (1984) Twin concordance for operationally defined schizophrenia: confirmation of familiality and heritability. Arch Gen Psychiatry 41:541–545

McGuffin P, Huckle P (1990) Simulation of Mendelism revisited. the recessive gene for attending medical school. Am J Hum Genet 46:994–999

McGuffin P, Sargeant M, Hett G, Tidmarsh S, Whatley S, Marchbanks RM (1990) Exclusion of a schizophrenia susceptibility gene from the chromosome 5q11–q13 region. New data and reanalysis of previous reports. Am J Hum Genet 47:524–535

Morton NE (1955) Sequential tests for the detection of linkage. Am J Hum Genet 7:277–318

Nakamura Y, Carlson M, Krapcho K, Kanamori M, White R (1988) New approaches for isolation of VNTR markers. Am J Hum Genet 43:854–859

Sherrington R, Brynjolfsson J, Petursson H, Potter M, Dadleston K, Barraclough B, Wasmuth J, et al (1988) Localisation of a 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. Psychol Med 15:455–462

Weber JL, May PE (1989) Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. Am J Hum Gen 44:388–396