Molecular Psychiatry (2000) 5, 22–31

www.nature.com/mp

MILLENNIUM ARTICLE

Psychiatric genetics: back to the future

MJ Owen[1,2], AG Cardno[1,2] and MC O'Donovan[1]

Divisions of [1]Psychological Medicine; and [2]Medical Genetics, University of Wales College of Medicine, Heath Park, Cardiff, CF14 4XN, UK

For the last decade or more geneticists have been predicting that advances in molecular genetics are going to revolutionise our understanding of psychiatric disorders and human behavior. However, with a few exceptions, these expectations have yet to be fulfilled. As the century draws to a close and we contemplate the prospect of the complete sequence of the human genome it seems timely to consider the state of the field and to consider carefully how it might advance, the problems to be faced and the resources required. *Molecular Psychiatry* (2000) **5,** 22–31.

Keywords: molecular; epidemiology; linkage; linkage disequilibrium; phenotype; genomics

Much of the optimism surrounding the application of molecular genetics to human disease has resulted from successes in applying these techniques to monogenic disorders. However, common illnesses pose much greater challenges for geneticists because, in the majority of cases, they result from the combined action of a number of different genes (each of which may result in only a modest increase or decrease in liability) as well as environmental influences; a witches' brew termed polygenic, multifactorial causation. Further complexities include the possibility of non-additive genetic effects, including gene–gene interactions (epistasis), and also potential gene–environment interactions. In psychiatry, genetic complexity is compounded by the nosological complexity of the phenotypes, which are best viewed as syndromes rather than diseases. It is rarely possible to validate psychiatric diagnoses on the basis of physical examination or laboratory tests or even to confirm them postmortem, and for many disorders we have little idea of pathogenic mechanisms.

Current state of progress

Genetic epidemiology

In spite of these difficulties, genetic epidemiological studies based upon operational research diagnoses have shown that genes play a role in many of the syndromes defined by psychiatric nosology (Table 1). Increasingly, heritability estimates are based on a process of biometrical model fitting[1] which allows formal comparison of hypotheses concerning whether, and to what extent, genetic and environmental factors contribute to liability to a disorder. As well as showing that both genes and environment contribute to most specific diagnostic categories, genetic epidemiologists have begun to explore the extent to which there is overlap in genetic and environmental risk factors between disorders; whether genes influence clinical heterogeneity within disorders and the relationship between genetic susceptibility to disorders and putative environmental risk factors.

Table 1 Heritability estimates for selected psychiatric disorders

Disorder	*Heritability estimate (%)*
Schizophrenia[54]	80
Bipolar disorder[55]	80
Major depression[56]	40
Generalized anxiety disorder[56]	30
Panic disorder[56]	40
Phobia[56]	35
Alcohol problem or dependence[56]	60

Heritability is the proportion of liability to a disorder accounted for by genetic effects. The heritability estimates are based on twin studies which employed DSM-III-R research diagnoses. NB: the heritabilities should be regarded only as approximations.

Patterns of familial co-aggregation have been important in helping us to extend diagnostic boundaries. Examples include the schizophrenic and autistic 'spectrum' disorders that are mild, often sub-clinical phenotypes associated with genetic liability to schizophrenia and autism respectively.[2] There is also evidence of a considerable overlap between the genetic contributions to depression and anxiety.[3] Such findings have important implications for the classification of psychiatric disorders, and for establishing the most useful phenotypes for molecular genetic studies.

Correspondence: Professor M Owen, Neuropsychiatric Genetics Unit, University of Wales College of Medicine, Heath Park, Cardiff CF14 4XN, UK
Received and accepted 13 October 1999

Within a particular diagnostic category there is often considerable clinical heterogeneity which may result in large differences in the optimal treatment and likely prognosis for patients. It is often assumed that genetic influences on clinical heterogeneity are primarily related to the level of genetic liability to the disorder. This issue has been investigated in a range of studies of age at onset in schizophrenia. There is consistent evidence for a substantial genetic influence on age at onset in schizophrenia.[4] However, it is likely that susceptibility genes for schizophrenia have only a small effect at best on age at onset,[5] and that most of the effect is probably due to modifying genes. A similar pattern is emerging for the major groups of schizophrenia symptoms.[6]

The status of putative environmental risk factors for psychiatric disorders has become more complex in the light of genetic epidemiological studies. For example, episodes of depressive illness are often preceded by a cluster of adverse life events, which are traditionally regarded as environmental risk factors. However, there is increasing evidence that the experience or perception of life events themselves has a genetic basis.[7,8] Furthermore, people who suffer from depressive illnesses may not experience more life events than their unaffected relatives.[9] Therefore, the association between life events and depression, may be due at least partly to the co-inheritance of susceptibility to both.[10] Another example comes from the field of alcohol dependence, where age at first drink is traditionally thought of as an environmental risk factor. However, there is evidence from data on twins that the relationship is due to common familial factors, and that early drinking does not have a causal effect on alcohol dependence.[11] The implication is that interventions aimed at delaying age of first drinking are unlikely to be successful in reducing alcohol dependence later in life.

It is often suggested that gene–environment interactions are likely to be important in many psychiatric disorders. However, it has been difficult to identify clear examples because, for most conditions, we have not yet confidently identified specific risk factors. There is some evidence, albeit unconfirmed, from adoption studies that individuals who are at relatively high genetic risk of disorders such as schizophrenia[12] and adolescent conduct disorder[13] may be affected to a greater extent by dysfunction and stress within the family environment than individuals at lower genetic risk for these disorders. There is also preliminary evidence that possession of the apolipoprotein E (APOE) 4 allele is a risk factor for the presence and severity of chronic traumatic brain injury in boxers.[14] It seems highly likely that genetic epidemiological studies will play an increasingly important role in disentangling the complexities of disorders with multifactorial aetiologies as more specific risk factors are identified.

Molecular genetics

The first wave of molecular genetic studies in psychiatry focused on large multiply affected families and was based upon the hope that such pedigrees, or at least a proportion of them, result from the segregation of genes of major effect. This approach has been highly successful in identifying single gene forms of some common disorders such as breast cancer and non-insulin-dependent diabetes. However, in psychiatry its success has been restricted to the identification of three genes (APP, PS1 and PS2) in which mutations can cause familial early-onset forms of Alzheimer's disease (AD), and a handful of genes in which mutations can cause other rare, early onset dementias.[15] While mutations in APP, PS1 and PS2 do not appear to play a major role in the more common late-onset form of AD their identification has given important clues as to the pathogenesis of the disorder and has been profoundly influential in guiding neurobiological studies.[16]

Difficulties in replicating findings Initially it seemed that this approach might be successful in other psychiatric disorders such as schizophrenia and bipolar disorder where early positive findings were highly publicised.[2] Unfortunately these could not replicated. The reasons for this have become clear as data from a number of systematic genome searches have accumulated: highly penetrant mutations causing these disorders are at best extremely rare and quite possibly non-existent.[17] The 'false positive' findings were due largely to a combination of multiple testing and the use of statistical methodology derived from work on single gene disorders, which inappropriately assumed that a family with multiply-affected members was segregating a single dominant mutation. Some modestly significant evidence for linkage to several chromosomal regions has been found in bipolar disorder, schizophrenia and other disorders (Table 2). In each case either further replication is awaited or there are negative as well as positive findings; although it should be stressed that true positive linkages are difficult to replicate in disorders that are predisposed to by the combined action of several genes of moderate effect.[18] However, at present, it is doubtful whether the statistical evidence for linkage is sufficiently strong in any case to warrant large-scale and expensive efforts aimed at cloning putative linked loci. One way of resolving this issue might be to study much larger samples, of say 600–800 nuclear families, which should be sufficient to detect susceptibility genes of moderate effect size. However, if psychiatric disorders reflect the operation of many genes of small effect, then even these large-scale studies will be unsuccessful.

Mapping genes of small effect Once genes of small effect are sought (ie those conferring a risk in siblings relative to the population of <1.5), the number of families that are required for linkage studies becomes prohibitively large.[19] Allelic association studies offer a powerful means of identifying such genes in realistically sized samples of unrelated cases by the study of polymorphisms within candidate genes. However, they are prone to false positives, sometimes due to population stratification, but probably more often to a combination of multiple testing and the low prior odds of

Table 2 Selected linkage findings for psychiatric disorders

Disorder	*Chromosomal regions showing evidence of linkage*	
	Suggestive	*Significant*
Schizophrenia	5q[57] 6p[59] 8p[59] 18p[60] 22q[61]	13q[58]
Bipolar disorder	6p[62] 16p[64] 18q[65] 21q[67] 22q[68] Xq[69]	4p[63] 12q[39] 18p[66]
Autism	6q[70] 7q[71]	
Alcohol dependence	1, 7[72]	
Late onset Alzheimer's disease	1, 9, 10, 19[73] 12[74]	

Suggestive, statistical evidence greater than would be expected to occur once by chance in a genome scan; significant, statistical evidence greater than would be expected to occur 0.05 times by chance in a genome scan (following the criteria of Lander and Kruglyak[75]) including adjustments for multiple testing.

Table 3 Selected allelic association findings for psychiatric disorders

Disorder	*Evidence for allelic association*	
	Suggestive	*Strong*
Late onset Alzheimer's disease	A2M[76] ACE[77]	APOE[22]
Schizophrenia	5HT2A[25] DRD3[78]	
Bipolar disorder/ depression/neuroticism	SERT[79–81]	
Attention-deficit hyperactivity disorder	DAT[82] DRD4[83]	

Suggestive, evidence for association in some studies but not consistently and/or further replications awaited; strong, consistent evidence for association. Gene symbols: A2M, alpha-2 macroglobulin; APOE, apolipoprotein E; ACE, angiotensin converting enzyme; 5HT2A, 5HT2A receptor; DRD3, dopamine D3 receptor; SERT, serotonin transporter; DAT, dopamine transporter; DRD4, dopamine D4 receptor.

association.[20,21] Family-based association methods can be used to avoid the effects of population stratification, but these are difficult to collect, sensitive to non-random mating, lack power in many situations[21] and offer less potential for future studies of interaction between genetic and environmental risk factors than traditional case-control designs.

The most successful application of the association approach in psychiatric genetics has been the identification of the ϵ4 allele of the gene encoding apolipoprotein E (APOE) as a risk factor for Alzheimer's disease.[22] This gene was originally studied as a 'positional candidate' when linkage studies suggested that a susceptibility locus for late-onset AD might be found in the region of chromosome 19 to which APOE maps and APOE was found to be associated with senile plaques in AD brains. Possession of one copy of ϵ4 confers a moderate increase in risk of AD (odds ratio [OR] 2.2–4.4) and two copies confer a larger increase in risk (OR 5.1–17.9), although ϵ4 is neither necessary nor sufficient to cause the disease.[23] However, in other psychiatric disorders our relative ignorance of pathophysiology has meant that compelling candidates are hard to find. There have been some positive findings (Table 3), but once again conflicting results have been reported. It has usually been assumed that these reflect the tendency of association studies to produce false positives. However, we should not discount the possibility of false negatives since, where small effects (ORs approximately 1.2–1.5) are concerned, larger samples than those used in many studies may be required to have adequate power to achieve replication. For example, an association between schizophrenia and allele C of the T102C polymorphism in the 5HT2A-receptor gene was recently observed in a large collaborative study.[24] A number of small 'failures to replicate' have been published but a meta-analysis of all published data, which included over 3000 subjects, supports the original finding with an OR of 1.2 and suggests that publication bias is unlikely.[25] Thus, it is with some urgency that association studies which aim to test putative associations should be based upon large enough samples (usually several hundred to 1000 cases with as many, if not more, controls) to provide meaningful tests of association hypotheses. On the other hand, putative associations should be assumed to be false until replication, even if the *P* value is low.

The requirement for large sample sizes has practical implications for sampling strategies and study design. While in some cases, particularly very common childhood disorders, family-based association samples might be almost as easy to collect as case control samples, in general, large sample sizes will be difficult to collect. The place for family-based studies is therefore probably as part of a first or even second replication tier where the lower power, both intrinsic and due to smaller sample-size (due to difficulty of collection), will be offset by higher prior odds, the ability to make specific predictions concerning the specific associated allele and the direction of effect, and the fact that few markers will reach this stage of the sequential process.

However, there are problems in interpreting discrepancies between association studies that cannot simply be addressed by study design. Thus, if there are differences in the contribution of a given allele in different ethnic groups (arising from different allele frequencies or different allele frequencies at interacting loci), aetiological heterogeneity can always provide a convenient but untested explanation. Further potential for heterogeneity occurs if the association with the marker is a result of tight linkage with the true susceptibility allele. Similarly, in the face of aetiological heterogeneity in diseases for which there are no biologically meaningful diagnostic tests, it is possible that conflicting studies contain different balances of subtypes, known or unknown, of the disease which might again dramatically alter the power for excluding association. At present, there are no simple ways to take such possibilities into account other than to recognize their potential existence. Finally, samples are often ascertained in widely differing and unsystematic ways. The associations reported might therefore be with confounding variables such as comorbidity or with modifying factors such as those influencing disease severity or age at onset. The design of studies using only incident cases may be required before explanations of this sort can be excluded.

Some may question whether any priority should be given to finding genes of small effect. However, if a risk factor is common in the population then it can be associated with a high attributable risk even though the relative risk may be low. This means that a therapy targeting the immediate effects of the risk factor could have quite a large impact on the disorder at a population level. For example, in schizophrenia the attributable fraction associated with possession of allele C of the T102C polymorphism in 5HT2A is 0.35.[25] Furthermore even associations of small effect provide evidence that a particular pathway is involved in a disease process which is extremely important because historically, it has been difficult to distinguish between biological measurements that are a consequence of psychiatric disorders rather than a cause (cf APP and Alzheimer's disease). Independent findings from genetics can therefore provide a firm foundation for further research efforts into these pathways and their modulators in the search for other aetiological factors, genetic and environmental.

Other approaches to gene mapping may also prove valuable. For example, association studies in general populations are limited by the fact that the marker and disease loci must be very close together (<1 cM) for linkage disequilibrium (LD) to be detected. In contrast, LD may be detectable over larger distance in genetically isolated founder populations, and gene mapping might therefore be successfully performed using techniques such as shared-segment analysis,[26] which is based on extended haplotypes. However, the extent to which these and similar methods will be applicable is not yet clear, and recent simulation studies have suggested that the extent of LD is only greater in isolated populations when the founding bottleneck is very narrow or the frequency of the variant is low (<5%).[27]

The future

Genome-wide association studies

Recently there has been increasing interest in the possibility of systematic, genome-wide association studies.[19,28,29] Optimism has been fuelled by the fact that the most abundant form of genetic variation, the single nucleotide polymorphism (SNP), is usually biallelic and potentially amenable to binary, high-throughput genotyping assays, the most promising of which are at present micro-arrays (so-called DNA chips). Moreover as the amount of sequence data accumulates, it has become possible to contemplate the construction and application of very dense maps of hundreds of thousands of SNPs.[30]

Essentially two types of genome-wide association study have been proposed; direct and indirect (see Figure 1). In the former, association is sought between the disease and a comprehensive catalogue of every variant that can alter the structure, function or expression of every single gene. In contrast indirect studies seek associations between markers and disease that are due to linkage disequilibrium between the markers and susceptibility variants. The hope is that if dense enough marker maps can be applied, the whole genome can be systematically screened for evidence of LD without the requirement of actually screening every single functional SNP in the genome. However several uncertainties remain.

First, genome-wide direct studies depend upon the identification of all functionally significant polymorphisms relevant to every possible susceptibility gene. While this is feasible in principle with regard to cSNPs that are common in the target population, an exhaustive catalogue is still economically impossible. Furthermore functional polymorphisms outside the coding sequence such as those in regulatory regions that alter gene expression or those that effect splicing, or RNA

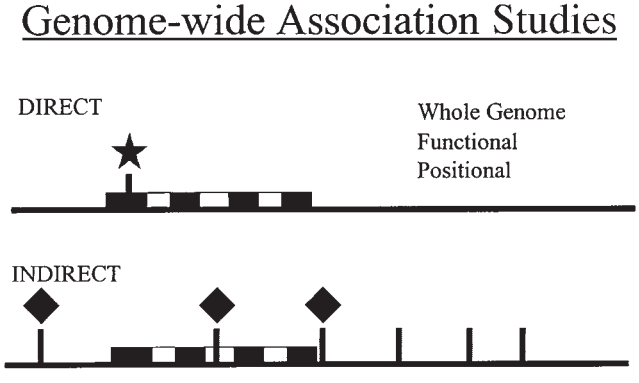

Figure 1 In direct association studies intragenic polymorphisms likely to alter structure, function or expression of the gene product are studied. In indirect studies a dense map of markers is studied in the hope of detecting linkage disequilibrium with a susceptibility locus.

editing might be very hard to find. Indeed, it is quite plausible that variants that influence gene expression in quite subtle ways are important in psychiatric disorders. In addition therefore to the screening programmes of candidate genes that are under way in many centres, systematic attempts to detect SNPs in regulatory regions of genes would therefore appear to be timely. Thus, we are currently undertaking an analysis of the promoter regions of approximately 2000 genes prioritised by function, expression or map position as of broad neuropsychiatric interest. To date, in the first 200 promoters we have studied, SNPs have been found in around 30%. We anticipate that this work will provide a useful resource for indirect association studies of the promoter region and adjacent coding sequence of positional and functional candidate genes well before systematic genome-wide association studies become possible. Furthermore, as we are undertaking functional analyses of all variants using a high throughput reporter gene assay we will identify a subset of the SNPs that are functional thus permitting rational selection of SNPs for direct association studies.

Another important unknown is the extent to which susceptibility to psychiatric disorders results from relatively common variants conferring small relative risks such as APOE ϵ4 or from rare variants with larger effect sizes implying that a certain amount of heterogeneity must exist. The safest assumption is that both sorts of variant play a role, but there are important differences in the optimum strategies to detect the two types. Thus, to achieve power of 0.8 for detecting relatively common alleles present at a frequency of 0.3 in the general population, only three unselected individuals need be screened whereas the corresponding figure for a rare variant with an allele frequency of 0.01 is 80 subjects. Common variants are therefore likely to be detected by generic SNP harvesting approaches whereas rarer ones will most readily be detected in disease cases and preferably those with a strong family history. However, both will require large samples to detect association. Rarer variants of larger effect will be more readily detected if the samples can be stratified on the basis of genetically valid variables such as family history, age at onset, severity and symptom profile, though the latter three might reflect modifying loci rather than being an index of loading at a specific susceptibility locus.

Next, it is not clear over what distances sufficient degrees of LD are typically maintained in the genome and the extent to which this varies from one region of the genome to another and from one population to another. This means that we do not know the density of markers required for systematic, indirect studies. The problem here is that empirical data are scarce. The available evidence suggests that LD between SNPs is sometimes maintained over distances of approximately 50–500 kb in outbred populations but that there is considerable variation from one genomic region to another.[31] A 60-kb map, which would result in no locus being more than 30 kb from the nearest marker, would require at least 50000 markers. However, this differs by an order of magnitude from recent estimates (3 kb and 500000 markers) made on theoretical grounds.[27] If this estimate is correct, indirect genome-wide association studies using SNPs might not offer much of a saving in genotyping as compared to direct studies since recent estimates of the total number of functional polymorphisms in the genome are of a similar magnitude.[32] However, theoretical estimates depend upon several assumptions and there is clearly a pressing need for more empirical data. Such studies should also compare the extent to which microsatellites can detect LD. Although no direct comparisons have been reported, the evidence suggests that microsatellites might detect LD over larger distances,[33–35] and therefore be useful in preliminary studies to identify LD prior to fine mapping by SNP analysis.

Finally, we are as yet still some way from having a sufficiently rapid, cheap and accurate method of genotyping SNPs. Microarray technology is advancing quickly, but several hurdles have to be overcome before fast, accurate genotyping of tens of thousands of SNPs is possible.[36] In the meantime methods based on DNA pooling may well be applicable. Pooling has been used successfully with microsatellites in model experiments[37–40] and to detect evidence of putative QTL for cognitive ability.[41] Pooled genotyping is even more readily applicable to SNPs where problems relating to stutter and differential amplification are irrelevant. Thus we have developed a pooling method based on single-nucleotide extension and DHPLC which should allow large samples to be screened rapidly and cheaply.[42]

Given the above considerations, it seems clear that the era of genome-wide association studies for psychiatric disorders, direct or indirect, is not yet at hand. Instead, studies in the next few years should probably focus mainly upon the direct approach utilising non-synonymous cSNPs in a wide range of functional and positional candidate genes. Preferably, complete functional systems should be dissected by a combination of the application of sensitive methods for mutation detection, followed by association studies in appropriately sized samples. We should also use our knowledge of functional pathways to make predictions about likely epistatic interactions between genes. However, given our ignorance of pathophysiology, the expectation should be that most reported associations will be false, and will only be resolved by replication in large well-characterised samples. At present although the indirect approach is not widely applicable at a genome-wide level, smaller scale studies focusing upon specific regions indicated by the results of linkage studies may allow loci to be mapped, and will have the added benefit of generating the sort of data concerning patterns of linkage disequilibrium in typical 'association samples' that will be required in order to determine whether genome-wide studies are likely to be feasible and what density of map will be required.

If the results are encouraging, that is, linkage disequilibrium exists across useful distances in the genome, then until genotyping technology has sufficient

capacity to permit mass genotyping at low-cost, then linkage disequilibrium analyses at the genome-wide level are most likely to be based upon DNA pooling technologies.

Refining phenotypes

One clear implication of the foregoing is that we should be collecting and studying larger samples. However, the effectiveness of molecular genetic studies will depend upon the genetic validity of the phenotypes. Perhaps if we were better at defining phenotypes we would be better at finding genes. The first thing to note at this point is that the commonly used diagnostic criteria define phenotypes with moderate to high heritability[2] (Table 1). In principle, therefore, it should be possible to identify the genes predisposing to these syndromes if sufficiently large samples are studied. However, perhaps genetic validity could be improved by focusing upon aspects of clinical variation such as age of onset or symptom profiles, or by identifying biological markers that predict degree of genetic risk or define more homogeneous subgroups.

Despite much work, it has not been possible to identify genetically distinct sub-types within the major diagnostic categories for most mental disorders. Instead, clinical variation is likely to reflect in part at least a combination of quantitative variation in genetic risk for the disorder and the effect of modifying genes which influence illness expression rather than the risk of illness *per se*. Examples of this phenomenon relating to age at onset and symptom pattern in schizophrenia have already been discussed.

The search for trait markers aims to move genetic studies beyond the clinical syndrome by identifying indices of genetic risk that can be measured in asymptomatic individuals and/or by identifying markers of pathophysiological processes that are closer to the primary effects of susceptibility genes than clinical symptoms. Work in this area is developing fast; for example, candidate trait markers for schizophrenia include schizotypal personality traits, measures of cognitive processing, brain evoked potentials and abnormalities in eye movements.[43] It is also hoped that advances in brain imaging will lead to the identification of genetically valid trait markers. However, it seems unlikely that these will provide a quick fix to the problem. First, we will need measures that can be practicably applied to a sufficient number of families or unrelated patients and second, we will need to ensure that the traits identified are highly heritable, which will itself require a return to classic genetic epidemiology and model fitting.

Efforts to improve the selection of phenotypes are also concerned with enhancing the traditional categorical approach to defining psychiatric disorders, by identifying genetically valid phenotypes that can be measured quantitatively. These can be used in quantitative trait locus (QTL) approaches to gene mapping. QTL linkage[44] is generally carried out in sib-pairs and is based on the principle that at a locus influencing a trait or at a linked marker locus, sibs with more similar phenotypic scores should share more alleles identical-by-descent (ibd). This approach has been used to map a locus for reading disability on chromosome 6,[45] and similar approaches have been applied, for example, to linkage studies of alcohol dependence, and symptom severity in schizophrenia. QTL association[46] is based on samples of unrelated individuals, and seeks evidence of differences in phenotypic values according to allelic or genotypic differences at a locus. This approach has been used, for example, to investigate a possible association between the dopamine transporter gene (DAT) and severity of attention-deficit hyperactivity disorder (ADHD).[47]

Animal models

A further approach that may facilitate gene mapping in the future is the use of animal models. These are particularly difficult to construct for psychiatric disorders, which are predominantly defined in terms of subjective experiences described by patients. However, it may be possible to identify behavioural correlates of some more objective aspects of psychiatric disorders, or of physiological trait markers. Perhaps the most promising example is currently rodent emotionality, for which evidence from genetic, neuropharmacological, neurophysiological and neuroanatomical studies suggests that it may be a reasonable homologue of human neuroticism.[48] Mapping studies of emotionality in mice suggest that the genetic architecture may be simpler than many expected, with only three loci accounting for virtually all the genetic variance in a mouse strain selected for high emotionality.[49] This is in agreement with QTL analysis in a number of complex traits in both animals and plants where a small number of loci account for the majority of genetic variance. Of course the final validation of rodent emotionality as a model of human anxiety will only come with the demonstration that variation in the DNA sequence of the same genes influences the phenotype in both human and model species. Other examples of psychiatric disorders for which plausible animal models currently exist include depression, substance abuse, and hyperactivity.

Disorders that predominantly involve higher cognitive function such as schizophrenia are likely to prove more difficult to model in animals. However, there are features of the human phenotype such as subtle abnormalities of cell migration, enlarged cerebral ventricles and information processing abnormalities including defects in pre-pulse inhibition that can be detected in animals. In fact, a possible approach to producing a mouse model for at least some of the schizophrenia phenotype is suggested by the finding that people with velo-cardio-facial syndrome due to deletions of chromosome 22q11 have very high rates of schizophrenia.[50] Currently attempts are underway to produce transgenic mice deleted for the VCFS syntenic region of mouse chromosome 16.[51] These animals should be investigated closely for neuroanatomical and behavioural phenotypes of possible relevance to schizophrenia.

Animal models for human psychiatric disorder might also emerge from large mutagenesis programmes such as that being carried out at the MRC Mammalian Genetics Unit, Harwell, UK (http://www.mgc.har.mrc.ac.uk). This is aimed at generating large numbers of new mouse phenotypes, many of which will carry disorders that model human genetic disease. The challenge here will be to develop rapid phenotypic screening protocols allowing models of possible relevance to psychiatry to be selected for more detailed behavioural analysis.

Identifying modifying genes

Genes affecting the clinical features of a disease or that influence treatment response in many instances may simply reflect pleiotropic effects of aetiological risk factors. However there may also be specific genes that influence clinical features and treatment response independent of those affecting liability. Thus another potentially fruitful line of inquiry might be to design studies aimed at seeking modifying rather than causative genes since these may themselves allow novel drug targets to be identified. An example here is the association between the low activity allele of a polymorphism in COMT and rapid cycling in bipolar affective disorder.[39] An obvious potential example is the possibility that there are genes for age of onset in Alzheimer's disease. A treatment that delayed the mean age of onset of AD by 10 years would virtually eliminate it from the general population. It is therefore important that we design and implement appropriate pharmacogenetic studies. However, as for trait markers, a phase of clinical genetic epidemiology is in most cases indicated prior to embarking upon large-scale pharmacogenetic studies. It should also be borne in mind that there is no *a priori* reason why response to behavioural and psychological treatments should be less influenced by genetic factors than pharmacological treatments.

Implications of identifying genes

Functional studies

What would be the implications of identifying genetic risk factors for the major psychiatric disorders? The most important, and the most obvious, is that it will inspire a new wave of neurobiological studies from which new and more effective therapies will hopefully emerge. However, while the unequivocal identification of associated genetic variants will represent a great advance, many years of work will still be required before this is likely to translate to the bedside. An early problem will be to determine exactly which genetic variation amongst several in LD in a given gene is actually responsible for the functional variation. This is clearly illustrated by recent work on the gene encoding the Angiotensin Converting Enzyme.[84] Even where a specific variant within a gene can be identified as the one of functional importance, functional analysis, in terms of effect at the level of the organism is likely to be particularly difficult for behavioural phenotypes in the absence of animal models. An extra level of complexity is that we will need to be able to produce model systems, both *in vivo* and *in vitro*, that allow gene–gene and gene–environment interaction to be studied.

Genetic nosology

While the development of new therapies will take time, it is likely that the identification of susceptibility genes will have an earlier impact on psychiatric nosology. By correlating genetic risk factors with clinical symptoms and syndromes it should be possible to study heterogeneity and co-morbidity in order to improve the diagnosis and classification of mental disorders. The prospects will also be enhanced for identifying clinically useful biological markers of psychiatric disorders as an aid to diagnosis, thus moving beyond the current situation of making diagnoses based entirely on clinical signs and symptoms. Improvements in the validity of psychiatric diagnosis will clearly facilitate all avenues of research into these disorders. However, for our purposes, perhaps we should limit ourselves to speculating that improvements in diagnosis and classification will enhance our ability to detect susceptibility factors genetic and otherwise and thus a positive feedback between nosology, epidemiology and molecular genetics can be envisaged.

Molecular epidemiology

The identification of genetic risk factors should allow us to investigate the ways in which genes and environment interact. Studies of this kind will require large epidemiologically based samples together with the collection of relevant environmental data. This work could start now with DNA being banked for future use, although in many instances the identification of plausible environmental measures might require clues from the nature of the genetic risk factors yet to be identified. A major theme in relation to this will be the bringing together of methodologies from genetics and epidemiology which have traditionally adopted somewhat differing analytical approaches.[52] Treating susceptibility alleles as risk factors in an epidemiological context will allow estimates of effect sizes within a population to be made. Accounting for specific genetic effects will also facilitate the search for independent environmental factors, and the investigation of potential gene–environment interactions. Scientific validity is likely to be enhanced by ensuring as far as possible that control samples are drawn from the same base population as controls. In addition, the use of incident cases, rather than cases who happen to have the disorder at the time of a study, should guard against the risk of identifying loci related to chronicity of illness rather than susceptibility. Phenotypic assessment is likely to benefit from a prospective element to studies, to counteract the tendency of patients to forget historical details, and the difficulty of making observed ratings retrospectively from case records. However, the price of improved scientific rigor is likely to be considerably more expensive studies, due to the longer period and larger number of investigators that will be required to ascer-

tain the detailed data on the thousands of subjects which will probably be required.

Genetic testing

A further implication concerns genetic testing. This is a complex area that raises a number of ethical issues which have been discussed elsewhere.[53] However, the potential for predictive testing has probably been overstated, given that susceptibility to most psychiatric disorders almost certainly depends upon the combined effects of predisposing and protective alleles at a number of loci. Consequently, in many instances the predictive value of tests for such genes will be low. This applies, for example, to APOE testing for late onset Alzheimer's disease, and has led to a recommendation not to perform such testing in asymptomatic individuals.[12] Indeed, even when all susceptibility genes for a given disorder have been identified, it will still not be possible to predict the development of disease with certainty until the relevant environmental risk factors have also been identified and the nature of the various interactions understood. Such interactions may be as complex as chaotic systems like the weather, which is notoriously unpredictable over even the relatively short term (5 days plus). However, on a more optimistic note, genetic testing could have other roles that are likely to be of more value to patients and clinicians, for example, in helping to optimise treatment choices by testing genes that are found to influence treatment responses in psychiatric disorders, leading to a greater individualisation of treatment.

Conclusion

Despite the undoubted complexities of psychiatric genetic research, significant advances have been made, most notably in the field of AD. The prospects for determining the genetic basis of other common psychiatric disorders are good, but important difficulties still have to be overcome. Technical and methodological advances are certainly required. However, ultimately the identification of genetic risk factors for mental disorders and the translation of this into improved patient care will depend as much upon the traditional medical disciplines of clinical description and epidemiology and our ability to combine these with genetic, psychological and sociological approaches in the integrative science of psychiatry.

Acknowledgements

We are supported by a Cooperative Group Grant from the Medical Research Council.

References

1 Neale MC, Cardon LR. *Methodology for Genetic Studies of Twins and Families*. NATO ASI Series: Kluwer Academic Publishers: Dordrecht, 1992.
2 McGuffin P, Owen MJ, O'Donovan MC, Thapar A, Gottesman I. *Seminars in Psychiatric Genetics*. Gaskell: London 1994.
3 Kendler KS, Neale MC, Kessler RC, Heath AC, Eaves LJ. Major depression and generalized anxiety disorder: same genes, (partly) different environments? *Arch Gen Psychiatry* 1992; **49**: 716–722.
4 Kendler KS, Tsuang MT, Hays P. Age at onset in schizophrenia: a familial perspective. *Arch Gen Psychiatry* 1987; **44**: 881–890.
5 Neale MC, Eaves LJ, Hewitt JK, MacLean CJ, Meyer JM, Kendler KS. Analysing the relationship between age at onset and risk to relatives. *Am J Hum Genet* 1989; **45**: 226–239.
6 Kendler KS, Karkowski-Shuman L, O'Neill FA, Straub RE, MacLean CJ, Walsh D. Resemblance of psychotic symptoms and syndromes in affected sibling pairs from the Irish study of high-density schizophrenia families: evidence for possible etiologic heterogeneity. *Am J Psychiatry* 1997; **154**: 191–198.
7 Kendler KS, Neale M, Kessler R *et al*. A twin study of recent life events and difficulties. *Arch Gen Psychiatry* 1993; **50**: 789–796.
8 Thapar A, McGuffin P. Genetic influences on life events in childhood. *Psychol Med* 1996; **26**: 813–820.
9 McGuffin P, Katz R, Bebbington P. The Camberwell Collaborative Depression Study III. Depression and adversity in the relatives of depressed probands. *Br J Psychiatry* 1988; **152**: 775–782.
10 Owen MJ, McGuffin P. Genetics and psychiatry. *Br J Psychiatry* 1997; **171**: 201–202.
11 Precott CA, Kendler KS. Age at first drink and risk for alcoholism: a noncausal association. *Alc Clin Exp Res* 1999; **23**: 101–107.
12 Tienari P, Lyman CW, Moring J, Lahti I, Naarala M, Sorri A *et al*. The Finnish adoptive family study of schizophrenia: implications for family research. *Br J Psychiatry* 1994; **164** (suppl 23): 20–26.
13 Cadoret RJ, Yates WR, Troughton E, Woodworth G, Stewart E. Gene-environment interaction in the genesis of aggressivity and conduct disorder. *Arch Gen Psychiatry* 1995; **52**: 916–924.
14 Lendon C, Ashall F, Goate AM. Exploring the etiology of Alzheimer disease using molecular genetics. *JAMA* 1997; **277**: 825–831.
15 Hardy J, Gwinn-Hardy K. Genetic classification of primary neurodegenerative disease. *Science* 1998; **282**: 1075–1079.
16 Hardy J, Israël A. Alzheimer's disease: in search of γ-secretase. *Nature* 1999; **398**: 466–467.
17 McGuffin P, Owen MJ. Molecular genetic studies of schizophrenia. *Cold Spring Harb Symp Quant Biol* 1996; **61**: 815–822.
18 Suarez BK, Hampe CL, Van Eerdewegh P. Problems of replicating linkage claims in psychiatry. In: Gershon ES, Cloninger CR (eds). *Genetic Approaches to Mental Disorders*. American Psychiatric Press: Washington, DC, 1994, pp 23–46.
19 Risch N, Merikangas K. The future of genetic studies of complex human diseases. *Science* 1996; **273**: 1516–1517.
20 Owen MJ, Holmans P, McGuffin P. Association studies in psychiatric genetics. *Mol Psychiatry* 1997; **2**: 270–273.
21 Risch N, Teng J. The relative power of family-based and case-control designs for linkage disequilibrium studies of complex human diseases: I. DNA pooling. *Genome Res* 1998; **8**: 1273–1288.
22 Corder EH, Saunders AM, Strittmatter WJ, Schmechel DE, Gaskell PC, Small GW *et al*. Gene dose of apolipoprotein E type-4 allele and the risk of Alzheimer's disease in late-onset families. *Science* 1993; **261**: 921–923.
23 National Institute on Aging/Alzheimer's Association Working Group. Apolipoprotein E genotyping in Alzheimer's disease. *Lancet* 1996; **347**: 1091–1095.
24 Williams J, Spurlock G, McGuffin P *et al*. Association between schizophrenia and the T102C polymorphism of 5-hydroxytryptamine type 2a receptor gene. *Lancet* 1996; **347**: 1294–1296.
25 Williams J, McGuffin P, Nothen M, Owen MJ and the EMASS Collaborative Group. Meta-analysis of association between the 5HT2a receptor T102C polymorphism and schizophrenia. *Lancet* 1997; **349**: 1221.
26 Houwen RHJ, Baharloo S, Blankenship K *et al*. Genome screening by searching for shared segments: mapping a gene for benign recurrent intrahepatic cholestasis. *Nature Genet* 1994; **8**: 380–386.
27 Kruglyak L. Prospects for whole-genome linkage disequilibrium mapping of common disease genes. *Nature Genet* 1999; **22**: 139–144.
28 Owen MJ. Will schizophrenia become a graveyard for molecular geneticists? *Psychol Med* 1992; **22**: 289–293.
29 Collins FS, Guyer MS, Chakravarti A. Variation on a theme: cataloging human DNA sequence variation. *Science* 1997; **278**: 1580–1581.
30 Wang DG, Fan JB, Siao CJ *et al*. Large-scale identification, mapping,

and genotyping of single-nucleotide polymorphisms in the human genome. *Science* 1998; **280**: 1077–1082.
31 Jorde LB, Watkins WS, Carlson M, Groden J, Albertson H, Thliveris A *et al*. Linkage disequilibrium predicts physical distance in the adenomatous polyposis coli region. *Am J Hum Genet* 1994; **54**: 884–898.
32 Cargill M, Altshuler D, Ireland J, Sklar P, Ardlie K, Patil N *et al*. Characterization of single-nucleotide polymorphisms in coding regions of human genes. *Nature Genet* 1999 **22**: 231–238.
33 Kendler KS, MacLean CJ, Ma YL, O'Neill FA, Walsh D, Straub RE. Market-to-marker linkage disequilibrium on chromosomes 5q, 6p and 8p in the Irish high-density schizophrenia pedigrees. *Am J Med Genet* 1999; **88**: 29–33.
34 Petruhkin K, Fischer SG, Pirastu M, Tanzi RF, Chrnov I, Devoto M *et al*. Mapping cloning and genetic characterization of the region containing in Wilson disease gene. *Nature Genet* 1993; **5**: 338.
35 Peterson AC, Di Rienzo A, Lehesjoki A-E, de la Chapelle A, Slatkin M, Freimer NB. The distribution of linkage disequilibrium over anonymous genome regions. *Hum Mol Genet* 1995; **4**: 887–894.
36 Lander ES. Array of hope. *Nature Genet* 1999; **21** (suppl): 3–4.
37 Barcellos LF, Klitz W, Field L, Tobias R, Bowcock AM, Wilson R *et al*. Association of mapping of disease loci by use of a pooled DNA genomic screen. *Am J Hum Genet* 1997; **61**: 734–747.
38 Biomed European Bipolar Collaborative Group. No association between bipolar disorder and alleles at a functional polymorphism in the COMT gene. *Br J Psychiatry* 1997; **170**: 526–528.
39 Barden N, Morissette J, Shink E, Rochette D, Gagne B, Bordeleau L *et al*. Confirmation of bipolar affective disorder susceptibility locus on chromosome 12 in the region of the Darier disease gene. *Am J Med Genet* 1998; **81**: 475.
40 Kirov G, Murphy KC, Arranz MJ, Jones I, McCandless F, Kunugi H *et al*. Low activity allele of catecho-o-methyl transferase gene associated with rapid cycling bipolar affective disorder. *Mol Psychiatry* 1998; **3**: 342–345.
41 Fisher PJ, Turic D, Williams NM, McGuffin P, Asherson P, Ball D *et al*. DNA pooling identifies QTLs for general cognitive ability in children on chromosome 4. *Hum Mol Genet* 1999; **8**: 915–922.
42 Hoogendoorn B, Owen MJ, Oefner PJ, Williams N, Austin J, O'Donovan MC. Genotyping single nucleotide polymorphisms by primer extension and high performance liquid chromatography. *Hum Genet* 1999; **104**: 89–93.
43 DeLisi LE. A critical overview of recent investigations into the genetics of schizophrenia. *Curr Opin Psychiatry* 1999; **12**: 29–39.
44 Kruglyak L, Lander ES. Complete multipoint sib-pair analysis of qualitative and quantitative traits. *Am J Hum Genet* 1995; **57**: 439–454.
45 Cardon LR, Smith SD, Fulker DW, Kimberling WJ, Pennington BF, DeFries JC. Quantitative trait locus for reading disability on chromosome 6. *Science* 1994; **266**: 276–279.
46 Page GP, Amos CI. Comparison of linkage-disequilibrium methods for localization of genes influencing quantitative traits in humans. *Am J Hum Genet* 1999; **64**: 1194–1205.
47 Waldman ID, Rowe DC, Abramowitz A *et al*. Association and linkage of the dopamine transporter gene and attention-deficit hyperactivity disorder in children: heterogeneity owing to diagnostic subtype and severity. *Am J Hum Genet* 1998; **63**: 1767–1776.
48 Flint J. Freeze! *Nat Genet* 1997; **17**: 250–251.
49 Flint J, Corley R, DeFries JC *et al*. A simple genetic-basis for a complex psychological trait in laboratory mice. *Science* 1995; **269**: 1432–1435.
50 Murphy KC, Jones LA, Owen MJ. Elevated rates of schizophrenia in adults with velo-cardio-facial syndrome. *Arch Gen Psychiatry* 1999; **56**: 940–945.
51 Lindsay EA, Botta A, Jureuc V *et al*. Congenital heart disease in mice deficient for the DiGeorge Syndrome region. *Nature* 1999; **40**: 379–383.
52 Sham PC. Statistical methods in psychiatric genetics. *Stat Meth Med Res* 1998; **7**: 279–300.
53 Nuffield Council on Bioethics. *Mental Disorders and Genetics: the Ethical Context*. Nuffield Council on Bioethics: London, 1998.
54 Cardno AG, Marshall EJ, Coid B, Macdonald AM, Ribchester TR, Davies NJ *et al*. Heritability estimates for psychotic disorders. *Arch Gen Psychiatry* 1999; **56**: 162–168.
55 Kendler KS, Pedersen NL, Neale MC, Mathé AA. A pilot Swedish twin study of affective illness including hospital- and population-ascertained subsamples: results of model fitting. *Behav Genet* 1995; **3**: 217–232.
56 Kendler KS, Walters EE, Neale MC, Kessler RC, Heath AC, Eaves LJ. The structure of the genetic and environmental risk factors for six major psychiatric disorders in women: phobia, generalized anziety disorder, panic disorder, bulimia, major depression and alcoholism. *Arch Gen Psychiatry* 1995; **52**: 374–383.
57 Straub RE, MacLean CJ, O'Neill FA, Walsh D, Kendler KS. Support for a possible schizophrenia vulnerability locus in a region of 5q22–31 in Irish families. *Mol Psychiatry* 1997; **2**: 148–155.
58 Blouin JL, Dombroski BA, Nath SK, Lasseter VK, Wolyniec PS, Nestadt G *et al*. Schizophrenia susceptibility loci on chromosomes 13q32 and 8p21. *Nature Genet* 1998; **20**: 70–73.
59 Schizophrenia Linkage Collaborative Group for Chromosomes 3, 6, and 8. Additional support for schizophrenia linkage on chromosomes 6 and 8: a multicentre study. *Am J Med Genet* 1996; **67**: 580–594.
60 Wildenauer DB, Schwab SG, Hallmayer J, Lerer B, Albus M, Borrman M *et al*. Genome scan for autosomal genes conferring risk to schizophrenia in a German/Israeli sample. *Am J Med Genet* 1998; **81**: 454.
61 Schizophrenia Collaborative Linkage Group. A combined analysis of D22S278 marker alleles in affected sib-pairs: support for a susceptibility locus for schizophrenia at chromosome 22q12. *Am J Med Genet* 1996; **67**: 40–45.
62 Ginns EI, Ott J, Egeland JA, Allen CR, Fann CSJ, Pauls DL *et al*. A genome-wide search for chromosomal loci linked to bipolar affective disorder in the Old Order Amish. *Nature Genet* 1996; **12**: 431–435.
63 Blackwood DHR, He L, Morris SW, McLean A, Whitton C, Thomson M *et al*. A locus for bipolar affective disorder on chromosome 4p. *Nature Genet* 1996; **12**: 427–430.
64 Ewald H, Mors O, Flint T, Koed K, Eiberg H, Kruse TA. A possible locus for manic depressive illness on chromosome 16p13. *Psychiatr Genet* 1995; **5**: 71–81.
65 Stine OC, Xu J, Koskela R, McMahon FJ, Gschwend M, Friddle C *et al*. Evidence for linkage of bipolar disorder to chromosome 18 with a parent-of-origin effect. *Am J Hum Genet* 1995; **57**: 1384.
66 Berrettini WH, Ferraro TN, Goldin LR *et al*. Pericentromeric chromosome 18 DNA markers and manic-depressive illness: evidence for a susceptibility gene. *Proc Natl Acad Sci* 1994; **91**: 5918–5921.
67 Straub RE, Lehner T, Luo Y, Loth JE, Shao W, Sharpe I *et al*. A possible vulnerability locus for bipolar affective disorder on chromosome 21q22.3. *Nature Genet* 1994; **8**: 291–296.
68 Kelsoe JR, Loetscher E, Spence MA, Foguet M, Sadovnick AD, Remick RA *et al*. A genome survey of bipolar disorder indicates a susceptibility locus on chromosome 22. *Am J Med Genet* 1998; **81**: 461–462.
69 Pekkarinen P, Bredbacka PE, Terwilliger J *et al*. Evidence for a susceptibility locus for manic depressive disorder in Xq26. *Am J Hum Genet* 1994; **55**: 133.
70 Philippe A, Martinez M, GuilloudBataille M, Gillberg C, Rastam M, Sponheim E *et al*. Genome-wide scan for autism susceptibility genes. *Hum Mol Genet* 1999; **8**: 805–812.
71 International Molecular Genetic Study of Autism Consortium. A full genome screen for autism with evidence for linkage to a region on chromosome 7q. *Hum Mol Genet* 1998; **7**: 571–578.
72 Reich T, Edenberg HJ, Goate A, Williams JT, Rice JP, VanEerdewegh P *et al*. Genome-wide search for genes affecting the risk for alcohol dependence. *Am J Med Genet* 1998; **81**: 207–215.
73 Kehoe PG, Wavrant De Vrieze F, Crook R, Wu WS, Holman P, Fenton I *et al*. A full genome scan for late onset Alzheimer's disease. *Hum Mol Genet* 1999; **8**: 237–245.
74 Pericak-Vance MA, Bass ML, Yamaoka LH, Gaskell PC, Scott WK, Terwedow HA *et al*. Complete genomic screen in late-onset familial Alzheimer's disease. *Neurobiol Aging* 1998; **19**: S39–S42.
75 Lander E, Kruglyak L. Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. *Nature Genet* 1995; **11**: 241–247.
76 Blacker D, Wilcox MA, Laird NM, Rodes L, Horvath SM, Go RCP *et al*. Alpha-2 macroglobulin is genetically associated with Alzheimer disease. *Nature Genet* 1998; **19**: 357–360.
77 Kehoe PG, Carsten R, Mcllroy S, Williams H, Holmans P, Holmes

C *et al*. Variation in DCP1, encoding ACE, is associated with susceptibility to Alzheimer's disease. *Nature Genet* 1999; **21**: 71–72.
78 Williams J, Spurlock G, Holmans P, Mant R, Murphy K, Jones L *et al*. A meta-analysis and transmission disequilibrium study of association between the dopamine D3 receptor gene and schizophrenia. *Mol Psychiatry* 1998; **3**: 141–149.
79 Collier DA, Arranz MJ, Sham P, Battersby S, Vallada H, Gill P *et al*. The serotonin transporter is a potential susceptibility factor for bipolar affective disorder. *Neuroreport* 1996; **7**: 1675–1679.
80 Ogilvie AD, Battersby S, Bubb VJ, Fink G, Harmar AJ, Goodwin GM *et al*. Polymorphism in serotonin transporter gene associated with susceptibility to major depression. *Lancet* 1996; **347**: 731–733.
81 Lesch KP, Bengel D, Heils A, Sabol SZ, Greenberg BD, Petri S *et al*. Association of anxiety-related traits with a polymorphism in the serotonin transporter gene regulatory region. *Science* 1996; **274**: 1527–1531.
82 Cook EH, Stein MA, Krasowski MD, Cox NJ, Olkon DM, Kieffer JE *et al*. Association of attention-deficit disorder and the dopamine transporter gene. *Am J Hum Genet* 1995; **56**: 993–998.
83 LaHoste GJ, Swanson JM, Wigal SB, Glabe C, Wigal T, King N *et al*. Dopamine D4 receptor gene polymorphism is associated with attention-deficit hyperactivity disorder. *Mol Psychiatry* 1996; **1**: 121–124.
84 Rieder MJ, Taylor SL, Clark AG, Nickerson DA. Sequence variation in human Angiotensin Converting Enzyme. *Nature Genet* 1999; **22**: 59–62.