METHODS IN MOLECULAR MEDICINE™

Psychiatric Genetics

Methods and Reviews

Edited by

Marion Leboyer, MD, PhD
Frank Bellivier, MD, PhD

Humana Press

1

Psychiatric Genetics

Overview on Achievements, Problems, Perspectives

Wolfgang Maier

1. The Progress of Psychiatric Genetics

Psychiatric genetics is a relatively new term for an old research question: "Are behavioral and psychological conditions and deviations inherited?" The systematic empirical inquiries in this field started in the late nineteenth century with the work of F. Galton and his monograph *Talent and Character*, which was motivated by Darwin's theory and the concept of degeneration. During the twentieth century, the methodological standard of the field was improved by the development of epidemiological, biometrical, and clinical research tools. This was the precondition to perform valid family, twin, and adoption studies. These methods revealed that all psychiatric disorders aggregate in families, and that genes influence the manifestation of these disorders. It became clear that the degree of familiality and extent of genetic influence varies among diseases, with schizophrenia showing the strongest genetic background and disorders such as obsessive-compulsive and borderline personality disorder showing the weakest genetic background. Although there is some overlap, the familial patterns of diagnoses reveal a surpris-

From: *Methods in Molecular Medicine, vol. 77: Psychiatric Genetics: Methods and Reviews*
Edited by: M. Leboyer and F. Bellivier

ingly high specificity, which was considered an argument for the appropriateness of diagnostic definitions. Considering the limitations in the pathophysiological understanding of psychiatric disorders, "breeding true" of diagnosis in families became the hallmark indicator of clinical validity ***(1)***.

Segregation analyses of the specific mode of transmission were performed in many family samples over an extended period of time. One major goal was to find Mendelian patterns. It took decades to rule out the theory that the familial pattern of aggregation does not fit into the Mendelian mode of transmission. Environmental influences on the manifestation of all psychiatric disorders were also unequivocally demonstrated. Thus, like other common diseases, all psychiatric disorders revealed a complex genetic and a multifactorial etiology rather than a monogenic etiology.

Since about 1980, developments in molecular genetics made is possible to systematically map genes on the DNA strand by so-called linkage studies without any knowledge of the "true" pathophysiology and of the gene products (proteins) involved. This strategy required:

- systems of positional DNA markers placed densely on the whole genome—first restriction-fragment-length polymorphism (RFLP), then microsatellite and now single-nucleotide polymorphism (SNP) markers; and
- samples of genetically informative families, each with more than one affected case (e.g., extended pedigrees with multiple cases or pairs of affected siblings).

Linkage analysis identifies regions on the genome that hosts disease genes through the position of markers that segregate together with the disease in the families. This method is most conclusive when the disease is transmitted in a Mendelian fashion. Thus, monogenic diseases were the first target for this method. Thousands of disease genes for monogenic (Mendelian) diseases were successfully mapped and subsequently identified by stepwise application of this strategy during the last two decades. The detection of disease genes and etiologically relevant proteins using only

positional information (positional cloning) became the major tool in revealing the etiology of Mendelian diseases.

Simultaneously, the success of the positional cloning approach in monogenic diseases motivated hopes and optimism that the genetic basis of more complex diseases (with a genetic component but without Mendelian transmission) would be revealed, including the most common chronic diseases. Their etiology is not as fully understood as that of monogenic diseases, presumably because of phenotypic and genetic heterogeneity; this is particularly true for all psychiatric disorders. Therefore, the positional cloning strategy offers an especially promising method to reveal the unknown etiology of psychiatric disorders, because other strategies have failed to fully elucidate the etiology and pathophysiology.

The positional cloning strategy based on linkage analysis was first applied two decades ago to complex diseases, yet the early hopes for a new success story of the linkage strategy failed. Until now, the search for genes was disappointing for all complex diseases, particularly for psychiatric disorders. Frustration initiated a process of revising the most appropriate strategy. Arguments and proposals can be subdivided into two lines of reasoning:

1. What is the most appropriate analytic strategy to detect disease genes for complex diseases?
2. How can phenotypes be properly defined in order to detect disease genes? How should the etiological heterogeneity of common diseases be approached?

This book addresses these important questions with a series of articles. This outlines the current status of progress in psychiatric genetics and discusses perspectives on these questions on a more general level.

2. The Search for Genes: Current Status

2.1. Linkage Studies

Genome-wide linkage studies are the key to finding the genes that carry mutations causative for monogenic diseases. Is this strat-

egy as useful for complex diseases? Although a positive answer to this crucial question is not guaranteed—especially with regard to psychiatric disorders—genome-wide linkage studies in specific psychiatric disorders were also initially claimed to be the success strategy. A series of chromosomal regions with at least suggestive linkage to the disease emerged in the various genome-wide scans. These positive results are contrasted by an unexpected pattern of findings:

- The linkage signals were only modest, and a very broad interval on the genome was implicated independently of the structure of the family sample under study (extended family or affected sibs).
- Even the strongest linkage results were not consistently replicable. In general, some of the initial reports with at least suggestive evidence for linkage to schizophrenia, manic-depressive illness, or alcoholism were replicable with similar magnitude of the linkage signal, but neither of the initially positive linkage results were consistently replicable in four or more scans.
- Linkage strategy in large extended pedigrees with a Mendelian-like pattern of familial loading did not produce linkage signals that were clearly more distinct and pronounced than in samples of affected siblings up to now (e.g., the most distinct signal in schizophrenia observed by Brzustowicz et al. ***(2)*** in an inbred sample in contrast to an outbred population with signals up to 6.5). In particular, there is no single extended family with a known influential disease gene.
- In light of this scenario, meta-analyses were offered as a consensus strategy. However, even after combining multiple samples with approx 1000 families, the magnitude of the linkage signals did not exceed the magnitude observed in the first positive result ***(3)***.

The analogy to monogenic diseases would recommend first to replicate and then to systematically sharpen the linkage signal (i.e., increase the magnitude of the signal, and reduce the length of the linked region) in a stepwise manner by extension to other informative families. Finally, the disease gene can be identified in this stepwise manner. Linkage in monogenic diseases is powerful, and recombination events between marker and disease loci can be iden-

tified, which is impossible in complex traits. Thus, it does not come as a surprise that this strategy does not work in common diseases using the available tools, as the extension of the sample size does not increase the magnitude of linkage. This constellation has been anticipated on theoretical grounds (*see* **ref. *4***).

From the mutiple genome-wide scans in schizophrenia, bipolar disorder, alcoholism, or late-onset Alzheimer's disease, we can conclude that:

1. No single gene causes any of these disorders. Thus, susceptibility genes rather than causal disease genes are operating; otherwise, a sharp, consistently replicable linkage signal should have been detected.
2. There is no evidence that a major gene contributes most to the genetic variance.
3. Multiple susceptibility genes account for each of these disorders; neither of these contributing genes is necessary and/or sufficient for the manifestation of the disorder (vulnerability or susceptibility genes in complex disorders in contrast to causal genes in monogenic diseases).
4. Each of these multiple genes contributes only modest effects. Some authors speculate that in schizophrenia, for example, the contribution of each susceptibility gene is limited to an odds ratio of less than 2.0 ***(5)***.
5. The genetic heterogeneity cannot be decomposed to more homogeneous subtypes. In particular, subtypes of the major psychiatric disorders that are influenced by a single gene or a major gene have not been found by linkage studies (although postulated on the basis of segregation analysis).

Thus, although a few susceptibility genes have been suggested, no susceptibility gene has been clearly identified in major psychiatric disorders with more than 50% heritability (such as schizophrenia or bipolar disorder). Given the difficulties of narrowing down a candidate region in complex diseases in a systematic manner (as in monogenic diseases), additional opportunities and tools are required to find vulnerability genes. The few successful examples of identifying susceptibility genes for complex diseases reveal the need for additional strategies or favorable conditions:

• *Good luck*: In late-onset Alzheimer's disease, a candidate gene ApoE was located in a linked region, and was confirmed first by association and finally by functional studies ***(6)***.

• *A combination with association or linkage disequilibrium strategy*: In diabetes type 2, a promising candidate gene (calpain-10) was detected in a linked region using a combined linkage-association approach ***(7)***.

Thus, although some progress has been made, the speed needed for disease-gene discoveries is substantially slower than expected when the first linkage studies with DNA markers began nearly 20 yr ago. Several factors may contribute to the lack of replicability of positive linkage findings and to other disappointments and may challenge our initial assumptions, but these may also stimulate new, more promising approaches:

• *Magnitude of gene effects:* Given the results of genome scans in psychiatric disorders, the susceptibility genes are likely to contribute only with small or modest effects. Thus far, linkage analysis has been enormously successful in detecting causal or major gene effects, but not for small effects. In addition, model-based considerations have demonstrated that association studies are usually far more powerful in detecting minor or modest gene effects.

• *Non-additive interaction of susceptibility genes:* Biometrical analysis of the familial pattern of aggregation of diagnoses make it possible to draw conclusions on the putative number of underlying interacting genes and on the mode of interaction. The analysis of cumulative family studies by Risch ***(8)*** suggested the non-additive interaction of multiple genes in schizophrenia. Risch et al. ***(9)*** also concluded from the extended and widespread weak linkage signals detected in a genome scan in autism that more than 20 different loci are interacting. Linkage analysis may also identify interacting loci, but only with a distinct loss of power. Thus, in the presence of non-additive interaction, the required sample size is even higher.

• *Strength of the magnitude of linkage signals across populations:* Some linkage signals were found to be only replicable with comparable genetic background, but not in other populations. Indeed, some susceptibility genes (such as ApoE4 for late-onset

Alzheimer's disease) are only influential in some populations (ApoE4 is mainly relevant in Caucasian but not in black populations) ***(6)***. In schizophrenia, some linkage findings on 8p, 9q, and 15q were exclusively replicable in African populations, whereas 10p was until now only replicable among Caucasian populations ***(5)***.

- *Sample size problem:* Small effect sizes as odds ratios (OR) of about 1.5 require unrealistically large numbers of informative families (e.g., affected sib-pairs). Risch and Merikangas ***(10)*** calculated for OR = 1.5 the number of families required to detect the gene by linkage analysis as 18.000 and more depending on the model; the currently available family sample sizes (~200) are at best able to identify genes with an OR of approx 4. It is evident from these considerations that narrowing down the candidate region to the disease gene cannot be accomplished by linkage analysis alone.
- The sample size required for replication of a specific true linkage finding in complex disorders is substantially higher than for detecting one among many susceptibility genes ***(11)***. Thus, considering the available sample sizes and the previously mentioned complicating factors, replication of "true" linkage findings cannot regularly be expected. Even a single replication of a reported linkage among 10 replication tests is a non-random event that argues for the validity of the initial positive result.

Currently, the positional cloning approach through linkage analysis has also proven disappointing in non-psychiatric complex diseases. The human genome project produced millions of polymorphic genetic markers for fine-mapping of candidate regions, which will improve the power to detect linkage and to refine the candidate regions (*see* Chapter 3). However, there appear to be serious inherent limitations of linkage analysis in complex diseases. It even remains doubtful that the application of most informative marker systems such as SNPs will be able to identify susceptibility genes with modest effects ***(12)***. Therefore, the skeptical attitudes on the utility of linkage analysis in complex diseases are gaining more and more acceptance ***(12,13)***. Thus, alternatives to linkage analyses are receiving growing attention.

2.2. Association Studies

Another strategy to identify susceptibility genes is association studies in case-control samples. This strategy has thus far focused on candidate genes, which are either hypothetically involved in the pathophysiology of the disease (functional candidates), or located in candidate regions identified by linkage analysis (positional candidates). Association studies are based on linkage disequilibrium, which refers to correlations among neighboring alleles in the genome. "True" associations between a disease and a marker allele may have two meanings:

- the marker allele impacts on the risk for the disease, or
- a genetic variant near the marker allele is the actual determinant and is in linkage disequilibrium with the disease allele.

Generally, many studies have followed this approach. The association approach was only clearly successful in psychiatric disorders in identifying the two functional candidates—the ADH-2 and ALDH-2 genes—as susceptibility genes for alcoholism (with the ADH-2*2 and ALDH-2*2 alleles shown to be less common among Asian alcoholics ***(14)***. Similarly, an association between ApoE4 and late-onset Alzheimer's disease has been proven with no negative report in Caucasian populations after linkage analysis identified ApoE as a positional candidate. Functional studies have shown that the identified ADH/ALDH alleles and the ApoE4 allele are susceptibility alleles that directly increase disease risk.

In other diseases and candidate genes, the results are very diverse and difficult to interpret. Reported associations were followed by some positive replications. But there is no claim for association without non-replication. Thus, the association strategy was blamed as the cause of a very high number of false-positives. However, this limitation is not a result of the association technique, but of the inappropriate chosen levels of significance ***(12)***. Meta-analyses for particularly promising associations covering several thousand patients and controls—e.g., 5-HT-2a-receptor or D3-receptor gene in schizophrenia ***(15,16)***—were performed to clarify this diversity;

relative risks for susceptibility alleles of 1.2 to 1.5 were suggested for a very limited number of claimed associations in schizophrenia.

A major advantage of association compared to linkage studies is their relatively high efficiency in the detection of genes with small effect size. Thus, it was suggested that testing every gene in the genome for association may be more feasible than detecting a susceptibility gene by linkage analysis ***(10)***.

Although thousands of cases and controls are needed, this strategy is *a priori* more realistic. However, difficulties and warnings with the association strategy should not be ignored and have to be weighed against the prospects and limitations of linkage analysis ***(12,17,18)***. As there is not a convincingly optimal decision for either of two strategies, both must be considered as complementary.

There are several unresolved problems with the association strategy. One problem is the selection of the most appropriate study group: Are all cases with a specific diagnosis appropriate, or only those with a secondary case in the family? Should probands with comorbidity for two disorders, each with a genetic determination, also be included? A related problem: Should the non-genetic influences on the manifestation of the disorder being studied be taken into consideration? Would an adjustment for impacting environmental factors increase the power of analysis, or even decrease the power ***(19)***?

Valid answers depend on the knowledge of underlying etiological mechanisms, which are largely unknown for psychiatric disorders. Currently, decisions must be based on the most plausible assumptions.

2.3. Combination of Linkage and Association

It has already been demonstrated that a combination of the linkage and the association strategy may overcome the limitations of either strategy alone: the identification of ApoE as a susceptibility gene for late-onset Alzheimer's disease, and calpain-10 as a susceptibility gene for non-insulin-dependent diabetes. In both cases, linkage analysis identified a candidate region. Either:

• association studies with candidate genes in the candidate region, or

• stepwise narrowing down the candidate region by a combination of association and linkage analysis with markers in the candidate region proposed susceptibility genes, which must prove their impact on the disease risk in subsequent functional studies.

However, only a few examples have succeeded by stepwise application of linkage and association studies. Although other examples may follow, it is still to be demonstrated that most of the relevant susceptibility genes, particularly those with only modest effect, can be detected by "combination" strategies. Particularly, it may be difficult to detect susceptibility genes without a replicable linkage signal (i.e., those with an OR of 2.0 and lower). Considering that the realistic sample sizes available for linkage studies are only able to identify susceptibility genes with strong effects with certainty, the stepwise approach may fail to detect linkage signals for genes with only modest or mild effects. Therefore, alternative and complementary strategies are needed.

3. Promising Future Analytic Strategies

Until now, case-control association studies were limited:

• By focus on a candidate-gene approach in the absence of sufficient knowledge of the pathophysiology and etiology of the disease. A positional cloning, genome-wide approach was technically not feasible because the available marker systems could not cover the genome densely enough.

• By uncertain ethnic comparability between cases and controls, which is decisive to avoid false-positives; however, beyond family-based controls comparability is difficult to demonstrate.

Recently, the progress of the human genome project in combination with the detection of the broad variability on the genome has opened new prospects, particularly for association studies:

• Single-nucleotide polymorphisms were found to occur so densely on the genome that in each population each specific SNP variant seemed to be in linkage disequilibrium with SNP variants

nearby (mean linkage disequilibrium ~60 kb in European populations ***[20]*** and one SNP per 2 kb [mean] ***[21]***). Eighty-five percent of the exons of genes are within 5 kb of the nearest SNP (*see* Chapter 3). Thus, using these dense-marker-system "hypothesis"-free genome-wide association studies may detect disease genes through a positional cloning approach ***(12)***.

- Another recent development of molecular genetic control ensures ethnic comparability, and offers stratification techniques to adapt for non-comparability ***(22)***.
- Recently developed analytic techniques enable the consideration of case-control studies—not only differential frequencies of single markers, but also haplotypes (combination of markers) increasing the informativeness of this strategy ***(23)***.

Taken together, genome-wide case-control association studies for a hypothesis-free search for susceptibility genes will be feasible in the near future. Theoretically, this linkage-disequilibrium-based approach can be expected to reveal increased power compared to linkage studies in detecting modest gene effects (RR of 2 and lower) ***(12)***. A series of arguments can be found in favor of as well as against the putative success of this new perspective in excellent reviews (*see* **ref. *18***). Clearly, this controversy can only be solved by doing. As this genome-wide association strategy is only beginning to be set up, its practical utility has not yet been demonstrated. One foreseeable practical problem is that power analyses suggest that very high sample sizes are needed to overcome the multiple testing problem. Although the required sample sizes as calculated can still be achieved in multicenter recruitment programs, the appropriateness of this strategy is still under discussion.

These association studies can be performed in case-control as well as nuclear family samples. Although there is no advantage of family samples in terms of power, nuclear family samples were considered the preferred strategy, as they provide a perfect ethnic matching between cases and the family-based controls. However, the reputation of case-control studies recently gained major support for the following reasons:

- Ethnic comparability of the case and controls can now be tested and achieved by restratification; thus, false-positives can usually be avoided, even with external controls.
- The case sample and the control sample can both be pooled, whereas family-based samples require an individualized genotyping; thus, the recent achievements of high-throughput techniques can best be utilized in case-control samples.
- It is far easier to recruit a well-characterized control sample than a family sample; for late-onset disorders nuclear family samples are impossible to obtain.

Thus, in the future, more rigorously designed case-control samples can be expected to become an optimal study design.

4. Optimal Phenotype Definition

Diagnostic definitions of psychiatric disorders are clinical conventions supported by some external validation criteria. The diagnostic criteria cover a broad range of behavioral and experiental phenomena. The first approach to define the phenotype in searching for susceptibility genes was based on clinical diagnoses. Many efforts were undertaken to develop techniques to maximize reliability and validity and to guarantee comparability across samples and studies of the clinical phenotypes (e.g., interview techniques and polydiagnostic assessments). Some attempts were initiated to refine the clinical diagnoses and maximize the magnitude heritability, with the ultimate goal of limiting the number of false-positive cases. However, it is now evident that the power of linkage analyses in complex diseases remains limited, although the complex phenotype can be defined both in a reliable and valid manner.

Another putative strategy is to decompose complexity into a series of more homogeneous and genetically less complex subtypes. Thus, the phenotypic heterogeneity may result from the mixture of more homogeneous clinical subtypes. However, although some homogeneous subtypes defined by candidate symptoms ***(24)*** (e.g., periodic catatonia in schizophrenia) were postulated, none could finally be validated, with one exception: Alzheimer's disease with several monogenic subtypes among the early-onset variant.

Clearly, alternative approaches to define the phenotype must be explored. Alternative phenotypes should avoid disadvantages of the diagnostic phenotype:

- by reduction of the phenotypic complexity;
- by moving the phenotype to be studied closer to the gene (i.e., from the diagnostic level of behavior and experience to the underlying neurobiology, which may be closer to the gene with less mediating factors); and
- by a more simple genetic transmission than the disease itself.

The more basic and genetically determined abnormality of a disorder was first introduced by Gottesman ***(25)*** into psychiatry and was called "endophenotype." Subsequently, the term "intermediate phenotype" also became familiar. Modern versions of this concept ***(24,26)*** are based on three well-established observations:

- *Each psychiatric disorder is characterized by neurobiological deficits.* These deficits may exist before the manifestation of the disorder. Growing evidence on the neuropathological, physiological, and biochemical basis of psychiatric disorders proposed basic neurobiological deficits as basic characteristics of the disease. Several psychiatric disorders have presented with stable abnormalities in multiple domains, some under genetic control. Thus, the disorder can be considered as a series of distinct deficits, and each of these alone does not present in a disorder. Only the combination of most of these deficits results in the disorder, and only one or a few of the deficits present as subthreshold condition. For example, schizophrenia is associated with deficits in information processing (indicated by P50) or frontal-brain cortical structure. Both indicators are genetically influenced, and may therefore contribute to the genetic impact on schizophrenia. Assuming that brain structure and functioning are closer to the gene function than diagnostically relevant behavior, these neurobiological deficits appear to be more appropriate, simpler phenotypes.
- *Neurobiological heterogeneity:* Multiple pathophysiological pathways are believed to be optionally involved. Given this variability, the clinically defined diagnostic categories present as "final common pathology" defined in behavioral terms emerging from

very different individual basic neurobiological constellations (phenotypical heterogeneity).

- *Etiological heterogeneity:* Genetic and non-genetic determinants have been demonstrated that propose etiological heterogeneity; in addition, all psychiatric disorders are genetically heterogeneous, with multiple genes contributing (genetic heterogeneity).
- *Genetic heterogeneity:* The results of genome-wide linkage studies available now for schizophrenia, bipolar affective disorders, panic disorder, alcoholism, bulimia, and late-onset Alzheimer's disease clearly demonstrate the absence of a causal or major gene for any disorder, but suggest that multiple vulnerability genes are operating in each of these disorders.

The concept of endophenotypes assumes that heterogeneity will map the phenotype on the genetic heterogeneity:

- The endophenotype (i.e., neurobiological deficit) is genetically influenced with a lower number of genes than the disorder itself.
- The endophenotype-genotype relationship is less complex.
- The genes influencing the endophenotype also influence the manifestation of the disease.

First screening of neurobiological correlates of the disorder for endophenotypes is possible in family studies: elevated frequency in high-risk subjects, familial-genetic determination, and stability over time can be used as criteria.

Endophenotypes offer a major advantage. In contrast to the categorical clinical phenotype (disorder present or absent), they are mainly quantitative traits (quantitative trait loci—QTL). Genes for quantitative traits can be more easily detected, because the analyses are more powerful than with categorical traits. Indeed, there is evidence from insulin-dependent diabetes that mutations contributing to the disease risk (VNTRS polymorphism near the insulin gene) are impacting on the disease in a quantitative manner ***(27)***.

The concept of endophenotypes has become successful in targeting susceptibility genes for diseases in some medical diseases beyond psychiatry but also in schizophrenia (P50 abnormality) or in Alzheimer's disease (early age at onset), as discussed in Chapter 6.

This book is unique because it includes the most comprehensive contribution to intermediate phenotypes in psychiatric disorders.

The chapters are organized according to the method of defining the alternative phenotype. Sometimes, such behavioral features as personality are considered as alternative phenotypes. In contrast to neurobiological traits, it is difficult to assume a more direct relationship to the genotype and a less complex genetic determination than for the disorder itself.

5. Ethical Issues

One of the founders of psychiatric genetics, F. Galton, observed the familiality of wanted and unwanted behavioral and mental properties. Motivated by this observation, he proposed an eugenic program of birth control. Driven by the concept of degeneration, the intention was to increase the prevalence of wanted and to decrease the unwanted traits in the general population. Subsequently, Galton and his scholars noticed that their practical conclusion was unjustified because of the possibility of polygenic transmission. However, this reevaluation of the family-study literature did not restrain others such as German and certain Scandinavian psychiatrists from recommending a forced eugenic birth-control program. As a result, about 200,000 ill subjects were forcibly sterilized in Germany before 1945. Since these times, psychiatric genetics has had an uncertain reputation. Until today, as psychiatric geneticists we must always be careful to protect our patients and to recognize and prevent the misuse of our knowledge. The field of psychiatric genetics is sensitized for misuse. Thus, we must face the ethical challenge both today and in the future.

In the past, practical eugenics has tried to increase the wanted and decrease the unwanted elements in the population by forced birth control in population-wide programs. It was soon recognized that those programs could not decrease the frequency of common, genetically influenced disorders because of their polygenic nature. But there are concerns today that eugenic thinking may re-emerge on a voluntary basis: Parents may screen for the occurrence of known susceptibility alleles, and may decide on abortion because of this information. These decisions would ignore the fact that com-

mon diseases can be treated more and more successfully once their etiology is elucidated, and that protective environmental factors may prevent complex diseases, even among high-risk persons. The current ethical concerns focus on the putative misuse of genetic information on common diseases.

The two major areas of concern are discrimination of carriers of susceptibility alleles by employers and insurance companies, and prenatal testing for susceptibility alleles and birth control.

1. Once specific susceptibility genes are known, new targets for the development of more efficient treatments become available. Yet discrimination of carriers of susceptibility alleles may be a likely scenario, as the risk of disorders with major psychosocial impairment, lost working days, and early retirement can be estimated on the basis of genetic predisposition.
2. A universal consensus and data protection laws must be developed before the knowledge of susceptibility genes becomes widely available.

The possibility that the majority of citizens may carry several of the common susceptibility alleles should encourage a general consensus to prevent misuse.

References

1. Robins, E., and Guze, S.B. (1970) Establishment of diagnostic validity in psychiatric illness: its application to schizophrenia. *Am. J. Psychiatry* **126,** 983–987.
2. Brzustowicz, L.M., Hodgkinson, K.A., Chow, E.W., Honer, W.G., and Bassett, A.S. (2000) Location of a major susceptibility locus for familial schizophrenia on chromosome 1q21-q22. *Science* **288,** 678–682.
3. Levinson, D.F., Holmans, P., Straub, R.E., Owen, M.J., Wildenauer, D.B., Gejman, P.V., et al. (2000) Multicenter linkage study of schizophrenia candidate regions on chromosomes 5q, 6q, 10p, and 13q: schizophrenia linkage collaborative group III. *Am. J. Hum. Genet.* **67,** 652–663.
4. Boehnke, M. (1994) Limits of resolution of genetic linkage studies: implications for the positional cloning of human disease genes. *Am. J. Hum. Genet.* **55,** 379–390.
5. Riley, B.P. and McGuffin, P. (2000) Linkage and associated studies of schizophrenia. *Am. J. Med. Genet.* **97,** 23–44.

6. Roses, A.D. (1998) Alzheimer diseases: a model of gene mutations and susceptibility polymorphisms for complex psychiatric diseases. *Am. J. Med. Genet.* **81,** 49–57.
7. Horikawa, Y., Oda, N., Cox, N.J., Li, X., Orho-Melander, M., Hara, M., et al. (2000) Genetic variation in the gene encoding calpain-10 is associated with type 2 diabetes mellitus. *Nat. Genet.* **26,** 163–175.
8. Risch, N. (1990) Linkage strategies for genetically complex traits. I. Multilocus models. *Am. J. Hum. Genet.* **46,** 222–228.
9. Risch, N., Spiker, D., Lotspeich, L., Nouri, N., Hinds, D., Hallmayer, J., et al. (1999) A genomic screen of autism: evidence for a multilocus etiology. *Am. J. Hum. Genet.* **65,** 493–507.
10. Risch, N. and Merikangas, K. (1996) The future of genetic studies of complex human diseases. *Science* **273,** 1516–1517.
11. Suarez, B.K., Hampe, C.L., and Van Eerdewegh, P. (1994) Problems of replicating linkage claims in psychiatry, in *Genetic Approaches to Mental Disorders*, Gershon, E.S. and Cloninger, R.C. American Psychiatric Press, Washington, DC, pp. 23–46.
12. Risch, N. (2000) Searching for genetic determinants in the new millennium. *Nature* **405,** 847–856.
13. Stoltenberg, S.F. and Burmeister, M. (2000) Recent progress in psychiatric genetics—some hope but no hype. *Hum. Mol. Genet.* **9,** 927–935.
14. Shen, Y.C., Fan, J.H., Edenberg, H.J., Li, T.K., Cui, Y.H., Wang, Y.F., et al. (1997) Polymorphism of ADH and ALDH genes among four ethnic groups in China and effects upon the risk for alcoholism. *Alcohol Clin. Exp. Res.* **21,** 1272–1277.
15. Williams, J., Spurlock, G., McGuffin, P., Mallet, J., Nothen, M.M., Gill, M., et al. (1996) Association between schizophrenia and T102C polymorphism of the 5-hydroxytryptamine type 2a-receptor gene. European Multicentre Association Study of Schizophrenia (EMASS) Group. *Lancet* **347,** 1294–1296.
16. Williams, J., Spurlock, G., Holmans, P., Mant, R., Murphy, K., Jones, L., et al. (1998) A meta-analysis and transmission disequilibrium study of association between the dopamine D3 receptor gene and schizophrenia. *Mol. Psychiatry* **3,** 141–149.
17. Malhotra, A.K. and Goldman, D. (1999) Benefits and pitfalls encountered in psychiatric genetic association studies. *Biol. Psychiatry* **45,** 544–550.
18. Baron, M. (2001) The search for complex disease genes: fault by linkage or fault by association? *Mol. Psychiatry* **6,** 143–149.

19. Rijsdijk, F.V., Sham, P.C., Sterne, A., Purcell, S., McGuffin, P., Farmer, A., et al. (2001) Life events and depression in a community sample of siblings. *Psychol. Med.* **31,** 401–410.
20. Reich, D.E., Cargill, M., Bolk, S., Ireland, J., Sabeti, P.C., Richter, D.J., et al. (2001) Linkage disequilibrium in the human genome. *Nature* **411,** 199–204.
21. Sachidanandam, R., Weissman, D., Schmidt, S.C., Kakol, J.M., Stein, L.D., Mullikin, J.C., et al. (2001) A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms. *Nature* **409,** 928–933.
22. Devlin, B. and Roeder, K. (1999) Genomic control for association studies. *Biometrics* **55,** 997–1004.
23. Hoehe, M.R., Kopke, K., Wendel, B., Rohde, K., Flachmeier, C., Kidd, K.K., et al. (2000) Sequence variability and candidate gene analysis in complex disease: association of myopioid receptor gene variation with substance dependence. *Hum. Mol. Genet.* **9,** 2895–2908.
24. Leboyer, M., Bellivier, F., Nosten-Bertrand, M., Jouvent, R., Pauls, D., and Mallet, J. (1998) Psychiatric genetics: search for phenotypes. *Trends Neurosci.* **21,** 102–105.
25. Gottesman, II (1991) *Schizophrenia genesis: The Origins of Madness*. Freeman, New York.
26. Freedman, R., Adler, L.E., and Leonard, S. (1999) Alternative phenotypes for the complex genetics of schizophrenia. *Biol. Psychiatry* **45,** 551–558.
27. Kennedy, G.C., German, M.S., and Rutter, W.J. (1995) The minisatellite in the diabetes susceptibility locus IDDM2 regulates insulin transcription. *Nat. Genet.* **9,** 293–298.

2

Analytical Methods Applied to Psychiatric Genetics

Elena L. Grigorenko and David L. Pauls

1. Introduction

The development of gene-mapping methodology has not been a linear process. Instead, this development has been multidimensional, culminating in the creation of a powerful and heterogeneous collection of tools. A description of the history of the development of this would include words such as "opportunistic" (i.e., capitalizing on the newest developments in computer technology and genomics) and "problem-solving oriented" (i.e., constantly addressing issues (such as the spotted nature of linkage disequilibrium) that arose during the development of the methodology). Therefore, the following presentation is method-oriented rather than problem-oriented. In describing the modern methodology of gene mapping, attempts will be made to describe the origin of a given methodology, the problems it was designed to address, and its known strengths and weaknesses.

There are several ways to categorize current approaches to gene mapping. One possible subdivision is whether a given methodology is a linkage approach or an association approach. A second possible

From: *Methods in Molecular Medicine, vol. 77: Psychiatric Genetics: Methods and Reviews*
Edited by: M. Leboyer and F. Bellivier

division would focus on whether a methodology deals with related individuals (e.g., family members) or unrelated individuals. A third possible division would consider the approaches dealing with related individuals only, summarizing the methods on the basis of the unit of analysis employed (i.e., the type and size of family units—sib-ships, nuclear and extended families, distant relatives, and so on). By necessity, these subdivisions are not exact because of the nature of data collected from families. And as would be expected, there are modern methods that simultaneously evaluate linkage and association, combine information from samples of related and unrelated individuals, and utilize multiple types of relatives.

This chapter is organized as follows. First, linkage methods are reviewed. Then, association study methods are summarized. And finally, the strengths and weaknesses of both approaches (pitfalls unique and common to both) are discussed.

2. Linkage Methods

Newton Morton is generally credited with initiating modern gene-mapping methodology with the publication of the classic paper in which he first introduced the lod-score method ***(1)***. The lod-score method allowed an estimate of the position of a disease gene on a map of markers by examining the likelihood of linkage given a specific genetic model and a specific recombination fraction. In later modifications, it was possible to incorporate incomplete disease allele penetrance and/or the absence of some key individuals in the analyzed pedigrees. Lod (log of the odds) scores consist of the base 10 logarithm of the likelihood ratio of two hypotheses. The first hypothesis postulates that a hypothetical gene is linked to a genetic marker at a given distance determined by the recombination fraction. The second hypothesis postulates no linkage (i.e., the recombination fraction is assumed to be 0.5). The base 10 logarithm of the ratio of the likelihoods of these two hypotheses is defined as the lod score. A separate lod is calculated for a range of recombination fractions. The test for linkage is conducted by examining the maximum value of the lod score for this range of recombination fractions.

The first lod-score test took the form of a sequential probability ratio test ***(1)***. This test was ideally suited for a Mendelian, single-gene mode of inheritance. In the early seventies, the method was extended with the introduction of the Elston-Steward ***(2)*** algorithm that allowed for complex inheritance (e.g., reduced penetrance) in large extended pedigrees. This algorithm was incorporated into the computer program LIPED ***(3)***. The development of LIPED and the advent of faster computers transformed linkage analyses from a time-consuming sophisticated "ordeal" into a common research tool. A major limitation of LIPED was its capacity to deal with only one marker at a time. Thus, a new set of programs ***(4)*** was developed that allowed linkage analyses of multiple markers simultaneously.

At the present time, most linkage analyses utilize multipoint strategies. It is well-known that these methods increase power when analyzing both Mendelian ***(4)*** and non-Mendelian (so-called complex traits) ***(5)***. A number of additional methods have been developed that facilitate the analysis of the multipoint data that are generated by studies performed at today's accepted marker density (10–25 cm marker spacing) ***(6)***. These methods include the exact enumeration of multi-locus genotype probabilities in small pedigrees ***(7)***; estimation of such probabilities for pedigrees of any size and of some complexity ***(8–10)***; and approximation of such probabilities for pedigrees of arbitrary size ***(11)***.

Yet, the lod-score method is preferred for Mendelian traits with (approximately) known inheritance parameters. However, the power of lod-score methods is reduced (sometimes dramatically) when the mode of inheritance ***(12–14)***, penetrance ***(15)***, and disease allele frequency ***(16–17)*** are not known and therefore possibly misspecified. Although this is a potential shortcoming of this method, it has been shown that when lod-score methods are applied many times with different modes of inheritance (e.g., dominant and recessive), a correct approximation of the mode results in lod scores that are generally superior to those obtained through other types of analyses ***(15)***.

Moreover, researchers have developed statistical methods that appear to be robust to misspecification of selected parameters. For example, a likelihood-based efficient score statistic ***(18)*** permits

testing the null hypothesis of no trait locus in a given chromosomal region. This statistic is asymptotically equivalent to the lod score, and it generalizes to a class of statistics developed for a non-parametric approach that examines only affected members of a pedigree ***(7,19–21)***. One advantage of this approach is that in the absence of complete information about the genetic model parameters, this statistic is easier to compute than the exact lod score. It does not require likelihood maximization with respect to the unknown parameters.

Although parametric linkage approaches are continually developed and remain heavily used in the field, the main disadvantage of these methods is that genetic model parameters (i.e., disease allele frequency, mode of inheritance, and penetrance) must be specified. By definition, this is not possible for complex (non-Mendelian) traits. To overcome this dilemma, non-parametric linkage methods have been developed.

Non-parametric linkage methods allow for the study of linkage between a marker (or a set of markers) and a disease without the need to specify the genetic model parameters for the trait under investigation. In classical statistics, non-parametric methods refer to methods in which observed values are replaced by their ranks. In human linkage analysis, non-parametric methods refer to methods in which parameters of disease inheritance are replaced by parameters of inheritance of markers hypothesized to be close to disease loci. An entire constellation of computer software has been developed since the 1990s (for review, *see* http:\\linkage.rockefeller.edu). This development capitalized on and was stimulated by progress in methods for likelihood calculations ***(7,9,22,23)***. Considering that the development of non-parametric methods started significantly later than that of parametric methods, most of them have developed the capacity to analyze both single and multipoint linkage data. For example, methods implemented in programs such as ASPEX ***(24)***, GENEHUNTER ***(7,25)***, and ALLEGRO ***(26)*** can utilize information from all markers on a chromosome and render any point along the chromosome as informative as possible.

It is important to remember, however, that the distinction between parametric and non-parametric methods is not sharp. In fact, it has

been shown that the affected sib-pair paradigm, a clearly non-parametric method in which the only connection to the disease is through the ascertainment scheme (i.e., families are studied in which there are at least two affected siblings) and which bases all calculations on the sharing of markers between these two affected siblings (i.e., no assumptions about parameters such as mode of inheritance or disease penetrance are necessary), is equivalent to the lod-score method when the latter is carried out under assumptions of recessive inheritance with full penetrance and all parental phenotypes are taken to be unknown ***(27,28)***. This implicit similarity is apparent in the use of the ANALYZE program, which emulates affected sib-pair analysis through lod-score analysis.

Whether parametric or non-parametric, linkage approaches utilize family data (its various configurations—siblings, nuclear families, or extended families) with the purpose of estimating the relevant parameters such as recombination fractions (map distances) in intervals between gene loci given certain sets of allele frequencies. These estimations are accomplished by maximum likelihood methods with recursive, family-based calculations of likelihood.

The most common procedures for numerical likelihood evaluation are the Elston-Steward ***(2)*** and second the Lander-Green (1987) ***(29)*** algorithms. The Elston-Steward algorithm (and its extensions) is based on pedigree traversing ("peeling") algorithms. With this approach, pedigrees are split into portions that are handled recursively, resulting in the evaluation of the full pedigree likelihood. Procedures of this type have been implemented in such programs as LIPED, LINKAGE, MENDEL, and VITESSE. The Lander-Green algorithm carries out peeling over loci; this algorithm is implemented in MAPMAKER, CRI-MAP, and GENEHUNTER. Thus, the methods have reciprocal profiles—the first method allows for the analysis of large pedigrees, but the number of gene loci that can be analyzed simultaneously is currently limited (the computational burden increases linearly with family size but exponentially with the number of loci), whereas the second method allows for the analysis of a relatively large number of loci in small pedigrees (the computational burden increases linearly with the number of loci and

exponentially with pedigree size). In addition, the development of the Markov chain Monte-Carlo methods of estimation of likelihoods ***(9,30)*** has allowed the analysis of large families and large numbers of markers (disease genes).

The common assumption for all these methodologies is that there are genes of major effect that "cause" the disease in question. Although this assumption has been modified to some degree in some of the software packages (e.g., the assumptions of heterogeneity within families (for example, as implemented in HOMOLOG and HOMOGM) and varying penetrance), it has limited investigators in the range of genetic systems that can be examined. For the most part, all analytic models are restrained to isolated chromosomes, treating multiple disease loci as if they were independent of each other.

This limitation has been recently addressed by a number of researchers interested in understanding the genetic etiology of complex traits. As noted here, by definition, complex traits are non-Mendelian, and thus are most likely influenced by multiple genetic and non-genetic factors. It is hypothesized that susceptibility to disease results from gene-gene and gene-environment interactions. In fact, the majority of medically and developmentally interesting traits are complex traits that are best conceptualized as quantitative rather than categorical. Methods developed to facilitate the identification of genomic locations of loci contributing to quantitative traits attempt to estimate the variance components associated with individual loci. Usually, such estimations are carried out using the concept of measured-locus heritability. There has been some debate in the literature as to whether there is a universally unbiased estimate of heritability and whether this estimate can be obtained ***(31–33)***. At the present time, there are no universally accepted measured-locus heritability estimates. The choice of an ideal estimator is a function of the sample size and magnitude of the locus-specific contribution to the overall phenotypic variance. Fortunately, the observed biases resulting from the use of different estimators are small, and, thus, this shortcoming should not be viewed as endangering overall outcomes of quantitative trait-linkage analyses.

There are two major classes of methods used for the identification of quantitative trait loci (QTLs), although arguably, the dividing line is artificial. The first class of methods is based on the regression of trait differences between sib-pairs on the number of alleles shared identical by descent (IBD) at a locus being tested ***(34)***. As noted, this approach is confined to sib-pairs and is not applicable to data collected from larger pedigrees.

The second class of approaches is based on classical variance-component analysis. This technique simply separates the total variance into components because of genetic and environmental effects ***(35)***. The first application of this approach to linkages analysis was developed by Hopper and Matthews (1982) ***(36)***. The focus of the method is in modeling an additional variance component for a hypothesized QTL near a marker site and establishing linkage to the marker in the presence of a statistically significant nonzero value for the QTL component (a relative size of the component is interpreted as an indicator of the magnitude of the effect of a detected locus).

Early implementations of the variance-component methodology were based on analysis of only one or two markers at a time ***(37–39)***. Then the methodology was extended to multipoint applications ***(11)*** and further strengthened by the added power of an exact multipoint approach ***(40)***. A number of simulation studies have demonstrated that the variance-components approach appears to be more powerful than the Haseman-Elston regression approach ***(11,41–44)***.

Demonstrating linkage between a disease gene and a marker is only the first (and, sometimes the smallest) step in the process of cloning the gene of interest. Traditionally, after establishing linkage, further recombination mapping techniques have been applied to narrow the region of interest. However, recombination mapping has not yielded significant success for complex traits in refining the region once it has been reduced to one or two megabases, since it is improbable that recombinants will be observed in extant family material ***(45)***. To address this challenge, researchers have developed a number of other methods. One successful approach is based on the observation that ancestral recombinants can produce a

predictable pattern of linkage disequilibrium between the disease gene and a set of markers spanning the critical region ***(46–48)***.

3. Association Methods

Whereas linkage analysis focuses merely on the position of a tested marker, association methodology tests whether a particular allele of a marker, a specific genotype, or a haplotype is enriched in (or statistically associated with) affected individuals compared with unaffected controls. In other words, genetic association studies evaluate the relationship between genetic variants and trait differences in a general population.

Association is observed either because the genetic variant being examined is a functional variant of a gene or the marker is in linkage disequilibrium with a susceptibility gene. When two markers are in linkage disequilibrium (LD), alleles at one locus will show a strong statistical association with alleles at a nearby locus, whereas alleles at distant loci will show no association. If one of these loci is a susceptibility gene, an association between an allele at the first locus and the disease being investigated will be observed. This circumstance forms the basis of LD mapping. The intuitive basis of this method is that specific alleles at loci that were immediately adjacent to the disease locus when it arose (through mutation) will tend to remain on the same chromosome as the disease locus (because of the paucity of recombination events), and thus will be transmitted together with the disease locus from generation to generation.

The genetic association study design has a controversial history in genetic research. Nevertheless, its popularity has grown remarkably during the last few years. The major reason for this growth is the increased number of genetic polymorphisms available to investigators. Ten years ago, the paucity of markers available to researchers made association studies tenuous at best. However, technological advances over the last 2–3 yr have resulted in the identification of nearly 2,000,000 DNA polymorphisms ***(49–50)*** and LD mapping studies are now becoming more feasible. Furthermore, with the

development of more efficient high-throughput genotyping methods, a growing understanding of the underlying structure of the complex phenotypes and the continued development of statistical methods, association approaches have become even more attractive.

The analysis of LD has been widely used for fine-genome mapping and has proven to be fruitful (*see* **ref. *51*** for theoretical support for the empirical success). These successful applications have included (but have not been limited to) simple disequilibrium mapping, examination of the pattern of pairwise disequilibrium between the disease gene and each of a set of markers ***(48,52)***, likelihood-based analyses ***(46,53,54)***, and haplotype fine mapping ***(55)***.

The goal of all these methods is to identify the precise disease-causing DNA variant(s) in a region that is known to be linked and associated with a disease. Within a targeted region, two association strategies are common: a positional candidate approach and a positional cloning approach. Within the positional candidate approach, specific genes or variants are examined on the basis of proposed relationships with the phenotype. Within the positional cloning approach, markers are selected for evaluation purely on the basis of their proximity to one another on a chromosome. These two types of positional searches are usually preceded by replicated linkage data, which typically narrow a region of interest to 1–10 cm. Both positional strategies have been successfully employed in the searches for genes in fully penetrant gene disorders such as cystic fibrosis and Huntington's disease ***(48,56,57)***. However, the application of these strategies has been less useful in complex disorders. A possible reason for this lack of success is that complex disorders are likely to be caused by multiple genes of moderate/small effects, making identification of the underlying genes more difficult. One of the pitfalls of the research on complex disorders using the LD method is our limited understanding of the extent to which LD occurs across the genome ***(58)***. Specifically, there may be a region in which only one functional variant may be relevant to the disorder, but LD could be present across multiple markers in the region,

making the task of "closing in on" the variant of interest much more challenging ***(59)***.

Two design strategies are employed in most association linkage-disequilibrium studies: population case-control designs and family-based association designs.

3.1. Case-Control Studies

The case-control design is the most frequently used design of association studies. The advantage of this design lies in the fact that cases are readily obtained, and can be efficiently genotyped and compared with control populations. The disadvantage of this approach is the difficulty in identifying an appropriate group of matched control cases. It is essential to establish an appropriate control sample, because any systematic allele frequency differences between cases and controls can appear as disease associations—although these may actually result from a number of other factors including but not limited to evolutionary history, group (e.g., ethnicity and gender) differences, and cultural traditions (e.g., mating customs).

The case-control design has been widely used, and its weaknesses are well-known. Specifically:

1. Association studies are often characterized by high rates of Type I (false-positive) errors—a statistically significant association between a phenotype and a polymorphism resulting from randomness in ascertainment of the case and control individuals. The danger of Type I error is increased in situations of multiple tests and relatively small sample sizes of case and control individuals. One reason for a Type I error is population stratification—a characteristic of a population in which cases and controls differ, not only with respect to the phenotype of interest and its genetic etiology, but also with respect to their overall population genetic ancestry (i.e., their general range and frequency of polymorphisms). The result of population stratification is that many irrelevant markers appear to be disease-associated.
2. In the presence of genetic heterogeneity, in which there may be many distinct and potentially interacting environmental and genetic risk factors, it is likely that no single tested genetic marker will pre-

dict disease accurately enough to be statistically apparent within the cost-effective limitations of a single study. Thus, at the present time, sample sizes may be too small to detect real associations.

3. Since association studies usually test many polymorphisms, the majority of them utilize conservative multi-test corrections (e.g., Bonferroni correction for N tests with a target per-test statistical threshold of p-value). However, there is no clear understanding of the magnitude of the Type II error (missed signal error) imposed by such corrections. These corrections may be especially detrimental for alleles with small main but large interactive effects.
4. Another source of false-positive findings is "cryptic relatedness" ***(60)***—an association between affected individuals sharing a genetic disorder. In the presence of cryptic relatedness, test statistics for case-control studies are likely to be inflated, relative to expectations, under the assumption of an independent sample and no genetic association with the disease.
5. Since LD appears to be variable over the genome, the current statistical procedures may not be sensitive enough to allow for the adequate evaluation of statistical significance of specific regions of interest.

Although the limitations of association studies are well-recognized, the association design represents an essential step in the identification and description of disease-mediating genetic variants. In the last several years, a number of proposals in the literature have been made, which should help to overcome some of the limitations of case-control studies. These are summarized here.

Cardon and Bell ***(59)*** suggest that the most appropriate way to ascertain a control sample is through a prospective cohort study. This approach requires the ascertainment of a large population sample of individuals, selected before the onset of disease, who are then followed prospectively until onset of the disease of interest. After the disease has manifested in some individuals, a group of affected individuals would be chosen and matched to a group of unaffected individuals who are part of the same original population sample. Although this approach may be feasible for disorders with relatively early onset, it would be prohibitively expensive for diseases of late onset.

Another possible way to approach the problem of stratification would be the recruitment of several control populations reflecting the various substructures that may exist in the case population. For example, one control population could be matched with the case population for age (to account for cohort-specific mating, migration, and other effects), whereas another control population could be matched with the case population for geographic location. The results of such multiple matching would be the comparison of the case population with a panel of subpopulations representative of the observed stratification.

Another very important consideration in designing an association study is that of power. Simply stated, for association studies to succeed, the samples should be large. This point has recently been vividly demonstrated in studies on the role of polymorphisms around the angiotensin l-converting enzyme (*ACE*) locus and its contribution to the risk of cardiovascular disease. One of the early publications on the role of this gene was conducted on samples of hundreds of men who had survived myocardial infarction and matched controls ***(61)***; it was reported that the *ACE* locus played a role in the risk of particular subgroups to cardiovascular disease. A series of replications, carried out with even smaller sample sizes, produced variable results ***(62)***. The hypothesis was then tested on samples involving thousands of individuals, and was not verified ***(63)***. Thus, for association studies aimed at identifying genes of moderate effects, samples should be comprised of thousands or even tens of thousands of individuals (also *see* **ref. *64***, for research on diabetes). There are very few association studies in which sample sizes approach the ones cited here. If samples of this magnitude were studied, it is likely that the number of unreplicated results would probably decrease ***(59)***.

One important advantage of case-control association studies is that DNA samples from cases and controls can be pooled and genotypes can be grouped together to determine differences in allele frequency across groups of affected and unaffected individuals. This technological advancement, recently applied in a number of contexts ***(65–67)***, must be extremely precise—the difference in

allele frequencies can be quite small and an experimental error of 1–2% can be high enough to jeopardize the outcome. When it is accurate, this technology allows rapid processing of samples from many individuals. However, its application is limited because it does not lend itself to direct haplotype assessment.

Although much work has been devoted to the development of research designs and analytic strategies to minimize Type I errors, it should be noted that the best way to confirm results is through independent replication. For example, Emahazion et al. ***(68)*** argue that Type I errors should be accepted as inevitable. These researchers suggest that association studies should be viewed as a way to screen large numbers of genes or markers, and that statistical thresholds should be chosen that would help identify genes of moderate-to-large effects. They further propose that there should be widespread efforts to replicate these findings. In addition, in an attempt to minimize the false-positive load, the association studies should be designed to minimize the clinical and population heterogeneity and to maximize the utilization of markers with known functional importance.

Although it is inevitable that there will be false-positive results, efforts should be made to attempt to minimize them. One recent approach has been suggested by Devlin and Roeder ***(60)***. These investigators have described a population-based association method using what they describe as a "genomic control" (GC). This method should help to minimize Type I errors that are caused by inappropriate matching of cases and controls. This method is designed to address two major problems that are characteristic of association studies—population stratification and cryptic relatedness. The method requires the additional genotyping of markers that are unlikely to affect liability (null loci). Chi-square statistics are calculated for both null and candidate loci. Utilizing the information on the variability and magnitude of the test statistics observed at the null loci, which are inflated by the impact of population stratification and cryptic relatedness, a multiplier is derived to adjust the critical values for significance tests for candidate loci, permitting analysis of stratified case-control data without an increase rate of

false-positives. If population stratification and cryptic relatedness are not detected from null loci, then the GC method is identical to a standard test of independence for a case-control design.

As previously mentioned, there are limitations to the case-control design. Yet it is clear that this paradigm can be a powerful tool to demarcate the genetic region of a disease-predisposing gene. As Jorde et al. ***(69)*** have argued, the application of association methodologies is especially useful in the case of markers that are tightly linked to a disease gene, when other mapping techniques become difficult. Yet given the variability of LD across the genome, once recombination distances between marker and disease genes become very small, accurate estimates of map position may become very difficult or impossible ***(70)***.

In summary, case-control studies should be considered to be one of several tools that may be useful in identifying susceptibility loci. It is unlikely that they will allow the identification of all genes of interest without other tools. Yet they may be very helpful in combination with other approaches, and they could be particularly helpful in situations in which the disorder under investigation has relatively late onset, making it difficult to obtain the family materials that are essential for other strategies.

For investigators who are considering case-control design, certain recommendations should be considered. First, the study should be designed to minimize population substructure. Second, when highly stratified populations are chose, every effort should be made to describe the substructures as much as possible and account for them in the ensuing statistical analysis. Third, if there is any doubt as to whether the sample being investigated is stratified, investigators should select null loci with common alleles and genotype them so that the GC approach can be utilized.

3.2. Family-Based Studies

An alternative approach for association studies that uses nuclear-family data to estimate control-marker allele frequencies was introduced by Rubinstein and colleagues ***(71)***, Field and colleagues ***(72)***,

and Falk and Rubinstein ***(73)***. The main objective for the development of this approach was to address the problem of population stratification caused by the ethnic mismatching between patients and randomly ascertained controls.

This approach is sometimes referred to as AFBAC (affected family-based controls), and is based on the assumption that the parental marker alleles that are not transmitted to an affected child can be used as control alleles. This matched design for patient (parental transmitted) and "control" (parental non-transmitted) marker alleles avoids ethnic confounding in the case of a stratified population ***(74–75)***. Thomson ***(76)*** demonstrated that for any single-locus model of disease susceptibility and for any nuclear family-based ascertainment scheme, the family-based association tests are an appropriate method for mapping disease genes.

If the "control population" is constructed from the non-transmitted parental alleles, a statistic known as "haplotype relative risk" (HRR—the family-based equivalent of the odds ratio or relative risk for rare diseases in a case-control study) can be computed if it can be assumed that there is random mating and that the population is in Hardy-Weinberg equilibrium ***(71,73,75,77–83)***.

Ott ***(78)*** discussed the statistical properties of the HRR in relation to the null hypothesis being tested. When random mating is assumed, the HRR statistic is equal to 1.0 when (*1*) there is no association between the marker and disease loci at the population level, (*2*) the marker and disease loci are unlinked, or (*3*) both (*1*) and (*2*) are true. However, when HRR = 1, the application of the conventional chi-square test is valid only under the assumption of random mating and when both (*1*) and (*3*) are true. If mating is nonrandom, the valid test for the condition (*2*) is the McNemar test, a statistic used in the evaluation of the "the transmission/disequilibrium test" (TDT) discussed here.

There has been considerable debate in the literature as to whether tests by HRR, contingency table, or McNemar statistics are tests of linkage or association ***(84–86)***. Thomson ***(76)*** has argued that none of these tests are association or linkage tests, according to the traditional definitions of these terms. He stated that these family-based

analyses allow detection of associations of marker genes in the presence of linkage to a disease gene, and therefore necessitate both association and linkage. A number of researchers ***(69,87)*** have noted that the requirement of association at the population level is usually a much more stringent condition than a requirement of linkage. Moreover, when there is no recombination in a randomly mating population, the quantities evaluated by HRR and contingency-table statistics can be compared to those obtained in case-control association studies. Terwilliger and Ott ***(79)*** demonstrated that when random-mating assumptions can be made, the contingency-table statistic is slightly more powerful than the HRR or McNemar tests. Only with large population stratification effects is the power of the McNemar test larger than that of the contingency-table test ***(76)***.

The family-based association paradigm has been extended to allow the incorporation of additional family members. For example, Field ***(88)*** and Thomson et al. ***(89)*** extended this approach to nuclear pedigrees ascertained for the presence of at least two affected siblings. In this design, the alleles that are not transmitted to either sib in the affected sib-pair are used as "control" alleles. Using the AFBAC approach for families with two affected siblings, Thomson and colleagues ***(89)*** showed a significant association between the class 1 allele of the 5' flanking polymorphism of the insulin gene and insulin-dependent diabetes (IDDM). Notably, affected-sib-pair-haplotype-sharing data showed no evidence of linkage to this marker ***(90)***.

Another application of this general approach is the transmission disequilibrium test (TDT) ***(81–82)***. The development of the TDT was motivated by the need to have a test of linkage in the presence of LD. However, it has been primarily used as a test of LD ***(91–92)***. The TDT has gained tremendous popularity because of its low computational demand and the fact that it is applicable to the most common study design used in complex diseases—that of affected and discordant sibling pairs ***(93–98).*** Further developments in TDT approaches resulted in inclusion of a number of additional statistical tests allowing investigation of maternal vs paternal marker association effects; marker associations that are genotype-dependent,

and maternal/fetal interaction effects, both allele- and genotype-specific ***(76)***.

Seltman, Roeder, and Devlin ***(99)*** have developed a strategy known as "evolutionary tree-TDT" (ET-TDT) by combining the theory of TDT with that of measured haplotype analysis (MHA) ***(100)***. MHA utilizes the evolutionary relationships among haplotypes to produce a limited set of hypotheses with regard to a subset of haplotypes. Thus, ED-TDT screens available haplotypes, clusters them, and points to the ancestral ones, which are especially useful for the determination of which polymorphisms within the haplotype are related to disorder liability. Finally, another very recent extension of the TDT for discrete traits includes the genome-wide analyses of SNPs ***(101)***.

Researchers ***(102)*** have compared the efficiency of the GC approach and the TDT method in the presence and absence of population stratification. When population substructure is absent, GC is found to be more efficient than TDT. In the presence of stratification, the GC method is an effective way to control for false-positives. Yet another advantage of GC is its applicability to the data obtained from small isolated populations, in which cryptic relatedness is often present (kinship is often established even between apparent non-relatives).

One disadvantage of the TDT is its reliance on heterozygous parents. Because not all parents will meet this criterion, many may have to be eliminated from the analyses, and this can result in a substantial loss of statistical power. In addition, these family-based approaches (including the TDT) require parental data that may not always be available, especially for disorders with late onset. Thus, although they are more robust in the presence of population stratification, the family-based methodologies are often less practical. Furthermore, in the presence of high homozygosity in families of affected individuals, these approaches could require sample sizes even larger than those for case-control studies to achieve adequate power.

Another disadvantage of the family-based approaches in general is that transmissions are sometimes difficult to resolve when parents

and offspring are all heterozygous for the same bi-allelic marker. To address this problem and increase definitive transmissions, several authors have proposed the use of haplotypes ***(103-108)***. With the exception of cases in which the markers being tested are functional variants of the susceptibility gene, transmissions from parents to offspring are more informative for haplotypes than single markers. However, it should be noted that using haplotypes increases the degrees of freedom of the test and thus reduces the power of the test.

In addition to the HRR and TDT, researchers have developed a number of statistical techniques to test for a marker/disease association by using nuclear-family data. In all of these approaches, contingency table analyses are used to examine the distribution of specific parental alleles among affected individuals.

Assuming random mating and no marker association with disease, a contingency table of parental transmitted vs non-transmitted alleles can be compared by means of the chi-square statistic ***(72,79,81,88,89).*** However, when there is evidence for non-random mating, the McNemar test can be applied to test deviations from the expected 50% transmission ratios of marker alleles from heterozygous parents ***(74,75,79,81,82,88,109–111)***.

Ott ***(78)*** and Knapp et al. ***(77)*** have demonstrated that the utilization of nuclear family-based data in the framework of association studies confounds tests of association and linkage. Family-based association studies will detect marker/disease associations only if the marker and disease genes are in LD. A number of comprehensive statistical packages have been developed that combine parametric and non-parametric linkage and disequilibrium analyses ***(112)***. For example, Göring and Terwilliger ***(16-17)*** estimate a test statistic that consists of three components: (1) linkage within sib-ships, (2) linkage between sib-ships, and (3) association between pedigrees. Unfortunately, at the present time, most of these methods are limited to studies in which the phenotypes are categorical.

As is the case for other analytic methods, the development of the association methodology for quantitative traits has lagged behind ***(32,113)***. Yet several developments should prove helpful in the

study of complex quantitative phenotypes. Allison ***(114)*** proposed a method for detecting linkage disequilibrium in proband/parent pairs for quantitative traits, and Rabinowitz ***(115)*** has extended this method to incorporate data from families. Subsequently, Fulker and colleagues ***(116)*** described a variance component model for the analyses of quantitative data generated from sib-pairs (in the absence of parental data). This method provides tests of linkage and association separately. Cardon ***(117)*** extended the model developed by Fulker et al. by describing a regression model for the analysis of LD in quantitative traits. One advantage of this extension is its relative ease and speed of application. And finally, Abecasis, Cardon, and Cookson ***(118)*** have extended Fulker's method to allow for sib-ships of any size, with or without parental data. With this approach, association is partitioned into two categories: between and within family components. One advantage of this method is that using families with multiple siblings can increase power. This extension is quite useful from a practical point of view. It is to be expected that in any study there will be families of variable sib-ship sizes and occasional missing parents. This method allows the use of all data collected.

In sum, association studies (whether case-control material or family-based) have both strengths and weaknesses. The eventual success of such studies is dependent on a more complete understanding of the distribution of LD across the genome, among other things. Given the information that has become available from the Human Genome Project, it is clear that more challenges remain in our attempts to identify genes of import for complex psychiatric traits. It is quite possible that new discoveries may challenge or strengthen some assumptions regarding association methodology. Nevertheless, association studies can be a valuable tool in identifying susceptibility genes, and can also help us to understand how the genome is organized and how it functions. However, as with any approach, this method must be applied with care. Investigators must be aware of the potential weaknesses in the results obtained and interpret their data accordingly. Caution and careful interpretation should be the mantra of all scientists, and this is especially true for researchers who study the genetics of complex psychiatric disorders.

3.3. Association Approaches Using Single-Nucleotide Polymorphisms (SNPs)

As noted, in order for association studies to be successful, a large number of closely linked markers spanning the regions of interest must be genotyped in order to demonstrate LD with the susceptibility gene. And this must be done inexpensively. Single-nucleotide polymorphisms (SNPs) ***(119–120)*** are a recently discovered class of polymorphisms that have been suggested as the markers of choice for such endeavors. SNPs are the most frequent type of variation in the human genome; the SNP refers to a position at which two alternative bases occur at appreciable frequency (>1%) in the human population. SNPs can be powerful tools for a variety of medical genetic studies (although individual SNPs, which have only two alleles, are less informative than currently used genetic markers (SSLPs—simple sequence-length polymorphisms), which are mostly multi-allelic), since they are much more abundant and the automatization of their processing can be done more easily than that of SSLPs ***(121)***.

SNP-based studies can be completed on either case-control or family data. The typical design of such a study relies on genotyping of a number of SNPs from candidate genes or regions (particularly those with hypothesized functional importance) in relatively large samples of affected and control participants gathered from families or specific populations. By genotyping many SNPs in a small region (or gene), it is likely that LD will be observed. It has been suggested that this approach should have the potential to identify common alleles that confer a twofold increased risk of disease. However, a number of investigators have suggested that this may be an optimistic prediction ***(122–127)***. The major concerns are: whether such common pathogenic variants exist for diseases of interest, and if so, whether sufficiently dense and powerful scans could be conducted given the diverse nature of human populations and the variability in the nature and extent of linkage disequilibrium across the genome ***(68)***.

As mentioned here, a generally accepted strategy in the mapping of a disease gene is to initially apply linkage analysis for an approx-

imate estimate of the location of the trait gene and to subsequently make use of linkage disequilibrium (association) for a more accurate localization. This general strategy is based on the assumption that disequilibrium extends over much shorter distances from a disease gene than linkage. The efficacy of this strategy has recently been challenged by the suggestion that, with a large number of SNPs available, it would be possible to localize disease genes with the disequilibrium mapping approach alone (e.g., by means of case-control studies). This assumption has not yet been empirically supported—no studies have used SNP LD strategy to map a disease gene. However, a number of theoretical investigations have explored efficiency, cost-effectiveness, and methods for this strategy.

One of the lines of such theoretical investigations involves the question of how many such markers exist on a genome-wide basis. This question can be reformulated in terms of the extent of LD in the genome—how rapidly does disequilibrium decay with the distance from the disease gene growing longer? An early estimate ***(128)*** was that, in large outbred populations, disequilibrium should be detectable within 100 kb of a disease locus. A later study that was based on a review of the published literature presented a more positive approach, suggesting that the distance is 300–500 kb ***(129)***. A recent computer simulation predicted an extremely short range of useful disequilibrium—3 kb ***(124)***. Such dramatic differences can be directly translated into associated costs—according to the first two estimates the required number of SNPs would be 30,000–100,000, and results from the third study suggest that 500,000 of SNPs would be needed.

One possible solution to the problem of not knowing the number of markers necessary to map a gene may be to select affected individuals from populations in which the extent of disequilibrium is greater than average. The literature contains some evidence suggesting that isolated populations are more advantageous for association mapping ***(130–131)***. However, this assumption has been challenged. Several examples have been published in which it appears that the extent of LD is either the same or only slightly higher in small, isolated populations as compared to large, outbred

populations ***(132–133)***. Although many factors may contribute to variability of the extent of LD in isolated populations (e.g., their histories, size, and current status), Ott ***(134)*** suggests that they appear to be of great inportance in association studies, especially when candidate genes are available.

3.4. Pitfalls of Current Gene-Mapping Methodologies

Whether conducting linkage or association studies, a number of factors make modern methodologies vulnerable to error.

3.4.1. Genotyping Errors

With the increased availability of markers (both SNPs and more conventional markers), the impact of genotyping error on the outcomes of different analytic methods is significant. Several authors have proposed methods to identify pedigrees and/or individuals with marker errors ***(135–140)***. Usually, an error is identified if it leads to a Mendelian inconsistency. However, Gordon et al. ***(141–142)*** have shown that under these conditions error detection rates are quite low, ranging between 25% and 30% (the detection rate is lowest when the two marker alleles have equal frequencies) when the true error rate is actually 3–4× higher. Clearly, better error detection is needed. On the other hand, a more economic approach may be to incorporate the allowance for errors into the analysis ***(143)*** as originally proposed by Keats et al. ***(144)***.

3.4.2. Map Misspecification

One of the major difficulties in the field at present is the degree of uncertainty in estimates of between-marker distances. Moreover, even when the distance is known, its estimate usually comes from a single source, and thus is usually a sex-averaged estimate. It is well-known that recombination rates differ in males and females ***(145)*** and they vary across different regions of the genome ***(146)***. Obviously, map misspecification can lead to lod score bias. Several studies have investigated the impact of map misspecification on linkage

and LD analyses ***(147–149)***. Most recently, Daw, Thompson, and Weijsman ***(150)*** have investigated map-misspecification bias (the discrepancy between the true lod score and the score estimated under incorrect map) in multipoint linkage analysis. These investigators reported that, in the presence of true linkage, any map misspecification causes a negative bias in lod scores, resulting in a loss of power to detect linkage. In the absence of linkage, map misspecification can cause positive or negative bias. Specifically, the utilization of the sex-average map results in a positive bias; so does overestimation of the distance. Underestimation of the distance results in a negative bias.

3.4.3. Allele-Frequency Misspecification

Genetic linkage and association analyses are highly sensitive to estimates of allele frequencies. Specifically, underestimation of allele frequencies can lead to false linkages, whereas overestimation can lead to reduced power ***(13,151–152)***. Allele frequency misspecification is especially dramatic for association and linkage studies performed on conglomerates of families (or case/control participants) of different ethnic origins. Several approaches have been proposed to address this problem. Specifically, when the analyzed sample is comprised of participants/families of different origins, one solution is to utilize published allele frequencies for each source population (i.e., populations whose representatives are present in the stratified sample), and, assuming that these published frequencies contain some error, "shrink" subpopulation estimates toward some common values ***(153–154)***. Lange's approach utilizes the empirical Bayes estimator for allele distributions and estimates the degree of allele frequency heterogeneity for a locus of interest, shrinking subpopulation-specific allele frequencies toward their pooled estimates as a function of the estimated subpopulation heterogeneity. Lockwood et al. ***(155)*** have extended Lange's approach by incorporating prior information about allele frequencies and interpopulation divergence into empirical Bayes analysis. This approach is implemented in the program ALLDIST.

3.5. Proportion of SNPs to Other Markers in Genome Screens

Chapman and Wijsman ***(156)*** have carried out a simulations analysis comparing diallelic markers and multiallelic markers in terms of sample sizes required for detection of LD (with the utilization of a single-marker locus in a case-control study for rare monophyletic diseases with Mendelian inheritance). These authors have demonstrated that multiallelic markers are more powerful for the detection of LD compared to diallelic markers, and that the ratio of the number of diallelic to the number of multiallelic markers, needed for equivalent power increases with mutation age and complexity in the mode of inheritance. In short, it takes many more diallelic markers than multiallelic markers to detect LD in a reasonable sample size.

3.6. Multiple Comparisons

The issue of multiple comparisons has long been the focus of attention of statisticians and epidemiologists, and there has been considerable debate about which corrections are appropriate. At one extreme is the position that the concern over multiple comparisons should not be an issue at all, and therefore does not require consideration ***(157)***. At the other extreme is the position that correction for multiple comparisons should be performed in all analyses, whether or not there were multiple comparisons in the reported analyses ***(158–159)***. And, of course, there are many positions between the two extremes.

This issue is important because of the rapidly approaching period when thousands of markers (SNPs and others) will be typed on the same sample. Thus, although it will be possible to significantly increase the number of markers typed, there are only a finite number of cases available for research. And as Lander ***(122)*** has pointed out, significantly increasing the number of comparisons requires a significant increase in the sample size studied. Although the magnitude of the increase in sample size is being debated ***(160)***, it is clear that more cases will be needed as the number of markers being genotyped increases.

3.7. Significance Level

Because a large number of markers are tested in genome-wide screens, Lander and Kruglyak ***(158)*** suggested that the point-wise (marker-specific) significance should be set at 0.000023 to correct for multiple comparisons. This *p*-value corresponds to a lod score of 3.6. The threshold results from the assumption that there is no disease gene when calculating the expected rate of false-positives. Thus, this approach has been challenged as being over-conservative, since it has been argued that researchers do not undertake linkage analyses unless there is strong evidence that genetic factors are important in the expression of the disorder being studied.

Several investigators have proposed methods for more accurately determining significance levels. Lucek and colleagues ***(161)*** introduced a methodology to investigate the inheritance of all markers jointly over the whole genome. The methodology unfolds as follows. For each parent in a set of affected sib-pair families, it is determined whether the parents pass on the same or a different allele to the two offspring. Under Mendelian inheritance, without influence of any disease loci, these two events have equal probabilities. However, under the assumption that the disease marker is located close to a disease locus underlying the trait in the two siblings, allele sharing is expected to occur with a probability higher than $^1/_2$. The goal of the methodology is to compare two sets of data, the observed allele-sharing data and randomly generated data that are known not to contain disease loci. The comparison is carried out by means of nonlinear discriminant analysis. The resulting weights are then used to construct a measure identifying the set of marker loci that jointly show deviations from random allele sharing.

Hoh and Ott ***(162)*** have developed a so-called scan statistic that is designed to combine information on multiple contiguous genetic markers used in a genome screen for susceptibility loci for various types of patients (e.g., sib-pairs, nuclear families, or case and control participants). This statistic can be calculated for a given length, and its significance can be assessed by a Monte Carlo permutation test. Multiple significance values are computed for statistics of given

lengths and then compared so that the smallest observed *p*-value is treated as the statistic of interest (for which the overall level can be determined). Illustrating this statistic, Hoh and Ott analyzed a 324-marker dataset obtained through a 10-cm-wide genome screen of autism affected-affected and affected-unaffected sibpairs. The initial single-marker screen did not result in any significant *p*-values. Thus, having used a set of statistics of a number of lengths (10, 20, 30, ..., 100 cm), the smallest observed value was obtained for the screen statistics calculated at 60 cm ($p = .015$). The overall significance level for the statistic of this length was .038. The scan statistic provides additional support for linkage above and beyond what is conveyed by the maximum lod score; it is especially useful when a susceptibility locus appears to be associated with multiple-marker loci (a situation frequently observed in genome-wide searchers). The scan statistic appears to be a useful tool in a number of designs, but its application may vary depending on the population investigated and/or the analysis utilized (e.g., it can be carried out with larger genomic regions in the context of linkage analyses and smaller genomic regions in the context of association analyses). The statistic appears to have more power as a method to detect linkage; however, once linkage has been established, it does not appear to be as useful for narrowing a candidate region.

4. Back to the Beginning

As stated at the outset, the development of appropriate methodologies for the genetic analysis of complex neuropsychiatric disorders has been, and will continue to be, challenging. This challenge arises because over the last decade the technology for gene mapping has developed exponentially, and the volumes of genomic data now available have required the development of new and innovative computing capabilities. There is considerable ongoing development in this area at the present time. Thus, it is very likely that the methods reviewed in this chapter will become outdated very quickly. A review of the major genetic journals—the *American Journal of Human Genetics*, *Neuropsychiatric Genetics*, and

Genetic Epidemiology, among others—reveals that there is at least one significant methodological advance each month. It is anticipated that there will now be an exponential increase in the analytic tools necessary to understand the genotypic and phenotypic data that will be generated.

References

1. Morton, N.E. (1955) Sequential tests for the detection of linkage. *Am. J. Hum. Genet.* **7,** 277–318.
2 Elston, R.C. and Steward, J. (1971) A general model for the analysis of pedigree data. *Hum. Hered.* **21,** 523–542.
3. Ott, J. (1974) Estimation of the recombination fraction in human pedigrees: Efficient computation of the likelihood for human linkage studies. *Am. J. Hum. Genet.* **26,** 588–597.
4. Lathrop, G.M., Lalouel, J.M., Julier, C., and Ott, J. (1984) Strategies for multilocus linkage analysis in humans. *PNAS* **81,** 3443–3446.
5. Wijsman, E.M. and Amos, C. (1997) Genetic analysis of simulated oligogenic traits in nuclear and extended pedigrees: summary of GAW10 contributions. *Genet. Epidemiol.* **14,** 719–735.
6. Yuan, B., Vaske, D., Weber, J.L., Beck, J., and Sheffield, V.C. (1997) Improved set of short-tandem-repeat polymorphisms for screening the human genome. *Am. J. Hum. Genet.* **60,** 459–460.
7. Kruglyak, L., Daly, M.J., Reeve-Daly, M.P., and Lander, E.S. (1996) Parametric and nonparametric linkage analysis: a unified multipoint approach. *Am. J. Hum. Genet.* **58,** 1347–1363.
8. Daw, E.W., Health, S.C., and Wijsman, E.M. (1999) Multipoint oligogenic analysis of age-at-onset data with applications to Alzheimer's disease pedigrees. *Am. J. Hum. Genet.* **64,** 839–851.
9. Heath, S.C. (1997) Markov chain Monte Carlo segregation and linkage analysis for oligogenic models. *Am. J. Hum. Genet.* **61,** 748–760.
10. Health, S.C., Snow, G.L., Thompson, E.A., Tseng, C., and Wijsman, E.M. (1997) MCMC segregation and linkage analysis. *Genet. Epidemiol.* **14,** 1011–1016.
11. Almasy, L. and Blangero, J. (1998) Multipoint quantitative-trait linkage analysis in general pedigrees. *Am. J. Hum. Genet.* **62,** 1198–1121.
12 Martinez, M., Khlat, M., Leboyer, M., and Clerget-Darpoux, F. (1989) Performance of linkage analysis under misclassification error when the genetic model is unknown. *Genet. Epidemiol.* **6,** 253–258.

13. Risch, N. and Giuffra, L. (1992) Model misspecification and multipoint linkage analysis. *Hum. Hered.* **42,** 77–92.
14. Schork, N.J., Boehnke, M., Terwilliger, J.D., and Ott, J. (1993) Two-trait-locus linkage analysis: a powerful strategy for mapping complex genetic traits. *Am. J. Hum. Genet.* **53,** 1127–1136.
15. Abreu, P.C., Greenberg, D.A., and Hodge, S.E. (1999) Direct power comparisons between simple lod scores and NPL scores for linkage analysis in complex diseases. *Am. J. Hum. Genet.* **65,** 847–857.
16. Göring, H.H.H. and Terwilliger, J.D. (2000) Linkage analysis in the presence of errors III: marker loci and their map as nuisance parameters. *Am. J. Hum. Genet.* **66,** 1298–1309.
17. Göring, H.H.H. and Terwilliger, J.D. (2000) Linkage analysis in the presence of errors. IV. Joint pseudomarker analysis of linkage and/or linkage disequilibrium on a mixture of pedigrees and singletons when the mode of inheritance cannot be accurately specified. *Am. J. Hum. Genet.* **66,** 1310–1327.
18. Whittemore, A.S. (1996) Genome scanning for linkage: an overview. *Am. J. Hum. Genet.* **59,** 704–716.
19. Fimmers, R., Seuchter, S.A., Neugebauer, M., Knapp, M., and Baur, M.P. (1989) Identity-by-descent analysis using all genotype solutions, in *Multipoint mapping and linkage based on affected pedigree members: Genetic Analysis Workshop 6,* (Elston, R.C., Spence, M.A., Hodge, S.E., and MacCluer, J. W., eds.), Alan R. Liss, New York, NY, pp. 123–128.
20. Kruglyak, L. and Lander, E.S. (1995) Complete multipoint sib-pair analysis of qualitative and quantitative traits. *Am. J. Hum. Genet.* **57,** 439–454.
21. Whittemore, A.S. and Halpern, J. (1994) A class of tests for linkage using affected pedigree members. *Biometrics* **50,** 117–127.
22. Irwin, M., Cox, N., and Kong, A. (1994) Sequential imputation for multilocus linkage analysis. *Proc. Natl. Acad. Sci. USA* **91,** 11,684–11,688.
23. O'Connell, J.R. and Weeks, D.E. (1995) The VITESSE algorithm for rapid exact multilocus linkage analysis via genotype set-recoding and fuzzy inheritance. *Nat. Genet.* **11,** 402–408.
24. Schwab, S.G., Albus, M., Hallmayer, J., Honig, S., Borrmann, M., Lichtermann, D., et al. (1995) Evaluation of a susceptibility gene for schizophrenia on chromosome 6p by multipoint affected sib-pair linkage analysis. *Nat. Genet.* **11,** 325–327.

25. Friddle, C., Koskela, R., Ranade, K., Hebert, J., Cargill, M., Clark, C.D., et al. (2000) Full-genome scan for linkage in 50 families segregating the bipolar affective disease phenotype. *Am. J. Hum. Genet.* **66,** 205–215.
26. Gudbjartsoon, D.F., Jonasson, K., Frigge, M.L., and Kong, C.A. (2000) Allegro, a new computer program for multipoint linkage analysis. *Nat. Genet.* **25,** 12–13.
27. Hyer, R.N., Julier, C., Buckley, J.D., Trucco, M., Rotter, J., Spielman, R., et al. (1991) High-resolution linkage mapping for susceptibility genes in human polygenic disease: insulin-dependent diabetes mellitus and chromosome 11q. *Am. J. Hum. Genet.* **48,** 243–257.
28. Knapp, M., Seuchter, S.A., and Baur, M.P. (1994) Two-locus disease models with two marker loci: the power of affected-sib-pair tests. *Am. J. Hum. Genet.* **55,** 1030–1041.
29. Lander, E.S. and Green, P. (1987) Construction of multilocus genetic maps in humans. *Proc. Natl. Acad. Sci. USA* **84,** 2363–2367.
30. Guo, S.W. and Thompson, E.A. (1992) A Monte Carlo method for combined segregation and linkage analysis. *Am. J. Hum. Genet.* **51,** 1111–1126.
31. Borewinkle, E., Chakraborty, R., and Sing, C. (1986) The use of measured genotype information in the analysis of quantitative phenotypes in man. *Ann. Hum. Genet.* **50,** 181–194.
32. Boerwinkle, E.R.C. and Sing, C.F. (1986) Bias of the contribution of single-locus effects to the variance of a quantitative trait. *Am. J. Hum. Genet.* **39,** 663–676.
33. Wijsman, E.M. and Nur, N. (2001) On estimating the proportion of variance in a phenotypic trait attributable to a measured locus. *Hum. Hered.* **51,** 145–149.
34. Haseman, J.K. and Elston, R.C. (1972) The investigation of linkage between a quantitative trait and a marker locus. *Behav. Genet.* **2,** 3–19.
35. Lange, K., Westlake, J., and Spence, M.A. (1976) Extensions to pedigree analysis. III. Variance components by the scoring method. *Ann. Hum. Genet.* **39,** 485–491.
36. Hopper, J.L. and Matthews, J.D. (1982) Extensions to multivariate normal models for pedigree analysis. *Ann. Hum. Genet.* **46,** 373–383.
37. Amos, C.I. (1994) Robust variance-components approach for assessing genetic linkage in pedigrees. *Am. J. Hum. Genet.* **54,** 535–543.
38. Goldgar, D.E. (1990) Multipoint analysis of human quantitative genetic variation. *Am. J. Hum. Genet.* **47,** 957–967.

39. Schork, N.J. (1993) Extended multipoint identity-by-descent analysis of human quantitative traits: efficiency, power, and modeling considerations. *Am. J. Hum. Genet.* **53,** 1306–1319.
40. Pratt, S.C., Daly, M.J., and Kruglyak, L. (2000) Exact multipoint quantitative-trait linkage analysis in pedigrees by variance components. *Am. J. Hum. Genet.* **66,** 1153–1157.
41. Amos, C.I., Zhu, D.K., and Boerwinkle, E. (1996) Assessing genetic linkage and association with robust components of variance approaches. *Ann. Hum. Genet.* **60,** 143–160.
42. Amos, C.I., Krushkal, J., Thiel, T.J., Young, A., Zhu, D.K., Boerwinkle, E., et al. (1997) Comparison of model-free linkage mapping strategies for the study of a complex trait. *Genet. Epidemiol.* **14,** 743–748.
43. Pugh, E.W., Jaquish, C.E., Sorant, A.J.M., Doetsch, J.P., Bailey-Wilson, J.E., and Wilson, A.F. (1997) Comparison of sib-pair and variance components methods for genomic screening. *Genet. Epidemiol.* **14,** 867–872.
44. Williams, J.T., Duggirala, R., and Blangero, J. (1997) Statistical properties of a variance components method for quantitative trait linkage analysis in nuclear families and extended pedigrees. *Genet. Epidemiol.* **14,** 1065–1070.
45. Boehnke, M. (1994) Limits of resolution of genetic linkage studies: implications for the positional cloning of human genetic disease. *Am. J. Hum. Genet.* **55,** 379–390.
46. Hästbacka, J., de la Chapelle, A., Kaitila, I., Sistonen, P., Weaver, A., and Lander, E. (1992) Linkage disequilibrium mapping in isolated founder populations: diastrophis dysphasia in Finland. *Nat. Genet.* **2,** 204–211.
47. Hästbacka, J.K., de la Chapelle, A., Mahanti, M. M., Clines, G., Reeve-Daly, M.P., Daly, M., et al. (1994) The diastrophic dysphasia gene encodes a novel sulfate transporter: positional cloning by fine-structure linkage disequilibrium mapping. *Cell* **78,** 1073–1087.
48. Kerem, B., Rommens, J.M., Buchanan, J.A., Markiewicz, D., Cox, T.K., Chakravarti, A., et al. (1989) Identification of the cystic fibrosis gene: genetic analysis. *Science* **245,** 1073–1080.
49. Altshuler, D., Pollara, V.J., Cowles, C.R., Van Etten W.J., Baldwin, J., Linton, L., et al. (2000a) An SNP map of the human genome generated by reduced representation shotgun sequencing. *Nature* **407,** 513–516.

50. Mullikin, J.C., Hunt, S.E., Cole, C.G., Mortimore, B.J., Rice, C.M., Burton, J., et al. (2000) An SNP map of human chromosome 22. *Nature* **407,** 516–520.
51. Devlin, B. and Risch, N. (1995) A comparison of linkage disequilibrium measures for fine scale mapping. *Genomics* **29,** 311–322.
52. Feder, J.N., Gnirke, A., Thomas, W., Tsuchihashi, Z., Ruddy, D.A., Basava, A., et al. (1996) A novel MHC class I-like gene is mutation in patients with hereditary haemochromatosis. *Nat. Genet.* **13,** 399–408.
53. Kaplan, N., Hill, W.G., and Weir, B.S. (1995) Likelihood methods for locating disease genes in nonequilibrium populations. *Am. J. Hum. Genet.* **56,** 18–32.
54. Devlin, B., Risch, N., and Roeder, K. (1996) Disequilibrium mapping: composite likelihood for pairwise disequilibrium. *Genomics* **36,** 1–16.
55. Lam, J.C., Roeder, K., and Devlin, B. (2000) Haplotype fine mapping by evolutionary trees. *Am. J. Hum. Genet.* **66,** 659–673.
56. Riordan, J.R., Rommens, J.M., Kerem, B., Alon, N., Rozmahel, R., Grzelczak, Z., et al. (1989) Identification of the cystic fibrosis gene: cloning and characterization of complementary DNA. *Science* **245,** 1066–1073.
57. Rommens, J.M., Iannuzzi, M.C., Kerem, B., Drumm, M.L., Melmer, G., Dean, M., et al. (1989) Identification of the cystic fibrosis gene: chromosome walking and jumping. *Science* **245,** 1059–1065.
58. Abecasis, G.R., Nogushi, E., Heinzmann, A., Traherne, J.A., Bhattacharyya, S., Leaves, N.I., et al. (2001) Extent and distortion of linkage disequilibrium in three genomic regions. *Am. J. Hum. Genet.* **68,** 191–197.
59. Cardon, L.R. and Bell, J.I. (2001) Association study designs for complex diseases. *Nature Reviews, Genetics* **2,** 91–99.
60. Devlin, B. and Roeder, K. (1999) Genomic control for association studies. *Biometrics* **55,** 997–1004.
61. Cambien, F., Poirier, O., Lecerf, L., Evans, A., Cambou, J.P., Arveiler, D., et al. (1992) Deletion polymorphism in the gene for angiotensin-converting enzyme is a potent risk factor for myocardial infarction. *Nature* **359,** 641–644.
62. Jeron, A., Hengstenberg, C., Engel, S., Lowel, H., Riegger, G.A., Schunkert, H., et al. (2001) The D-allele of the ACE polymorphism is related to increased QT dispersion in 609 patients after myocardial infarction. *Eur. Heart J.* **22,** 663–668.

63. Keavney, B., McKenzie, C., Parish, S., Palmer, A., Clark, S., Youngman, L., et al. (2000) Large-scale test of hypothesized associations between the angiotensin-converting-enzyme insertion/deletion polymorphism and myocardial infarction in about 5000 cases and 6000 controls. International Studies of Infarct Survival (ISIS) Collaborators. *Lancet* **355,** 434–442.
64. Altshuler, D., Hirschhorn, J.N., Klannemark, M., Lindgren, C.M., Vohl, M.C., Nemesh, J., et al. (2000b) The common PPARgamma Pro12Ala polymorphism is associated with decreased risk of type 2 diabetes. *Nat. Genet.* **26,** 76–80.
65. Barcellos, L.F., Klitz, W., Field, L.L., Tobias, R., Bowcock, A.M., Wilson, R., et al. (1997) Associating mapping of disease loci by use of a pooled DNA genomic screen. *Am. J. Hum. Genet.* **61,** 734–747.
66. Daniels, J., Holmans, P., Williams, N., Turic, D., McGuffin, P., Plomin, R., et al. (1998) A simple method for analyzing microsatellite allele image patterns generated from DNA pools and its application to allelic association studies. *Am. J. Hum. Genet.* **62,** 1189–1197.
67. Shaw, S.H., Carrasquillo, M.M., Kashuk, C., Puffenberger, E.G., and Chakravarti, A. (1998) Allele frequency distributions in poled DNA samples: applications to mapping complex disease gene. *Genome Res.* **8,** 111–123.
68. Emahazion, T., Feuk, L., Jobs, M., Sawyer, S.L., Fredman, D., St. Clair, D., et al. (2001) SNP association studies in Alzheimer's disease highlight problems for complex disease analysis. *Trends Genet.* **17,** 401–407.
69. Jorde, L.B., Watkins, W.S., Carlson, M., Groden, J., Albertson, H., Thliveris, A., et al. (1994) Linkage disequilibrium predicts physical distance in the adenomatous polyposis coli region. *Am. J. Hum. Genet.* **54,** 884–898.
70. Hill, W.G. and Weis, B.S. (1994) Maximum-likelihood estimation of gene location by linkage disequilibrium. *Am. J. Hum. Genet.* **54,** 705–714.
71. Rubinstein, P., Walker, M., Carpenter, C., Carrier, C., Krassner, J., Falk, C., et al. (1981) Genetics of HLA disease association: the use of the haplotype relative risk (HRR) and the "Haplo-Delta" (Dh) estimates in juvenile diabetes from three racial groups. *Hum. Immunol.* **3,** 384.
72. Field, L.L., Fothrgill-Payne, C., Bertrams, J., and Baur, M.P. (1986) HLD-DR effects in a large German IDDM data set. *Genet. Epidemiol.* (suppl.) **1**, 323–328.

73. Falk, C.T. and Rubinstein, P. (1987) Haplotype relative risks: an easy reliable way to construct a proper control sample for risk calculations. *Ann. Hum. Genet.* **51,** 227–233.
74. Khoury, M.J. (1994) Case-parental control method in the search for disease-susceptibility genes. *Am. J. Hum. Genet.* **55,** 414–415.
75. Schaid, D.J. and Sommer, S.S. (1994) Comparison of statistics for candidate-gene association studies using cases and parents. *Am. J. Hum. Genet.* **55,** 402–409.
76. Thomson, G. (1995) Mapping disease genes: family-based association studies. *Am. J. Hum. Genet.* **57,** 487–498.
77. Knapp, M., Seuchter, S.A., and Baur, M.P. (1993) The haplotype-relative risk (HRR) method for analysis of association in nuclear families. *Am. J. Hum. Genet.* **52,** 1085–1093.
78. Ott, J. (1989) Statistical properties of the haplotype relative risk. *Genet. Epidemiol.* **6,** 127–130.
79. Terwilliger, J.D. and Ott, J. (1992) A haplotype based 'haplotype relative risk' approach to detecting allelic associations. *Hum. Hered.* **42,** 337–346.
80. Schaid, D.J. and Sommer, S.S. (1993) Genotype relative risks: methods for design and analysis of candidate-gene association studies. *Am. J. Hum. Genet.* **53,** 1114–1126.
81. Spielman, R.S. and Ewens, W.J. (1993) Transmission/disequilibrium test (TDT) for linkage and linkage disequilibrium between disease and marker. *Am. J. Hum. Genet.* (suppl.) **53,** 863.
82. Spielman, R.S., McGinnis, R.E., and Ewens, W.J. (1993) Transmission test for linkage disequilibrium: the insulin gene region and insulin-dependent diabetes mellitus (IDDM). *Am. J. Hum. Genet.* **52,** 506–516.
83. Tiret, L., Nicaud, V., Ehnholm, C., Havekes, L., Menzel, H.J., Ducimetiere, P., et al. (1993) Inference of the strength of genotype-disease association from studies comparing off-spring with and without parental history of disease. *Ann. Hum. Genet.* **57,** 141–149.
84. Hodge, S.E. (1993) Linkage analysis versus association analysis: distinguishing between two models that explain disease-marker associations. *Am. J. Hum. Genet.* **53,** 367–384.
85. Hodge, S.E. (1994) Reply to Suarez and Hampe and Spielman et al.: cosegregation, association, and linkage. *Am. J. Hum. Genet.* **54,** 560–563.
86. Suarez, B.K. and Hampe, C.L. (1994) Linkage and association. *Am. J. Hum. Genet.* **54,** 554–559.

87. Klitz, W., Thomson, G., Borot, N., and Cambon-Thomsen, A. (1992) Evolutionary and population perspective of the human HLA complex, in *Evolutionary Biology* (vol. 26), (Hecht, M. K., ed.), Plenum, New York, NY, pp. 35–72.
88. Field, L.L. (1989) Genes predisposing to IDDM in multiplex families. *Genet. Epidemiol.* **6,** 101–106.
89. Thomson, G., Robinson, W.P., Kuhner, M.K., Joe, S., and Klitz, W. (1989) HLA and insulin gene associations with IDDM. *Genet. Epidemiol.* **6,** 155–160.
90. Spielman, R.S., Baur, M.P., and Clerget-Darpoux, F. (1989) Genetic analysis of IDDM: summary of GAW 5 IDDM results. *Genet. Epidemiol.* **6,** 43–58.
91. Bennet, S.T., Lucassen, A.M., Gough, S. C., Powell, E.E., Undlien, D.E., Pritchard, L.E., et al. (1995) Susceptibility to human type diabetes at IDDM2 is determined by tandem repeat variation at the insulin gene minisatellite locus. *Nat. Genet.* **9,** 284–292.
92. Merriman, T., Twells, R., Merriman, M., Eaves, I., Cox, R., Cucca, F., et al. (1997) Evidence by allelic association-dependent methods for a type I diabetes polygene (IDDM6) on chromosome 18q21. *Hum. Mol. Genet.* **6,** 1003–1010.
93. Boehnke, M. and Langefeld, C.D. (1998) Genetic association mapping based on discordant sib pairs: the discordant-alleles test. *Am. J. Hum. Genet.* **61,** 319–333.
94. Curtis, D. (1997) Use of siblings as controls in case-control association studies. *Ann. Hum. Genet.* **61**, 319–333.
95. Martin, E.R., Kaplan, N.I., and Weir, B.S. (1997) Tests for linkage and association in nuclear families. *Am. J. Hum. Genet.* **61,** 439–448.
96. Sham, P.C. and Curtis, D. (1995) An extended transmission/disequilibrium test (TDT) for multiallelic marker loci. *Ann. Hum. Genet.* **59,** 323–336.
97. Spielman, R.S. and Ewens, W.J. (1996) The TDT and other family-based tests for linkage disequilibirum and association (Editorial). *Am. J. Hum. Genet.* **59,** 983–989.
98. Spielman, R.S. and Ewens, W.J. (1998) A sibship test for linkage in the presence of association: the sib transmission/disequilibrium test. *Am. J. Hum. Genet.* **62,** 450–458.
99. Seltman, H., Roeder, K., and Devlin, B. (2001) Transmission/Disequilibrium Test meets measured haplotype analysis: family-based association analysis guided by evolution of haplotypes. *Am. J. Hum. Genet.* **68,** 1250–1263.

100. Templeton, A.R. (1995) A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping of DNA sequencing. V. Analysis of case/control sampling designs: Alzheimer's disease and the apoprotein E locus. *Genetics* **140,** 403–409.
101. Gordon, D., Heath, S.C., Liu, X., and Ott, J. (2001) A transmission/disequilibrium test that allows for genotyping errors in the analysis of single-nucleotide polymorphism data. *Am. J. Hum. Genet.* **69,** 371–780.
102. Bacanu, S.-A., Devlin, B., and Roeder, K. (2000) The power of genomic control. *Am. J. Hum. Genet.* **66,** 1933–1944.
103. Clayton, D.G. and Jones, H. (1999) Transmission/disequilibrium tests for extended marker haplotypes. *Am. J. Hum. Genet.* **65,** 1161–1169.
104. Clayton, D.G. (1999) A generalization of the transmission/disequilibrium tests for uncertain-haplotype transmission. *Am. J. Hum. Genet.* **65,** 1170–1177.
105. Lazzeroni, L.C. and Lange, K. (1998) A conditional inference framework for extending the transmission/disequilibrium test. *Hum. Hered.* **48,** 67–81.
106. Merriman, T.R., Eaves, I.A., Twells, R.C., Merriman, M.E., Danoy, P.A., Muxworthy, C.E., et al. (1998) Transmission of haplotypes of miscrosatellite markers rather than single marker alleles in the mapping of a putative type 1 diabetes susceptibility gene (IDDM6). *Hum. Mol. Genet.* **7,** 517–524.
107. Rabinowitz, D. and Laird, N. (2000) A unified approach to adjusting association tests for population admixture with arbitrary pedigree structure and arbitrary missing marker information. *Hum. Hered.* **50,** 211–223.
108. Zhao, H., Zhang, S., Merikangas, K.R., Trixter, M., Wildenauer, D.B., Sun, F., et al. (2000) Transmission/disequilibrium tests using multiple tightly linked markers. *Am. J. Hum. Genet.* **67,** 936–946.
109. Jin, K., Speed, T., Klitz, W., and Thomson, G. (1994) A statistical test of segregation distortion. *Biometrics* **50,** 1189–1198.
110. Maxwell, A.E. (1970) Comparing the classification of subjects by two independent judges. *Br. J. Psychiatry* **116,** 651–655.
111. Rothman, K.J. (1986) *Modern Epidemiology*, Little & Brown, Boston, MA.
112. Trembath, R.C., Clough, R.L., Rosbotham, J.L., Jones, A.B., Camp, R.D.R., Frodsham, A., et al. (1997) Identification of a major

susceptibility locus on chromosome 6p and evidence for further disease loci revealed by a two stage genome-wide search in psoriasis. *Hum. Mol. Genet.* **6,** 813–820.

113. George, V.T. and Elston, R.C. (1987) Testing the association between polymorphic markers and quantitative traits in pedigrees. *Genet. Epidemiol.* **4,** 193–201.
114. Allison, D.D. (1997) Transmission-disequilibrium tests for quantitative traits. *Am. J. Hum. Genet.* **60,** 676–690.
115. Rabinowitz, D. (1997) A transmission disequilibrium test for quantitative trait loci. *Hum. Hered.* **47,** 342–350.
116. Fulker, D.W., Cherny, S.S., Sham, P.C., and Hewitt, J.K. (1999) Combined linkage and association sib-pair analysis for quantitative traits. *Am. J. Hum. Genet.* **64,** 259–267.
117. Cardon, L.R. (2000) A sib-pair regression model of linkage disequilibrium for quantitative traits. *Hum. Hered.* **50,** 350–358.
118. Abecasis, G.R., Cardon, L.R., and Cookson, W.O.C. (2000) A general test of association for quantitative traits in nuclear families. *Am. J. Hum. Genet.* **66,** 279–292.
119. Brookes, A.J. (1999) The essence of SNPs. *Gene* **234,** 177–186.
120. Kwok, P.Y., Deng, Q., Zakeri, H., Taylor, S.L., and Nickerson, D.A. (1996) Increasing the information content of STS-based genome maps: identifying polymorphisms in mapped STSs. *Genomics* **31,** 123–126.
121. Wang, D.G., Fan, J.B., Siao, C.J., Berno, A., Young, P., Sapolsky, R., et al. (1988) Related articles, nucleotide large-scale identification, mapping, and genotyping of single-nucleotide polymorphisms in the human genome. *Science* **280(5366),** 1077–1082.
122. Lander, E.S. (1996) The new genomics: global views of biology. *Science* **274,** 536–539.
123. Jorde, L.B. (2000) Linkage disequilibrium and the search for complex disease genes. *Genome Res.* **10,** 1435–1444.
124. Kruglyak, L. (1999) Prospects for whole-genome linkage disequilibrium mapping of common disease genes. *Nat. Genet.* **22,** 139–144.
125. Risch, N. and Merikangas, K. (1996) Related Articles. The future of genetic studies of complex human diseases. *Science* **273(5281),** 1516–1517.
126. Terwilliger, J.D. and Weiss, K.M. (1998) Linkage disequilibrium mapping of complex disease: fantasy or reality? *Curr. Opin. Biotechnol.* **8,** 578–594.

127. Weiss, K.M. and Tewilliger, J.D. (2000) How many diseases does it take to map a gene with SNPs? *Nat. Genet.* **26,** 151–157.
128. Bodmer, W.F. (1986) Human genetics: the molecular challenge. *Cold Spring Harbor Symp. Quant. Biol.* **51,** 1–13.
129. Collins, A., Lonjou, C., and Morton, N.E. (1999) Genetic epidemiology of single-nucleotide polymorphisms. *PNAS* **95,** 15,173–15,177.
130. Laan, M. and Pääbo, S. (1997) Demographic history and linkage disequilibrium in human populations. *Nat. Genet.* **17,** 435–438.
131. Peltonen, L. (2000) Positional cloning of disease genes: advantages of genetic isolates. *Hum. Hered.* **50,** 66–75.
132. Jorde, L.B., Watkins, W.S., Kere, J., Nyman, D., and Eriksson, A.W. (2000) Gene mapping in isolated populations: new roles for old friends? *Hum. Hered.* **50,** 57–65.
133. Lonjou, C., Collins, A., and Morton, N.E. (1999) Allelic association between marker loci. *PNAS* **96,** 1621–1626.
134. Ott, J. (2000) Predicting the range of linkage disequilibrium. *PNAS* **97,** 2–3.
135. Brzustowicz, L.M., Mérette, C., Xie, X., Townsend, L., Gilliam, T.C., and Ott, J. (1993) Molecular and statistical approaches to the detection and correction of errors in genotype databases. *Am. J. Hum. Genet.* **53,** 1137–1145.
136. Ehm, M.G., Kimmel, M., and Cottingham, R.W., Jr. (1996) Error detection for genetic data, using likelihood methods. *Am. J. Hum. Genet.* **58,** 225–234.
137. Lathrop, G.M., Hooper, A.B., Huntsman, J.W., and Ward, R.H. (1983) Evaluating pedigree data. 1. The estimation of pedigree error in the presence of marker mistyping. *Am. J. Hum. Genet.* **35,** 241–262.
138. Lincoln, S.E. and Lander, E.S. (1992) Systematic detection of errors in genetic linkage data. *Genomics* **14,** 604–610.
139. Sasse, G., Müller, H., Chakraborty, R., and Ott, J. (1994) Estimating the frequency of nonpaternity in Switzerland. *Hum. Hered.* **44,** 337–343.
140. Stringham, H.M. and Boehnke, M. (1996) Identifying marker typing incompatibilities in linkage analysis. *Am. J. Hum. Genet.* **59,** 946–950.
141. Gordon, D., Heath, S., and Ott, J. (1999) True pedigree errors more frequent than apparent errors for single nucleotide polymorphisms. *Hum. Hered.* **49,** 65–70.

142. Gordon, D., Leal, S.M., Heath, S.C., and Ott, J. (2000) An analytic solution to single nucleotide polymorphism error-detection rates in nuclear families: implications for study design. *Pacific Symposium on Biocomputing* **1,** 663–674.
143. Ott, J. (2001) Major strengths and weaknesses of the lod score method. *Adv. Genet.* **42,** 125–132.
144. Keats, B.J., Sherman, S.L., and Ott, J. (1990) Human gene mapping 10.5—report of the Committee on Linkage and Gene Order. *Cytogenet. Cell Genet.* **55,** 387–394.
145. Broman, K.W., Murray, J.C., Sheffield, V.C., White, R.L., and Weber, R.L. (1998) Comprehensive human genetic maps: individual and sex-specific variation in recombination. *Am. J. Hum. Genet.* **63,** 861–869.
146. Li, W., Fann, C., and Ott, J. (1998) Low-order polynomial trends of female-to-male map distance ratios along human chromosomes. *Hum. Hered.* **48,** 266–270.
147. Cleves, M.A. and Elston, R.C. (1997) Alternative test for linkage between two loci. *Genet. Epidemiol.* **14,** 117–131.
148. Halpern, J. and Whittemore, A.S. (1999) Multipoint linkage analysis. A cautionary note. *Hum. Hered.* **49,** 194–196.
149. Ott, J., Lehner, T., and Sandkuyl, L.A. (1988) Fallacies of the technique of "ahplotyping" two apparently linked loci in multipoint linkage analysis. *Am. J. Hum. Genet.* **43,** A154.
150. Daw, E.W., Thompson, E.A., and Wijsman, E.M. (2000) Bias in multipoint linkage analysis arising from map misspecification. *Genet. Epidemiol.* **19,** 336–380.
151. Friemer, N., Sandkuijl, L., and Blower, S. (1993) Incorrect specification of marker allele frequencies: effects on linkage analysis. *Am. J. Hum. Genet.* **52,** 1102–1110.
152. Ott, J. (1992) Strategies for characterizing highly polymorphic markers in human gene mapping. *Am. J. Hum. Genet.* **51,** 293–290.
153. Devlin, B., Risch, N., and Roeder, K. (1991) Estimation of allele frequencies for VNTR loci. *Am. J. Hum. Genet.* **48,** 662–676.
154. Lange, K. (1995) Applications of the Dirichlet distribution to forensic match probabilities. *Genetica* **96,** 107–117.
155. Lockwood, J.R., Roder, K., and Devlin, B. (2001) A Bayesian hierarchical model for allele frequencies. *Genet. Epidemiol.* **20,** 17–33.
156. Chapman, N.H. and Wijsman, E.M. (1998) Genome screen using linkage disequilibrium tests: optimal marker characteristics and feasibility. *Am. J. Hum. Genet.* **63,** 1872–1885.

157. Rothman, K. (1990) No adjustment are needed for multiple comparisons. *Epidemiology* **1,** 43–46.
158. Lander, E. and Kruglyak, L. (1995) Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. *Nat. Genet.* **11,** 241–247.
159. Witte, J.S., Elston, R.C., and Schork, N.S. (1996) Genetic dissection of complex traits (correspondence). *Nat. Genet.* **12,** 355–356.
160. Witte, J.S., Elston, R.C., and Cardon, L.R. (2000) On the relative sample size required for multiple comparisons. *Stat. Med.* **19,** 369–372.
161. Lucek, P., Hanke, J., Reich, J., Solla, S., and Ott, J. (1998) Multilocus nonparametric linkage analysis of complex trait loci with neural network. *Hum. Hered.* **48,** 275–284.
162. Hoh, J. and Ott, J. (2000) Scan statistics to scan markers for susceptibility genes. *PNAS* **17,** 9615–9617.

3

Genetic Markers in Psychiatric Genetics

Thomas Bourgeron and Bruno Giros

1. Introduction

Are the strategies for the identification of susceptibility genes in psychiatric diseases the same as those that have been used to successfully identify genes in monogenic diseases? How many cases, families, and markers are needed? These methodological questions are still a matter of debate for all complex traits, including psychiatric diseases. Furthermore, new technologies and methodologies for studying human genetics are emerging, and current protocols may rapidly become obsolete. This chapter presents an overview of the classical and novel gene identification strategies in complex diseases, and focuses on the problems encountered in psychiatric diseases. For each method, we discuss their relevance and what will probably be available in the near future. This review is written for clinicians close to the patients and molecular biologists close to the genes, who are not always familiar with the methods used to detect the link between individual genetic variations and the susceptibility to psychiatric diseases.

From: *Methods in Molecular Medicine, vol. 77: Psychiatric Genetics: Methods and Reviews*
Edited by: M. Leboyer and F. Bellivier

2. From Patients to Chromosomal Regions

Two complementary approaches have been used to identify genes in monogenic diseases: taking advantage of a chromosomal rearrangement segregating with the investigated phenotype, or using polymorphic markers to identify genomic regions that segregate with the phenotype (**Fig. 1**). More often, a combination of these two strategies has been extremely successful. In addition, alternative linkage analyses have also been developed in order to further address the problem of complex traits.

2.1. Chromosomal Rearrangements

Balanced chromosomal rearrangements such as reciprocal translocations and inversions associated with genetic disorders have been very useful for the positional cloning of numerous human disease genes. Translocations have been identified in several psychiatric diseases by standard or high-resolution karyotypes ***(1–5)***. Identification of the sequence disrupted by the translocation is usually performed by fluorescent *in situ* hybridization (FISH) with yeast artificial chromosome (YAC) or cosmid probes overlapping breakpoints. When the translocated genomic sequence is known, the next step is to identify the genes disrupted or close to the breakpoint.

Deletions ***(6–9)*** or duplications ***(10,11)*** of chromosomal regions may also be present in patients with psychiatric diseases. However, they are often less informative than translocations in the identification of genes because they usually involve larger chromosomal regions that contain a relatively high number of genes. Large deletions are identified using standard or high-resolution karyotypes. To detect smaller deletions, two methods are used: telomere probes and comparative genomic hybridization (CGH).

The first method focuses on telomeres, since chromosome ends may represent a major source of human pathologies ***(12)***. Pilot studies estimate that at least 6% of unexplained mental retardation may be the result of these relatively small telomere abnormalities ***(13)***. If true, then subtelomeric rearrangements could be the second most

common cause of mental retardation after Down's syndrome ***(12)***. Therefore, representative YAC probes for all human chromosome ends have been generated ***(14)***, which are on average several megabases (Mb) away from the actual chromosome end. These YAC probes are chromosome-specific, and do not contain telomeric and subtelomeric repeats that are shared by multiple chromosomes. The absence of a signal at a particular chromosome telomere is detected by classical FISH experiments ***(14)***.

The second approach, CGH, is a molecular cytogenetic method for the detection of chromosomal imbalances ***(15)***. In a CGH experiment, genomic DNA from the proband is labeled with one fluorochrome, whereas an equal amount of control DNA (derived from cells with a normal karyotype) is labeled with another fluorochrome. Then, differentially labeled control and proband genomic DNA are simultaneously hybridized *in situ* to normal metaphase chromosomes. The intensity ratio of the two fluorescence signals yields a measure of the copy number ratio between the two genomic DNA samples. For the objective identification of deletions or duplications, quantitative fluorescence digital image analysis is necessary. CGH allows a comprehensive analysis of multiple DNA gains and losses in entire genomes within a single experiment. New technologies will soon be available, and CGH will be performed on DNA chips rather than using metaphase chromosomes ***(16)***. Using this protocol, CGH will be specific for each gene of the human genome, and deleted or duplicated genes will be directly identified. To our knowledge, these methods have not yet been used in psychiatric diseases.

2.2. Linkage Analyses

For most patients, the susceptibility genetic variations are not detectable by investigating the proband's chromosomes with classic or molecular cytogenetic methods. Thus, a linkage approach is used, and families with several affected individuals are needed to identify genetic markers segregating with the phenotype (**Fig. 1**). In monogenic diseases, the mode of inheritance of the disease can be determined by looking at the segregation of the phenotype within

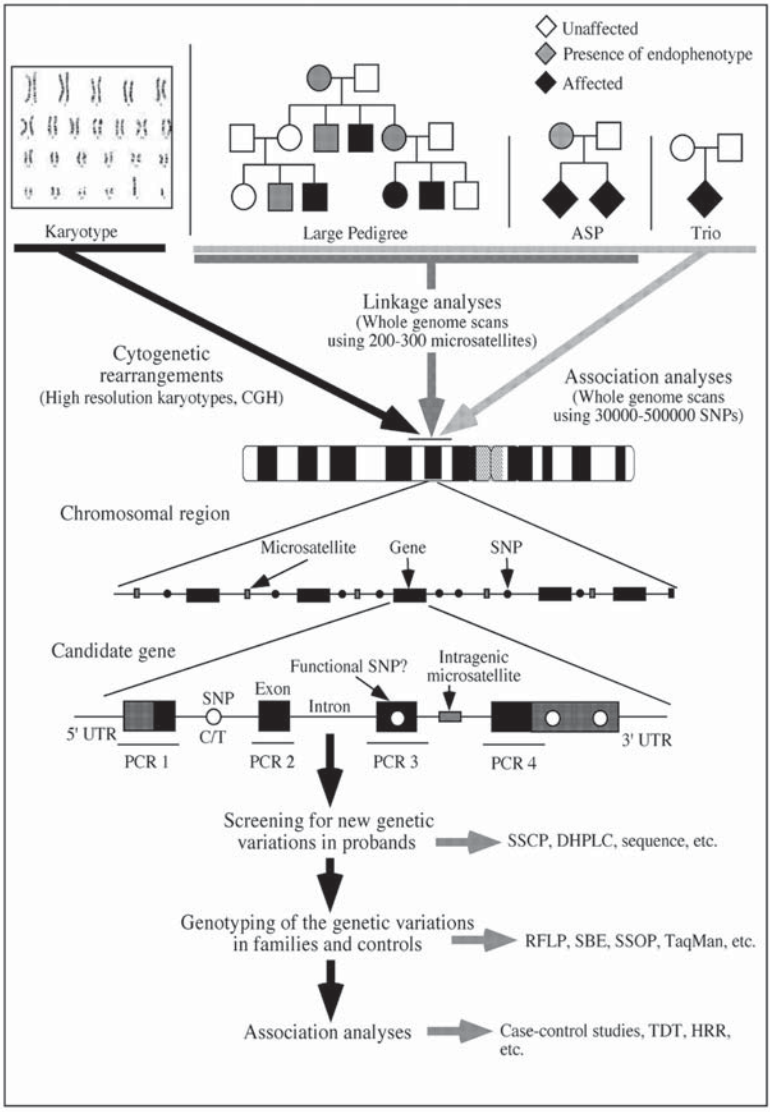

Fig. 1. Strategies to identify susceptibility genes in psychiatric diseases. Based on cytogenetic rearrangements, linkage or association studies, a chromosomal region is identified. In this region, a transcript map is generated, and potential candidate genes are identified. Each candidate gene is screened for new genetic variations in probands. Identified SNPs are genotyped in the families and controls and analyzed for their putative association or linkage disequilibrium with the disease.

the family. Therefore, parameters such as dominance, recessivity, allele frequency, and penetrance are available to perform linkage analyses. However, in the majority of psychiatric diseases, the number of genes involved, the interaction between them, the frequencies of the susceptibility alleles, and penetrance are unknown. Accordingly, non-parametric methods have been conceived to study linkage, in such complex diseases. Whatever the methods used, the constitution of the patients' collection and accurate phenotypic evaluation of the disorder probably represent the most challenging steps in psychiatric genetic studies.

2.2.1. Families

The collection of families is the first step in linkage studies. The quantity (number of patients and relatives) and the quality (clinical evaluation, phenotypes, and endophenotypes) of the samples are crucial for the following of the linkage analysis. Families (*see* **Note 1; Fig. 1**) can include a high number of affected subjects in one or more generations (large pedigrees), two or more affected sib-pairs (ASP), or one affected subject (trios). As discussed in several chapters of this book and in the literature, phenotypic evaluation is probably the most important step in the genetic analysis of complex disorders, which is the case for psychiatric disorders ***(17)***. One clinical challenge is therefore to detect endophenotypes, which may also be present in unaffected relatives, such as the blood serotonin level in autism ***(18)*** or the auditory-evoked response P50 in schizophrenia ***(19)***. Moreover, if these endophenotypes can be quantified, the statistical power of the linkage analyses will be greatly increased ***(20)***. The idea behind this specific or quantitative coding of phenotype is to detect isolated phenotypic traits, which are likely to involve a limited number of genes. As a result, linkage studies using these endophenotypes have more chance of success compared to those based on the broader phenotype involving too many gene interactions (epistasis). In this situation, the genotype of unaffected relatives is important and should be included in the analysis. Depending on the choice of sampling strategy (large pedigrees, ASP, or

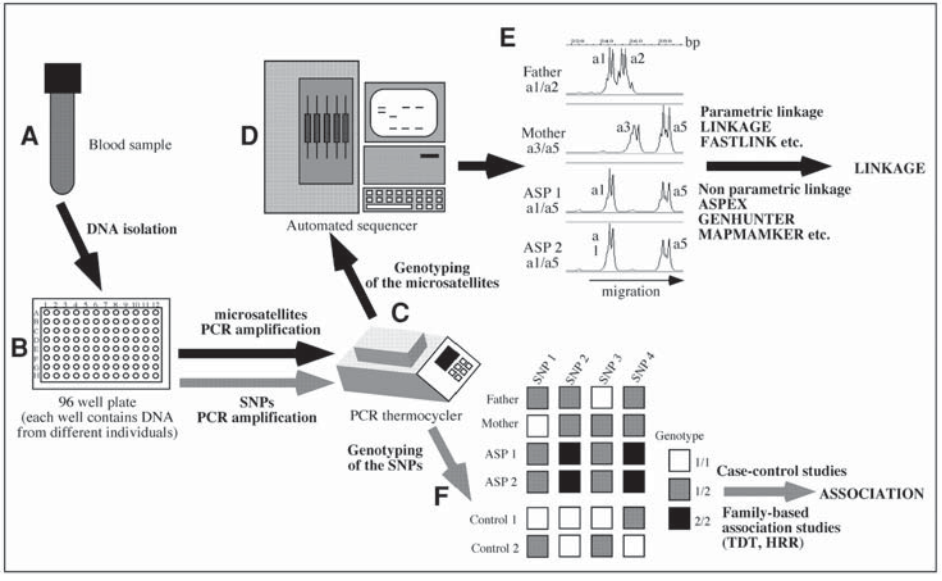

Fig. 2. Basic protocol for high-throughput genotyping of microsatellites and SNPs. **(A)** Genomic DNA is isolated from blood sample; **(B)** DNA from each proband, family members or controls is disposed in different wells of a microtiter plate. **(C)** The microsatellites or the fragments of DNA containing the SNPs are amplified. **(D)** Microsatellites are genotyped using an automated sequencer. **(E)** Different size of microsatellites indicates different alleles. The father is a1/a2, and transmits a1 to both his children. The mother is a3/a5, and transmits a5 to both her children. In this situation, the two ASP are IBD 2 sharing the same parental alleles. **(F)** SNPs are genotyped using various methods such as specific hybridization on oligonucleotide microarrays as described here. Allele 1 (homozygote 1/1) and allele 2 (homozygote 2/2) are detected by fluorescence, respectively. Heterozygotes 1/2 are detected by the presence of both fluorescent markers.

trios), the linkage analyses will differ, but the genetic markers used are the same: microsatellites or single-nucleotide polymorphisms (SNPs).

2.2.2. Genetic Markers

2.2.2.1. Microsatellites

These markers are polymorphic variations of di-, tri- or tetranucleotide repeats that are genotyped by polymerase chain reaction (PCR; **Fig. 2;** *see* **Note 2**). A complete genetic map of the human genome based on microsatellites is available, with all the experimental conditions needed to carry out the genotyping ***(21)***. Linkage studies use microsatellite markers because of their high frequencies in the human genome and their high polymorphic information content (PIC), which depends on the heterozygote frequency of the marker in the population and the number of alleles. Microsatellites are ideal markers to follow the segregation of a genomic region with its cognate phenotype or to calculate the allele-sharing in ASP. However, because of their relative high mutation rate (10^{-4} per generation), microsatellites are not the ideal markers to detect an association of one specific allele of a gene with the phenotype. For these studies, SNPs are more suitable markers.

2.2.2.2. Single-Nucleotide Polymorphisms (SNPs)

SNPs are single-base differences in the DNA sequence that can be observed between individuals in the population. The third-generation linkage map of the human genome is based on SNPs ***(22)***. A map of 1,42 million SNPs distributed throughout the human genome has been recently published, providing an average density of one SNP every 1900 bp ***(23)***. The reported collection is expected to comprise roughly 11–12% of human SNPs (**Table 1**). An estimated number of 60 000 SNPs fall within exons (coding and untranslated regions), and 85% of exons are within 5 kb of the nearest SNP. SNPs are binary (i.e., they have only two alleles), and thus are well-suited to automated, high-throughput genotyping. Further-

Table 1
Occurrence of SNPs in the Human Population and Their Representation in the Current Collection

Minimal allele frequency	Expected SNP number (million)	Expected SNP frequency (bp)	Expected % in collection
1%	11.0	290	11–12
5%	7.1	450	15–17
10%	5.3	600	18–20
20%	3.3	960	21–25
30%	2.0	1,570	23–27
40%	0.97	3,280	24–28

Adapted from **ref.** ***(22).***

more, in contrast to more mutable markers such as microsatellites, SNPs have a low rate of recurrent mutation (2×10^{-8} per bp per generation), making them stable landmarks of human history. These properties offer an advantage when performing haplotype-based association studies.

2.2.3. Whole-Genome Scans

Using microsatellites, the standard protocol for a whole-genome scan consists of genotyping one marker each 10 cM. The choice of microsatellites depends mainly on location, PIC, and ease of amplification by PCR. About 300 markers are needed to cover the entire genome. Additionally, specific candidate regions or genes can also be tested by intragenic (or close) markers. In practice, microsatellites are now mainly genotyped by automated sequencers using fluorescent primers (**Fig. 2**). Different fluorochromes and various sizes of PCR fragments allow the genotyping of many microsatellites in one run. Numerous softwares have been developed to collect the genotyping data, to transform the size of the PCR fragments in allele numbers, and to test the Mendelian inheritance of the markers in the pedigrees (*see* **Note 2**).

Using SNPs, the choice of "how many and where" is still a matter of debate ***(24–26)***. As discussed in the next section, SNPs can be

used to perform haplotype-based association studies (**Fig. 1**, **Fig. 2**). In such studies, the range of linkage disequilibrium (LD) between one specific SNP and the disease locus can vary from 30 to 500 kb ***(24,25)***. Thus, an estimation of the number of SNPs required for a genome scan can range from 30,000 ***(25)*** to 500,000 ***(24)***. One strategy, which provides at least a cost-effective solution, may be to test SNPs in or very near the candidate genes ***(26)***. SNPs with functional consequences may be given highest priority, but their PICs are usually lower, presumably as a result of selective pressure ***(27)*.** Additionally, a precise knowledge of the degree and pattern of fluctuation of recombination frequency in the human genome would allow us to choose the more adequate distribution and number of SNPs to test ***(28)***.

2.2.4. Statistical Analyses

Statistical methods to detect linkage involve both parametric (model-based) and non-parametric (model-free) analyses. These terms are used to distinguish between methods that do or do not require choosing a particular mode of inheritance for the trait phenotype under investigation.

2.2.4.1. Parametric Linkage

Usually, model-based linkage analyses are used when large pedigrees are available. The lod score (Z) is the logarithm of the odds that the loci are linked (with recombination fraction Θ) rather than unlinked (recombination fraction 0.5). A function of Θ, lod scores are calculated for a range of Θ values, and the maximum Z is estimated. If different families are available, lod scores can be added up across families. However, heterogeneity (different genes involved in different families) may drastically reduce the power to detect linkage, and unfortunately, this is expected in psychiatric diseases. Software that has been developed to implement these calculations includes LINKAGE or FASTLINK (**Table 2**). Most software packages require two input files: the data file and the pedigree file. The data file contains the information about the markers (number of

Table 2
Web-Based Resources Concerning the Human Genome Project and Linkage Analyses

Name	Address	Description
UK Human Genome Mapping Project Resource Centre	http://www.hgmp.mrc.ac.uk/	This site contains a very complete list of all the useful human genome web sites.
Ensembl (EMBL-EBI / Sanger Centre)	http://www.ensembl.org/	These sites are used to identify genes in a candidate region. Both sites maintain automatic annotation on the human genome.
Human Genome Project Working Draft at UCSC	http://genome.ucsc.edu/	
UniGene	http://www.ncbi.nlm.nih.gov/UniGene/	Each UniGene cluster contains sequences that represent a unique gene, as well as related information such as the tissue types in which the gene has been expressed and map location.
SNP database	http://www.ncbi.nlm.nih.gov/SNP/	Useful tool to find the SNPs in the human genome.
Web resources of genetic linkage analysis (Rockefeller University)	http://linkage.rockefeller.edu/	This site provides almost all software packages necessary for linkage and association analyses.
Cooperative Human Linkage Center (CHLC)	http://lpg.nci.nih.gov/CHLC/	Genetic maps showing the positions of genetic markers.

markers studied, names, number and frequency of alleles, and distance between markers) and the disease (dominant or recessive model, disease gene frequency, and penetrance). The pedigree file contains the sample names, sex, phenotypic status (affected or non-affected), and the genotyping data for each microsatellite, for each individual. As output, Z is calculated for each loci as a function of Θ. Because of the low prior probability that two randomly chosen loci should be linked, evidence yielding 1000:1 odds in favor of linkage (Z = 3) is required in order to yield overall 20:1 odds in favor of linkage. This corresponds to a posterior probability of Type I error (false-positive) that is equal to $p = 0.05$, the conventional threshold of statistical significance. To exclude a locus, a lod score of –2 is generally accepted.

In the case of complex diseases, recessive and dominant models are tested. The phenotype status of the individuals can also be modified—for example, using broader or narrow phenotypes or endophenotypes ***(29–31)***.

2.2.4.2. Non-Parametric Linkage

The model-free methods of linkage are commonly used in psychiatric diseases and were first derived for samples comprising sib-pairs ***(32)***, but can also use data on other relative pairs ***(33)***. The sib-pair analysis involves calculating the percentage of parental allele-sharing in ASP (**Fig. 3A**). The sharing status of each ASP is called identity by descent (IBD). For each tested locus of the genome, ASP can be either IBD 0 (no allele-sharing), IBD 1 (one allele, paternal or maternal, in common), or IBD 2 (ASP share the same paternal and maternal alleles). In the absence of linkage, the expected IBD distribution is IBD 0 25%, IBD 1 50%, and IBD 2 25%. The idea behind this test is that if both sibs are affected by a genetic disease, they will share the segment of chromosome carrying the disease locus. Thus, testing of a relatively high number of families with at least two ASP should allow the detection of markers and chromosomal regions with sharing above the level predicted by random segregation. The procedure of choice is to genotype the

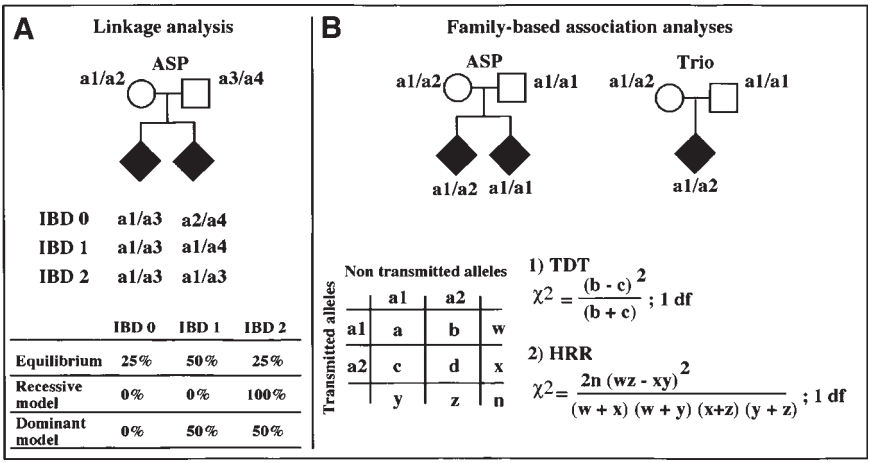

Fig. 3. The ASP method and family-based association studies. **(A)** Linkage analysis. When both parents are genotyped for one marker, ASP may share no allele (IBD0), one allele (IBD1), or two alleles (IBD2). In the absence of linkage, the IBD distribution is 25% IBD0, 50% IBD1, and 25% IBD2. In the presence of linkage, there is an excess of allele sharing. **(B)** Family-based association analyses. For both TDT and HRR, the number of transmitted and non-transmitted alleles is calculated. In TDT, only heterozygote parents are considered. In HRR, all parents are considered, and transmitted alleles represent the cases and non-transmitted alleles represent the HRR controls. For both tests, the statistical difference between transmitted and non-transmitted alleles is calculated using a χ^2 test with one degree of freedom.

ASP as well as their parents to ascertain the parental origin of each allele. However, if the parents' DNAs are not available, one alternative method is to calculate the allele-sharing using the identity by state (IBS). The main difference between these two approaches is that if two sibs have the same genotype for a given marker, their alleles are IBS, but they may or may not be IBD. Thus, in practice, microsatellites with multiple alleles are more efficient than two-allele markers such as SNPs for defining the IBD underlying the IBS. Softwares used to perform these calculations include SIBPAIR, GENEHUNTER, ASPEX, and Mapmaker/sibs (for a complete list *see* **ref.** ***34;*** **Table 2**). Most software programs generally use the

same data and pedigree files previously described for the parametric linkage analyses. A comparative analysis of these programs has been performed ***(34)*** and an estimated number of 400 ASP is needed to have >95% power to detect initially loci that increase risk by a factor of 2 ***(35,36)***. Non-parametric lod score (NPL) or maximum lod scores (MLS) are calculated for each single marker (single-point or two-point analyses) or for each region of the chromosomes using information from the flanking markers (multipoint analyses). Using the ASP method in a whole-genome scan, a lod = 2.2 ($p = 0.00074$) means suggestive linkage and a lod = 3.6 ($p = 0.00002$) means significant linkage ***(37)***, although the exact significance levels are still a topic of considerable debate.

3. From Chromosomal Regions to Candidate Genes

When a genomic region is identified, either by a chromosomal rearrangement or by linkage analysis, the next step is to identify the genes within that region and to choose the best candidates.

3.1. Candidate Genes in Candidate Regions

Until recently, this part of the work usually called "positional cloning" was considered to be a nightmare—especially for PhD students! In the chromosomal region studied, gene identification was performed by "brute force approach" techniques such as exon-trapping, cDNA selection, or high-throughput sequence of genomic clones in the genetic interval. Today, this part of the work has been greatly reduced and facilitated by the nearly complete human genome sequence ***(38,39)***. Therefore, gene identification is now performed by *in silico* database mining ***(40,41)***. Three major groups of genes are annotated in the human genome: genes already identified, which are now precisely localized and for which genomic structure (intron/exon organization) is known; "full-length mRNA," which correspond mainly to previously unidentified genes and originate from the alignment of different overlapping expressed sequence tags (ESTs); and "putative" genes, which are identified only *in silico* by software such as Genefinder or Grail ***(42)***. These softwares were

created to identify putative exons and introns in raw genomic sequences. Taken together, the estimated gene number (coding for protein) in the human genome is 30,000 ***(38,39)***. URLs of the main websites used to analyze the human genome are indicated in **Table 2**.

3.2. Candidate Genes in Functional Pathways

The candidate gene approach to psychiatric diseases has been widely used, but has not been successful thus far ***(43,44)***. The main reason for this is that, with some imagination, almost the entire genome can be more or less directly involved in the studied phenotype (i.e., in brain functions). Thus, one approach is to choose genes that are candidates in terms of their function and their localization. Information drawn from different sources (pharmacology, endophenotypes, or animal models) can transform any common gene into a functional candidate gene. For instance, genes involved in the dopaminergic ***(45)*** or the serotonergic ***(46)*** systems (including development, synthesis, metabolism, transport, and receptors) have been intensively studied in almost all psychiatric diseases or personality traits, such as schizophrenia, bipolar affective disorder, attention deficit-hyperactivity disorder, autism, panic disorder, and addiction. However, until now, genetic studies on the role of these neurotransmitter pathways in these phenotypes have been disappointing ***(47)***. Finally, animal models with behavioral phenotypes that may possess some face validity to a given phenotype have also motivated an intense screen of orthologous genes in patients with psychiatric diseases ***(48)***.

If there are no functional candidate genes in the studied genetic interval, one approach is to characterize the expression pattern of the "full-length mRNA" or the "putative genes." This can be achieved by *in silico* Northern blotting (**Table 2**), which takes advantage of the systematic screen of the mRNAs in different tissues such as ESTs and serial analysis of gene expression (SAGE). However, the data should be interpreted with caution and for most genes, the expression pattern must be well characterized at the bench using classical techniques such as Northern blot, RT-PCR, or *in situ* RNA hybridization.

4. From Candidate Genes to Susceptibility Genes

When the candidate gene has been selected, the following step is to test if the gene is associated with the disease. This analysis includes the search for genetic variation and association studies.

4.1. Screening for Mutation

This part of the work, together with the phenotypic and linkage analysis, represents the second limiting step. Although the sequence of the human genome is now almost fully available, it is still difficult to quickly and accurately identify a high number of sequence variations between different individuals. However, this problem will be greatly reduced as emerging technologies allow the identification of variations and the rapid genotyping of these variations.

Several techniques are commonly used to detect new sequence variations in patients compared to the control sequence (wild-type). In all protocols, the detection method is based on DNA fragments amplified by PCR. In practice, the first step is to define the genomic structure (introns/exons). This genomic organization may already be available in the databases (**Table 2**). Alternatively, it can be derived from the alignment of the cDNA with the genomic sequence. The coding regions of the candidate gene are then amplified by intronic primers in the 5' and 3' flanking region of each exon (**Fig. 1**). Four techniques are commonly used to detect new genomic variations: single-strand conformation polymorphism (SSCP), denaturing high-performance liquid chromatography (DHPLC), variant detector array (VDA), and direct sequencing.

4.1.1. Single-Strand Conformation Polymorphism (SSCP)

The principle of this method is that single-strand DNA has a tendency to fold up and form complex structures stabilized by weak intramolecular bonds. In most cases, especially for small DNA fragments (< 200 bp), the conformation depends on the nucleotide sequence. Experimentally, amplified DNA samples are denaturated by heat and loaded on a non-denaturing polyacrylamide gel.

Products can be radiolabeled (usually with ^{33}P) or silver-stained. The migration pattern of the new sequence is compared to the migration of the control sequence. When a difference in the DNA migration is detected, the sequence of the PCR product is needed to identify the nucleotide variation. This technique is simple and not prohibitively expensive, but the size of the DNA fragments should be less than 300 bp, and even with this condition, SSCP will detect only about 70–90% of the genetic variation. Examples of genetic variations identified by SSCP are described in Erdmann et al. ***(49)*** and Feng et al. ***(50).***

4.1.2. DHPLC

This method is currently the most rapid and efficient technique ***(51–54)***. Control and proband amplified DNA samples are mixed to create DNA heteroduplexes, which are then detected by the DHPLC. Whereas conventional heteroduplex analysis uses a polyacrylamide gel matrix to separate homo- and heteroduplex species in a non-denaturing environment, DHPLC uses partially denaturing conditions to amplify the difference between the two species. Thus, homo-duplexes (wt/wt and m/m) and heteroduplexes (wt/m) DNAs that are progressively and differently eluted from the column are detected using an ultraviolet detector. Although single-base mutations have been detected in 1.5-kb fragments, maximum sensitivity is achieved with fragments of 150–450 bp. The efficiency for detecting a mutation is close to 100%. The negative aspects are the expensive cost for the DHPLC apparatus and, as for SSCP, the need to sequence the variant samples to identify the nucleotide variation.

4.1.3. Variant Detector Array (VDA)

This technique allows the identification of SNPs by hybridization of a PCR product to oligonucleotides arrayed on a glass chip and measurement of difference in hybridization strength between matched and mismatched oligonucleotides. Recently, Halushka et al. used this technique to identify 874 SNPs in 75 candidate genes for hypertension ***(55)***.

4.1.4. Sequence

Direct sequencing of the PCR product is the ultimate technique to detect genetic variations between individuals. However, even with high-quality sequences, detection of heterozygotes from the sequence chromatography is still not as easy as it could be, and usually requires more time to analyze the data compared to the previous methods.

4.2. Genotyping of Identified SNPs

One alternative to the mutation screening mentioned here is to take advantage of the already identified SNP map, which may provide interesting variations in or near the coding region of the candidate gene. The previous methods (SSCP, DHPLC, and sequence) can be used to detect already identified SNPs, but several specific techniques more effectively genotype variations already identified ***(56)***. This is probably one of the most competitive areas for biotechnology companies, and it is likely that these methods will greatly progress in the near future. Fluorescent dye-based genotyping technologies are emerging as high-throughput genotyping platforms, including pyrosequencing ***(57)***, single-base extension with fluorescence detection ***(58)***, homogeneous solution hybridization such as TaqMan ***(59,60)***, and molecular beacon genotyping ***(61)***. Moreover, the Invader assay is now able to genotype directly from genomic DNA without PCR amplification ***(62)***.

Independently of the technology used, each method must be combined into an appropriate detection system allowing high-throughput genotyping. Several systems are available, such as fluorescent microarray-based systems (Affymetrix), fluorescent bead-based technologies (Luminex, Illumina, Q-dot), automated enzyme-linked immunosorbent assays (ELISAs) (Orchid Biocomputer), fluorescent detection of pyrophosphatase realease (Pyrosequencing), fluorescence resonance energy transfer (FRET)-based cleavase assay (Third Wave technologies), and mass spectroscopy detection techniques (Rapigene, Sequenom). In this competitive area, it is difficult to predict which protocols will be the most reliable and less

expensive. The economic aspect is crucial, since for SNP-based whole genome scans, at least 30,000 SNPs must be genotyped for each individual, and this represents millions of different tests. To reduce the number of assays, one interesting idea is to genotype equimolar pools of DNAs from various probands and to compare with corresponding pools of DNAs from different controls ***(51,63)***. However, this DNA pooling method requires precise quantification of samples and is not appropriate when haplotype construction is required.

4.2.1. Restriction Fragment-Length Polymorphism (RFLP)

RFLP was one of the first methods used in molecular genotyping. The principle of RFLP is to test the presence of a modified restriction site by enzymatic digestion. In order to use this test, the sequence variation must obliterate or create a restriction site in the new sequence. The presence of the variation will be tested by digestion of the PCR product (e.g., allele 1: digested, and allele 2: nondigested). If the variation does not modify a restriction site, one site can be created using specific oligonucleotides ***(64)***. This simple technique can be performed in any molecular biology laboratory, but is laborious for high-throughput genotyping.

4.2.2. Sequence-Specific Oligonucleotide Probe Hybridization (SSOP)

This technique involves annealing of a PCR product spanning the SNP to oligonucleotides complementary to each of the alternative SNP states and measuring relative hybridization efficiency ***(65)***. Using microarrays of oligonucleotides, this method was used to detect SNP in the mu opioid receptor gene ***(66)*** or to determine the distant history of SNPs in current human populations ***(67)***.

4.2.3. Molecular Beacon

Molecular beacon consists of a stem-loop DNA structure with a fluorophore and a quencher bound to the ends of the DNA probe. In

free solution, these probes are non-fluorescent because the stem hybrid keeps the fluorophore close to the quencher. When the probe sequence in the loop hybridizes to its target, forming a rigid double helix, a conformational reorganization occurs that separates the quencher from the fluorophore, restoring fluorescence. This technique was used for the detection of allelic differences in the human chemokine receptor 5 (CCR5) gene ***(68)***.

4.2.4. Single-Base Extension (SBE)

SBE or template-directed primer extension (TDI) measures the ability of a DNA polymerase to extend an oligonucleotide across the polymorphic site ***(58)***. This reaction is essentially a cycle sequencing reaction in which only dideoxynucleotides (terminators) are present. In the presence of proband DNA amplification product and the appropriate dideoxyribonucleoside triphosphate (ddNTP), the specific primer, close to the SNP (e.g., C/T), is extended by one base at the polymorphic site. By determining which ddNTP is incorporated (ddCTP or ddTTP), the alleles present in the target DNA can be inferred. Identification of the ddNTP is performed by using different fluorochromes for each ddNTP. This genotyping method has been widely used in different formats, and has proven to be highly sensitive and specific ***(69,70)***.

4.2.5. The 5' Nuclease Assay (TaqMan)

Lee et al. first demonstrated that the 5' nuclease assay could be used for allelic discrimination ***(59)***. In this assay, a hybridization probe included in the PCR is cleaved by the inherent 5' nuclease activity of Taq polymerase only if the probe is bound to the target sequence. The hybridization probe is co-labeled with one fluorescent reporter and one quencher dye. Thus, cleavage results in separation of both dyes and an increase in fluorescence emission intensity from the reporter dye. The TaqMan assay combines amplification and detection of specific PCR products in a single step. For allelic discrimination, two doubly labeled fluorescent probes are included in the reaction—one specific for each allele.

4.2.6. The Cleavase and the Invader Assay

Cleavase fragment-length polymorphism (CFLP) measures SNP status according to the ability of a cleavase enzyme to cut a matched or mismatched three-strand hybridization structure ***(71,72)***. The Invader assay is a PCR-independent methodology that uses a microtiter plate format. In the assay, a specific upstream Invader oligonucleotide and a downstream probe hybridize in tandem to a complementary DNA template and form a partially overlapping structure. The cleavase enzyme recognizes and cuts this structure to release the 5' flap of the probe. This flap then serves as an Invader oligonucleotide to direct cleavage of a fluorescence resonance energy transfer (FRET) probe in a second invasive cleavage reaction. Cleavage of this FRET probe generates a signal, which can be readily analyzed by fluorescence microtiter plate readers. This signal amplification permits identification of single-base changes directly from genomic DNA without prior amplification ***(62,73,74)***.

4.3. Linkage Disequilibrium and Association Studies

When SNPs are identified and genotyped in patients or in families, data must be analyzed to determine whether the association between the gene and the disease is significant ***(36)***. Just as linkage is a relationship between loci, association is a relationship between alleles. Thus, an alternative to linkage mapping in families is to search for statistical associations between one allele and the disease (**Fig. 1**, **Fig. 2**). The idea behind this approach is that if one allele m1 can directly cause susceptibility to the disease, generally, possession of m1 is generally not necessary or sufficient for someone to develop the disease, but its frequency should increase in the proband population compared to unaffected controls. The same should also be true for all alleles in LD with allele m1 (**Fig. 4**). Thus, SNP-based association studies can be performed in two ways: direct testing of a SNP with functional consequence for association, or using a SNP as a marker for LD. Therefore, a common core haplotype involving several SNPs should be more frequent in probands compared to controls.

LD mapping attempts to infer the location of a disease gene from observed associations between marker alleles and disease phenotype. This approach can be quite powerful when disease chromosomes are descended from a single founder mutation and the markers considered are tightly linked to the disease locus (**Fig. 4**). In addition to genetic heterogeneity (not the same genes involved in all families), the main problem for (LD) or association studies is allelic heterogeneity. In other words, susceptibility genes may be the same in different probands but mutations may be different in each individual or family. For linkage analyses, allelic heterogenity is not a problem because all families (independently of the mutation) will show linkage or sharing to the same chromosomal region. By contrast, for association studies, LD or association will be difficult to detect if allelic heterogeneity is present. For example, no association has been found between Duchenne muscular dystrophy (DMD) and alleles of any marker, however closely linked. Because of strong natural selection, the half-life of a DMD mutation is only two generations, and unrelated boys with DMD usually carry different independent mutations. On the other hand, striking results were obtained for diastrophic dysplasia, which was predicted to be 60 kb from the best marker, and the locus was indeed found at a distance of 70 kb ***(75)***. Finally, it appears that increasing population isolation is more or less correlated with increasing association. Thus, such studies have a better chance to succeed when performed in small, relatively isolated founder populations, where heterogeneity of the disease is less likely. Finally, population size, history, and structure should be considered before starting LD or association studies ***(76)***. To detect LD or association between one specific allele with the disease, two types of studies can be performed: case-control studies and family-based association studies.

4.3.1. Case-Control Studies

Case-control studies to detect association have been used extensively by epidemiologists, regardless of the cause of that association. This method compares the frequency of the susceptibility allele m1

in the proband population with its frequency in the control population. These studies have been strongly criticized for several reasons. The criticisms concern the sometimes non-rigorous statistical methods used to detect association, but the main problem concerns the selection of controls. Indeed, it is difficult to distinguish between LD or association and a difference in the case and control populations stratification. Lander and Schork provided a good example to illustrate the population stratification problem: The chopstick gene *(77)*. They found a significant association in the San Francisco Bay area between HLA-A1 haplotype and the ability to eat with chopsticks. The obvious bias in the study was that HLA-A1 is more frequent among Chinese than among Caucasians.

Despite this obvious pitfall, case-control studies cannot be rejected systematically and may still be highly useful, especially when parents are not available to conduct family-based association studies, as is the case for several psychiatric diseases with a relatively late age of onset. One approach, called the "genomic control" method (GC), was developed to ascertain the level of stratification between proband and control populations *(78)*. This method has the robustness of family-based designs, although it uses population-based data. GC uses the genome itself as a "control" to determine appropriate corrections for population-based association tests. In practice, allelic distribution in different regions of the genome (supposedly not involved in the disease) are analyzed in cases and in controls, in order to detect an eventual difference of these "control" genomic regions in the two populations. The GC method may be more powerful than family-based association studies, and when a disease becomes more prevalent, the discrepancy in power becomes more extreme *(78)*. When population substructure is present, however, the results are still more complex. Economically, GC is at least comparable to and often less expensive than family-based methods. Therefore, GC methods should prove a useful complement to family-based methods for the genetic analysis of complex traits.

4.3.2. Family-Based Association Studies

In order to avoid the problems related to the choice of the control group in association studies, several methods have been developed which use internal controls. These methods are based on the transmission rather than on the frequency of the studied allele (**Fig. 2B**). This chapter focuses on two family-based methods to detect the association between one gene and the disease: the transmission disequilibrium test (TDT) and the haplotype relative risk method (HRR).

4.3.2.1. Transmission Disequilibrium Test (TDT)

This test uses families with one or more affected offspring, in which at least one parent is heterozygous at the marker locus ***(79,80)***. Using only heterozygous parents and a bi-allelic marker (a1 and a2), the number of transmission (b) and no transmission (c) of the allele a1 to the affected offspring is calculated (**Fig. 3B**). Thus, in the absence of LD or association between a1 and the disease gene, the expected number of transmission should be equal to the number of non-transmission ($b = c$). Alternatively, if there is LD between the marker allele m1 and the disease, b will be higher than c ($b > c$). The statistical difference between b and c is calculated using the McNemar's test $(b - c)^2 / (b + c)$ and a P value can be obtained using a χ^2 distribution with one degree of freedom (df). Although designed as a test of linkage, the TDT is also valid as a test of association in simplex families ***(80,81)***. This test can also include unaffected relatives to rule out the presence of a segregation distortion (meiotic drive) of the chromosomal region independently of the phenotype studied.

4.3.2.2. Haplotype Relative Risk Method (HRR)

In the HRR test ***(82)***, all parents (heterozygous and homozygous) are used, all the transmitted alleles are considered as the case group, and all the non-transmitted alleles are considered as the internal HRR "control" group (**Fig. 3B**). The statistical difference between

cases (transmitted alleles) and HRR controls (non-transmitted alleles) is calculated as described in **Fig. 3B**, and a *P* value can be obtained using a χ^2 distribution with one df.

4.4. Functional Studies

The ultimate proof to ascertain the association of one genetic variation with the disease is to study the functional role of this variation and to determine the consequence of the modification at the level of the behavioral phenotype. In psychiatric diseases, this task is obviously very difficult, but relies heavily upon our knowledge of fundamental neuroscience and on the possibility generating mice models with engineered genetic modifications. Depending on the gene, different methods can be used to study in vitro the functional role of genetic variation ***(83–87)***. Functional polymorphism may be located in the non-coding ***(83)*** or coding ***(86,87)*** region of the gene. For instance, a polymorphic tandemly repeated sequence in the 5' flanking region of the human serotonin transporter gene modifies the promoter activity ***(83)***, thus affecting transcription and serotonin uptake in blood platelets ***(84,85)*** depending on the size of the repeat. Another example is the catechol-O-methyltransferase (COMT), which metabolises catecholamines such as dopamine, noradrenaline and adrenaline, and exists as common high and low activity alleles in the population ***(86)***. These functional studies are crucial for association studies and will surely expand along with the systematic screen for genetic variation in patients with psychiatric diseases.

Additionally, animal models can be used to evaluate the role of candidate genes and the functional consequences of different allelic forms of the gene. As reviewed by Tarantino and Bucan, a constant effort is being dedicated to the screening of mouse mutant for specific behavioral phenotypes such as sensorimotor gating, learning and memory, anxiety, and sleep and circadian rhythms ***(48)***. Furthermore, the use of conditional or point mutations instead of the complete knockout of the gene make these approaches promising. One example is the knockout mouse for the NMDAR1 (NR1)-receptor

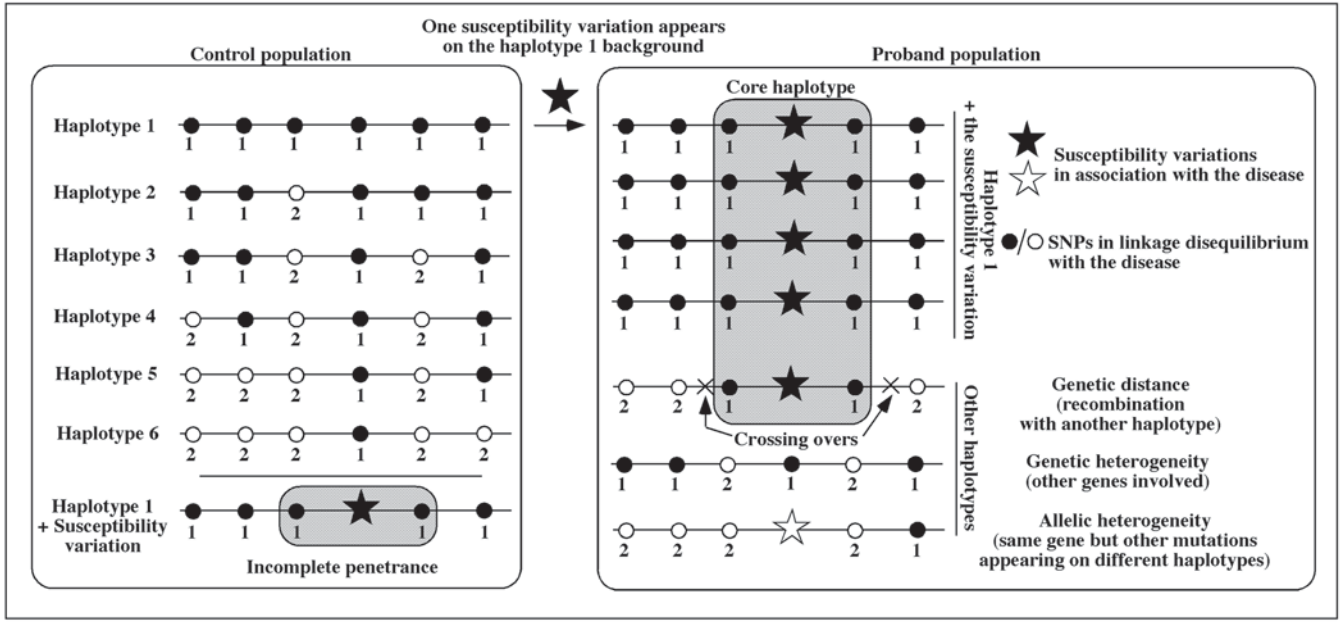

Fig. 4. Theoretical haplotype distribution in control and proband populations. The susceptibility variations in association with the disease (black and white stars) and allele 1 of the closest SNPs are more frequent in the proband population. However, genetic distance between the SNPs and the susceptibility variation, heterogeneity (different genes or alleles involved), and incomplete penetrance may reduce the difference of this core haplotype distribution between the two populations.

subunit. Whereas complete loss of function of this gene is associated with lethality, mice that express only 5% of normal levels of the essential NR1 subunit survive to adulthood and display behavioral abnormalities, including increased motor activity and stereotypy, as well as deficits in social and sexual interactions. These behavioral alterations are similar to those observed in pharmacologically induced animal models of schizophrenia and can be alleviated by treatment with haloperidol or clozapine, antipsychotic drugs widely used in the clinic, which antagonize dopaminergic and serotonergic receptors ***(88)***.

5. From Susceptibility Genes to Patients

Because of the complexity of the genetic susceptibility to psychiatric diseases (including the number of genes, heterogeneity, and unknown penetrance), diagnosing or predicting the disease by examining the genes may seem impossible. However, genotype information may already be used to understand the inter-individual sensitivity of differential drug response in psychiatric patients. In this aspect, genetically controlled levels of drug metabolism ***(89)*** and individual variation in genetic vulnerability to the disease may be important information ***(90)***. For instance, a polymorphism in cytochrome P450 CYP2D enzyme accounts for the adverse response to tricyclic antidepressant observed in certain patients ***(91)***. Recently, it was reported that clozapine response in schizophrenic patients can be predicted with about 75% success on the basis of typing six polymorphisms in four genes: two serotonin receptors, 5-HT2A and 5-HT2C, the serotonin transporter, and the histamine H2 receptor ***(92)***.

A second aspect is to develop new molecules based on the functional understanding of the products encoded by susceptibility genes. Thus, targeting the individual susceptibility to the disease may reduce the secondary effects of the drugs that are used today. Depending on the susceptibility gene products (such as receptors, hormones, and enzymes involved in neurotransmitter biosynthesis), these new therapeutic molecules could be specific transcriptional activators or repressors, agonists, or antagonists.

6. Perspectives

This chapter does not mention the fashionable chips technology used to analyze gene expression. The microarray technology is obviously an attractive method to efficiently identify individual differences in gene expression. This approach may also be one alternative to the genetic approaches (linkage or association analyses) to identify directly differences in the gene-expression pattern between probands and controls. However, at least three limiting steps of this technology are still present: (i) the relatively high degree of variation in gene transcription between individuals (independently of the susceptibility genes); (ii) the relative difficulty in achieving access to brain tissues from probands (the best "candidate organ" for psychiatric diseases); and (iii) susceptibility genes may not alter the gene expression pattern in probands.

However, one exciting perspective offered by chips and other high-throughput technologies is its ability to analyze the "complete" genetic diversity of each individual and to consider different genotypes at different loci. Even if this is a very long-term goal, this comparison of the overall genetic diversity of each individual would allow the detection of interacting susceptibility genes involved in psychiatric diseases—a promising future.

In conclusion, as described in this chapter, methods and protocols used to identify and to test genetic markers in complex diseases are growing very rapidly. Therefore, this review should be considered a methodological introduction to the field, rather than a definitive guideline. It is critical to stay informed of the new possibilities available in psychiatric genetics and more generally in the field of complex traits.

7. Notes

1. Blood samples should be obtained from family members in order to isolate DNA and to generate B lymphoblastoid cell lines (BLCL). BLCL are Epstein-Barr virus (EBV) transformed lymphocytes, and therefore represent an unlimited source of DNA. Approximately 0.3–0.6 mg of DNA is obtained from 10 mL of blood.

2. Using proband DNA, each microsatellite is amplified by PCR with one fluorescent primer. Usually, 30 ng of DNA are used for each PCR. A pool of the different amplified microsatellites is then loaded in the gel or in the column of the automated sequencer. Alternatively, several microsatellites can be amplified simultaneously (multiplex PCR) and loaded directly into the gel. As output, the image of the gel shows the various PCR fragments and software programs used to analyze the data (**Fig. 2**). Precise protocols to collect the data depend on the apparatus used and are given by the manufacturer (Applied Biosystems, Amersham, and others).

References

1. Ishikawabrush, Y., Powell, J.F., Bolton, P., Miller, A.P., Francis, F., Willard, H.F., et al. (1997) Autism and multiple exostoses associated with an X-8 translocation occurring within the GRPR gene and 3' to the Sdc2 gene. *Hum. Mol. Genet.* **6,** 1241–1250.
2. Gecz, J., Barnett, S., Liu, J.J., Hollway, G., Donnelly, A., Eyre, H., et al. (1999) Characterization of the human glutamate receptor subunit 3 gene (GRIA3), a candidate for bipolar disorder and nonspecific X-linked mental retardation. *Genomics* **62,** 356–368.
3. Vincent, J.B., Herbrick, J.A., Gurling, H.M.D., Bolton, P.F., Roberts, W., and Scherer, S.W. (2000) Identification of a novel gene on chromosomes 7q31 that is interrupted by a translocation breakpoint in an autistic individual. *Am. J. Hum. Genet.* **67,** 510–514.
4. Millar, J.K., Wilson-Annan, J.C., Anderson, S., Christie, S., Taylor, M.S., Semple, C., et al. (2000) Disruption of two novel genes by a translocation co-segregating with schizophrenia. *Hum. Mol. Genet.* **9,** 1415–1423.
5. Petek, E., Windpassinger, C., Vincent, J.B., Cheung, J., Boright, A.P., Scherer, S.W., et al. (2001) Disruption of a novel gene (immp21) by a breakpoint in 7q31 associated with tourette syndrome. *Am. J. Hum. Genet.* **68,** 848–858.
6. Papolos, D.F., Faedda, G.L., Veit, S., Goldberg, R., Morrow, B., Kucherlapati, R., et al. (1996) Bipolar spectrum disorders in patients diagnosed with velo-cardio-facial syndrome—does a hemizygous deletion of chromosome 22q11 result in bipolar affective disorder. *Am. J. Psychiatry* **153,** 1541–1547.
7. Bassett, A.S., Hodgkinson, K., Chow, E.W.C., Correia, S., Scutt,

L.E., and Weksberg, R. (1998) 22q11 deletion syndrome in adults with schizophrenia. *Am. J. Med. Genet.* **81,** 328–337.

8. Thomas, N.S., Sharp, A.J., Browne, C.E., Skuse, D., Hardie, C., and Dennis, N.R. (1999) Xp deletions associated with autism in three females. *Hum. Genet.* **104,** 43–48.
9. Smith, M., Filipek, P.A., Wu, C., Bocian, M.,Hakim, S., Modahl, C., et al. (2000) Analysis of a 1-megabase deletion in 15q22-q23 in an autistic patient: identification of candidate genes for autism and of homologous DNA segments in 15q22-q23 and 15q11-q13. *Am. J. Med. Genet.* **96,** 765–770.
10. McGillivray, B.C., Bassett, A.S., Langlois, S., Pantzar, T., and Wood, S. (1990) Familial 5q11.2-q13.3 segmental duplication cosegregating with multiple anomalies, including schizophrenia. *Am. J. Med. Genet.* **35,** 10–13.
11. Repetto, G.R., White, L.M., Bader, P.J., Johnson, D., and Knoll, J.H.M. (1998) Interstitial duplications of chromosome region 15q11q13—clinical and molecular characterization. *Am. J. Med. Genet.* **79,** 82–89.
12. Knight, S.J.L. and Flint, J. (2000) Perfect endings: a review of subtelomeric probes and their use in clinical diagnosis. *J. Med. Genet.* **37,** 401–409.
13. Flint, J., Wilkie, A.O.M., Buckle, V.J., Winter, R.M., Holland, A.J., and McDermid, H.E. (1995) The detection of subtelomeric chromosomal rearrangements in idiopathic mental retardation. *Nat. Genet.* **9,** 132–140.
14. Knight, S.J.L., Regan, R., Nicod, A., Horsley, S.W., Kearney, L., Homfray, T., et al. (1999) Subtle chromosomal rearrangements in children with unexplained mental retardation. *Lancet* **354,** 1676–1681.
15. Tanner, M.M., Tirkkonen, M., Kallioniemi, A., Isola, J., Kuukasjarvi, T., Collins, C., et al. (1996) Independent amplification and frequent co-amplification of three nonsyntenic regions on the long arm of chromosome 20 in human breast cancer. *Cancer Res.* **56,** 3441–3445.
16. Pinkel, L., Seagraves, R., Sudar, D., Clark, S., Poole, I., Kowbel, D., et al. (1998) High resolution analysis of DNA copy number variation using comparative genomic hybridization to microarrays. *Nat. Genet.* **20,** 207–211.
17. Leboyer, M., Bellivier, F., Nostenbertrand, M., Jouvent, R., Pauls, D., and Mallet, J. (1998) Psychiatric genetics—search for phenotypes. *Trends Neurosci.* **21,** 102–105.

18. Leboyer, M., Philippe, A., Bouvard, M., Guilloud-Bataille, M., Bondoux, D., Tabuteau, F., et al. (1999) Whole blood serotonin and plasma beta-endorphin in autistic probands and their first-degree relatives. *Biol. Psychiatry* **45,** 158–163.
19. Freedman, R., Coon, H., Myles worsley, M., Orrurtreger, A., Olincy, A., Davis, A., et al. (1997) Linkage of a neurophysiological deficit in schizophrenia to a chromosome 15 locus. *Proc. Natl. Acad. Sci. USA* **94,** 587–592.
20. Brzustowicz, L.M., Honer, W.G., Chow, E.W.C., Hogan, J., Hodgkinson, K., and Bassett, A.S. (1997) Use of a quantitative trait to map a locus associated with severity of positive symptoms in familial schizophrenia to chromosome 6p. *Am. J. Hum. Genet.* **61,** 1388–1396.
21. Dib, C., Faure, S., Fizames, C., Samson, D., Drouot, N., Vignal, A., et al. (1996) A comprehensive genetic map of the human genome based on 5,264 microsatellites. *Nature* **380,** 152–154.
22. Kruglyak, L. and Nickerson, D.A. (2001) Variation is the spice of life. *Nat. Genet.* **27,** 234–236.
23. Sachidanandam, R., Weissman, D., Schmidt, S.C., Kakol, J.M., Stein, L.D., Mullikin, J.C., et al. (2001) A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms. *Nature* **409,** 928–933.
24. Kruglyak, L. (1999) Prospects for whole-genome linkage disequilibrium mapping of common disease genes. *Nat. Genet.* **22,** 139–144.
25. Collins, A., Lonjou, C., and Morton, N.E. (1999) Genetic epidemiology of single-nucleotide polymorphisms. *Proc. Natl. Acad. Sci. USA* **96,** 15,173–15,177.
26. Ott, J. (2000) Predicting the range of linkage disequilibrium. *Proc. Natl. Acad. Sci. USA* **97,** 2–3.
27. Cargill, M., Altshuler, D., Ireland, J., Sklar, P., Ardlie, K., Patil, N., et al. (1999) Characterization of single-nucleotide polymorphisms in coding regions of human genes. *Nat. Genet.* **22,** 231–238.
28. Jeffreys, A.J., Ritchie, A., and Neumann, R. (2000) High resolution analysis of haplotype diversity and meiotic crossover in the human TAP2 recombination hotspot. *Hum. Mol. Genet.* **9,** 725–733.
29. Straub, R.E., Maclean, C.J., Oneill, F.A., Burke, J., Murphy, B., Duke, F., et al. (1995) A potential vulnerability locus for schizophrenia on chromosome 6p24-22—evidence of genetic heterogeneity. *Nat. Genet.* **11,** 287–293.

30. Adams, L.J., Mitchell, P.B., Fielder, S.L., Rosso, A., Donald, J.A., and Schofield, P.R. (1998) A susceptibility locus for bipolar affective disorder on chromosome 4q35. *Am. J. Hum. Genet.* **62,** 1084–1091.
31. Brzustowicz, L.M., Hodgkinson, K.A., Chow, E.W.C., Honer, W.G., and Bassett, A.S. (2000) Location of a major susceptibility locus for familiar schizophrenia on chromosome 1q21–q22. *Science* **288,** 678–682.
32. Penrose, L.S. (1935) The detection of automsomal linkage in data which consist of pairs of brother and sisters of unspecified parent age. *Ann. Eugen.* **6,** 133–138.
33. Olson, J.M. and Wijsman, E.M. (1993) Linkage between quantitative trait and marker loci: methods using all relative pairs. *Genet. Epidemiol.* **10,** 87–102.
34. Davis, S. and Weeks, D.E. (1997) Comparison of nonparametric statistics for detection of linkage in nuclear families—single-marker evaluation. *Am. J. Hum. Genet.* **61,** 1431–1444.
35. Hauser, E.R. (1998) Genetic linkage analysis of complex genetic traits by using affected sibling pairs. *Biometrics* **54,** 1238–1246.
36. Risch, N.J. (2000) Searching for genetic determinants in the new millennium. *Nature* **405,** 847–856.
37. Lander, E. and Kruglyak, L. (1995) Genetic dissection of complex traits—guidelines for interpreting and reporting linkage results. *Nat. Genet.* **11,** 241–247.
38. Lander, E.S., Linton, L.M., Birren, B., Nusbaum, C., Zody, M.C., Baldwin, J., et al. (2001) Initial sequencing and analysis of the human genome. *Nature* **409,** 860–921.
39. Venter, J.C., Adams, M.D., Myers, E.W., Li, P.W., Mural, R.J., Sutton, G.G., et al. (2001) The sequence of the human genome. *Science* **291,** 1304–1351.
40. Wolfsberg, T.G., McEntyre, J., and Schuler, G.D. (2001) Guide to the draft human genome. *Nature* **409,** 824–826.
41. Birney, E., Bateman, A., Clamp, M.E., and Hubbard, T.J. (2001) Mining the draft human genome. *Nature* **409,** 827–828.
42. Singh, G.B. (2000) Computational approaches for gene identification. *Methods Mol. Biol.* **132,** 351–364.
43. Stoltenberg, S.F. and Burmeister, M. (2000) Recent progress in psychiatric genetics—some hope but no hype. *Hum. Mol. Genet.* **9,** 927–935.
44. Evans, K.L., Muir, W.J., Blackwood, D.H., and Porteous, D.J. (2001) Nuts and bolts of psychiatric genetics: building on the Human Genome Project. *Trends Genet.* **17,** 35–40.

45. Paterson, A.D., Sunohara, G.A., and Kennedy, J.L. (1999) Dopamine D4 receptor gene: novelty or nonsense? *Neuropsychopharmacology* **21,** 3–16.
46. Veenstra-VanderWeele, J., Anderson, G.M., and Cook, E.H. (2000) Pharmacogenetics and the serotonin system: initial studies and future directions. *Eur. J. Pharmacol.* **410,** 165–181.
47. Collier, D.A., Curran, S., and Asherson, P. (2000) Mission: Not impossible? Candidate gene studies in child psychiatric disorders. *Mol. Psychiatry* **5,** 457–460.
48. Trantino, L.M. and Bucan, M. (2000) Dissection of behavior and psychiatric disorders using the mouse as a model. *Hum. Mol. Genet.* **9,** 953–965.
49. Erdmann, J., Nothen, M.M., Shimronabarbanell, D., Rietschel, M., Albus, M., Borrmann, M., et al. (1996) The human serotonin 7 (5-HT7) receptor gene—genomic organization and systematic mutation screening in schizophrenia and bipolar affective disorder. *Mol. Psychiatry* **1,** 392–397.
50. Feng, J.N., Sobell, J.L., Heston, L.L., Cook, E.H., Goldman, D., and Sommer, S.S. (1998) Scanning of the dopamine D1 and D5 receptor genes by Ref in neuropsychiatric patients reveals a novel missense change at a highly conserved amino acid. *Am. J. Med. Genet.* **81,** 172–178.
51. Hoogendoorn, B., Norton, N., Kirov, G., Williams, N., Hamshere, M.L., Spurlock, G., et al. (2000) Cheap, accurate and rapid allele frequency estimation of single nucleotide polymorphisms by primer extension and DHPLC in DNA pools. *Hum. Genet.* **107,** 488–493.
52. Spiegelman, J.I., Mindrinos, M.N., and Oefner, P.J. (2000) High-accuracy DNA sequence variation screening by DHPLC. *Biotechniques* **29,** 1084–1090.
53. Speight, G., Turic, D., Austin, J., Hoogendoorn, B., Cardno, A.G., Jones, L., et al. (2000) Comparative sequencing and association studies of aromatic L-amino acid decarboxylase in schizophrenia and bipolar disorder. *Mol. Psychiatry* **5,** 327–331.
54. Austin, J., Hoogendoorn, B., Buckland, P., Speight, G., Cardno, A., Bowen, T., et al. (2000) Comparative sequencing of the proneurotensin gene and association studies in schizophrenia. *Mol. Psychiatry* **5,** 208–212.
55. Halushka, M.K., Fan, J.B., Bentley, K., Hsie, L., Shen, N.P., Weder, A., et al. (1999) Patterns of single-nucleotide polymorphisms in

candidate genes for blood-pressure homeostasis. *Nat. Genet.* **22,** 239–247.

56. Shi, M.M. (2001) Enabling large-scale pharmacogenetic studies by high-throughput mutation detection and genotyping technologies. *Clin. Chem.* **47,** 164–172.
57. Ronaghi, M. (2001) Pyrosequencing sheds light on DNA sequencing. *Genome Res.* **11,** 3–11.
58. Syvanen, A.C., Aalto-Setala, K., Harju, L., Kontula, K., and Soderlund, H. (1990) A primer-guided nucleotide incorporation assay in the genotyping of apolipoprotein E. *Genomics* **8,** 684–692.
59. Lee, L.G., Connell, C.R., and Bloch, W. (1993) Allelic discrimination by nick-translation PCR with fluorogenic probes. *Nucleic Acids Res.* **21,** 3761–3766.
60. Livak, K.J. (1999) Allelic discrimination using fluorogenic probes and the 5' nuclease assay. *Genet. Anal.* **14,** 143–149.
61. Tapp, I., Malmberg, L., Rennel, E., Wik, M., and Syvanen, A.C. (2000) Homogeneous scoring of single-nucleotide polymorphisms: comparison of the 5'-nuclease TaqMan (R) assay and molecular beacon probes. *Biotechniques* **28,** 732–738.
62. Fors, L., Lieder, K.W., Vavra, S.H., and Kwiatkowski, R.W. (2000) Large-scale SNP scoring from unamplified genomic DNA. *Pharmacogenomics* **1,** 219–229.
63. Breen, G., Harold, D., Ralston, S., Shaw, D., and Clair, D.S. (2000) Determining SNP allele frequencies in DNA pools. *Biotechniques* **28,** 464–466.
64. Haliassos, A., Chomel, J.C., Grandjouan, S., Kruh, J., Kaplan, J.C., and Kitzis, A. (1989) Detection of minority point mutations by modified PCR technique: a new approach for a sensitive diagnosis of tumor-progression markers. *Nucleic Acids Res.* **17,** 8093–8099.
65. Hacia, J.G. (1999) Resequencing and mutational analysis using oligonucleotide microarrays. *Nat. Genet.* **21,** 42–47.
66. LaForge, K.S., Shick, V., Spangler, R., Proudnikov, D., Yuferov, V., Lysov, Y., et al. (2000) Detection of single nucleotide polymorphisms of the human mu opioid receptor gene by hybridization or single nucleotide extension on custom oligonucleotide gelpad microchips: potential in studies of addiction. *Am. J. Med. Genet.* **96,** 604–615.
67. Hacia, J.G., Fan, J.B., Ryder, O., Jin, L., Edgemon, K., Ghandour, G., et al. (1999) Determination of ancestral alleles for human single-nucleotide polymorphisms using high-density oligonucleotide arrays. *Nat. Genet.* **22,** 164–167.

68. Kostrikis, L.G., Tyagi, S., Mhlanga, M.M., Ho, D.D., and Kramer, F.R. (1998) Molecular beacons—spectral genotyping of human alleles. *Science* **279,** 1228–1229.
69. Nikiforov, T.T., Rendle, R.B., Goelet, P., Rogers, Y.H., Kotewicz, M.L., Anderson, S., et al. (1994) Genetic bit analysis—a solid phase method for typing single nucleotide polymorphisms, *Nucleic Acids Res.* **22,** 4167–4176.
70. Pastinen, T., Kurg, A., Metspalu, A., Peltonen, L., and Syvanen, A.C. (1997) Minisequencing—a specific tool for DNA analysis and diagnostics on oligonucleotide arrays. *Genome Res.* **7,** 606–614.
71. Rossetti, S., Englisch, S., Bresin, E., Pignatti, P.F., and Turco, A.E. (1997) Detection of mutations in human genes by a new rapid method—cleavage fragment length polymorphism analysis (CFLPA). *Mol. Cell. Probes* **11,** 155–160.
72. Oldenburg, M.C. and Siebert, M. (2000) New cleavase (R) fragment length polymorphism method improves the mutation detection assay. *Biotechniques* **28,** 351–357.
73. Ledford, M., Friedman, K.D., Hessner, M.J., Moehlenkamp, C., Williams, T.M., and Larson, R.S. (2000) A multi-site study for detection of the factor V (Leiden) mutation from genomic DNA using a homogeneous invader microtiter plate fluorescence resonance energy transfer (FRET) assay. *J. Mol. Diag.* **2,** 97–104.
74. Mein, C.A., Barratt, B.J., Dunn, M.G., Siegmund, T., Smith, A.N., Esposito, L., et al. (2000) Evaluation of single nucleotide polymorphism typing with invader on PCR amplicons and its automation. *Genome Res.* **10,** 330–343.
75. Hastbacka, J., de la Chapelle, A., Kaitila, I., Sistonen, P., Weaver, A., and Lander, E. (1992) Linkage disequilibrium mapping in isolated founder population diastrophic dysplasia in Finland. *Nat. Genet.* **2,** 204–211.
76. Laan, M. and Paabo, S. (1997) Demographic history and linkage disequilibrium in human populations. *Nat. Genet.* **17,** 435–438.
77. Lander, E.S. and Schork, N.J. (1994) Genetic dissection of complex traits. *Science* **265,** 2037–2048.
78. Bacanu, S.A., Devlin., B., and Roeder, K. (2000) The power of genomic control. *Am. J. Hum. Genet.* **66,** 1933–1944.
79. Spielman, R.S., McGinnis, R.E., and Ewens, W.J. (1993) Transmission test for linkage disequilibrium: the insulin gene region and insulin-dependent diabetes mellitus (IDDM). *Am. J. Hum. Genet.* **52,** 506–516.

80. Spielman, R.S. and Ewens, W.J. (1996) The TDT and other family-based tests for linkage disequilibrium and association. *Am. J. Hum. Genet.* **59,** 983–989.
81. Spielman, R.S. and Ewens, W.J. (1998) A sibship test for linkage in the presence of association—the sib transmission/disequilibrium test. *Am. J. Hum. Genet.* **62,** 450–458.
82. Terwilliger, J.D. and Ott, J. (1992) A haplotype-based 'haplotype relative risk' approach to detection allelic associations. *Hum. Hered.* **42,** 337–346.
83. Heils, A., Teufel, A., Petri, S., Stober, G., Riederer, P., Bengel, D., et al. (1996) Allelic variations of human serotonin transporter gene expression. *J. Neurochem.* **66,** 2621–2624.
84. Lesch, K.P., Bengel, D., Heils, A., Sabol, S.Z., Greenberg, B.D., Petri, S., et al. (1996) Association of anxiety-related traits with a polymorphism in the serotonin transporter gene regulatory region. *Science* **274,** 1527–1531.
85. Greenberg, B.D., Tolliver, T.J., Huang, S.J., Li, Q., Bengel, D., and Murphy, D.L. (1999) Genetic variation in the serotonin transporter promoter region affects serotonin uptake in human blood platelets. *Am. J. Med. Genet.* **88,** 83–87.
86. Lachman, H.M., Papolos, D.F., Saito, T., Yu, Y.M., Szumlanski, C.L., and Weinshilboum, R.M. (1996) Human catechol-O-methyltransferase pharmacogenetics—description of a functional polymorphism and its potential application to neuropsychiatric disorders. *Pharmacogenetics* **6,** 243–250.
87. Cravchik, A. and Gejman, P.V. (1999) Functional analysis of the human D-5 dopamine receptor missense and nonsense variants: differences in dopamine binding affinities. *Pharmacogenetics* **9,** 199–206.
88. Mohn, A.R., Gainetdinov, R.R., Caron, M.G., and Koller, B.H. (1999) Mice with reduced NMDA receptor expression display behaviors related to schizophrenia. *Cell* **98,** 427–436.
89. Wolf, C.R., Smith, G., and Smith, R.L. (2000) Science, medicine, and the future—pharmacogenetics. *Br. Med. J.* **320,** 987–990.
90. Roses, A.D. (2000) Pharmacogenetics and the practice of medicine. *Nature* **405,** 857–865.
91. Coutts, R.T. (1994) Polymorphism in the metabolism of drugs, including antidepressant drugs—comments on phenotyping. *J. Psychiatry* **19,** 30–44.

92. Arranz, M.J., Munro, J., Birkett, J., Bolonna, A., Mancama, D., Sodhi, M., et al. (2000) Pharmacogenetic prediction of clozapine response. *Lancet* **355,** 1615–1616.

4

Clinical Assessment in Psychiatric Genetics

Philip Gorwood

1. Introduction

The use of reliable instruments to collect relevant data in psychiatry is required before studying determinants of psychiatric disorders. Although genetic research in psychiatry devotes large sums of money to biological studies, the investment made in time and effort to describe the clinical features of the subjects being studied is often lacking. Variation of types and amount of clinical information available, psychopathological explanations of the information collected, and differences in definitions of disorders probably explain great discrepancies of results before the 70s.

The use of structured and semistructured interviews may increase reliability by controlling the amount and quality of the information collected. Sharing international diagnostic criteria—such as International Statistical Classification of Diseases (ICD) and Diagnostic and Statistical Manual (DSM) allows comparisons between the results of research performed in different settings, countries, and/or cultures. The quality of clinicians, importance of training programs, detailed glossaries, and use of computer programs to generate diagnosis may reduce bias in interpretation. Choosing the most relevant clinical instrument is certainly one other major factor.

From: *Methods in Molecular Medicine, vol. 77: Psychiatric Genetics: Methods and Reviews*
Edited by: M. Leboyer and F. Bellivier

There is no perfect clinical instrument; thus, the choice of the relevant instrument is dependent on many factors. The type and size of the population studied (complexity of the disorder, time of evaluation possible, and capacity to understand the questions), frequency and characteristics of the disorder (such as time-stability), recruitment modalities (inpatients or outpatients), budget devoted to clinical assessment, level of practice of interviewers (psychiatrist or lay interviewers), need to recruit exhaustively or not and materials available could be proposed. It is thus understandable that a large variety of instruments are being used.

Metrological qualities are also important when choosing an instrument, such as inter-rater reliability (concordance between two evaluators who rate identical interview material), or test-retest reliability (concordance between two evaluations separated by a certain period of time). These measures are classically evaluated with a Kappa test, which assesses the percentage of agreement that controls for the effect of chance finding. It is difficult to evaluate the quality of an instrument, because no instrument can claim to be perfect. Thus, instruments are compared to each other, or compared to a LEAD (longitudinal observation by experts using all available data as sources of information) standard, which may include interviews, chart review, and longitudinal observation. Specificity, sensitivity and positive and negative predictive values are sometimes used to describe an instrument. A balance between reliability and validity must be achieved. Reliability is high in structured interviews (with good inter-rater K), but validity is frequently poor. Validity is good in face-to-face interviews with non-structured instruments, but reliability is frequently low. Furthermore, some diagnostics are exposed to low reliability (such as social phobia), compared to others that may be easier to detect (such as bipolar disorder).

Because no research in psychiatry genetics can guarantee perfect specificity, sensitivity, and validity of the psychiatric diagnoses collected, it is necessary to establish priorities. The requirement for selecting, training and maintaining and supervising raters—as well as checking and clarifying the data collected as proposed by Endicott ***(1)***—is often difficult to achieve with large samples. Nev-

ertheless, time-consuming laboratory experiments and complex statistical procedures are very sensitive to this first step in the research process. Selecting the clinical interview that will provide the most detailed information on the topic of research, without being overwhelmed by useless data, is part of the good practice of any clinical study, including psychiatric genetics. The importance of a relevant interview—which is more specific to research in psychiatric genetics—is related to the problem of phenotype definition as well as reliability. Multifactorial disorders with complex mechanisms—including the vast majority of psychiatric syndrome—may clearly benefit from a closing-in of phenotype measures and specific mechanisms. Ming Tsuang ***(2)*** proposed that one of the major tasks of psychiatric genetics will be to construct a "psychiatric genetic nosology" capable of classifying individuals in ways that correspond to distinct genetic entities. Difficulties in identifying homogeneous phenotypes can be linked to the limitations of existing assessment methods that do not cover a variety of diagnostic classification systems, ignore differential diagnosis and comorbidity, or omit many items that are suitable for providing broad descriptions of phenotypic subforms. It is thus important to measure recurrence or persistence of symptoms, impairment or incapacitation caused by symptoms, and age of symptom onset and offset, which may contribute to variability of the phenotype ***(3)***. Some examples that highlight the importance of clinical assessment in psychiatric genetics—not only for a reliable assessment, but also for an approach that allows analyses beyond the initial phenotype—are provided. Complexity of psychiatric syndromes systematically raises the unresolved question of what is inherited with a particular genetic marker.

2. Clinical Assessment Beyond the Initial Phenotype

Research on the gene that codes for the dopamine D3 receptor gene (DRD3) in schizophrenia may have a heuristic value concerning the importance of focusing beyond the initial phenotype. Homozygosity for the DRD3 gene (Ser9Gly BalI polymorphism) was found in excess in schizophrenic patients as compared to healthy controls,

according to initial studies ***(4–7)***. However, most results could not replicate this finding, and the odds ratios computed in the meta-analyses were low ***(8–9)***. A further analysis of the phenotype that may be associated with BalI DRD3 gene polymorphism proposed that tardive dyskinesia may be more specifically involved, rather than schizophrenia *per se*. Tardive dyskinesia (TD) develops in approx 20% of patients during long-term treatment with typical antipsychotics. The dopamine D3 receptor is of primary interest in tardive dyskinesia, as dopamine D3-receptor knockout mice show locomotor hyperactivation resembling extrapyramidal side-effects of neuroleptic treatment. Steen ***(10)*** initially found a high frequency (22–24%) of homozygosity for the Ser9Gly variant (allele 2) of the DRD3 gene among subjects with TD in both a cross-sectional and a longitudinal evaluation, as compared to the relative under-representation (4–6%) of this genotype in patients with no or fluctuating TD. This author proposed that correlation between a serious motor side-effect and a genetic marker could lead to selection bias in the sampling of schizophrenic patients for genetic studies, and may therefore explain the apparent association reported between susceptibility for schizophrenia phenotype and homozygosity for the DRD3 gene. In a stepwise multiple-regression analysis, the DRD3gly allele significantly contributed 5% to the variance in orofacial dyskinesia ***(11)***. In another independent study, homozygosity for the Ser9Gly variant of the DRD3 gene was connected to an 88% incidence of acute akathisia as compared with a significantly lower 47% incidence of acute akathisia in schizophrenic patients who were not homogenous for the 2 allele ***(12)***. The difference of distribution of the DRD3ser-gly genotypes among three groups (schizophrenia patients with TD, schizophrenia patients without TD but similar exposure and normal controls) was highly significant ($p = 0.0008$), and this was caused by an excess of the DRD3ser-gly genotype in the schizophrenia patients with TD ***(13)***. The glycine allele of DRD3 was also found to be associated with typical neuroleptic-induced TD ($p < 0.0005$) in another set of 112 schizophrenic patients. Higher mean AIMS scores were found in patients who were homozygous for the glycine variant of the DRD3

gene, as compared to both heterozygous and serine homozygous patients ***(14)***.

In conclusion, including TD in the post-hoc phenotype analyses of schizophrenia allowed the detection of a more direct and specific impact of a gene previously considered as a candidate gene in schizophrenia (i.e., the DRD3). The relationship between vulnerability to TD and vulnerability to schizophrenia is not well-known, but neurological soft-signs are frequently found in excess in patients with schizophrenia ***(15)***, which favors a partly common substratum. Numerous examples may be given for schizophrenia, reminding researchers that schizophrenia may be approached in different ways. For example, age of onset distinguishes various forms of schizophrenia, and late-onset schizophrenia may be associated with the gene coding for the dopamine D2-receptor gene ***(16)***. The therapeutic response to specific antipsychotics ***(17)***, the importance of positive symptoms ***(18)***, or the distinction of some severe subtypes of schizophrenia such as catatonia ***(19–20)*** may also pinpoint the role of some key receptors, such as the D4 dopamine receptor, as vulnerability genes.

Alcohol-dependence may provide another good example of the importance of defining more specific phenotypes with a relevant clinical tool. It should be considered that the genes believed to play a minor but significant role in alcoholism are, in their majority, limited to certain subgroups. For example, some authors have proposed that the D2 dopamine receptor (DRD2) gene is involved in severe alcoholism ***(21)***, the dopamine transporter (DAT) gene in alcohol withdrawal ***(22–23)***, the serotonin transporter gene (SERT) in the suicidal behavior of alcohol-dependent subjects ***(24)***, the alcohol dehydrogenase gene in alcohol tolerance ***(25)***, and the 5-HT_{1B} in antisocial alcoholism ***(26)***. The majority of these genes were revealed by post-hoc analyses, and benefited from extensive clinical evaluations on the related conditions of alcohol-dependence. Namely, somatic complications, type and severity of withdrawal symptoms, number and lethality of suicidal behavior, lifetime maximum number of drinks in 24 h, or antisocial personality disorder and comorbid intermittent explosive disorder (IED) were the

phenotypes to which the detected association found in alcohol-dependence syndrome was attributed.

Thus, whether it is for the quality of the clinical information gathered or the ability to specifically analyze which is the endophenotype that more specifically explains association or linkage, the way clinical assessment is organized has a major impact. A possible strategy is to combine the use of two different instruments to increase both sensitivity and specificity. For example, agreement was highest (55% sensitivity and 90% specificity) when both the SCAN and DIS threshold were set at the level of depression syndrome instead of diagnosis ***(27)***. A description of the various clinical interviews used to screen psychiatric diagnoses is already available ***(28)***, but the next section describes only those used in psychiatric genetics research.

3. Clinical Instruments in Psychiatric Genetics

3.1. The Diagnostic Interview for Genetic Studies (DIGS)

The DIGS ***(29)*** is the result of a National Institute of Mental Health (NIMH) initiative, and was specifically based on the problem of the definition and the assessment of the phenotype in psychiatric genetics, regarding inconsistent and non-replicated findings. The complex relationship between mood disorders and schizophrenia, the impact of substance abuse on the reliability of other psychiatric disorders, and the need for quantitative phenotypes (such as age at onset, phenomenologic details, severity, or course of disorders) are a particular focus. Up to six diagnostic criteria may be generated by the DIGS, such as DSM-III, III-R and IV, RDC, OPCRIT, and ICD-10. The major diagnostic system of the DIGS are DSM-III-R and DSM-IV for the revised version, with 31 diagnoses screened, mostly from axis I, but also some personality disorders (antisocial, schizotypal, schizoid, paranoid, and depressive personality disorder). Schizophrenia, affective disorders and substance abuse are the most intensively studied by the DIGS. Some anxious disorders (excluding generalized anxiety disorder), eating disorders, alcohol and substance abuse and dependence, pathologi-

cal gambling, and suicide attempts are also assessed. The choice of screened phenotypes is thus clearly based on phenotypes, which are usually analyzed in psychiatric genetics. Special sections are also based on the psychiatric genetics requirement, such as a large section for comorbidity, if any, and SANS and SAPS scores for schizophrenia symptoms. Within-site and between-site reliability was globally excellent for major disorders (Kappas around 0.8), with the exception of schizoaffective disorder (Kappa = 0.4). The interviewer must be a mental health professional with clinical experience. The length of the interview is usually about 1 h, but can be much longer when there is comorbidity or when the proband has low informativity.

3.2. The Semi-Structured Assessment for the Genetics of Alcoholism (SSAGA)

The SSAGA ***(30)*** is a semistructured, polydiagnostic psychiatric interview, initially devoted to genetic studies on alcoholism, but also used for a variety of family studies on psychiatry. The SSAGA assesses physical, psychological, social, and psychiatric manifestations of alcohol abuse and dependence and related psychiatric disorders in adults. Reliability of the SSAGA was high for DSM-III-R substance dependence, but less for substance abuse. The concurrent diagnostic validity of the SSAGA across alcohol and drug dependencies, major depression, anxiety disorders, and ASPD was examined and compared to the SCAN and a section of the SCID ***(31)***. Kappas for alcohol dependence were in the acceptable range (0.63). Kappas were lower for sedative dependence (0.48) and for cannabis dependence (0.53), but were higher for cocaine and stimulant dependence (0.85) and for opioid dependence (0.73). Kappa for major depression and the ASPD diagnoses were high (0.71 and 0.70), but slightly lower agreement was found for panic disorder (0.62). Kappa for social phobia was 0.47. Companion interviewers are also proposed, such as those for children (C-SSAGA-C) and adolescents (C-SSAGA-A). DSM-IV, DSM-III-R, RDC, and ICD-10 are the main classification systems covered by the SSAGA. Apart from all

types of substance abuse and dependence, the SSAGA screens 11 psychiatric diagnoses (eating disorders, mood disorders and different anxious disorders), one personality disorder (ASPD) and three conditions of interest (tobacco use, psychosis, and suicidal behavior).

3.3. Structured Clinical Interview for DSM Axis I Disorders (SCID-I)

This instrument *(32)* covers a broad set of psychiatric diagnosis, includes nine modules (and now 51 specific DSM-IV diagnoses), and is a clinician-administered semistructured interview. The SCID requires special training (10-h training video available through Biometric Research). The SCID-I was designed to be simple to use and reduce the time needed for training and administration. The SCID-I is devoted to research *(33)*, whereas the SCID-CV is simpler and thus specifically designed for clinicians *(34)*. The SCID-II is a separate interview for axis II personality disorders, and the SCID-NP is devoted to nonpatients. Different versions are available on the web (www.appi.org, www.mhs.com). The interview takes between 1 h and 3 h, depending on the number, severity, and complexity of the disorders. The overall reliability *(35)* was fair to good in patient samples (weighted Kappa 0.61) but poor in nonpatients (weighted Kappa 0.37). Good to excellent reliability was obtained for diagnosis of bipolar disorder, substance abuse, or dependence and eating disorders (Kappas around 0.7–0.8), but relatively poor for dysthymia (0.40) and social phobia (0.47). The validity has not been exhaustively measured, but 85% of patients with known psychotic symptoms revealed all or some of their symptoms with the SCID. Interestingly, European (German) vs American raters had a similar evaluation, as overall weighted Kappas were good (Kappas = 0.6).

3.4. The Schedules for Clinical Assessment in Neuropsychiatry (SCAN)

The SCAN *(36)* is a set of instruments and manuals that screens psychopathology and behavior associated with a broad range of major psychiatric diseases according to ICD as well as DSM crite-

ria, among others. It is derived from one of the very first clinical interviews in psychiatry, the PSE (Present State Examination), and allows dimensional rating of symptoms and syndromes. The CATEGO-5 is a set of computer programs that process SCAN data and provide output. Interviewers must be well-trained, with SCAN-specific training. Lay interviewers without previous clinical experience may also have good reliability. Two hours are usually needed to administer the SCAN, but the time may be considerably longer for patients with complex disorders. The instrument has been translated into 13 languages, and is distributed on the internet (www.who.int). A training course at a WHO training center is recommended.

Test-retest reliability is fairly good (Kappa 0.41–0.64) for both inpatients and outpatients. Agreement between SCAN and SADS was disappointingly low ***(37)***. SCAN offers certain advantages: it is not devoted to a single diagnostic system, and it deeply screens psychopathology that may allow more subtle clinical analyses. On the other hand, it is long, cumbersome, and contains a long glossary (the code book contains 1,700 separate ratings for each episode of disorder).

3.5. The Schedule for Affective Disorders and Schizophrenia (SADS)

The SADS ***(38)*** is a NIMH initiative made of a semistructured clinician-administered interview based on Research Diagnostic Criteria (RDC). The SADS covers lifetime mental symptoms and 23 major disorders, according to RDC criteria. This instrument covers current mental disorders as well as lifetime history (of disease and comorbidity), according to their most severe level. In this context, it is not surprising that the first psychiatric genetic studies largely used this instrument. The interviewers must have clinical experience and must receive special training. The interview, which is delivered by a trained clinician, takes an average of 1 h for healthy subjects and up to 3 h for psychiatric patients. The SADS has been developed for lifetime evaluation (SADS-L), for bipolar disorders (SADS-LB),

and for anxious disorders (SADS-LA). The SADS has been translated into 10 languages.

3.6. The Comprehensive Assessment of Symptoms and History (CASH)

The Comprehensive Assessment of Symptoms and History (CASH) is an instrument for assessing diagnosis and psychopathology *(39)*. It is designed to provide a comprehensive information base, mainly for schizophrenia and affective spectrum conditions, concerning current and past signs and symptoms, premorbid functioning, sociodemographic status, cognitive functioning, treatment, and course of illness. The RDC, DSM, and ICD diagnostic criteria are used. The main characteristic of the CASH is its focus on the problems regarding the boundaries of major psychiatric disorders, and is able to describe in detail common psychiatric symptoms, to make diagnoses with multiple criteria, to explore new criteria, and to generate alternative criteria derived from homogenous subgroups of patients. Another important point is that items are usually rated on a six-point Likert-like scale, for which anchor points are provided. The reliability is good for inter-rater (Kappa between 0.45 for schizoaffective disorder and 1.00 for bipolar disorder) and test-retest assessments (between 0.52 for schizoaffective disorder and 1.00 for schizotypal personality disorder [SPD]) *(39)*. The CASH was rarely used in psychiatric genetics studies, probably because of the length of time required and the complexity of the instrument.

3.7. The Diagnostic Interview Schedule (DIS) and the Composite International Diagnostic Interview (CIDI)

These highly structured instruments screen present and lifetime psychiatric disorders. Trained nonclinical or lay interviewers may administer these two instruments. The DIS was written at the request of the NIMH for use in the Epidemiological Catchment Area (ECA) project, so it now has a considerable history of experience. Thirty-one diagnoses are screened by the DIS—the average that occurs during an interview of 45–75 min. The Kappa between a psychia-

trist and a lay interviewer is generally good (above 0.5), except for panic disorder according to DSM diagnosis (Kappa = 0.4). The CIDI ***(40)*** is an extended DIS for use across cultures, for both ICD and DSM criteria (for 40 diagnoses). One potential weakness of both the DIS and the CIDI is their reliance on the judgments and insights of the respondent instead of a psychiatrist, although strong agreement between the DIS and the clinical interview was found in medical settings ***(41)***. Some feel that the DIS-style interviews capture a wider range of trivial or less severe symptoms and disorders than a psychiatrist would ***(42)***. For example, only two-thirds of depressive disorders detected with the DIS by lay interviewers were confirmed to be major depressive disorders by psychiatrists who used the SCAN ***(27)***. Nevertheless, cross-cultural reliability was good ***(43)***, which may be an important characteristic for international collaborative studies in psychiatric genetics. Predictors of false-negative depressive disorder using the DIS included: age over 45 yr (OR = 6.12), with less than six depressive symptoms (OR = 4.42), the depressive episode being passed for more than 1 yr (OR = 3.69), being male (OR = 3.41) or black (OR = 2.80), and with little impairment (OR = 1.48) ***(27)***.

3.8. The Primary Care Evaluation of Mental Disorders (PRIME-MD) and the System-Driven Diagnostic System for Primary Care (SDDS-PC)

These two instruments combine a self-report with a follow-up clinical examination designed for use by nonpsychiatric physicians in primary care centers. PRIME-MD ***(44)*** assesses mental disorders through a 10-page questionnaire, including one page of self-report items. The training is rapid (from 1 to 3 h), the instruction manual is relatively compact (13 pages), and the time required by patients is considerably shorter than other instruments (8 min), yet this system cannot screen the same amount of psychiatric symptoms and diagnoses. The Kappa values are not as poor as they could be (global Kappa around 0.7), even if the PRIME-MD has low reliability for some disorders such as minor depressive episode (Kappa = 0.16).

The Symptom-Driven Diagnostic System-PC ***(45)*** is a computerized instrument for the detection, diagnosis, and management of mental disorders in primary care practice, on the basis of 29 items (patient self-report) and a structured diagnostic interview guide with six diagnostic modules. The time required to use this instrument is equal to the PRIME-MD—5 min for the autoquestionnaire and 5–10 min for each diagnostic module. Validation studies showed large variations in sensitivity and specificity when the PRIME-MD is compared to the SCID interview conducted by trained clinicians. Some studies showed good sensitivity (between 0.62 and 0.9) and correct specificity (from 0.54–0.98), but others showed less convincing quality, with moderate sensitivity (from 0.24 to 0.85) and low Kappa compared to the SCID diagnoses (from 0.11 to 0.48). Some diagnoses are especially likely to be undetected by this instrument, such as generalized anxiety disorder and obsessive-compulsive disorder (OCD) with a positive predictive value of 5%.

3.9. The Mini International Neuropsychiatric Interview (MINI)

The MINI ***(46)*** is a brief structured questionnaire used by health technicians in clinical and research settings. A total of 120 questions cover 17 current DSM axis I disorders. With an administration time of approx 15 min, the MINI was designed for an accurate structured psychiatric interview for multicenter clinical trials and epidemiological studies. The MINI-screen (close to the PRIME-MD), the MINI-plus (much more detailed) and the MINI-kid (for child and adolescent psychiatry) were developed later ***(47)***. When the MINI (rated by clinicians) is compared to the SCID-P (for patients), the Kappas are fairly good (from 0.43 for current drug dependence to 0.9 for anorexia nervosa), but Kappas are poorer when rated by the patient (from 0.11 for dysthymia to 0.66 for anorexia nervosa). The worst inter-rater Kappa is 0.79; thus, the instrument has a good inter-rater reliability. Test-retest Kappa is fairly good on average, except for current mania episode (Kappa = 0.35). Different translations are available through the internet at no cost (http://www.medical-outcomes.com).

3.10. The OPerational CRITeria Diagnostic System (OPCRIT)

The OPerational CRITeria diagnostic system ***(48)*** is a series of computer programs that allow data entry and generate diagnoses according to 12 operational diagnostic systems (including DSM, ICD, and RDC), and is comprised of a checklist for major psychiatric classifications. The OPCRIT has a wide range of psychiatric research applications, including both European Science Foundation and NIMH research initiatives in the molecular genetics of mental disorders. The OPCRIT checklist is based on 90 items that involve psychopathology and background information, and has been included within the DIGS. The OPCRIT checklist is designed to be completed by trained clinicians and can be based on information from diagnostic interviews but also case records, which is sometimes useful in psychiatric genetics for which a precise phenotype definition is required, although the available clinical information is often limited. An international study including 30 clinicians from both sides of the Atlantic who rated 30 case studies demonstrated good reliability within all classification systems (e.g., RDC, Kappa = 0.71 and ICD-10, Kappa = 0.70) ***(49)***. Furthermore, good agreement is found between lifetime diagnoses generated by OPCRIT on the basis of rating by a single rater as compared with consensus best-estimate procedures, with Kappas between 0.7 and 0.8 according to the classification system used ***(50)***. Similar trends are found for Kappas between OPCRIT diagnoses carried out by a trained interviewer, and OPCRIT diagnoses made by the best-estimate lifetime consensus procedure (0.83 for DSM-III-R and 0.81 for ICD-10) ***(51)***.

3.11. The Psychiatric Research Interview for Substance and Mental Disorders (PRISM)

The PRISM ***(52)*** provides detailed diagnostic information on disorders involving substance use, and also includes sections on mood, anxiety and eating disorders, antisocial and borderline personality disorders, and psychotic symptoms. The interview takes 50–150 min

for a trained clinician. Test-retest reliabilities are globally good in current and lifetime dependence (Kappa around 0.8), but this is not the case for substance abuse (many Kappas below 0.4). Major depressive (current or past) episodes are reliably detected in different periods (Kappa 0.64-0.81). Kappa scores are moderately lower for anxious disorder (0.68 for current panic disorder, 0.45 for lifetime agoraphobia without panic); good for psychotic symptoms, anorexia, and bulimia nervosa (around 0.80 for lifetime assessment); and correct for antisocial and borderline personality disorder (0.60 and 0.83, respectively).

3.12. Familial History

Psychiatric illness in biological relatives is frequently screened in psychiatric genetic studies, whether for aggregation studies or for phenotype definitions (i.e., sporadic vs familial cases). Systematic family history methods are thus frequently used. The problems of refusal, death, or prohibitive cost explain why indirect interviews are proposed. Each time a relative is unreachable, one or more informants are questioned about their relative. There are different instruments available, such as the FH-RDC (Family History Method for Research Diagnostic Criteria), the FISC (Family Informant Schedule and Criteria), the FIGS (Family Interview for Genetic Studies), the FHAM (Family History Assessment Module) and the FHS (Family History Screen).

The first systematic family history method, the Family History Method for Research Diagnostic Criteria (FH-RDC) has good test-retest reliability and an acceptable level of validity as compared to direct interviews ***(53)***. The FISC is an extended version of the FH-RDC ***(54)***, with DSM-III-R criteria, RDC diagnoses with level of confidence, and more anxiety disorders and mood-related disorders. The FIGS is associated with the DIGS, and has the same structure ***(29)***. The FHAM ***(55)*** was developed from a collaborative study on alcoholism. The specificity, sensitivity, and positive predictive values of this instrument were 98%, 39%, and 45%, respectively. The FHS ***(56)*** collects information on 15 psychiatric disorders and sui-

cidal behavior from informants and their first-degree relatives. Each question is posed only once, and treats all family members as a group. The administrative time is thus relatively short (5–20 min). The FHS has a median sensitivity (0.676 for proband, 0.711 for relative) and specificity (0.876 for proband, 0.894 for relatives). Test-retest agreement of FHS administration 15 mo apart is acceptable, ranging from 0.38 for agoraphobia to 0.74 for any substance dependence, and the median Kappa is 0.56. As for most instruments, the use of more than one informant substantially improves sensitivity, increasing sensitivity from 35.2% with one informant (mostly the proband) to 68.2% with multiple informants.

The list of diagnoses screened by each of these instruments may be a good indicator of which instrument is best for a specific study. For example, antisocial personality disorder is assessed in the FISC and the FHS and not in the FIGS, and suicidal attempts are not evaluated in the FISC but are present in the FIGS. Length of time, the possibility of recruiting clinicians for face-to-face interviews, and the availability of relatives are also major factors in the choice of the relevant familial interview. The need for a first quick global screening would present an argument for using the FHS, especially if the global notion of "positive familial history" of the proband is sufficient. A more complete evaluation—especially if a face-to-face interview is not agreed upon or possible—may require a more exhaustive indirect evaluation, such as that provided by the FIGS.

4. The Present Use of Clinical Instruments in Psychiatric Genetics

The articles published on psychiatric genetics during the year 2000 in the main journals devoted to psychiatry, or even specifically devoted to psychiatric genetics, have been collected. We have limited our investigation to the articles that meet five conditions, because the boundaries between psychiatric genetics and other research topics in psychiatry are not always clear-cut, because of specificities and size limitation. These conditions are (1) human populations, with (2) molecular genetics (mainly PCR) and/or familial studies (including aggregation and twin and adoption studies),

(3) giving original results, whether the sample is original or not, and (4) including psychiatric patients (except for Alzheimer's disease). The description of the study is focused and limited to psychiatric genetics. The table of results is not meant to be exhaustive, because the material is also dependent on personal availability of the various journals. Thus, all journals are not included, and all the articles in each journal were not systematically assessed (although each journal, when available, was carefully analyzed). The idea is that having more than 120 articles (corresponding to the clinical investigation of 37.442 patients, 24.557 relatives and 19.416 controls all together) on the topic would provide a good idea of the clinical instruments used, and indirectly, in which country, for which phenotype, with which genetic approach, according to which diagnostic criteria, and for which type of recruitment. The majority of articles (*see* Table1) are sourced in the United States (N = 42), Canada (N = 12), Germany (N = 10), Japan and the UK (N = 9), France and Israel (N = 5), and Ireland and Italy (N = 4). The SADS (36%) and the SCID (32%) are the most frequently used clinical interviews, followed by the DIGS (9%), DIS (7%), OPCRIT (7%), and SSAGA (4%). The SCAN, CASH, and CIDI have only been detected once.

When the type of journal is the criteria selected, the data shows that the OPCRIT is more frequently quoted in journals that specialize in psychiatric genetics (RR = 1.59), with the same trend for the SADS (RR = 1.17). The SCID is used more often in general titles (RR = 1.50), and the DIGS is almost equally distributed (RR = 1.08).

The method used in the studies also has an impact on the type of clinical interview chosen. Case-control-association studies are more likely to use the OPCRIT (RR = 2.26) or the DIGS (RR = 1.36). The TDT approach is more often associated with the choice of the SCID (RR = 1.59), whereas linkage studies more frequently use the SADS (RR = 1.61). There is no specific pattern for aggregation studies (used in similar scale DIGS, SADS, and SCID), but twin studies are based mainly on the SCID (RR = 1.59). FISH and RED methods are exclusively based on the DIGS (RR = 8.6).

In the same respect, the type of recruitment permits specific time and availability. The selections are clear-cut for the DIGS when

in- and out-patients are recruited (RR = 1.57), for the SCID for twin register (OR = 2.54), and for the SADS for outpatients (OR = 1.35).

Some clinical interviews have focused more specifically on certain psychiatric disorders. For example, ADHD is essentially analyzed through the SADS (RR = 1.79), or more precisely, the K-SADS for children. Alcohol and substance abuse and dependence use the SCID in the vast majority (RR = 2.59), whereas anxiety disorders are based on the SADS (OR = 1.43), and eating disorder use the SADS (RR = 2.15). Some disorders are less clearly based on one type of interview. For example, the SADS (R=1.51) and the DIGS (RR = 2.58) are over-represented in affective disorder, and schizophrenia is associated with OPCRIT (RR = 1.70) and the SCID (OR = 1.20).

The diagnoses are based on clinical criteria, which partly refer to specific interviews, such as RDC for SADS (OR = 2.25) and DSM for SCID. Some interviews have multiple criteria, but the most frequent choice is DSM, for the DIGS (RR = 1.5) and the OPCRIT (RR = 1.88).

Finally, the country in which the research is conducted also plays a role in the use of a clinical interview, whether it is for availability of translated interview, training, or cultural habits. For example, articles from the United States predominantly use the SCID (OR = 1.3), as do Japan, Sweden, and Finland (OR = 2.6). The DIGS is preferred by Canadian studies (OR = 2.35), as well as French (OR = 2.35) and Portuguese (OR = 9.4) studies. The OPCRIT is clearly dominant in the United Kingdom (OR = 7.83), Italy (OR = 11.7), and Germany (OR = 1.96), and the SADS is chosen by studies conducted in Germany (OR = 1.96), Israel, Turkey, Belgium, and Spain (OR = 2.35), as well as France (1.18).

5. Conclusion

The choice of a clinical interview is a complex yet important part of a research protocol, and is limited by numerous factors that may not be modified (such as availability in local language of interviews, type of interviewers, and the time available for each patient). The

Table 1
Review of Clinical Assessment Used in Psychiatric Genetics in Journals Specifically Devoted to This Topic

Author	Ref	Journal[a]	Main method	Phenotype	Recruitment of proband	Interviewer	Sample (N=) Pro-band	Rela-tives	Con-trols
Froehlich	24:265-277	ACPR	Twin study	Alcoholism	Twin registry	0	88	88	0
Niethammer	157: 272-274	AJP	Twin study	Schizophrenia	?	?	30	30	34
Conklin	157: 275-277	AJP	Aggregation	Schizophrenia	Inpatients (psychiatry)		52	56	73
Strober	157: 393-402	AJP	Aggregation	Eating disorder	Inpatients (psychiatry)	PhD	323	1831	181
Kendler	157: 402-408	AJP	Linkage	Schizophrenia	Inpatients (psychiatry)	Psychiatrists	265	1143	0
Wade	157: 469-471	AJP	Twin study	Anorexia	population-based female twins		1030	1030	0
Kendler	157: 506-513	AJP	Twin study	Mental health	population-based female twins		794	794	0
Preisig	157: 948-955	AJP	Association	BP	in and outpatients (psychiatry)	MD ?	272	0	122
Malaspina	157: 994-1003	AJP	Aggregation	Schizophrenia	inpatients (psychiatry)	Researchers?	99	?	0
McMahon	157: 1058-1064	AJP	Association	BP	?	Psychiatrists	93	18	63
Ross	157: 1071-1076	AJP	Aggregation	Schizophrenia	Inpatients (psychiatry)	Psychiatrists	466	347	0
Faraone	157: 1077-1083	AJP	Aggregation	ADHD	in and outpatients (psychiatry)	MD	140	786	122
Mundo	157: 1160-1161	AJP	TDT	OCD	Anxiety clinic	?	67		0
Herbst	157: 1285-1290	AJP	Association	Temperament	Global population		946	0	0
Nishiguchi	157: 1329-1331	AJP	Association	Schizophrenia	?	?	164	0	171
Rosa	157: 1511-1513	AJP	Twin study	Psychosis	NIMH twins	?	50	50	0
Reichenberg	157: 1514-1516	AJP	Twin study	Psychosis	Draft board registry		10	10	2218
Jacobsen	157: 1700-1703	AJP	Association	DAT availability	?	?	30	0	0
Kendler	157: 1843-1846	AJP	Twin study	Sexual orientation	Twins and siblings	None	1806	1162	0
Tsuang	157: 1955-1959	AJP	Aggregation	Psychosis	Genetic clinics	Psychiatrists	22	0	22
Constantino	157: 2043-2044	AJP	Twin study	Reciprocal social behavior	Global population	None	232	232	0
Lichtermann	157: 2045-2047	AJP	TDT	Alcohol-dependence	Inpatients (psychiatry)	?	92	184	0
Kendler	57:261-269	AGP	Twin study	Substance abuse	Twin registry	Students	1198	1198	0
Nestadt	57:358-363	AGP	Aggregation	OCD	in and outpatients ?	MD or psychiatrist	99	643	73
Maziade	57:1077-1083	AGP	Aggregation	Autism	in and outpatients	Psychiatrists	78	251	521
Slutske	57:666-673	AGP	Twin study	Alcohol-dependence and pathological gambling	Twin registry	students	3372	3372	0
Prescott	57:803-811	AGP	Twin study	Alcoholism & MDD	Twin registry	Students	3755	3755	0
Kendler	57:886-892	AGP	Twin study	Tobacco dependence	Twin registry	0	778	778	0
Cubells	5 (1): 56-63	MP	Association	Cocaine dependence	?	?	45	0	66
Okuyama	5 (1): 64-69	MP	Association	Novelty seeking	Students		86	0	0
Li	5 (1): 77-84	MP	TDT	Schizophrenia		Psychiatrists	198	396	0
Edgar	5 (1): 85-90	MP	Proteome	Schizophrenia	Brain		7	0	7
Hranilovic	5 (1): 91-95	MP	TDT and sibs	Schizophrenia		Psychiatrists	61	133	0
Benjamin	5 (1): 96-100	MP	association	Novelty seeking	Students	0	455	0	0
Franke	5 (1): 101-104	MP	TDT	Heroin dependence	Inpatients (psychiatry)		396	222	197
Ibañez	5 (1): 105-109	MP	Association	Pathological gambling	Outpatients	Psychiatrists	68	0	68
Arranz	5 (2): 124-125	MP	Association	Clozapine response	?	?	180	88	0
Li	5 (2): 128-130	MP	Association	Heroin abuse	Inpatients (psychiatry)	?	375	0	198
Ito	5 (2): 159-164	MP	Association	Schizophrenia	Inpatients (psychiatry)	Psychiatrists	88	0	105
Yoshikawa	5 (2): 165-171	MP	Association	Bipolar	NIMH initiatives	?	96	0	59
Sjoholt	5 (2): 172-180	MP	Association	Bipolar	?	?	23	0	2
Hu	5 (2): 181-188	MP	Association	Smoking cessation	NIMH initiatives	0	759	0	
Lerman	5 (2): 189-192	MP	Association	Smoking cessation	Outpatients	Advertisement	185	0	0
Bondy	5 (2): 193-195	MP	Association	Suicide	Suicide victims	?	58	0	110
Moutsatsou	5 (2): 196-202	MP	Association	Bipolar	?	?	12	3	12
Austin	5 (2): 208-212	MP	Association	Schizophrenia	Inpatients (psychiatry)	?	203	0	203
Osher	5 (2): 216-219	MP	Linkage	Anxiety trait	Advertisement	Advertisement	148	0	0
Damberg	5 (2) : 220-224	MP	Association	Personality traits	Population register	MD	137	0	
Di Bella	5 (3) : 233-241	MP	Association	Eating disorder	Outpatients	?	106	0	120
Chowdari	5 (3) : 237-238	MP	TDT	Schizophrenia	?	?	101	131	0
Zhang	5 (3) : 239-240	MP	Association	Hallucination	?	?	84		70
Serretti	5 (3) : 270-274	MP	Association	Schizophrenia symptoms	Inpatients (psychiatry)	Psychiatrists	1182	0	267
Grünhage	5 (3) : 275-282	MP	Association	Bipolar	Inpatients (psychiatry)	Psychiatrists	45	0	46
Vandenbergh	5 (3) : 283-292	MP	Association	ADHD, AD, Tourette	Various	?	188	0	150
Loh	5 (3) : 301-307	MP	Association	Alcohol dependence	?	?	189	0	152
Iwata	5 (3) : 316-319	MP	Association	Alcohol dependence	Inpatients (psychiatry)	Psychiatrists	196	0	188
Auranen	5 (3) : 320-322	MP	Sibpairs	Autism	?	?	17	14	0
Joober	5 (3) : 323-326	MP	Association	Neuroleptic response	?	?	105	0	90
Speight	5 (3) : 327-331	MP	Association	SZ and BP	?	?	422	0	283
Oliveira	5 (4) : 348-349	MP	Association	Mood disorders	in and outpatients (psychiatry)	?	192	0	196
Tahir	5 (4) : 396-404	MP	TDT	ADHD	Outpatients	Psychiatrists	111	215	0
Barr	5 (4) : 405-409	MP	TDT	ADHD	?	?	97	194	0
Basile	5 (4) : 410-417	MP	Association	Tardive dyskinesia in SZ		?	85	0	0
Kaiser	5 (4) : 418-424	MP	Association	Antipsychotic response	in and outpatients (psychiatry)	Psychiatrists	638	0	0

(continued)

Table 1

Evaluation[b]	Psychiatric assessment: Proband	Psychiatric assessment: Relatives	Psychiatric assessment: Controls	Main instrument	Criteria	Country
EP	0	0		0	0	USA (Indianapolis)
DE	SCAN	FHRDC	SCAN	SCAN	ICD-10	Germany (Eur)
DE	SCID	SCID	SCID	SCID	DSM-IV	USA (Minn)
DE	Direct interview	90% of direct interview, 2/3	SADS	SADS	DSM-IV	USA (LA)
DE	SCID (modified) and SIS	SCID (modified) and SIS		SCID	DSM-III-R	Ireland (Eur)
FI	?	?			DSM-III-R	USA (Virginia)
TI	SCID adapted	SCID adapted		SCID	DSM-III-R	USA (Virginia)
FI	DIGS	FIGS	DIGS	DIGS	DSM-IV	France & Switzerland (Eur)
FI	DIGS and SDS	FIGS		DIGS	DSM-III-R	USA (NY)
FI	SADS	interview	interview	SADS	RDC	USA (San Diego)
FI	SCID	face-to-face interview		SCID	DSM-III-R	Ireland (Eur)
FI	K-SADS-E	SCID	K-SADS-E	K-SADS-E	DSM-III-R / DSM-IV	USA (Massachusetts)
FI	SCID	scid	0	SCID	DSM-iV	CAN (Toronto)
SQ	self-questionnaire	0	self-questionnaire	TCI	0	USA (Baltimore)
?	?	0	?	?	DSM-IV	Japan (Kobe)
?	NIMH twins	?	0	?	DSM-III-R	USA (Bethesda)
HR	Hospitalisation registry	Hospitalisation registry	hospitalisation registry	0	ICD10	Israel (Ramat Gan)
FI	SCID	0	0	SCID	DSM-IV	USA (Conn)
SQ	Self-questionnaire	self-questionnaire	0	0		USA (Virginia & Boston)
FI	SCID	0	SCID	SCID	DSM-IV	USA (Seattle)
Q	Questionnaire	questionnaire	0		0	USA (Missouri)
FI	SSAGA	0	0	SSAGA	DSM-IV	Germany (Eur)
I	SCID	SCID	0	SCID	DSM-IV	USA (Virginia)
I	SADS	SADS, FISC, SCID-P	SADS	SADS	DSM-IV	USA (Baltimore)
DE	ADI	questionnaire	Questionnaire	ADI	DSM-III-R	Canada (Quebec)
TI	DIS	DIS	0	DIS	DSM-III-R	USA (Missouri)
TI	DIS	DIS	0	DIS	DSM-IV	USA (Missouri)
Q	Questionnaire	Questionnaire	0		0	Sweden (Eur)
?	?	0	?	CEQ	DSM-IV	USA (CT)
Q	TCI (questionnaire)	0	0		0	Japan (Irabaki)
FI	scid	0	0	SCID	DSM-IV	Japan (Sishuan)
R	file	0	File	none	0	New Zealand (Auckland)
I and files	SADS	0	SADS, SCID, OPCRIT	0	RDC	Germany (Eur)
Q	TCI (questionnaire)	0	0		0	Israel (Ramat Gan)
FI	?	0	?		DSM-III-R	Germany (Eur)
I	Interview	FH-RDC	interview		DSM-IV and ICD-10	Spain (Madrid)
?	GAS	0	GAS	0	DSM-IV	UK
I	?	0	?	0	DSM-IV	Japan (Sishuan)
I	SCID	0	SCID	SCID	DSM-III-R	France (Eur)
I	DIGS ?	0	0	DIGS ?	?	USA (Bethesda)
?	?	?	?	?	?	Norway (Eur)
I	NEO-PI-R	0	0	NEO-PI-R	0	USA (Bethesda)
Q	EPI	0	0	EPI	0	USA (Washington)
Q	0	0	OPCRIT	OPCRIT	0	Germany (Munich)
?	?	?	?	?	DSM-IV	Greece (Athens)
I	OPCRIT	0	none	OPCRIT	DSM-IV	UK (Cardiff)
Q	TPQ & NEO-PI-R	0	0	TPQ	0	Israel (Beer Sheva)
Q	KSP	0	0	KSP		Sweden (Stockholm)
I	?	0	?	?	DSM-IV	Italy (Milan)
I	DIGS	?	0	?	DSM-IV	USA (Pittsburgh)
?	HDS	?	?	HDS	ICD-10	China (Beijing)
I	OPCRIT	0	Interview	OPCRIT	DSM-IV	Italy (Milan)
I	SADS	FISC	?	SADS	DSM-III-R	Germany (Bonn)
I	DIS	?	?	DIS	DSM-III-R	USA & Canada
I	SCID	0	?	SCID	DSM-III-R	Japan (Tokyo)
I	SCID	0	SCID	SCID	DSM-III-R	Finland (Eur)
I	CARS, ASSQ, ASDI	CARS, ASSQ, ASDI	0	CARS	DSM-IV	Finland (Eur)
I	DIGS	0	DIGS	DIGS	DSM-IV	Canada (Montreal)
I	OPCRIT, SCAN, PSE	0	?	OPCRIT	DSM-IV	UK (Cardiff)
?	?	0	?	?	DSM-IV	Brazil (Sao Paolo)
I	K-SADS	none	0	K-SADS	DSM-IV	Turky (Istanbul)
I	PICS, TTI	0	0	PICS, TTI	DSM-IV	Canada (Toronto)
I	AIMS	0	0	AIMS	DSM-III-R	Canada (Ontario)
I	PANSS	0	0	PANSS	DSM-IV	Germany (Berlin)

(see next page)

Table 1
Review of Clinical Assessment Used in Psychiatric Genetics in Journals Specifically Devoted to This Topic

Author	Ref	Journal[a]	Main method	Phenotype	Recruitment of proband	Interviewer	Sample (N=) Pro-band	Rela-tives	Con-trols
Ishiguro	5 (4) : 433-438	MP	Association	Schizophrenia	in and outpatients (psychiatry)	?	252	0	274
McInnis	5 (4) : 439-442	MP	TDT	BP, SZ and ataxia	in and outpatients (psychiatry) ?	?	618	0	644
Hamilton	5 (5) : 465-466	MP	TDT	Panic disorder	?	?	81	162	0
Toyota	5 (5) : 489-494	MP	Association	Affective disorder	in and outpatients (psychiatry)	Psychiatrists	254	0	299
Holmes	5 (5) : 523-530	MP	TDT	ADHD	in and outpatients (psychiatry) ?	?	137	239	442
McCracken	5 (5) : 531-536	MP	TDT	ADHD	in and outpatients (psychiatry) ?	Psychiatrists	197	174	
Quist	5 (5) : 537-541	MP	TDT	ADHD	in and outpatients (psychiatry) ?	?	115	208	0
Jorm	5 (5) : 542-547	MP	Association	Anxiety trait	Global population survey	0	660	0	0
Barr	5 (5) : 548-551	MP	TDT	ADHD	?	?	97	194	0
Austin	5 (5) : 552-557	MP	Association	Schizophrenia	in and outpatients (psychiatry)		203	0	230
Krebs	5 (5) : 558-562	MP	Association	Schizophrenia	in and outpatients (psychiatry)	Psychiatrists	88	0	52
Lakatos	5 (6) : 633-637	MP	Association	Attachment disorganisation	Community sample		103	0	0
Schwab	5 (6) : 638-649	MP	Linkage	Schizophrenia	in and outpatient ?	Psychiatrists	71	67	0
Pulver	5 (6) : 650-653	MP	Linkage	Schizophrenia	Epidemiological sample	?	28	28	0
Wassink	5 (6) : 678-682	MP	TDT	Schizophrenia	in and outpatient ?	?	140	197	46
Wei	96 4-7	NG	TDT	Schizophrenia	Outpatients	Psychiatrist	70	70	0
Lin	96 12-14	NG	Association	Mood disorders	In and outpatients	?	132	0	88
Ogilvie	96 15-17	NG	FISH	Autism	National Autistic Society		77	26	0
Weissman	96 24-35	NG	Linkage	Panic disorder	In and outpatients ?	Clinicians	34	248	0
Ho	96 36-42	NG	Association	Mood disorder	In and outpatients	Clinicians	270		
Serreti	96 84-87	NG	Association	Szchizophrenia	Inpatients (psychiatry)	Psychiatrist	188	0	0
Henderson	96 102-107	NG	Association	Personality traits	Community sample	0	2752	0	0
Barr	96 114-117	NG	TDT	ADHD	?	?	100	184	0
Ohmori	96 118-122	NG	Association	Schizophrenia	In and outpatients ?	?	129	0	140
Massat	96 136-140	NG	Association	Bipolar disorder	?	?	309	0	309
Greenberg	96 202-216	NG	Association	Anxiety trait	Advertisement	0	397	0	0
Vogt	96 217-221	NG	TDT	SZ and BP	?	Psychiatrist	196	210	46
DeLisi	96 235-239	NG	Linkage	Schizophrenia	?	Clinicians	389	389	0
Eisenberg	96 258-261	NG	TDT	ADHD	Outpatients	Psychiatrist	46	92	0
Barr	96 262-267	NG	TDT	ADHD	Outpatients	?	82	154	0
Hawi	96 268-272	NG	TDT	ADHD	Outpatients	?	97	177	0
Muglia	96 273-277	NG	TDT	ADHD	Outpatients	Psychiatrist	111	72	66
Kotler	96 278-281	NG	TDT	ADHD	Outpatients	Psychiatrist	49	98	0
Hawi	96 282-284	NG	TDT	ADHD	Outpatients	?	94	174	0
Tahir	96 285-288	NG	TDT	ADHD	Outpatients	Psychiatrist	72	144	0
Jiang	96 289-292	NG	TDT	ADHD	Outpatients	Psychiatrist	72	144	0
Willcutt	96 293-301	NG	Twin	Reading disorder	Epidemiological sample	?	313	313	0
Lawford	96 592-598	NG	Association	Substance abuse	Outpatients	Medical officers	95	0	50
Noble	96 622-631	NG	Association	Alcohol-dependence	?	?	92	0	85
Saccone	96 632-637	NG	Linkage	alcohol-dependence	In and outpatients	?	1138	492	428
Vanyukov	96 654-658	NG	Association	ASPD	Outpatients	?	42	0	114
Blomqvist	96 659-664	NG	TDT	Dependence	Outpatients	?	85	170	0
Vandenbergh	96 678-683	NG	Association	Substance abuse	In and outpatients ?	?	184	0	122
Young	96 684-695	NG	Twin study	behavioral dishinibition	Epidemiological sample	?	334	334	0
Ishiguro	96 716-720	NG	Association	Schizophrenia	Inpatients (psychiatry)	?	417	0	189
Yan	96 749-753	NG	FISH	Schizophrenia	Outpatient	?	1	0	0
Persico	96 784-790	NG	TDT	Autism	In and outpatients	?	135	258	152
Hwu	96 797-800	NG	Association	Alcohol Dependence		?	160	0	70
Schulze	96 801-803	NG	Association	MDD	?	Psychiatrist	146	0	101
Alda	96 804-807	NG	Aggregation	BP	?		132	29	310
Buervenich	96 808-813	NG	Association	SZ and BP	In and outpatients	Psychiatrist	324	0	226
Bondy	96 831-835	NG	Association	Suicide	Inpatients (psychiatry)	Psychiatrist	215	0	125
Goizet	96 839-844	NG	FISH	Autism	Outpatient	Psychiatrist	1	0	0
Pato	96 854-857	NG	RED method	Bipolar disorder	In and outpatients	Psychiatrist	24	0	53
Haider	96 870-872	NG	Association	Schizophrenia	Inpatients (psychiatry)	Psychiatrist	80	0	114
Vincent	96 873-876	NG	RED method	Major psychosis	Multicentric sample	?	608	0	219
Sivaqnanasunc	279:13-16	NL	Association	Schizophrenia	?	?	73	0	56
Tsai	44:177-181	SR	Association	Schizophrenia	In and outpatients	Psychiatrist	90	0	104
Brzustowicz	288:678-682	Sc	Linkage	Schizophrenia	?	?	22	275	0
Goldstein	26(2):323-334	SZB	Aggregation	Schizophrenia	Epidemiological sample	Students and psychiatrists	242	?	117
Cannon	26(2):351-366	SZB	Aggregation	Schizophrenia	Epidemiological sample	Students, Psychiatrist, PhD	72	63	7941
Saoud	26(4):893-902	SZB	Aggregation	Schizophrenia	Outpatients	Psychiatrists	18	18	15

[a] ACPR = Alcohol Clin Exp Res; AJP = Am J Psych; AGP = Arch Gen Psychiatry; MP = Molec Psychiat; NP = Neuropsych genet; NL = Neurosc Letter; SR = Schizophrenia Res; SC = Science; SZB = SZ Bulletin.

[b] EP = Experimental procedure; DE = Direct examination; FI = Face-to-face interview; TI = Telephone Interview; SQ = Self-questionnaire; HR = Hospitalization registry; Q = Questionnaire; I = Interview; R = Retrospective; II = Indirect interview; PS = Postal survey; DI = Direct interview; T = Test; CR = Charts review.

(continued)

Table 1

	Psychiatric assessment					
Eval-uation[b]	Proband	Relatives	Controls	Main instrument	Criteria	Country
?	?	0	none	?	DSM-III-R or ICD -10	Japan (Tokyo)
I	SADS or DIGS	0	none	SADS	RDC	Various
?	?	?	?	?	?	USA (New York)
I	interview		Interview	?	DSM-IV	Japan (Tokyo)
II	indirect interview	0	none	CAPA	DSM-IV and ICD-10	UK (Manchester)
I	K-SADS	SADS	0	K-SADS	DSM-IV	USA (LA)
II	PICS, TTI	none	0	PICS, TTI	DSM-IV	Canada (Toronto)
PS	questionnaire (STS,I-T-C)	0	0	STSI	0	Australia
I	PICS, TTI	0	0	PICS, TTI	DSM-IV	Canada (Toronto)
I	?	0	none	?	DSM-IV	UK (Cardiff)
I	SADS	0	SCID	SADS	DSM-III-R	France (Paris)
EP	ASS	0	0	ASS	0	Hongary (Budapest)
I	SADS, SCID, OPCRIT	SADS, SCID, OPCRIT	0	SADS, SCID, OPCRIT	RDC	Germany (Munich)
I	Interview	Interview	0		DSM-IV	USA (Maryland)
I	CASH	none	0	CASH	DSM-IV or DSM-III-R	USA (Iowa)
0	0					UK (Bangor)
?	?		?	?	DSM-III-R	China (Shangai)
DI	ADI	ADI		ADI	?	UK (London)
DI	SADS-LA	SADS + FISC	0	SADS	DSM-III-R	USA (New York)
DI	SADS			SADS	RDC	UK (Cambridge)
DI	OPCRIT		OPCRIT		DSM-III-R	Itay (Milan)
Q	EPQ-R			EPQ-R	0	Australia (Camberlia)
I	PICS, TTI	0	0	PICS, TTI	DSM-IV	Canada (Toronto)
?	?	?	?	?	DSM-IV	Japan (Kitukyushu)
DI	SADS or SCAN			SADS or SCAN	RDC and DSM-IV	Belgium (Brussel)
Q	Questionnaire	0	0	NEO-PI-R	0	USA (Bethesda)
DI	SADS-L			SADS-L	DSM-IV	Germany (Bonn)
DI	SADS & DIGS	SADS & DIGS	0	SADS & DIGS	DSM-III-R	USA (New York)
DI	KSADS			KSADS	DSM-IV	Israel (Jerusalem)
II		PICS-IV		PICS-IV	DSM-IV	Canada (Ontario)
?	CBCL	0	0	CBCL	DSM-IV	Ireland (Dublin)
DI	SCID	0	0	SCID	DSM-IV	Canada (Toronto)
DI	?		?	?	DSM-IV	Israel (Petak Tikvah)
?	CBCL	0	0	CBCL	DSM-IV	Ireland (Dublin)
DI	K-SADS			K-SADS	DSM-IV	Turkey (Istanbul)
DI	?			?	DSM-III-R	China (Shangai)
TI	PIAT test	PIAT test		PIAT test	DSM-IV	USA (Colorado)
DI	?	?	?	?	DSM-IV	Australia (Brisbane)
?	?		?		DSM-III-R	USA (LA)
I	SSAGA	SSAGA	SSAGA	SSAGA	DSM-III-R	USA
I	SCID		SCID	SCID	DSM-III-R	USA (Pittsburgh)
DI	SCID			SCID	DSM-IV	USA (Connecticut)
?	DUS +/- DIS		?	DIS	DSM-III-R	USA (Baltimore)
DI	DISC, CIDI-SAM	DISC, CIDI-SAM		CIDI	DSM-IV	USA (Colorado)
?	?		0	?	DSM-IV	Japan (Ibaraki)
DI	?	0	0	?	?	USA (Bethesda)
?	?	?	?	?	DSM-IV	Italy (Rome)
DI	DIS	0	DIS	DIS	DSM-III	Taiwan (Taipei)
DI	SADS-L & OPCRIT		?	SADS-L	DSM-IV	Germany (Bonn)
DI	SADS-L	SADS-L	SADS-L	SADS-L	RDC	Canada (Halifax)
DI	SCID		SCID	SCID	DSM-IV	Sweden (Stockholm)
Q	HAMD			HAMD	DSM-IV	
DI	VABS & CARS			VABS & CARS	DSM-IV	France (Bordeaux)
DI	DIGS & OPCRIT	DIGS		DIGS & OPCRIT	DSM-III-R	Portugal (Coumbra)
DI	0	0	0	0	ICD-10	Kuwait
?	?	?	?	?	?	Canada (Toronto)
?	?		?	?	?	UK (London)
DI	BPRS		?	BPRS	DSM-IV	Taiwan (Taipei)
DI	SCID I and II	SCID 1 and II	0	SCID I and II	DSM-III-R	USA (Newark)
DE	DIS +/- SCID	FIGS	medical records	SCID	DSM-IV	USA (Boston)
CR	Charts review	Charts review	charts review	0	DSM-IV	USA (Philadelphia)
I	SADS	SADS	SADS	SADS	DSM-IV	France (Lyons)

large variety of interviews available nevertheless justifies at least some choices. This chapter emphasizes the importance of high-quality clinical assessment for genetic studies of psychiatric disorders, for reasons of reliability, as well as the role of clinical diversity in the assessment of psychiatric disorders, for reasons of sensitivity. These two problems (collecting reliable and detailed material) may seem contradictory. For example, the use of clinical interviewers may be more expensive, and testing inter-rater and test-rest agreement may be time-consuming. Furthermore, collecting a large amount of data on phenotypes (and endophenotypes) encourages multiple testing and increases the number of subgroups, which means that the size of the total sample must be increased accordingly.

However, results indicate that collecting high-quality clinical data is essential. A focus on specific clinical symptoms, syndromes, or clusters may be more directly related to biological findings—for example, hallucination in schizophrenia and dopaminergic transmission. Furthermore, the number and severity of symptoms may be considered as quantitative traits analyzable through a quantitative trait loci (QTL) approach, which is statistically more powerful than qualitative analyses. Besides, the risk of chance finding may be decreased by the use of different methodological or statistical techniques, such as sharing the sample in two looking for internal replication, limiting the tests to post-hoc analyses, or measuring the impact of the association with attributable risk. Finally, the investment in time and money in clinical assessment more directly reflects our lack of knowledge of the phenotype boundaries of the psychiatric disorders presently considered.

A compromise must be achieved between limiting the analyses to the precise initial hypothesis and the possibility to further define the phenotype involved and its related traits.

References

1. Endicott, J. (2001) Good diagnoses require good diagnosticians: collecting and integrating the data. *Am. J. Med. Genet. (Neuropsychiatr. Genet.)* **105,** 48–49.
2. Tsuang, M., Faraone, S., and Lyons, M. (1993) Identification of the phenotype in psychiatric genetics. *Eur. Arch. Psychiatry Clin. Neurosci.* **243,** 131–142.
3. Bucholz, K.K., Cadoret, R., Cloninger, C.R., Dinwiddie, S.H., Hesselbrock, V.M., Nurnberger, J.I., et al. (1994) A new, semi-structured psychiatric interview for use in genetic linkage studies: a report on the reliability of the SSAGA. *J. Stud. Alcohol.* **55,** 149–158.
4. Crocq, M.A., Mant, R., Asherson, P., Williams, J., Hode, Y., Mayerova, A., et al. (1992) Association between schizophrenia and homozygosity at the dopamine D3 receptor gene. *J. Med. Genet.* **29,** 858–860.
5. Nanko, S., Fukuda, R., Hattori, M., Sasaki, T., Dai, X.Y., Yamaguchi, K., et al. (1994) Further evidence of no linkage between schizophrenia and the dopamine D3 receptor gene locus. *Am. J. Med. Genet.* **15,** 264–267.
6. Kennedy, J.L., Billett, E.A., Macciardi, F.M., Verga, M., Parsons, T.J., Meltzer, H.Y., et al. (1995) Association study of dopamine D3 receptor gene and schizophrenia. *Am. J. Med. Genet.* **60,** 558–562.
7. Griffon, N., Crocq, M.A., Pilon, C., Martres, M.P., Mayerova, A., Uyanik, G., et al. (1996) Dopamine D3 receptor gene: organization, transcript variants, and polymorphism associated with schizophrenia. *Am. J. Med. Genet.* **67,** 63–70.
8. Williams, J., Spurlock, G., Holmans, P., Mant, R., Murphy, K., Jones, L., et al. (1998) A meta-analysis and transmission disequilibrium study of association between the dopamine D3 receptor gene and schizophrenia. *Mol. Psychiatry* **3,** 141–149.
9. Dubertret, C., Gorwood, P., Adès, J., Feingold, J., Schwartz, J.C., Sokoloff, P. (1998) A meta-analysis of DRD3 gene and schizophrenia: ethnic heterogeneity and significant association in caucasians. *Am. J. Med. Genet. (Neuropsychiat. Genet.*) **81,** 318–322.
10. Steen, V.M., Lovlie, R., MacEwan, T., and McCreadie, R.G. (1997) Dopamine D3-receptor gene variant and susceptibility to tardive dyskinesia in schizophrenic patients. *Mol. Psychiatry* **2,** 139–145.
11. Segman, R.H., Heresco-Levy, U., Finkel, B., Inbar, R., Neeman, T.,

Schlafman, M., et al. (2000) Association between the serotonin 2C receptor gene and tardive dyskinesia in chronic schizophrenia: additive contribution of 5-HT2Cser and DRD3gly alleles to susceptibility. *Psychopharmacology (Berl)* **152,** 408–413.

12. Eichhammer, P., Albus, M., Borrmann-Hassenbach, M., Schoeler, A., Putzhammer, A., Frick, U., et al. (2000) Association of dopamine D3-receptor gene variants with neuroleptic induced akathisia in schizophrenic patients: a generalization of Steen's study on DRD3 and tardive dyskinesia. *Am. J. Med. Genet.* **96,** 187-191.
13. Segman, R., Neeman, T., Heresco-Levy, U., Finkel, B., Karagichev, L., Schlafman, M., et al. (1999) Genotypic association between the dopamine D3 receptor and tardive dyskinesia in chronic schizophrenia. *Mol. Psychiatry* **4,** 247–253.
14. Basile, V.S., Masellis, M., Badri, F., Paterson, A.D., Meltzer, H.Y., Lieberman, J.A., et al. (1999) Association of the MscI polymorphism of the dopamine D3 receptor gene with tardive dyskinesia in schizophrenia. *Neuropsychopharmacology* **21,** 17–27.
15. Browne, S., Clarke, M., Gervin, M., Lane, A., Waddington, J.L., Larkin, C., et al. (2000) Determinants of neurological dysfunction in first episode schizophrenia. *Psychol. Med.* **30,** 1433–1441.
16. Dubertret, C., Gorwood, P., Gouya, L., Deybach, J.C., Feingold, J., and Adès, J. (2001) Association and excess of transmission of a DRD2 haplotype in a sample of French schizophrenic patients. *Schizophr. Res.* **49(1–2):** 203–212.
17. Van Tol, H.H.M., Bunzow, J.R., Guan, H.C., et al. (1991) Cloning of the gene for human D4 receptor with high affinity for the antipsychotic clozapine. *Nature* **350,** 610–619.
18. Serretti, A., Macciardi, F., Catalano, M., Bellodi, L., and Smeraldi, E. (1999) Genetic variants of dopamine receptor D4 and psychopathology. *Schizophr. Bull.* **25,** 609–618.
19. Kaiser, R., Konneker, M., Henneken, M., Dettling, M., Muller-Oerlinghausen, B., Roots, I., et al. (2000) Dopamine D4 receptor 48-bp repeat polymorphism: no association with response to antipsychotic treatment, but association with catatonic schizophrenia. *Mol. Psychiatry* **5,** 418–424.
20. Meyer, J., Huberth, A., Ortega, G., Syagailo, Y.V., Jatzke, S., Mossner, R., et al. (2001) A missense mutation in a novel gene encoding a putative cation channel is associated with catatonic schizophrenia in a large pedigree. *Mol. Psychiatry* **6,** 302–306.

21. Noble, E.P., Zhang, X., Ritchie, T.L., and Sparkes, R.S. (2000) Haplotypes at the DRD2 locus and severe alcoholism. *Am. J. Med. Genet.* **96,** 622-631.
22. Sander, T., Harms, H., Podschus, J., Finckh, U., Nickel, B., Rolfs, A., et al. (1997) Allelic association of a dopamine transporter gene polymorphism in alcohol dependence with withdrawal seizures or delirium. *Biol. Psychiatry* **41,** 229–304.
23. Schmidt, L.G., Harms, H., Kuhn, S., Rommelspacher, H., and Sander, T. (1998) Modification of alcohol withdrawal by the A9 allele of the dopamine transporter gene. *Am. J. Psychiatry* **155,**474–478.
24. Gorwood, P., Batel, P., Adès, J., Hamon, M., and Boni, C. (2000) Serotonin transporter gene polymorphisms, alcoholism and suicidal behaviour. *Biol. Psychiatry* **48,** 259–264.
25. Saccone, N.L., Kwon, J.M., Corbett, J., Goate, A., Rochberg, N., Edenberg, H.J., et al. (2000) A genome screen of maximum number of drinks as an alcoholism phenotype. *Am. J. Med. Genet.* **96,** 632–637.
26. Lappalainen, J., Long, J.C., Eggert, M., Ozaki, N., Robin, R.W., Brown, G.L., et al. (1998) Linkage of antisocial alcoholism to the serotonin 5-HT1B receptor gene in 2 populations. *Arch. Gen. Psychiatry* **55,** 989–994.
27. Eaton, W., Neufeld, K., Chen, L.S., and Cai, G. (2000) A comparison of self-report and clinical diagnostic interviews for depression. *Arch. Gen. Psychiatry* **57,** 217–222.
28. American Psychiatric Association. (2000) *Handbook of Psychiatric Measures*. American Psychiatric Association, ed., Washington, DC.
29. Nurnberger, J.I., Blehar, M.C., Kaufmann, C.A., York-Cooler, C., Simpson, S.G., Harkavy-Friedman, J., et al. (1994) Diagnostic Interview for Genetic Studies: rational, unique features, and training. *Arch. Gen. Psychiatry* **51,** 849–862.
30. Bucholz, K.K., Cadoret, R., Cloninger, C.R., Dinwiddie, S.H., Hesselbrock, V.M., Nurnberger, J.I., Jr., et al. (1994) A new, semi-structured psychiatric interview for use in genetic linkage studies: a report on the reliability of the SSAGA. *J. Stud. Alcohol.* **55,** 149–158.
31. Hesselbrock, M., Easton, C., Bucholz, K., Schuckit, M., and Hesselbrock, V. (1992) A validity study of the SSAGA-a comparison with the SCAN. *Addiction* **94,** 1361–1370.
32. Spitzer, R.L., Williams, J.B., Gibbon, M., and First, M.B. (1992) The Structured Clinical Interview for DSM-III-R (SCID). I: History, rationale, and description. *Arch. Gen. Psychiatry* **49,** 624-629.

33. First, M., Spitzer, R., Williams, J., et al. (1995) Structured clinical interview for DSM-IV (SCID-I) (User's guide and interview) research version. Biometrics Research Department, New York Psychiatric Institute, New York, NY.
34. First, M., Spitzer, R., Williams, J., et al. (1997) Structured clinical interview for DSM-IV-Clinician Version (SCID-CV) (User's guide and interview) American Psychiatric Press, Washington, DC.
35. Williams, J.B., Gibbon, M., First, M.B., Spitzer, R.L., Davies, M., Borus, J., et al. (1992) The Structured Clinical Interview for DSM-III-R (SCID). II. Multisite test-retest reliability. *Arch. Gen. Psychiatry* **49,** 630–636.
36. Wing, J.K., Babor, T., Brugha, T., Burke, J., Cooper, J.E., Giel, R., et al. (1990) SCAN. Schedules for Clinical Assessment in Neuropsychiatry. *Arch. Gen. Psychiatry* **47,** 589–593.
37. Farmer, A., Cosyns, P., Leboyer, M., et al. (1993) A SCAN-SADS comparison study of psychotic subjects and their first-degree relatives. *Eur. Arch. Psychiatry Clin. Neurosci.* **242,** 352–356.
38. Endicott, J. and Spitzer, R.L. (1978) A diagnostic interview: the schedule for affective disorders and schizophrenia. *Arch. Gen. Psychiatry* **35,** 837–844.
39. Andreasen, N.C., Flaum, M., and Arndt, S. (1992) The Comprehensive Assessment of Symptoms and History (CASH). An instrument for assessing diagnosis and psychopathology. *Arch. Gen. Psychiatry* **49,** 615–623.
40. Robins, L.N., Wing, J., Wittchen, H.U., Helzer, J.E., Babor, T.F., Burke, J., et al. (1988) The Composite International Diagnostic Interview. An epidemiologic instrument suitable for use in conjunction with different diagnostic systems and in different cultures. *Arch. Gen. Psychiatry* **45,** 1069–1077.
41. Robins, L.N., Helzer, J.E., Croughan, J., and Ratcliff, K.S. (1981) National Institute of Mental Health Diagnostic Interview Schedule. Its history, characteristics, and validity. *Arch. Gen. Psychiatry* **38,** 381–389.
42. Frances, A. (1998) Problems in defining clinical significance in epidemiological studies. *Arch. Gen. Psychiatry* **55,** 119.
43. Wittchen, H.U., Robins, L.N., Cottler, L.B., Sartorius, N., Burke, J.D., and Regier, D. (1991) Cross-cultural feasibility, reliability and sources of variance of the Composite International Diagnostic Interview (CIDI). The Multicentre WHO/ADAMHA Field Trials. *Br. J. Psychiatry* **159,** 645–658.

44. Spitzer, R.L., Williams, J.B., Kroenke, K., Linzer, M., deGruy, F.V., Hahn, S.R., et al. (1994) Utility of a new procedure for diagnosing mental disorders in primary care. The PRIME-MD 1000 study. *JAMA* **272,** 1749–1756.
45. Broadhead, W.E., Leon, A.C., Weissman, M.M., Barrett, J.E., Blacklow, R.S., Gilbert, T.T., et al. (1995) Development and validation of the SDDS-PC screen for multiple mental disorders in primary care. *Arch. Fam. Med.* **4,** 211–219.
46. Lecrubier, Y., Sheehan, D., Weiller, E., et al. (1997) The Mini-International Neuropsychiatric Interview (MINI): a short diagnostic structured interview: reliability and validity according to the CIDI. *Eur. Psychiatry* **12,** 224–231.
47. Sheehan, D.V., Lecrubier, Y., Sheehan, K.H., Amorim, P., Janavs, J., Weiller, E., et al. (1998) The Mini-International Neuropsychiatric Interview (M.I.N.I.): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. *J. Clin. Psychiatry* **59S,** 22–33.
48. McGuffin, P., Farmer, A., and Harvey, I. (1991) A polydiagnostic application of operational criteria in studies of psychotic illness. Development and reliability of the OPCRIT system. *Arch. Gen. Psychiatry* **48,** 764–770.
49. Williams, J., McGuffin, P., et al. (1996) A multicenter reliability study of the OPCRIT computerised diagnostic system. *Psychol. Medicine* **26(4),** 775–783.
50. Craddock, M., Asherson, P., Owen, M.J., Williams, J., McGuffin, P., and Farmer, A.E. (1996) Concurrent validity of the OPCRIT diagnostic system. Comparison of OPCRIT diagnoses with consensus best-estimate lifetime diagnoses. *Br. J. Psychiatry* **169,** 58–63.
51. Azevedo, M.H., Soares, M.J., Coelho, I., Dourado, A., Valente, J., Macedo, A., et al. (1999) Using consensus OPCRIT diagnoses. An efficient procedure for best-estimate lifetime diagnoses. *Br. J. Psychiatry* **175,** 154–157.
52. Hasin, D.S., Trautman, K.D., Miele, G.M., Samet, S., Smith, M., and Endicott, J. (1996) Psychiatric Research Interview for Substance and Mental Disorders (PRISM): reliability for substance abusers. *Am. J. Psychiatry* **153,** 1195–1201.
53. Andreasen, N.C., Endicott, J., Spitzer, R.L., and Winokur, G. (1977) The family history method using diagnostic criteria. Reliability and validity. *Arch. Gen. Psychiatry* **34,** 1229–1235.

54. Chapman, T.F., Mannuzza, S., Klein, D.F., and Fyer, A.J. (1994) Effects of informant mental disorder on psychiatric family history data. *Am. J. Psychiatry* **151,** 574–579.
55. Rice, J.P., Reich, T., Bucholz, K.K., Neuman, R.J., Fishman, R., Rochberg, N., et al. (1995) Comparison of direct interview and family history diagnoses of alcohol dependence. *Alcohol Clin. Exp. Res.* **19,** 1018–1023.
56. Weissman, M., et al. (2000) Brief screening for family psychiatric history. The family history screen. *Arch. Gen. Psychiatry* **57,** 675–682.

5

Genetic Association Studies

Definition of Cases and Controls

Frank Bellivier

1. Introduction

The case-control association study is one of the most frequently used study designs for investigating the effect of environmental risk factors on disease pathogenesis. This study design is also now increasingly used to investigate the influence of genetic variability on disease susceptibility. This type of study has a number of theoretical advantages over linkage studies. Association studies do not require knowledge of the genetic model of the disease, and the phenotype being studied is controlled, making it possible to study homogeneous subgroups that are selected on the basis of a candidate symptom. These studies can also detect genes that contribute only a small proportion of the overall likelihood of developing the disorder, which would require prohibitively large samples for detection in linkage studies.

Yet association studies do not have the best of images. They have been criticized for non-replication, which may be attributed to chance findings, bias, and confounding. Such criticism is often well-founded—especially in many early studies performed on small num-

From: *Methods in Molecular Medicine, vol. 77: Psychiatric Genetics: Methods and Reviews*
Edited by: M. Leboyer and F. Bellivier

bers of poorly characterized cases compared with inappropriate groups of controls, who were often not screened, incorrectly matched with patients, and recruited in the absence of strict criteria from the medical staff or friends of the researchers. A new family-based study design that considers some of these problems has been proposed—in particular, the transmission-disequilibrium test. However, the traditional format of case-control studies may be improved by the incorporation of refinements developed over the past 10 yrs in genetic epidemiology and in the phenotypic analysis of psychiatric illnesses (*see* **ref. *1***). Environmental epidemiology has also provided key methodological tools. With these methodological improvements, the case-control study remains a powerful tool for identifying genetic vulnerability factors for complex diseases. The ability of association studies to contribute to the identification of genetic vulnerability factors for behavioral phenotypes in psychiatry will depend on the incorporation of these methodological refinements. This chapter provides an overview of these critical issues and guidelines for genetic association analyses, with a focus on the selection of cases and controls.

2. The Case-Control Study: Definition, Advantages, and Limitations

2.1. Definition

The basic principle underlying case-control studies is the comparison of a group of individuals with a disease to an unaffected group of individuals. Association with a factor is observed if the frequency of this factor differs in the two populations. Statistical tests, with a given risk of false-positive findings, can be used to assess the significance of differences between the two populations for a particular factor. If exposure to this factor is more frequently observed in the case population, it is referred to as a vulnerability factor. The odds ratio of exposure provides an estimate of the risk ratio. The risk ratio is a measure of how much more likely the disease is to occur in an individual exposed to the factor than in an individual who is not exposed to the factor. Only prospective cohort

studies can calculate the true risk ratio. The deviation of the odds ratio from the true underlying risk ratio depends on three contributing factors: bias, confounding, and chance ***(2)***. The assessment of these limitations differs according to whether the exposure variable is genetic or environmental.

2.2. Genetic Case-Control Studies: Limitations

Bias in genetic case-control studies of complex diseases arises from the selection of cases and controls and the accuracy of the information obtained from them. This type of bias is known as ascertainment bias. It results in the dilution of the case group by misclassified controls, or of the control group by misclassified cases. Bias in genetic case-control studies may also arise from ethnic differences between the two groups compared. This type of bias is known as stratification bias, and is a major problem in association studies.

Confounding has been investigated mostly for environmental risk factors, as many lifestyle exposures are correlated. An equivalent situation exists in genetic case-control studies because of linkage disequilibrium (LD). If association is observed with a polymorphic marker, the functional variant may be located either at the polymorphic site under study or at another polymorphic site somewhere else in the same gene or in a nearby gene.

This chapter explores ways of dealing with ascertainment bias in the subsection concerning sampling procedures for patients and controls. Stratification bias, confounding, and the role of chance are discussed in the section on the comparison of patients and controls.

2.3. Genetic Case-Control Studies: Advantages

Case-control studies are less expensive than cohort studies, and have greater power if the disease has a low frequency or the genetic risk factors have low penetrance. In addition, the case-control design is of particular value in genetic epidemiology because genetic risk factors do not change with time, are not affected by disease status or progression, and are easier than environmental factors to measure retrospectively ***(3)***.

Genetic association studies are more effective than family studies (linkage or affected sib-pair studies) at detecting genetic vulnerability factors that have only a small effect. They also have the advantage of controlling for the disease phenotype under study. This is a critical issue in psychiatric research, because the genetic validity of psychiatric diagnosis is unclear (with poor phenotype-genotype correlation)—as is the distinction between affected and non-affected individuals—and because etiologic and genetic heterogeneity are very likely. Thus, the study of subgroups is a valid choice for the investigator because it may reduce the underlying genetic heterogeneity, particularly if patients are selected on the basis of a candidate symptom (*see* Chapter 3). In addition, the interaction between several genetic markers and between environmental and genetic factors can be investigated in case-control studies.

Case-control studies including large numbers of subjects facilitate *a posteriori* analyses. For example, if an association with a candidate gene is found, the phenotypic characteristics of the patients who carry the vulnerability genotype can then be studied. This strategy may facilitate the identification of candidate symptoms with good genetic validity. *A posteriori* analyses of conflicting association results for bipolar affective disorder and the tryptophan hydroxylase gene polymorphism have revealed that suicidal behavior is a potentially strong candidate symptom ***(4,5)***—a finding confirmed by subsequent analyses ***(6–9)***.

Case-control studies have been criticized because of the high risk of false-positive and false-negative results, which leads to conflicting results. As conflicting results may arise from heterogeneity in sampling strategies, in the clinical profiles of patients, and in population genetic history, the analysis of conflicting data may be a useful way to reveal the etiologic heterogeneity of the disease.

3. Selection of Cases and Controls

The choice of study populations from which cases and controls are sampled may reflect some practical constraints. However, this represents a major decision for the investigators, who may decide to

use a specific criterion for the sampling procedure for a specific study. The use of strict inclusion criteria for cases and controls is a basic principle for preventing dilution of the case group by misclassified controls or of the control group by misclassified cases. Such misclassification errors are particularly common in psychiatry.

3.1. Selection of Cases

Both the selection of the study subjects from the underlying study population and the quality of the information obtained from them may be sources of bias.

The cases selected should correspond to all the individuals within a population who develop a disease, or a representative sample of them ***(10)***. Practical constraints often lead to the use of more convenient groups. Essentially, the researcher must decide whether to select incident or prevalent cases, each of which has various limitations and advantages. Incident cases, defined as new affected subjects who appear during a certain period of time, have a recent onset of disease. Prevalent cases are defined as all affected subjects at a given time. Incident and prevalent cases are often mixed in case-control studies. However, incident cases offer strong advantages over prevalent cases. Prevalent cases tend to be biased toward those with diseases of longer duration. In psychiatry, this is a critical point, because patients who commit suicide or those with long-term remission (good lithium response in some bipolar patients, for example) are undersampled when prevalent cases are used. The mode of recruitment may increase such biases in prevalent cases. For example, hospital recruitment of prevalent cases may oversample patients with a poor response to treatment, comorbid conditions, or poor social resources leading to frequent rehospitalization, whereas patients with suicidal behavior may be undersampled. Several of these characteristics (suicide, disease duration, treatment response, and comorbidity) are likely to be associated with the candidate gene rather than the disease itself. Survival has been suggested to bias prevalent cases in the association between the ApoE ε4 allele and Alzheimer's disease, as the ApoE ε2 allele is believed to be a

vulnerability factor for cardiovascular disease *(11)*. However, it should be emphasized that in most psychiatric diseases, incident cases are often associated with uncertainties surrounding the diagnosis. Indeed, confidence in the diagnosis increases with progression of the disease.

The selection of cases on the basis of candidate symptoms with good genetic validity may also be considered. Narrowing the definition of a phenotype through the use of such characteristics as severity, age at onset, associated symptoms, or family history, has already been fruitful in genetic studies of complex somatic diseases such as colon cancer *(12)*, hypertension *(13)*, and coronary heart disease *(14)*, making a significant contribution to the identification of relevant genes. For example, subdivision according to age of onset and mode of inheritance has been particularly effective in clarifying genetic heterogeneity in Alzheimer's disease *(15)*. In psychiatry, subdivision according to age of onset, severity, and family history may also facilitate the identification of more homogeneous subtypes. Early onset is associated with increased familial risk in schizophrenia *(16)*, major depressive disorder *(17)*, and obsessive compulsive disorder *(18)*. In bipolar affective disorder, clinical and familial heterogeneity according to age at onset have been demonstrated *(19,20)*, as well as the relevance of this phenotypic indicator for genetic studies *(21–23)*.

3.2. Selection of Controls

The selection of control groups for case-control studies is one of the most difficult tasks. Traditionally, the control group should be randomly sampled from the general population. More precisely, controls are subjects who are not ill, but if they were affected by the disease under study, would be eligible as cases *(24)*. As stated by these authors, this definition of the controls, offers the advantage of emphasizing that disease status does not affect (or is not associated with) the sampling procedure.

The use of controls screened for a personal or family history of the disorder under study has been advocated for case-control stud-

ies in psychiatry, in order to reduce the risk of misclassification (affected subjects or subjects carrying the vulnerability factors included in the control group) ***(25,26)***. Indeed, uncertainties surrounding the genetic validity of psychiatric nosology make it difficult to define the boundary between affected and non-affected subjects. Psychiatric disorders occur frequently, and are often comorbid. Genetic vulnerability factors of low penetrance for psychiatric disorders are thus likely to be very frequent in non-affected subjects. In addition, standardized interviews of control subjects may detect histories of psychiatric disorders that are not spontaneously reported by subjects because of the low social desirability of psychiatric diseases.

Additional evidence for the importance of defining criteria for control selection is provided by meta-analyses of association studies with candidate genes, which have shown homogeneous allele frequencies between centers for patients and heterogenous allele frequencies for controls. This has been demonstrated in association studies that examine the association of bipolar affective disorder with the monoamine oxidase A gene polymorphism ***(27)*** and for the tyrosine hydroxylase gene polymorphism ***(25)***. This finding clearly suggests that the lack of strict inclusion criteria for the controls is associated with heterogeneity in allele frequency.

All these arguments suggest that screened controls with the following characteristics should be used: (1) same genetic background as the affected subjects, (2) older than the mean age at onset plus one standard deviation for the disease under study, (3) with no personal or family history of the disease investigation (spectrum).

However, it is important to note that the use of such "super-normal" controls may result in a different interpretation of the results of the association study: the existence of a difference between patients and controls may reveal a genetic factor with a protective effect. Indeed, interpretation is sometimes impossible. If patients with suicidal behavior are compared with controls selected on the basis of their having no personal or family history of any psychiatric disorder or of suicidal behavior, then differences between patients and controls cannot be interpreted.

Table 1
The Risk of Including Vulnerability Factor Carriers in the Control Group

Psychiatric disorder concealed because of its low social desirability
Frequent and low penetrant genetic vulnerability factors, resulting in a high frequency in non-affected subjects
Poorly defined boundary between affected and non-affected subjects in psychiatry

4. Comparison of Cases and Controls

The interpretation of the results of a genetic case-control study is limited by several factors. This section presents a review of these problems and includes a short description of potential solutions.

4.1. Stratification

For genetic studies, stratification bias is mainly related to the ethnic origins of the two groups. Patients and controls should clearly originate from the same population. If not, there may be genetic differences between patients and controls at many loci because of the "genetic distance" between the two populations ***(28)***. This makes it difficult to prove that a difference observed for only one polymorphism reflects a causal impact on the pathogenesis of the disease rather than a simple overall population-level genetic difference between patients and controls ***(29)***. Historical information concerning the place of birth of family members over at least one generation is a minimal criterion for matching the two groups. If immigration is suspected, patterns of family migration over several generations may be very useful in defining subgroups ***(30)***.

Even if ethnic inclusion criteria are used, case-control studies remain subject to stratification bias, which may conceal a true relationship or lead to a false relationship between the genetic vulnerability factor tested and the disease under study. Several methods in addition to the use of strict ethnic inclusion criteria have been proposed to deal with the risk of stratification bias. Patients and con-

trols can be typed for a number of random, anonymous, and biologically inert polymorphisms, to test whether there is a "genetic distance" between cases and controls ***(29,31)***. This method can be used to assess and control for possible stratification bias, making it possible to perform an unbiased test ***(32)***.

4.2. Genotype Sensitivity and Specificity

Little attention has been given to the sensitivity and specificity of various genotyping methods, and problems may occur if the gene frequencies of cases are compared to those of controls determined in the past by inferior methods ***(10,33)***.

4.3. Locus Heterogeneity, Allele Heterogeneity, and Haplotype Diversity

A complex disease may have several independent genetic determinants, which may be polymorphisms, mutations, or defective genes. It is believed that, to express the disease or the trait, an individual must have a certain number of these genetic components. This may give rise to genetic subgroups among cases. Such subgrouping may result from "locus heterogeneity," with the various polymorphisms located at different sites around the genome. Alzheimer's disease provides a good example of "locus heterogeneity," as the Mendelian subgroups of the disease have been shown to be linked to point mutations in the Amyloid Precursor Protein gene, the Presenilin 1 gene, or the Presenilin 2 gene. Subgrouping may also result from "allelic heterogeneity," in which multiple mutations at the same locus segregate in different populations. Alzheimer's disease also provides a good example of "allelic heterogeneity," as several different mutations have been identified in the Presinilin 1 gene. Allelic heterogeneity may generate situations in which various alleles are associated with the disease in different populations, rather than the existence of a single, very specific allele. In such situations, only haplotype analysis is likely to detect the association, and several haplotypes have a higher frequency among

cases than controls, because each of these haplotypes may represent the "signature pattern" of alleles surrounding a locus that harbors a disease allele ***(29)***. Haplotype analysis is also a powerful tool if the polymorphisms under study display only weak or moderate linkage disequilibrium with the unknown functional variant.

4.4. The Role of Chance

A difference may also be observed between patients and controls because of chance. A major disadvantage of the candidate gene approach in psychiatry is that all genes expressed in the brain are potential candidates. Indeed, too little is currently known about the biology of psychiatric diseases for strong candidate genes to be proposed. In addition, very little is known about DNA polymorphisms in these candidate genes that may confer functional differences. Therefore, it is important to note that polymorphisms in exons, resulting in amino acid changes, are not the only sequences of interest. Non-coding regions may also influence the level of expression of an unchanged protein (promoter) or a modified protein (intron, alternative splicing). The testing of several polymorphisms in a given population increases the risk of chance positive findings.

4.5. Multiple Testing: To Correct or Not to Correct?

Both the testing of several polymorphisms and subgroup analyses increase the number of tests performed in a given population. Correction for multiple testing (lowering the level of significance) has been recommended to reduce the risk of false-positive findings (type I errors). Unfortunately, reducing type I errors increases the risk of allowing an existing association to go undetected (type II errors). For exploratory analyses, it is preferable to adopt a policy of non-adjustment for multiple testing, using replication as a confirmatory criterion ***(34)***. The risks associated with exploring hypotheses that may turn out to be wrong are less serious than those of missing potentially important findings.

5. Transmission/Disequilibrium Test (TDT) vs Case-Control Design

The case-control design has long been criticized because of its lack of reproducibility, which is believed to be a result of the high risk of chance findings, bias, and confounding. TDT has been recommended as the method of choice, mainly because this study design is less prone to false-positives arising from inadequate case-control matching, which may occur with population stratification. The other limitations of case-control studies described in this chapter (the choice of candidate genes, the risk of false-positives resulting from multiple testing, the variable LD between the functional variant and the polymorphism under study) also apply to the TDT design. In particular, TDT studies do not overcome the problem of ethnic differences in disease etiology or allelic association caused by tight linkage. As much as possible, parent-offspring trios should be drawn from an ethnically "homogeneous" population; if not, power is lost in this study design. Thus, the selection of ethnically heterogeneous populations increases the risk of false-positive findings in case-control studies (particularly if cases and controls have different ethnic backgrounds) and the risk of false-negative results in TDT studies.

TDT studies also have a few limitations in addition to those that apply to case-control studies. In this design, only probands with at least one heterozygous parent can be included in the analysis. The informativity of the genetic marker under study therefore depends on allele frequencies. In addition, simulation studies have clearly demonstrated that the total number of cases (not only those with at least one heterozygous parent) needed to achieve a given level of power is, in most situations, smaller for case-control studies than for TDT studies ***(30)***. TDT analysis also requires the recruitment of parents, which may be difficult or even impossible, as in late-onset diseases.

Selection bias is very likely to occur when using trios, as patients with two participating parents may differ from patients with no participating parents. Indeed, case-control and trio designs may

Table 2
Selection Bias in Trio Studies

Undersampled:
Patients with late age of onset
Patients with one or two affected parents (if the disease is associated with a high mortality rate [suicide, alcohol or drug use, personality characteristics influencing risk-taking behavior])
Parents of advanced age at the patient's birth
In cases in which the disease is associated with a gender difference in fertility, patients with the low-fertility gender-affected parent will be undersampled

potentially include patients who differ in both the genetic and environmental risk factors involved (summarized in **Table 2**).

Thus, traditional case-control studies remain a powerful and easily achieved alternative to TDT studies in some situations ***(30)*.**

6. Conclusion

Several limitations plague the interpretation of positive or negative association results in psychiatry genetics. This chapter reviews these limitations and proposes some methodological guidelines to reduce the risk of false-positive and false-negative findings. The usefulness of this strategy depends on the selection criteria used for cases and controls, the selection of true candidate genes, the analysis of LD with other nearby polymorphisms, haplotype construction, and ideally, measurement of the "genetic distance" between cases and controls before the association test. With these methodological keys, case-control association studies remain a powerful tool for detecting genetic vulnerability factors of moderate or small effect, using the candidate gene strategy. Case-control studies should also be considered as an exploratory tool for narrowing down the region of interest after preliminary mapping of a gene by linkage study. As psychiatric diseases are likely to be heterogeneous, such studies may also help to narrow the phenotype of interest for genetic studies by identifying the clinical characteristics of patients who carry the vulnerability genotype (or allele).

References

1. Leboyer, M., Bellivier, F., Nosten-Bertrand, M., et al. (1998) Psychiatric genetics: search for phenotypes. *Trends Neurosci.* **21,** 102–105.
2. Rothman, K.J. (1986) *Modern Epidemiology*, Little Brown and Co., Boston, MA.
3. Khoury, M.J. and Beaty, T.H. (1994) Applications of the case-control method in genetic epidemiology. *Epidemiol. Rev.* **16,** 134–150.
4. Bellivier, F., Leboyer, M., Courtet, P., et al. (1998) Association between the tryptophan hydroxylase gene and manic-depressive illness. *Arch. Gen. Psychiatry* **55,** 33–37.
5. Bellivier, F., Leboyer, M., Courtet, P., et al. (1999) TPH gene polymorphism and suicidal behavior. *Arch. Gen. Psychiatry* **56,** 101.
6. Abbar, M., Courtet, P., Bellivier, F., et al. (2001) Suicide attempts and tryptophan hydroxylase gene. *Mol. Psychiatry* **6,** 268–273.
7. Faludi, G., Du, L., Palkovits, M., et al. (2000) Serotonin transporter, serotonin-2A receptor and tryptophan hydroxilase gene polymorphisms in depressed suicide victims. *Neurobiology* **8,** 269–271.
8. Geijer, T., Frisch, A., Persson, M.L., et al. (2000) Search for association between suicide attempt and serotonergic polymorphisms. *Psychiatric Genetics* **10,** 19–26.
9. Roy, A., Rylander, G., Forslund, K., et al. (2001) Excess tryptophan hydroxylase 17 779C allele in surviving cotwins of monozygotic twin suicide victims. *Neuropsychobiology* **43,** 233–236.
10. Wacholder, S., McLaughlin, J.K., Silverman, D.T., et al. (1992) Selection of controls in case-control studies. I. Principles. *Am. J. Epidemiol.* **135,** 1019–1028.
11. van Duijn, C.M., de Knijff, P., Wehnert, A., et al. (1995) The apolipoprotein E epsilon 2 allele is associated with an increased risk of early-onset Alzheimer's disease and a reduced survival. *Ann. Neurol.* **37,** 605–610.
12. Kinzler, K.W., Nilbert, M.C., Su, L.K., et al. (1991) Identification of FAP locus genes from chromosome 5q21. *Science* **253,** 661–665.
13. Williams, R.R., Hopkins, P.N., Hunt, S.C., et al. (1990) Population-based frequency of dyslipidemia syndromes in coronary-prone families in Utah. *Arch. Intern. Med.* **150,** 582–588.
14. Cambien, F., Poirier, O., Lecerf, L., et al. (1992) Deletion polymorphism in the gene for antiotensin-converting enzyme is a potent risk factor for myocardial infarction. *Nature* **359,** 641–644.

15. Sandbrink, R., Hartmann, T., Masters, C.L., et al. (1996) Genes contributing to Alzheimer's disease. *Mol. Psychiatry* **1,** 27–40.
16. Sham, P.C., Jones, P., Russell, A., et al. (1994) Age at onset, sex, and familial psychiatric morbidity in schizophrenia. Camberwell Collaborative Psychosis Study. *Br. J. Psychiatry* **165,** 466–473.
17. Weissman, M.M., Wickramaratne, P. Merikangas, K.R., et al. (1984) Onset of major depression in early adulthood. Increased familial loading and specificity. *Arch. Gen. Psychiatry* **41,** 1136–1143.
18. Pauls, D.L., Alsbrook, J.P., 2nd, Goodman, W., et al. (1995) A family study of obsessive-compulsive disorder. *Am. J. Psychiatry* **152,** 76–85.
19. Schurhoff, F., Bellivier, F., Jouvent, R., et al. (2000) Early and late onset bipolar disorders: two different forms of manic-depressive illness? *J. Affect. Discord.* **58,** 215–221.
20. Strober, M., Morrell, W., Burroughs, J., et al. (1988) A family study of bipolar I disorder in adolescence. Early onset of symptoms linked to increased familial loading and lithium resistance. *J. Affect. Disord.* **15,** 255–268.
21. Bellivier, F., Laplanche, J.L., Schurhoff, F., et al. (1997) Apolipoprotein E gene polymorphism in early and late onset bipolar patients. *Neurosci. Lett.* **233,** 45–48.
22. Bellivier, F., Golmard, J.L., Henry, C., et al. (2001) Admixture analysis of age at onset in bipolar I affective disorder. *Arch Gen. Psychiatry* **58,** 510–512.
23. Leboyer, M., Bellivier, F., McKeon, P., et al. (1998) Age at onset and gender resemblance in bipolar siblings. *Psychiatry Res.* **81,** 125–131.
24. Clavel, J. and Hemon, D. (1993) Fundamentals of case-control studies. *Pharmacogenetics* **3,** 160–165.
25. Bellivier, F., Schurhoff, F., Nosten-Bertran, M., et al. (1998) Methodological problems in meta-analysis of association studies between bipolar affective disorders and the tyrosine hydroxylase gene. *Am. J. Med. Genet.* **81,** 349–352.
26. Baron, M. (1997) Association studies in psychiatry: a season of discontent. *Molecular Psychiatry* **2,** 278–281.
27. Preisig, M., Bellivier, F., Fenton, B.T., et al. (2000) Association between bipolar disorder and monoamine oxidase. A gene polymorphisms: results of a multicenter study. *Am. J. Psychiatry* **157,** 948–955.
28. Cavalli-Sforza, I.L. and Menozzi, P. (1994) *The History and Geography of Human Genes.* Princeton University Press, Princeton, NJ.

29. Schork, N.J., Fallin, D., Thiel, B., et al. (2001) The future of genetic case-control studies. *Adv. Genet.* **42,** 191–212.
30. Khoury, M.J. and Yang, Q. (1998) The future of genetic studies of complex human diseases: an epidemiologic perspective. *Epidemiology* **9,** 350–354.
31. Nei, M. (1978) Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics* **89,** 583–590.
32. Pritchard, J.K. and Rosenberg, N.A. (1999) Use of unlinked genetic markers to detect population stratification in association studies. *Am. J. Hum. Genet.* **65,** 220–228.
33. Brennan, P. (1999) Chapter 12. Design and analysis issues in case-control studies addressing genetic susceptibility. IARC Scientific Publications, Lyon, France, 123–132.
34. Rothman, K.J. (1990) No adjustments are needed for multiple comparisons. *Epidemiology* **1,** 43–46.
35. Brunn, T.G. and Ewald, H. (1999) Selection bias of susceptibility genes possible when using parent-offspring trios in genetic association studies. *Molecular Psychiatry* **4,** 415–416.

6

Searching for Alternative Phenotypes in Psychiatric Genetics

Marion Leboyer

1. Introduction

Ten years ago, Tsuang et al. ***(1)*** stated that "psychiatric genetics had reached a point where the sophistication of available experimental tools such as molecular genetics technologies and statistical procedures has surpassed the ability to describe relevant phenotypes." Indeed, the early enthusiasm about the role that molecular genetics would play in our understanding of mental disorders was tempered in the late 1980s by the failure to reproduce the previously observed linkage between bipolar affective disorder and schizophrenia ***(2,3)***. This may be partly attributed to the obstacles that hampered efforts to identify the genes responsible for a particular complex disorder, including unknown mode of inheritance, genetic heterogeneity, phenocopies, incomplete penetrance, and variable expressivity ***(4,5)***. However, an even larger obstacle for the identification of the genes underlying genetic vulnerability of psychiatric disorders is our inability to define the heritable phenotype. Although reliable diagnostic criteria and structured psychiatric interviews have been used to identify probands, little is known about their

From: *Methods in Molecular Medicine, vol. 77: Psychiatric Genetics: Methods and Reviews*
Edited by: M. Leboyer and F. Bellivier

genetic validity. Within apparently affected subjects, various types of phenotypic misclassification reduce the power of linkage studies because of phenocopies or genetic heterogeneity. Furthermore, our inability to identify non-affected subjects carrying vulnerable genes from a population of apparently unaffected subjects or controls, as a result of incomplete penetrance, also reduces the power of association studies. The failure to identify genetic susceptibility loci in psychiatric disorders may thus be partly responsible for inadequate phenotypic definition.

This chapter describes how phenotypic uncertainties contribute to the current difficulties in psychiatry genetics and describes how alternative strategies may prove useful for identifying phenotypes that are more suitable for genetic studies on psychiatric disorders.

2. Different Levels of Phenotypic Uncertainties

In the absence of external validators, psychiatrists must rely on clinical symptoms and diagnostic schemes that—although highly reliable—have no proven biological validity. Standard diagnostic criteria and structured psychiatric interviews are currently used in linkage studies of psychiatric disorders ***(6)***. This approach is based on the demonstrated inter-rater reliability of the assessment procedures ***(7)***. However, reliability does not ensure that a measure is a valid indication of a clinical construct. Furthermore, Kendler ***(8)*** emphasized that giving priority to reliability in the evaluation of a psychiatric disorder might decrease the concurrent validity. Thus, it is not known whether the current definitions of clinical syndromes that are considered to be phenotypes in linkage and association studies accurately identify the underlying genetic substrate.

What are the different levels of phenotypic uncertainty in affected subjects? First, various classification systems may have different stringencies for measuring clinical entities. However, little is known about how to choose the diagnostic system that best describes the most "familial" form of the illness. Secondly, clinical, etiological, and genetic heterogeneity, and phenocopies are thus far impossible to take into account by any of the diagnostic classification criteria.

For example, within each diagnostic category, there may be variation in age at onset, symptom patterns, comorbidity, familial risk, response to treatment, course, and outcome. Third, family studies of schizophrenic probands have revealed that the familial/genetic aspects of schizophrenia are more apparent when the affected status is widened to include sub-syndromal variants, such as schizophrenia-like personality disorders (schizotypal, paranoid, and schizoid personality disorders), leading to the definition of spectrum disorders. However, broadening the definition of affected status may increase the risk of false-positives, and we have no valid indicators of precisely how to define the threshold between affected and non-affected individuals within the several diagnoses described under the term "spectrum of schizophrenia or affective disorders."

In addition, within unaffected subjects, the criteria currently used cannot distinguish between unaffected subjects who carry predisposing loci, and unaffected subjects who are non-carriers. This type of misclassification error may be caused by incomplete penetrance, as shown by the fact that the risk of developing schizophrenia is the same for the offsprings of affected or non-affected monozygotic co-twins ***(9)***. These classification errors in patients and controls can also reduce the power of association studies.

Altogether, the various types of phenotypic misclassification, resulting from invalid diagnostic criteria and etiological heterogeneity among affected and unaffected subjects reduces the power of the analytical approach. Labeling family members as "unaffected" when they are in fact affected can reduce the apparent penetrance. Thus, a larger sample size is required to obtain meaningful results. Labeling family members as "affected" when they are in fact unaffected can also mask the presence of linkage because it appears that there is a recombination when in fact there is none. This makes it more difficult to detect linkage, and almost impossible to localize specific susceptibility genes.

Several strategies are currently being considered in order to bridge the gap between susceptible genotypes and psychiatric disorders. The goal of these methods is to elucidate psychiatric disorders by identifying basic phenotypes for which a more homogeneous

etiology may be expected. Such characteristics may thus be more suited to genetic analysis. Here we describe two complementary strategies that have focused either on describing affected subjects looking for "*candidate symptom"* or looking for "*endophenotypes*," i.e., vulnerability traits in unaffected relatives of affected individuals ***(10)***.

3. Candidate Symptom Approach

The candidate symptom approach aims to identify, among affected subjects, narrow clinical characteristics and/or homogeneous subgroups of the illness, which are probably associated with a disease genotype and show a simpler pattern of inheritance. This strategy assumes that clinical heterogeneity reflects etiological heterogeneity.

Specifically, a candidate symptom, also defined as a diagnostic phenotypic indicator ***(1)***, should fulfill the following criteria: it should show good concordance rates among affected monozygotic twins and should be correlated in pairs of affected siblings, and it should help us to identify a subform of the illness, characterized by a specific clinical pattern, high familial risk, and/or a specific therapeutic response. It should also follow a particular mode of inheritance and be validated by association or linkage with a candidate gene.

The extent to which restriction of the phenotype redefinition simplifies the task of genetic mapping can be measured by the resulting increase in the relative risk (defined by the risk for a relative of a patient divided by the risk in the general population) ***(5,11)***.

Narrowing the definition of the phenotype using clinical characteristics such as age at onset, family history, severity, and associated symptoms has already proven useful in several complex somatic diseases. Severity has been successfully used in several cases. For example, when the study of colon cancer was restricted to cases with extreme polyposis, transmission followed a simple autosomal pattern and facilitated positional cloning of the APC gene ***(12)***.

The study of associated symptoms and comorbid conditions has also proved to be helpful in the identification of subgroups. For

example, in the study of hypertension, it is possible to increase the relative risk by focusing on cases with combined hypertension and hyperlipidemia ***(13)***. The D allele of the gene encoding the angiotensin-converting enzyme (ACE) is a potent risk factor for coronary heart disease in patients with low body-mass index and without hypercholesterolemia ***(14)***.

Increased familial risk can also aid in the identification of further genetic subgroups. Hereditary nonpolyposis colon cancer was genetically mapped by limiting the study of families to those with at least two affected relatives ***(15)***. In Alzheimer's disease, point mutations in the genes encoding the amyloid protein precursor (Ch 21), pro-insulin 1 (Ch 14) and pro-insulin 2 (Ch 1) were only identified after the recognition of early-onset familial cases, showing autosomal dominant pattern of inheritance ***(16)***. Furthermore, subdivision according to age at onset and mode of inheritance has been particularly useful for clarifying genetic heterogeneity in Alzheimer's disease-like dementia. The amyloid precursor protein was not initially believed to be a susceptibility locus because it was incorrectly assumed that all cases of familial Alzheimer's disease are caused by the same gene. Direct sequence analysis of the amyloid precursor protein gene revealed mutations at this locus on chromosome 21q that segregate with the disease in pedigrees of Alzheimer's cases with onset prior to 60 yr. Other loci are also involved, and include most of the early-onset autosomal dominant forms of Alzheimer's disease that are linked to a defective gene on chromosome 14q ***(17)*** and the late-onset sporadic forms that are associated with the apolipoprotein E type 4 allele on chromosome 19 ***(18)***.

Age at onset, clinical dimensions, severity, and family history may also help us to further identify homogeneous subtypes in psychiatry. Early onset is associated with increased familial risk of schizophrenia ***(19)***, bipolar affective disorder ***(20)***, major depressive disorder ***(21)***, and obsessive-compulsive disorder (OCD) ***(22)***. Furthermore, early-onset bipolar patients share a particularly severe clinical pattern with psychotic features or comorbid panic disorder, which further supports the existence of a subgroup of bipolar patients ***(23)***. The age at onset of schizophrenia appears to be heavily

influenced by familial factors, since the correlation between age at onset in affected pairs of siblings ranges from 0.2 to 0.4 ***(24,25)*** and that of monozygotic twins ranges from 0.5 to 0.8 ***(24)***. Similarly, in bipolar disorder, there is a significant correlation for age at onset in affected siblings ***(25)***. Furthermore, the decrease in age at onset of schizophrenia in successive generations appears to be consistent with the phenomenon of genetic anticipation ***(27)***.

The occurrence of a constellation of specific symptoms may also help us to identify a subgroup of schizophrenia that is probably etiologically homogeneous. For example, promising data have been obtained when the clinical emphasis was shifted to the investigation of core manifestations of schizophrenia, such as anhedonia, blunted affect, poverty of speech, lack of a sense of purpose, and diminished social drive, which are considered to be enduring symptoms ***(28)***. Carpenter demonstrated that primary deficit syndrome constitutes a homogeneous subgroup that may share common vulnerability factors ***(29)***. Similarly, schizophrenic patients with periodic catatonia constitute a homogeneous subgroup with a high familial risk of psychosis (26.9% of first degree-relatives) and a significant pattern of anticipation ***(30)***. Cluster analysis helped to identify a subgroup of OCD with high family risk ***(31)***. OCD individuals who are aggressive, have sexual obsessions, and are obsessive about checking, symmetry, and exactness are more likely to have a positive family history of OCD than individuals with OCD characterized by contamination and compulsions. Thus, a clinical constellation of symptoms may help the identification of familial forms that can be selected for linkage studies.

To our knowledge, very few linkage studies using such candidate symptom have been performed. One such study was carried out in dyslexia. Two distinct reading-related phenotypes (phonological awareness and single-word reading) were found to be linked to chromosomes 6 and 15 ***(32)***. In contrast, exploratory studies found an association between candidate genes and subgroups of patients who were defined according to candidate symptoms. In manic-depressive disorder, polymorphisms at the apolipoprotein E gene have been found to be associated with the early onset sub-group ***(33)*** and

HUMTH01 polymorphisms of the tyrosine hydroxylase gene with the late onset ***(34)***. Several lines of evidence suggest that candidate genes are implicated in serotonergic neurotransmission and suicidal behavior. In order to help resolve the discrepancies among studies reporting a positive association result between bipolar disorder and the gene encoding tryptophan hydroxylase (TPH) ***(35)***, it was subsequently discovered that bipolar patients carrying the vulnerable TPH genotype were those who exhibited violent suicidal behavior ***(36)***. This observation was then confirmed in a population of suicidal patients, regardless of the primary psychiatric diagnosis ***(37)***. Similarly, transnosographical observations of an association between suicidal behavior and the serotonin transporter was demonstrated despite the psychiatric diagnosis, thus reinforcing the interest for a dimensional phenotype ***(38,39)***.

The response and adverse reactions to drugs have already been the focus of association studies, particularly in schizophrenia. Although the results remain highly controversial, the polymorphism of the dopamine D3 receptor has been shown to be associated with the development of tardive dyskinesia (TD) ***(40)***. Polymorphisms of the D4 receptor have also been reported to be associated to a favorable response to typical neuroleptic agents but not clozapine ***(41)***. Similar results were obtained for 5HT2-receptor polymorphisms ***(42)***.

These data are still preliminary and must be interpreted with caution. Nevertheless, prospective studies using such potential candidates are worth considering.

4. Endophenotypes

The endophenotype symptom approach aims to identify subtle clinical and biological differences in unaffected relatives of psychiatric patients. These subclinical traits, or endophenotypes, are neuropsychological, electrophysiological, biochemical, or other related variables that may help to unravel the various components of the genetic predisposition to a disorder. To be considered as a marker trait, it must be possible to measure an endophenotype in an objec-

tive and cost-effective manner in clinically unaffected relatives of patients, before the onset of illness. The endophenotype should therefore run in families and be associated with an increased risk of clinical illness.

The endophenotype approach emphasizes the need to use broader phenotypic approaches than the classical clinical distinction between affected and unaffected subjects. An endophenotype may be one variant of a disease genotype that has pleiotropic effects. Alternatively, it may be one of the several entities that contribute to the disease, but that is not required or sufficient for the disease outcome. In this case, an endophenotype may be valuable for identifying common alleles with non-specific and moderate effects on disease risk because psychiatric disorders are probably the consequence of the interaction of several vulnerability factors—each with good genetic validity but not necessarily disease-specific.

If such an endophenotype is a vulnerability trait for the illness, the identification of the genes that are important for the expression of this endophenotype will also identify genes that increase the susceptibility for the illness. The possibility that the endophenotype is associated with the illness should not be rejected because the responsible gene is in linkage disequilibrium (LD) with the disease gene(s). Thus, an endophenotype may help to identify a candidate region for the location of the illness susceptibility loci.

There are several examples of somatic diseases for which "endophenotypic" levels helped to define the genetic basis of the illness in molecular terms. For example, the mode of inheritance of idiopathic hemochromatosis was unknown until serum iron concentration was selected as a biological indicator of intrinsic predisposition to the disease. The inclusion of serum iron in the analysis revealed a link with the HLA-A locus ***(43)***. Epilepsy provides a good model for genetic studies of complex diseases. Epilepsy is a common disease for which the difficulties of nosology and phenotypic definition are similar to the difficulties encountered in psychiatric diseases. Scientists focusing on the families of patients with a juvenile myoclonic form of epilepsy chose a subclinical trait (abnormal electroencephalogram [EEG]) as an endophenotype in unaffected relatives and identified a positive linkage to chromosome 6 ***(44)***.

In adult psychiatry, schizophrenia is the most widely studied psychiatric disorder for the identification of potential endophenotypes. The first step of endophenotype identification involves the observation of group differences between relatives of schizophrenic patients and controls for many variables, such as eye-tracking dysfunction ***(45,46)***, attentional impairment ***(47,48)***, working memory ***(49)***, neuropsychological impairments ***(50,51)***, and auditory-evoked potentials ***(52,53)***. However, as demonstrated by epidemiological genetic studies ***(54)***, variables that significantly differ between relatives of patients and normal subjects do not necessarily constitute good potential endophenotypes: as for candidate symptoms, the extent to which an endophenotype will improve the power of the linkage study depends on its relative risk ***(5)***. For example, a high relative risk for impaired attention has been reported among parents and siblings of patients with schizophrenia ***(55,56)***. Linkage and/or association results are necessary to validate this approach. So far, this alternative phenotypic strategy has yielded positive linkage results with two different endophenotypes: eye-tracking ***(57)*** and measurements of P50-evoked potential ***(58)***.

To date, linkage analysis has only found one endophenotype, sensory gating dysfunction, which is a specific candidate mechanism for neuronal dysfunction in schizophrenia ***(53,59)***. When used as the affected phenotype, sensory gating inhibition (P50 deficit) was found to be linked with the α7-nicotinic receptor gene on the 15q14chromosome, a region that is known to be linked to schizophrenia. Similarly, there are several lines of evidence showing that eye movement dysfunction (EMD) potentially constitutes a good endophenotype for schizophrenia ***(46)***. Whereas 51–85% of schizophrenic patients exhibit EMDs, a higher prevalence is observed in non-affected relatives of schizophrenics (45%) than in the general population (8%) ***(60)***. In 1988, Holzman ***(61)*** suggested that in the offspring of monozygotic and dizygotic twins who were discordant for schizophrenia, a single autosomal dominant gene could account for EMDs and schizophrenia. It was only in 1996 that eye-tracking dysfunction was used as an endophenotype in a linkage analysis ***(57)***. EMDs were found to map to a locus on chromosome 6 situated

near regions that were previously highlighted in several linkage studies of schizophrenia. Although these preliminary data require replication in a larger sample, these results provide encouraging support for shifting the genetic analysis of psychiatric disorders to refined phenotypes, which should have a better diagnostic accuracy.

To overcome the absence of clear-cut boundaries within the spectrum of a categorical disorder and between affected and unaffected individuals, qualitative traits have been used to increase the power of linkage studies ***(1,11)*** The extent to which a quantitative phenotype increases the power of a categorical phenotype is illustrated by a recent linkage study ***(62)***. Schizophrenic families and sib-pairs were assessed by both diagnostic categories and quantitative-trait measures of positive and negative symptoms. Both parametric and sib-pair analysis failed to produce significant evidence of linkage with categorical disease definitions; however, they found significant evidence of linkage of positive symptom score to chromosome 6p. Similarly, a quantitative linkage analysis was carried out using a P300 event-related brain potential (ERP) as an endophenotype to identify individuals who were at genetic risk for alcoholism ***(63)***. An oddball paradigm showed that P300 is classically related to the post-perceptual updating of short-term working memory traces of expected environmental stimuli ***(64)***. The analysis revealed significant linkage on chromosomes 2 and 6, which suggests that candidate loci underlie a functional neuroelectric activity. Interestingly, in schizophrenic patients and their unaffected relatives ***(65)*** and in bipolar patients and their relatives ***(66)***, reduced P300 amplitude and increased latency were found during an oddball paradigm. Thus, using ERP abnormalities as endophenotypes may reveal transnosographical vulnerability factors.

In child psychiatry, Folstein and Rutter ***(67)*** have shown that the concordance rate for certain cognitive and language disorders was 82% among monozygotic autistic twins and 10% among dizygotic twins. This suggests that the familial aggregation of a broader range of cognitive disorders should be investigated to better define the phenotype in autism. More recent studies on the relatives of autistic children have revealed that certain cognitive, social, language, and

biochemical peculiarities may represent the variable expression of a gene related to autism ***(68–70)***.

5. Conclusion

Reshaping phenotypic definitions may stimulate the emergence of a different model for psychiatric disorders. It is tempting to consider classical psychiatric illnesses as multidimensional illnesses in which several discrete phenotypes, each with a simple genetic mechanism, interact. It is also conceivable that some of those narrow phenotypes are common to different nosographical entities. These transnosographical components may interact and result in a spectrum of diseases. For example, a growing body of data have shown that serotonin neurotransmission ***(71)*** and serotonergic markers are associated with suicidal behavior regardless of the psychiatric disorder. Considering that traditional psychiatric diagnoses are not unitary entities, but instead are the consequence of several interacting traits, the approach outlined here may also lead to the reshaping of our definition of validity criteria in clinical psychiatry.

Acknowledgments

This work was supported by the Assistance Publique des Hôpitaux de Paris (Délégation à la Recherche Clinique, PHRC) and INSERM.

References

1. Tsuang, M. (1993) Genotypes, phenotypes and the brain: a search for connections in schizophrenia. *Br. J. Psychiatry* **163,** 299–307.
2. Kelsoe, J.K., Ginns, E., Egeland, J., Gerhardt, D., Bale, S., Pauls, D., et al. (1989) Re-evaluation of the linkage relationship between chromosome 11 loci and the gene for bipolar affective disorder in the old order Amish. *Nature* **342**, 238–243.
3. Risch, N. and Botstein, D. (1996) A manic-depressive history (news). *Nat. Genet.* **12,** 351–353.
4. Pauls, D. (1993) Behavioural disorders: lessons in linkage. *Nat. Genet.* **3,** 4–5.

5. Lander, E.S. and Schork, N. (1994) Genetic dissection of complex traits. *Science* **265,** 2097–2045.
6. Merikangas, K.R., Spence, A., and Kupfer D.J. (1989) Linkage studies of bipolar disorder: methodologic and analytic issues. *Arch. Gen. Psychiatry* **46,** 1137–1141.
7. Nunberger, J.I., Blehar, M.C., Kaufman, C.A., York-Cooler, C., Simpson, S.G., Harkavy-Freedman, J., et al. (1994) Diagnostic interview for genetic studies. *Arch. Gen. Psychiatry* **51,** 849–864.
8. Kendler, K.S. (1990) Towards a scientific psychiatry nosology. Strengths and limitations. *Arch. Gen. Psychiatry* **47,** 969–973.
9. Gottesman, I. and Bertelsen, A. (1989) Confirming unexpressed genotypes for schizophrenia. *Arch. Gen. Psychiatry* **46,** 867–872.
10. Leboyer, M., Bellivier, F., Nosten-Bertrand, M., Jouvent, R., Pauls, D., and Mallet, J. (1998) Psychiatry genetics: search for phenotypes. *Trends Neurosci.* **21, 3,** 102–105.
11. Risch, N. (1990) Linkage strategies for genetically complex traits: I. multilocus models. *Am. J. Hum. Genet.* **46,** 222–228.
12. Kinzler, K.W., et al. (1991) Identification of FAP locus genes from chromosome 5q21. *Science* **253,** 661.
13. Williams, R.R., Hopkins, P.N., Hunt, S.C., Wu, L.L., Hasstedt, S.J., Lalouel, J.M., et al. (1990) Population-based frequency of dyslipidemia syndromes in coronary-prone families in Utah. *Arch. Intern. Med.* **150,** 582–588.
14. Cambien, F., Poirier, O., Lecerf, L., Evans, A., Cambou, J.P., Arveiler, D., et al. (1992) Deletion polymorphism in the gene for angiotensin-converting enzyme is a potent risk factor for myocardial infarction. *Nature* **359,** 641–644.
15. Fischel, R., Lescoe, M.K., RAO, M.R., Copeland, N.G., Jenkins, N.A., Garber, J., et al. (1993) The human mutator gene homolog MSH2 and its association with hereditary nonpolyposis colon cancer. *Cell* **75(5),** 1027–1038.
16. Sandbrink, R., Hartmann, T., Masters, C.L., and Beyreuther, K. (1996) Genes contributing to Alzheimer's disease. *Mol. Psychiatry* **1,** 27–40.
17. St Georges Hyslop, P., et al. (1992) Genetic evidence for a novel familial Alzheimer's disease locus on chromosome 14. *Nat. Genet.* **2,** 330–334.
18. Gaskell, P.C., et al. (1993) *Science* **261,** 921–923.
19. Sham, P., Jones, P., Russell, A., Gilvarry, K., Bebbington, P., Lewis,

S., et al. (1994) Age at onset, sex, and familial morbidity in schizophrenia. Camberwell Collaborative Psychosis Society. *Br. J. Psychiatry* **165,** 466–473.

20. Strober, M., Morrell, W., Burroughs, J., Lampert, C., Danforth, H., and Freeman, R. (1988) A family study of bipolar I disorder in adolescence. Early onset of symptoms linked to increased familial loading and lithium resistance. *J. Affect. Disord.* **15,** 255–268.
21. Weissman, M.M., Gershon, E.S., Kidd, K.K., Prusoff, B.A., Leckman, J.F., Dibble, E., et al. (1984) Psychiatric disorders in the relatives of probands with affective disorders. The Yale University ñ National Institute of Mental Health Collaborative Study. *Arch. Gen. Psychiatry* **41,** 13–21.
22. Pauls, D., Alsobrook, J.P., Goodman, W., Rasmussen, S., and Leckman, J.F. (1995) A family study of obsessive-compulsive disorder. *Am. J. Psychiatry* **152,** 76–84.
23. Schurhoff, F., Bellivier, F., Jouvent, R., Mouren-Simeoni, M.C., Bouvard, M., Allilaire, J.-F., et al. (2000) Early and late onset bipolar disorders: two different forms of manic depressive illness? *J. Affect. Disord.* **58,** 215–221.
24. Kendler, K.S, Tsuang, M., and Hays, P. (1987) Age at onset in schizophrenia. A familial perspective. *Arch. Gen. Psychiatry* **44,** 881–890.
25. Leboyer, M., Bellivier, F., McKeon, P., Albus, M., Borman, M., Perez-Diaz, F., et al. (1998) Age at onset and gender resemblance in bipolar siblings. *Psychiatry Res.* **81,** 125–131.
26. Kendler, K.S., Tsuang, M.T., and Hays, P. (1987) Age at onset in schizophrenia: a familial perspective. *Arch. Gen. Psychiatry* **44,** 881–890.
27. Gorwood, P., Leboyer, M., Falissard, B., Jay, M., and Feingold, J. (1996) Evidence of anticipation in schizophrenia. *Am. J. Psychiatry* **153,** 1173–1177.
28. Carpenter, W.T., Heinrich, D.W., and Wagman A.M. (1988) Deficit and non-deficit forms of schizophrenia: the concept. *Am. J. Psychiatry* **145,** 578–583.
29. Carpenter, W.T., Buchanan, R.W., and Kirkpatrick, B. (1995) New diagnostic issues in schizophrenic disorders. *Clin. Neurosc.* **3,** 57–63.
30. Stober, G., Franzek, E., Lesch, K.P., and Beckmann, H. (1995) Periodic catatonia: a schizophrenic subtype with major gene effect and anticipation. *Eur. Arch. Psychiatry Clin. Neurosci.* **245,** 135–141.
31. Eapen, V., Robertson, M.M., Alsobrook, J.P., and Pauls, D.L. (1997) Obsessive compulsive symptoms in Gilles de la Tourette syndrome

and obsessive compulsive disorder: differences by diagnosis and family history. *Am. J. Med. Genet.* **74,** 432–438.

32. Girgorenko, E.L., Wood, F.B., Hart, L.A., Speed, W.C., Shus, P., and Pauls, D.L. (1997) Susceptibility loci for distinct components of developmental dyslexia on chromosomes 6 and 15. *Am. J. Hum. Genet*. **60(1),** 27–39.
33. Bellivier, F., Laplanche, J.L., Schurhoff, F., Feingold, J., Feline, A., Jouvent, R., et al. (1997a) Apolipoprotein E gene polymorphisms in early and late onset manic depressive patients. *Neurosci. Lett.* **233,** 45–48.
34. Bellivier, F., Nosten-Bertrand, M., Leboyer, M., Schurhoff, F., Feingold, J., Meloni, R., et al. (1997b) Association between late onset bipolar affective disorder and tyrosine hydroxylase gene polymorphism. *Am. J. Med. Genet.* **74,** 614.
35. Bellivier, F., Leboyer, M., Courtet, P., Buresi, C., Beaufils, B., and Samolyk, D., et al. (1998) Association between the tryptophan hydroxylase gene and manic depressive illness. *Arch. Gen. Psychiatry* **55(1),** 33–37.
36. Bellivier, F., Leboyer, M., Courtet, P., Feingold, J., Buresi, C., and Malafosse, A. (1999) TPH gene polymorphisms and suicidal behaviour. *Arch. Gen. Psychiatry* **56(1),** 101.
37. Abbar, M., Courtet, P., Bellivier, F., Leboyer, M., Boulenger, J.-P., Castelnau, D., et al. (2001) Suicide attempts and the tryptophan hydroxylase gene. *Mol. Psychiatry* **6(3),** 1214–1220.
38. Bellivier F., Szoke A., Henry C., Lacoste J., Bottos, C., Nosten-Bertrand, et al. (2000) Possible association between serotonin transporter gene polymorphism and violent suicidal behavior in mood disorders. *Biol. Psychiatry* **48(4),** 319–322.
39. Gorwood, P., Batel, P., Ades, J., Hamon, M., and Bowi, C. (2000) Serotonin transporter gene polymorphism alcoholism and suicidal behavior. *Biology* **48(4),** 259–264.
40. Steen, V.M., Lovlie, R., Macewan, T., and McCreadie, R.G. (1997) Dopamine D3-receptor gene variant and susceptibility to tardive dyskinesia in schizophrenic patients. *Mol. Psychiatry* **2,** 139–145.
41. Cohen, B.M., Ennulat, D.J., Centorrino, Matthysse, S., Koniecza, H., and Cherkezian, S. (1999) Polymorphisms of the dopamine D4 receptor and response to antipsychotic drugs. *Psychopharmacology* **141,** 325–354.
42. Arranz, M.J., Munro, J., Sham, P., Kirov, G., Murray, R.M., Colier, D.A., et al. (1998) Meta-analysis of studies on genetic variation in 5-HT2A receptors and clozapine response. *Schizophr. Res*. **32,** 93–99.

43. Lalouel, J.M., et al. (1985) Genetic analysis of idiopathic hemochromatosis using both qualitative (disease status) and quantitative (serum iron) information. *Am. J. Hum. Genet*. **37,** 700–718.
44. Greenberg, D.A., Delgado-Escueta, A.V., Widelitz, H., Sparkes, R.S., Treiman, L., Maldonado, H.M., et al. (1988) Juvenile myoclonic epilepsy (JME) may be linked to the BF and HLA loci on human chromosome 6. *Am. J. Med. Genet.* **31(1),** 185–192.
45. Holzman, P.S., Solomon, C.M., Levin, S., and Waternuax, C.S. (1984) Pursuit eye movement dysfunctions in schizophrenia. *Arch. Gen. Psychiatry* **41,** 136–139.
46. Levy, D.L., Holzman, P.S., Matthysse, S., and Mendell, N.R. (1994) Eye tracking and schizophrenia: a selective review. *Schizophr. Bull.* **20(1),** 47–62.
47. Cornblatt, B. and Erlenmeyer-Kimling L. (1985) Global attentional deviance in children at risk for schizophrenia: specificity and predictive validity. *J. Abnorm. Psychol.* **94,** 470–486.
48. Chen, W.J. and Faraone, S.V. (2000) Sustained attention deficits as markers of genetic susceptibility to schizophrenia. *Am. J. Med. Genet.* **97,** 52–57.
49. Park, S., Holzman, P., and Goldman-Rakic, P.S. (1995) Spatial working memory deficits in the relatives of schizophrenic patients. *Arch. Gen. Psychiatry* **52(10),** 821–828.
50. Faraone, S.V., Seidman, L.J., Kremen, W.S., Pepple, J.R., Lyons, M.J., and Tsuang M.T. (1995) Neuropsychological functioning among the nonpsychotic relatives of schizophrenic patients: a diagnostic efficiency analysis. *J. Abnorm. Psychol.* **104(2),** 286–304.
51. Egan, M. and Godlberg, T. (2002) Intermediate cognitive phenotypes associated with schizophrenia, in Psychiatry Genetics: Methods and Protocols. (Leboyer, M. and Bellivier, F., eds.), Humana Press, Totowa, N.J., pp. 163–198.
52. Freedman, R., Adler, L.E., Waldo, M.C., Pachtman, E., and Francks, R.D. (1983) Neurophysiological evidence for a defect in inhibitory pathways in schizophrenia: comparison of medicated and drug-free patients. *Biol. Psychiatry* **18,** 537–551.
53. Freedman, R. (2002) Electrophysiological phenotypes, in *Psychiatry Genetics: Methods and Protocols* (Leboyer, M. and Bellivier, F., eds.), Humana Press, Totowa, N.J., pp. 215–226.
54. Faraone, S.V., Kremen, W.S., Lyons, M.J., Pepple, J.R., Seidman, L.J., and Tsuang, M. (1995) Diagnostic accuracy and linkage

analysis: how useful are schizophrenia spectrum phenoypes? *Am. J. Psychiatry* **152,** 1286–1290.

55. Chen, W.J., Liu, S.K., Chang, C.-J., Lien, Y.-J., Chang, Y.-H., and Hwu, H.-G. (1998) Sustained attention deficit and schizotypal personality features in nonpsychotic relatives of schizophrenic patients. *Am. J. Psychiatry* **155,** 1214–1220.
56. Egan, M., Goldberg, T., Gscheidle, T., Weirich, M., Bigelow, L., and Weinberger, D. (2000) Relative risk of attention deficits in siblings of patients with schizophrenia. *Am. J. Psychiatry* **157,** 1309–1316.
57. Arolt, V., Lencer, R., Nolte, A., Muller-Myhsok, B., Purmann, S., Shurmann, M., et al. (1996) Eye tracking dysfunction is a putative phenotypic susceptibility marker of schizophrenia and maps to a locus on chromosome 6p in families with multiple occurrence of the disease. *Am. J. Med. Genet*. **67,** 564–579.
58. Coon, H., Plaetke, R., Holik, J., Hoff, M., Myles-Worsley, M., Freedman, R., et al. (1993) Use of a neurophysiological trait in linkage analysis of schizophrenia. *Biol. Psychiatry* **34,** 277–289.
59. Freedman, R., Adams, C.E., Adler, L.E., Bickford, P.C., Gault, J., Harris, J.G., et al. (2000) *Am. J. Med. Genet. (Semin. Med. Genet.)* **97,** 58–64.
60. Blackwood, D., St Clair, D., Muir, W., and Duffy, J. (1991) Auditory P300 and eye tracking dysfunction in schizophrenic pedigrees. *Arch. Gen. Psychiatry* **48,** 899–909.
61. Holzman, P.S., Kringlen, E., Matthysse, S., Flanagan, S.D., Lipton, R.B., Cramer, G., et al. (1988) A single dominant gene can account for eye-tracking dysfunctions and schizophrenia in offspring of discordant twins (see comments). *Arch. Gen. Psychiatry* **45,** 641–647.
62. Brzustowicz, L., Honer, W., Chow, W., Hogan, J., Hodgkinson, K., and Bassett, A. (1997) Use of a quantitative trait to map a locus associated with severity of positive symptoms in familial schizophrenia to chromosome 6p. *Am. J. Hum. Genet.* **61,** 1388–1396.
63. Begleiter, H., Porejesz, B., Reich, T., Edenberg, H.J., Goate, A., Blangero, J., et al. (1998) Quantitative trait loci analysis of human event-related brain potentials: P3 voltage. *Electroencephalogr. Clin. Neurophysiol.* **108,** 244–250.
64. Donchin, E. (1981) Surprise!… Surprise? *Psychophysiology* **18,** 493–513.
65. Freedman, D. and Squires-Wheeler, E. (1994) Event-related potentials (ERPs) as indicators of risk for schizophrenia *Schizophr. Bull.* **20(1),** 63–74.

66. Pierson, A., Jouvent, R., Quintin, P., Perez-Diaz, F., and Leboyer, M. (2000) Information processing deficits in relatives of manic-depressive patients. *Psychol. Medicine* **30,** 545–555.
67. Folstein, S. and Rutter, M. (1977) Infantile autism: a genetic study of 21 twin pairs. *J. Child Psychol. Psychiat.* **18,** 297–321.
68. Piven, J. and Palmer, P. (1997) Psychiatric disorder and the broad autism phenotype. Evidence from a family study of multiple-incidence autism families. *Am. J. Psychiatry* **156,** 557–563.
69. Hughes, C., Plumet, M.H., and Leboyer, M. (1999) Towards a cognitive phenotype for autism: increased prevalence of executive dysfunction and superior spatial span amongst siblings of children with autism. *J. Child Psychol. Psychiatry* **40(5),** 705–718.
70. Leboyer, M., Philippe, A., Bouvard, M., Guilloud-Bataille, M., Bondoux, D., Tabuteau, F., et al. (1999) Whole blood serotonin and plasma beta-endorphin in autistic probands and their first-degree relatives. *Biol. Psychiatry* **15;45(2),**158–163.
71. Mann. J.J. (1998) The neurobiology of suicide. *Nat. Med*. **4,** 25–30.

7

Intermediate Cognitive Phenotypes Associated with Schizophrenia

Michael F. Egan and Terry E. Goldberg

1. Introduction

The search for intermediate phenotypes to inform genetic studies of psychiatric disorders is never as productive and exciting as in neuropsychological studies, particularly those of schizophrenia. A variety of cognitive impairments are frequently seen in patients with schizophrenia, and appear to be core features of this illness. In contrast, cognitive impairments seen in other psychiatric disorders, such as depression and mania, are state-related and may not be essential to these disorders ***(1).*** In studies of patients with schizophrenia, cognitive impairments do not appear to be caused simply by secondary factors associated with severe mental illness, such as poor cooperation or medications. They are present from an early stage of the illness and are stable over time, suggesting trait-like properties. Furthermore, emerging data suggest, but have not proven, that different profiles of impairment may be present in subgroups of patients. Another feature that makes these cognitive measures attractive targets for genetic studies is that they are closely related to clinical and functional outcome. Thus, the genes underlying cognitive phenotypes may be the most relevant to improving the quality

From: *Methods in Molecular Medicine, vol. 77: Psychiatric Genetics: Methods and Reviews*
Edited by: M. Leboyer and F. Bellivier

of patients' lives. Three domains of cognition are generally affected, including working memory/executive function, verbal/declarative memory, and attention.

Data from family studies of schizophrenia have created an increasingly compelling case for the use of such cognitive measures as intermediate phenotypes. Data from studies of parents, siblings, offspring, and twins have found cognitive deficits similar to—although less severe than—those seen in patients. Furthermore, the correlation between different types of impairments is very low, suggesting that they represent several independent phenotypes. This chapter reviews the nature of cognitive deficits associated with schizophrenia and focuses on data from family studies.

2. Cognitive Impairments in Schizophrenia

Only within the last 15 yr has it come to be appreciated that an important component of schizophrenia reliably involves compromises in various important facets of cognition. Early work in the field concentrated on reductions in intelligence and slowness in reaction time. More recently, the focus has shifted to other cognitive domains, including working memory, episodic memory, attention, and certain types of linguistic processing. In addition, studies have sought to more clearly define the onset and course of neuropsychological impairments. In general, the data suggest that subtle attenuations in cognitive function are apparent well before illness onset (in addition to delays in attaining developmental milestones) in patients-to-be. Marked reductions in at least several critical cognitive domains then occur over a 3–5-yr period beginning with the illness's prodromal period and extending beyond the appearance of overt clinical symptomatology. This is followed by a long period of stasis in which impairments, although evident, do not show progressive decline (at least not until late life) ***(2)***.

During the initial investigations of cognitive impairments in schizophrenia, it was critical to assess whether impairments were caused by secondary factors such as lack of cooperation, psychotic symptoms, or medications. The notion that cognitive impairments

are not the result of interference from symptoms has been supported by several studies showing a very low correlation between positive symptoms such as delusions or hallucinations with a degree of cognitive impairment ***(3)***. On the other hand, some negative symptoms are correlated with performance on prefrontal tasks, and disorganization is correlated with some attentional measures, such as Stroop performance and semantic processing measures (e.g., priming). However, negative symptoms do not produce a global impairment in all tests, and they do not prevent patients from cooperating, engaging, and performing well, at least with regard to some tasks. The associations between these symptoms and cognitive impairments are specific, suggesting a relationship based on underlying neurobiological deficits.

Cognitive impairments are also not related to duration of illness or medications. Prospective studies of patients at the onset of their illness, before treatment with medications, clearly show cognitive impairment similar to chronic patients ***(4)***. Furthermore, cross-sectional studies including subjects from all age groups, do not show increasing severity with increasing duration of illness ***(5,6)***. Regarding the effects of medications, a number of studies have shown that cognitive impairments are not markedly altered by treatment with either typical or atypical neuroleptics, even when negative symptoms are improved ***(7)***. Finally, Heaton has shown that cognitive impairments in schizophrenia are remarkably stable over periods of years ***(8)***. As such, test–retest reliability in such broad domains as memory, attention, language, visual-spatial ability, and abstraction/executive function were uniformly high and showed much less variability than symptoms. Overall, these data strongly suggest that, in contrast to symptoms, which may be state characteristics, cognitive impairments are trait-like, enduring features of this illness and are not, for the most part, simply artifacts related to nonspecific illness effects.

A second critical issue, which continues to be vexing, is whether schizophrenic cognitive impairment can be characterized by differential deficits in specific cognitive domains or whether it is generalized. This issue is particularly important, as discussed here,

regarding whether these impairments reflect one or several phenotypes useful for genetic studies. The argument that patients exhibit a generalized deficit has been made on psychometric grounds with the use of carefully selected tests ***(9)*** and by way of large normative databases for the Halstead-Reitan battery ***(10)***. These studies show that patients do equally poorly on tests of essentially all cognitive domains when these tests are standardized for degree of difficulty and ceiling or floor effects are removed. However, based on careful clinical analysis of tests with well-established brain-behavior relations, it is possible to come to a different conclusion using either a case study format ***(11)*** or double dissociation methodology, whereby schizophrenia is compared with other neuropsychiatric conditions ***(12)***. These studies have shown, for example, that patients with schizophrenia have a different pattern of cognitive impairment compared to patients with Alzheimer's disease, Huntington's disease, temporal lobe epilepsy, and affective disorders ***(13)***. Thus, impairments may not be diffuse but appear to affect several specific cognitive domains referable to prefrontal and mesial limbic regions. These results also suggest that patients perform relatively well on many cognitive domains referable to other neural systems, including those involving remote memory, procedural learning, and most elemental forms of sensory processing.

Studies of large groups of patients are beginning to emerge, and indicate that schizophrenia may represent a heterogeneous group, with a minority showing a diffuse pattern of impairment and others showing a modal relatively circumscribed pattern ***(14)***. These results are not subject to criticisms about test difficulty level (i.e., patients simply doing worse on more difficult tasks) and dispersion of test scores, considering that they derive from subgroups obtained from a single large sample. Some researchers have also argued that a small subgroup of patients are cognitively intact, and indeed there is a great deal of variance in the severity of impairment. However, data from discordant MZ twins found that even patients who perform in the normal range are impaired relative to where they would have been if not ill, suggesting that cognitive impairment may be a universal feature of the illness ***(15,16)***. In a broader context, it seems

likely that schizophrenia involves dysfunction in multiple "nodes" of at least several distributed cognitive systems. Regardless of this debate, even if patients as a group show generalized intellectual deficits, this still does not mean that intellectual decline is a unitary problem. The critical issue for genetic studies is whether specific cognitive measures are related to genetic risk for schizophrenia, and whether genes related to these cognitive domains increase the risk of schizophrenia.

One attractive feature of cognitive impairments as intermediate phenotypes is that they have been linked to social and vocational disabilities. This suggests that genes associated with impaired cognition may also account for some of the variance in functional outcome. Strong concurrent relations between global level functioning and intelligence and other specific neurocognitive tests have been repeatedly found in both twin studies and studies of schizophrenic singletons ***(17–19)***. In several meta-analyses, Green convincingly demonstrated that short- and long-term memory accounts for about 30–40% of the variance in a variety of specific functional domains and that executive and attentional impairments account for about 20% of functional outcome ***(18,19)***. Global or summary indicators of cognitive function exhibited even stronger relationships with outcome. Rather surprisingly, positive psychotic symptoms showed relatively weak relationships with outcome. Negative symptoms showed relatively strong relationships, possibly because of their overlap with cognitive impairments.

Although many questions remain about the meaning and prognostic utility of cognitive deficits, several cognitive tests have been increasingly used to explore the neurobiology of schizophrenia. Foremost among these are working memory/executive function tasks, which have been particularly useful in examining prefrontal function (*see* **refs. *20–29***). Deficits in verbal/declarative memory, as well as structural neuroimaging and postmortem findings, have also pointed to the mesial temporal lobe as a likely area of involvement and these tasks are increasingly being employed in clinical studies ***(30)***. Although such studies have not proven that cognitive deficits are critical to the pathophysiology of schizophrenia, at the

very least they strongly support the notion that such impairments are related to some alterations in biology, which in turn distinguish patients from other groups. Therefore, it seems increasingly evident that cognitive deficits are relevant to the biology of the illness. A more detailed description of the specific cognitive domains that are classically impaired in patients and how they are tested is described here.

2.1. Working Memory and Executive Function

Patients with schizophrenia may exhibit a wide range of working memory (WM) difficulties often involving numerous components of this cognitive system. Working memory/executive function, sometimes referred to as short-term memory, is a cognitive function whereby information is stored briefly (e.g., remembering a friend's phone number retrieved from the operator) and, if necessary, manipulated in some fashion (e.g., updating the city the friend now lives in). Simple and veridical recall of short sequences of letters and numbers is typically impaired in patients, and furthermore may be susceptible to both interference of a secondary task doing arithmatic operations on a check in "one's head" while figuring a restaurant bill, and longer periods of "delay" (in the Brown-Peterson test). Patients also have difficulty doing tasks that involve simultaneous storage and manipulation of material, as in the letter-number-span task in which random sequences of numbers and letters must be remembered over a short period and ordered (numerically and alphabetically). Interestingly, recent work has indicated that patients' impairments on tasks like this are not caused by qualitative abnormalities in sequencing, or in intrusion of responses, but rather are the result of capacity limitations on memory-set sizes and/or in the ability to sustain information over delays—two of the most basic aspects of any model of WM.

Other tests of WM have also been examined. These are generally called "executive" when a task requires recombination of existing knowledge in the service of a decision, shifting of various mental sets, planning and systematically monitoring responses, or use of

encoding strategies for future episodic memory recall. Two widely used executive tests that involve set shifting are the Wisconsin Card Sort Test (WCST) and the CANTAB Intra-Dimensional/Extra-Dimensional Set Shifting Test (ID/ED). Patients generally exhibit deficits on the WCST, in which they are asked to match a specific card to one of several target cards and receive feedback on their accuracy. A correct match can be made based on one of several "dimensions," such as color, number, or shape. The correct dimension is changed by the examiner, and the subject is expected to shift strategy on the basis of feedback. Patients characteristically perseverate on one dimension, even when they are told they need to shift and given explicit instructions to do so, similar to patients with prefrontal cortical brain damage. On the ID/ED task, patients have similar difficulty with shifting sets. In this more elaborate task, there are different types of dimensional shifts, including "intra-dimensional" (e.g., switching to a new exemplar of a specific constellation of lines) and "extra-dimensional" (e.g., responding to a previously irrelevant dimension that had been introduced earlier in the test, such as switching from lines to shapes) ***(31)***.

2.2. Episodic Memory

Memory is often considered to be the most severely impaired function in schizophrenia ***(4)***. The type of memory examined is usually called episodic memory, and it involves the acquisition of new representations of visual or verbal material that is language-based, object-based, or spatially based (i.e., lists of words, things, or maps). This type of memory is believed to involve several stages including encoding (constructing a memory trace with a spatiotemporal context), storage (consolidating a trace over time), and retrieval (accessing the trace). A number of studies have consistently found that schizophrenic patients perform worse than normals on a variety in memory paradigms; however, attempts to further dissect abnormalities at specific stages of memory as would be evidenced by group X manipulation interactions, have failed ***(32,33)***. For example, Paul et al. ***(34)*** examined the effects of deep and shallow

encoding on patients and normals. They found that both groups performed better and to an equivalent degree on memory for words that have been deeply encoded (i.e., "Is this entity living or nonliving?") as opposed to shallow encoding ("Does this word contain the letter A?"). Similarly, Elvevaag et al. ***(35)*** found that patients were not more susceptible to various types of tests that produce interference in episodic memory during the retrieval phase. This was examined using "paired associates" in which members of a pair of words are first learned and then "shuffled" so that the words are re-paired with different words from the same initial list of pairs. Patients and controls showed similar declines in performance. Finally, Elvevag and Goldberg (unpublished data) also looked at the capacity of patients to produce false memories in the context of semantic lures during encoding, a second type of interference. Patients made similar numbers of errors in comparison to normals once the overall degree of recall was controlled. Thus, patients with schizophrenia have a global impairment in episodic memory (i.e., overall performance was impaired on these tests), although the mechanism for this is unclear. It is also unclear whether subgroups of patients (and families) may be differentially affected at different stages of this memory system.

2.3. Attention

Attention, focusing "awareness" on a specific target and responding appropriately, is often measured in schizophrenia using the Continuous Performance Test (CPT). This test, which has several versions, requires a subject to attend to a series of letters or numbers and respond in some way (e.g., push a button) when a "target" appears. Using single digits (1 though 9) with one target number (e.g., "9") is very easy for most subjects, and ceiling effects reduce variance. Variations on the basic CPT make the task more difficult, but at the expense of involving additional computational demands. In the 1–9 version, for example, the subject only presses the response button when a 1 precedes a 9 (*see* **Table 1**). Overall, patients with schizophrenia perform about one standard deviation from the mean, but a large percentage of patients perform in the

normal range. This ceiling effect is reduced with harder CPT versions, including the "IP" and "degraded stimulus" CPT versions.

CPT tasks involve several computational demands including rapid encoding of stimuli, being ready to respond ("response readiness") and stimulus-response mapping. CPT tasks require sustained attention over time, typically 5–20 min. Performance in these tasks does not decrease over time, as error rate is increased even early in the test. Attempts to discern the stage at which patients with schizophrenia perform poorly on indicate difficulty with rapidly encoding briefly presented stimuli and coming to a decision when interstimulus intervals are short and targets are rare. Furthermore, complementary to this view, there is also evidence that this may reflect, at least in part, a failure in executive function in cognitive control of lower-level perceptual processes ***(36)***. In other words, there my be a failure in biasing sensory and/or perceptual areas to the target, so that signal-to-noise ratios leading to better discriminability are far from optimal.

In contrast to CPT paradigms, other tests of attention that involve covert orientation and shifts of attention have generally not revealed qualitative impairments in schizophrenia. Tests of selective attention such as the Stroop test, in which an overlearned response or prepotent response must be inhibited while a less salient aspect of this stimulus must be attended to, have indicated that patients with schizophrenia may be more susceptible to interference.

2.4. Other Cognitive Domains

Patients with schizophrenia have also been found to do poorly in tests of other cognitive domains, including psychomotor speed, reaction time, verbal fluency, and some specific aspects of semantic processing. Some tests, like trail-making tests, seem to demand a range of cognitive processing and the neuropsychological literature clearly demonstrates that this task is performed poorly by subjects with brain lesions in a variety of areas. Thus, impaired Trails B performance is a nonspecific but sensitive indicator of brain damage. However, although patients with prefrontal lesions do poorly on it,

Table 1
Description of Cognitive Tests Frequently Administered to Schizophrenic Individuals and Their Relatives

Cognitive test	Outcome measures	Abbreviations	Description
General intelligence			
Weschsler Adult Intelligence Scale	Intelligence quotient	WAIS, IQ	Intelligence based on a composite of language, visual processing, and memory abilities
Wide-Range Achievement Test	Standard score	WRAT	Putative measure of premorbid IQ in schizophrenia
Working memory and executive function			
Wisconsin Card Sort Test	1) Perseverative errors, 2) Categories	WCST 1) PE 2) Cat	Cards are matched based on different dimensions —e.g., color, shape, and number
CANTAB intradimensional/ extradimensional	Many	CANTAB ID/ED	Odd shapes and figures matched based on different dimensions
N-Back	Percent correct	N back	Series of numbers presented. Subject must recall the number presented: 0, 1, 2, or 3 back from the current number
Brown-Peterson task	Number correct; number of intrusions		Subject must remember 4 items for a 15-s delay period. During this time, simple naming tasks are performed
Declarative/verbal memory			
California Verbal List Learning Test	Trials 1-5 SS	CVLT	Lists of words recalled at several time points, items in lists related in various ways

Wechsler Memory Scale, revised	Logical memory I, II Visual reproductions I and II	WMS-R	Memory for 1) a story 2) geometric designs
Attention			
Continuous Performance tests	d', a measure from signal processing theory	CPT	Multiple versions of task. Subject attends to repetitive stimuli, such as numbers, and responds only when specific numbers or sequences appear
1–9 version	d'		Subject presses button when 9 follows 1 in a series of numbers
1–9 version with distractors	d'		In addition to the target stream of numbers, flanking distractors are presented
Degraded stimulus CPT	d'	DS CPT	Similar to 1–9 version, except numbers are visually degraded, and therefore difficult to perceive
Identical pairs CPT	d'	IP CPT	Similar to 1–9 version, except targets are pairs of numbers that appear sequentially
Psychomotor speed			
Trials A and B	Time needed (in s) to complete task		Subject connects spatially discrete printed numbers (A) or numbers alternating with letters (B)
Verbal ability/language			
Verbal fluency	Number correct		Subject lists as many words as possible in 60 s based on a prompt—e.g., "all words beginning with the letter K," or "animals in a zoo"

this does not mean that the only reason that patients with schizophrenia perform poorly is because of prefrontal deficits. Thus, impaired performance on such tasks could also possibly involve deficits in other cognitive domains. Extensive descriptions of neuropsychological deficits seen in schizophrenia are provided in several recent publications ***(13,14)***.

3. Family Studies of Cognitive Impairment

Data suggesting that neuropsychological deficits seen in patients are partly genetic are drawn from studies of family members, including parents, siblings, offspring, and affected and unaffected monozygotic (MZ) and dizygotic (DZ) twins. Because of the difficulty in analyzing twin samples, most studies have examined singleton first-degree relatives. The first reports focused on attention, and measures of other cognitive parameters soon followed. Methodological differences, such as which tests are used and ascertainment biases, complicate comparisons of these studies. Furthermore, most studies have examined only small samples, and often over-represent individuals with schizophrenia spectrum disorders. Some have even mixed parents with siblings, which is problematic because of age and generational differences. Nevertheless, the majority of studies show that first-degree relatives are impaired on several dimensions of cognition and that their profile is similar to, although less severe than, patients with schizophrenia ***(37,38)***. Furthermore, shared variance between tests of different cognitive domains is small, suggesting that several measures may be suitable for use as intermediate phenotypes. Finally, although limited, some data indicate that these familial traits are not wholly a result of the effects of shared environment, but that genetic factors are likely to play a role.

3.1. Non-Twin First-Degree Relatives

The earliest reports describing cognitive deficits in relatives came from "at risk" studies of children of schizophrenic mothers. This represented an effort to find possible antecedents of schizophrenia.

Descriptions of behavioral abnormalities included, among other things, impaired attention ***(39,40)***. Subsequent neuropsychological studies used increasingly standardized tests of attention, the majority of which were CPTs ***(41,42)***. These studies generally found that "at risk" children performed poorly only on more difficult versions of the CPT (e.g., *see* **refs. *43–46***). However, interpretation of "at risk" studies, is confounded because these deficits could either be antecedents of schizophrenia or secondary to environmental factors (e.g., having a psychotic parent) unrelated to the risk for schizophrenia. Studies of adult siblings, who have generally passed through much of the age of risk, avoid these problems to some degree. They have also reported impaired performance on the CPT ***(41,42)***. As with at-risk children, these differences are seen more often with difficult versions of the CPT ***(47,48)***. Results from studies using simpler versions—such as the "1–9" version (with or without distractors) or with degraded stimuli—have been mixed, with some findings of no overall differences ***(49)***, a trend for reduction ***(50,51)***, or marked impairments ***(37,52–55)***. One problem with interpreting results of the more difficult CPT tasks is that they inevitably involve additional cognitive demands beyond attention. For example, the IP version has a significant working memory load, and the degraded stimulus CPT increases the demands for perceptual processing. Regarding "pure attention" to the extent that it is a measurable and unique cognitive process, it remains unclear to what degree siblings are impaired. Overall, there may be at least a weak effect, perhaps in a subgroup of unaffected siblings ***(49)***, and these deficits are not simply antecedents of or secondary to psychiatric illness.

Tests of other cognitive domains have been included in family studies of schizophrenia and abnormalities similar to those seen in patients have been observed. As with studies using the CPT, comparisons are confounded by use of different tests, small sample sizes, and selective recruitment of sibs with schizophrenia spectrum disorders. Generally, most studies have included tests of working memory/executive function, verbal memory, and psychomotor speed. For example, Pogue-Geile et al. ***(56)***, in 40 non-schizo-

phrenic sibs, and Franke et al. ***(45,57)***, in 33 healthy sibs, found impaired performance on the WCST, Trails B, and verbal fluency compared to a matched control groups (*see* **ref.** ***48***). On the other hand, Scarone et al. ***(58)*** found no differences in WCST in 35 well siblings compared to matched controls. Yurgelun-Todd and Kinney ***(59)*** found lower scores on the WCST, but not Trails B, in a group of 15 healthy sibs. Shedlack et al. ***(60)*** found essentially no differences between 14 well siblings from multiplex families and controls on verbal memory (*see* **ref.** ***61***). In contrast, Cannon ***(37)*** found impaired performance on a broad battery of tests including attention, working memory/executive function, and verbal memory in 16 non-schizophrenic siblings, 6 of 16 of whom had definite or likely schizotypal personality disorder. Larger studies have more consistently found differences. Keefe et al. ***(62)***, in a cohort of 54 nonpsychotic first-degree relatives (sibs and parents), found impaired performance on Trails B and verbal fluency (both letter and category), but not on the WCST. Faraone et al. ***(63)*** found impairments in abstraction, verbal memory, and attention in a group of 35 nonpsychotic first-degree relatives, and, similarly, Toomey et al. ***(64)*** found impaired working memory (WCST) and verbal memory in 54 first-degree well relatives. More recently, several additional smaller studies have also found prefrontal cognitive deficits in siblings ***(55,65)***. Laurent, in a study of 25 parents and 22 siblings, found several cognitive deficits in both sets of relatives, but normal scores on the WCST ***(51)***. Egan et al., in a study of 193 siblings, found deficits on WCST, Trails B, and CVLT, whether or not sibs with schizophrenia spectrum disorders were included ***(66)***. Similarly, Goldberg et al. ***(67)*** found deficits in the N-back but not in reaction time in well siblings from this same cohort, suggesting that WM deficits were not secondary to upstream deficits in information processing speed.

Despite inconsistencies, theses studies support several conclusions. First, working memory/executive function is very likely impaired in first-degree relatives. This has been demonstrated using a variety of tests, most notably the WCST but also tests of verbal fluency and the "N-back." Second, tests of verbal or declarative

memory also appear to be impaired; these abnormalities have been demonstrated using tests such as the WMS-R and the CVLT. Third, scores on the Trail-making tests are reduced, including A and B versions, implicating volitional oculomotor scanning/psychomotor speed. Finally, attention may be impaired, at least in a subgroup of families, as evidenced by studies using various versions of the CPT. In contrast, reduced IQ has generally not been found in first-degree relatives, and covarying for IQ does not alter the results (*see* **refs.** ***62,63,66,68***). Additional abnormalities have also been reported on other tests (*see* **ref.** ***69***), but less consistently and only in small numbers of studies ***(38)***. Overall, cognitive deficits in first-degree family members are similar to—although less severe than—those seen in patients with schizophrenia ***(37,66)***, and implicate the same brain regions and cognitive domains. Furthermore, these deficits are probably not secondary to psychiatric disorders, as they are present even in otherwise psychiatrically healthy siblings ***(63)***. Thus, these data suggest that at least part of the variance in these cognitive deficits is familial.

If cognitive deficits are familial, a critical issue in the assessment of their suitability for genetic studies is the strength of the genetic effect (the genetic component of total phenotypic variance). Estimating the genetic portion of variance requires data from several types of relatives who share different degrees of genes and environment. Such data are lacking for cognitive phenotypes. Nevertheless, an upper limit of heritability can be derived from first-degree relatives. The first study to address this issue for CPT performance estimated heritability in 61 first-degree relatives using the ICC and found $h^2 = 0.79$ ***(70)***. This estimate, which was relatively large in comparison to most cognitive measures in normal subjects ***(71)***, was apparently estimated only in unaffected sib-pairs, and does not give a clear picture of the heritability of *impaired* attention. Estimating the relative risk for "impairment" is another approach that has been hampered by the small size of most cohorts, ascertainment that is typically biased toward those with schizophrenia spectrum disorders ***(37,50,53,62)*** and control groups selected for the lack of a psychiatric disorder—i.e., a "supranormal" control group that is not

representative of the general population ***(50,70)***. To the extent that siblings with spectrum disorders may have more cognitive deficits, the former overestimates impairs attention in the sib group and the latter underestimates it in the control group. Several recent studies of large cohorts of adult first-degree relatives and more representative controls have been conducted. An investigation of a Taiwanese cohort (mixing parents and siblings, often of low educational level) found high rates of impairment in first-degree relatives (n = 148) compared to controls using the 1–9 version with degraded stimuli ***(72)***. Relative risk, depending on the cutoff criteria, was elevated at 18–130, dramatically higher than the relative risk for schizophrenia itself! In contrast, in a study of 193 siblings in a U.S cohort ***(49)***, no overall group differences were seen compared to controls on the 1–9 version with distractor numbers. The latter study found that a subgroup of only siblings of patients with low CPT scores had lower scores than the control group and slightly increased relative risk. Comparing the two studies, the marked differences in their estimates of relative risk could be a result of a variety of factors, such as ethnicity, recruitment biases, education, and type of relatives (parent or sib) studied ***(49)***. With respect to other cognitive abnormalities, Egan et al. reported increased relative risk ranging from 2.0 to 4.0 for WCST, Trails B, and verbal memory scores in their large sibling cohort (*see* **ref. *63***). Relative risk values did not significantly change when subjects with schizophrenia spectrum disorders were excluded. These relative risk values are generally lower than those seen for schizophrenia, which raises questions about their value in genetic studies. On the other hand, confidence intervals are substantial, raising the possibility that true values are higher. Furthermore, the critical variable is genotypic relative risk for a specific gene and phenotype. Thus, a gene could have a substantially greater impact on CPT performance in patients, and thus be easier to detect, compared to its effect on the risk for schizophrenia. Many questions remain on the subject of relative risk for CPT, and there is clearly a need for additional data from large cohorts. Nevertheless, these results support the notion that cognitive impairment has at least a weak genetic component.

Table 2
Relative Risk

Relative risk, or λ, is an epidemiological term frequently used to estimate the strength of a genetic effect. Relative risk is calculated as a ratio of the risk to a relative divided by the risk to the general population. For schizophrenia, since the prevalence in the general population is 1%, the following are generally used (100):

Siblings (λs) = 10%/1% = 10
MZ twins (λmz) = 50%/1% = 50

The relative risk attributable to a specific genetic locus is the genotypic relative risk, which determines the power of specific study designs to detect susceptibility alleles. Therefore, its ability to find genes of an intermediate phenotype depends on its relative risk and its genetic architecture, which is usually unknown.

3.2. Twin Studies

Data on cognitive performance have been published on two cohorts of discordant twins. A caveat to interpreting these studies is that they have a limited ability to estimate the genetic portion of phenotypic variance for the entire population, because they focus primarily on differences between discordant pairs. Finding that unaffected MZ twins do not have an abnormality does not mean that the abnormality is not genetic. For example, the lack of concordance rate for schizophrenia itself (zero, by definition, for discordant pairs) does not mean that schizophrenia is not genetic. The reason for this is that such analyses exclude concordant pairs, which make up 40–50% of such pairs in the general population. A biological deficit seen in affected but not unaffected MZ twins could mean either that the abnormality is caused by environmental factors or that it is very closely linked to the illness, which itself is partly or largely genetic. Such a phenotype could have genetic determinants and penetrance characteristics very similar to the illness itself. Furthermore, such phenotypes could be useful for genetic studies, but would likely only add genetic information in subjects with the ill-

ness. A gene that encodes for visual hallucinations, for example, would only be found in schizophrenic subjects who have visual hallucinations, not in their own unaffected relatives nor in other schizophrenics without visual hallucinations. The lack of hallucinations in the unaffected MZ twin may be the result of incomplete penetrance or interactions with unique environment.

In the first study of discordant twins, Goldberg et al. ***(16,17,73)***, using a broad neuropsychological battery, found that unaffected MZ twins have impaired performance on subtests of the WMS-R , with trends for impairment on WCST PE and Trails A ($p < .05$) ***(17)***. Interestingly, looking at *concordance* rates for impairment (1 S.D. below control mean) on the WCST, 9 of 19 "unaffected" co-twins of those affected twins with impaired WCST results also scored in the impaired range. This 47% concordance rate for "impairment," although somewhat artificial, addresses more specifically the question of whether co-twins of probands with a specific impairment have abnormal function. It also avoids the trivial issue of whether prefrontal cognition itself is heritable, which is not the critical question. The fairly substantial rate of "impaired" WCST PE scores in MZ co-twins who are discordant for diagnosis suggests that cognitive phenotypes may have a higher penetrance than diagnosis.

A second, population-based study of 18 MZ pairs and 34 DZ discordant twin pairs examined the relationship between cognitive deficits and genetic risk for schizophrenia, using canonical discriminant analysis ***(74)***. This study used a different battery of tests, but included those that largely interrogate the same cognitive domains. Four tests contributed unique genetic variance to the increased risk for schizophrenia. These tests were spatial WM (visual span test of the WMS-R), WM/divided attention (using a Brown-Peterson dual-task paradigm), intrusions during recall of a word list (CVLT), and choice reaction time (using a CPT-like task). It is unclear whether the same group differences were seen with these MZ twins compared to the Goldberg sample. In Cannon's analysis, verbal memory was more impaired in affected MZ subjects, relative to cotwins—a finding that suggests that this was the result of an effect of unique environmental variance related to illness. As discussed previously,

this does not exclude the possibility that verbal memory has a significant genetic component. Furthermore, some of these tests—such as the spatial working memory test—demand a variety of compu-tations that may involve several neural systems and therefore may remove variance from more elemental tests. Considering the small number of subjects, additional cohorts are needed to clarify which tests are most useful. Nevertheless, the results of both twin studies are similar to studies from non-twin relatives in one important respect. Relatives of patients with schizophrenia appear to have impairments in several domains of cognition, including working memory/executive function, declarative/verbal memory, and attention.

3.3. How Many Cognitive Phenotypes?

Although neuropsychological tests can examine the function of neural systems that are somewhat independent, finding deficits on several tests does not necessarily mean that these measure independent traits. An alternative possibility is that impairments are found on different tests because of one underlying abnormality that affects a variety of neural systems. Generally, attempts to address this question, using several similar statistical approaches, suggest that this is not the case. First, correlations between measures (which often differ between studies) are usually low in these groups. For example, Yurgelun-Todd and Kinnney ***(59)*** found no correlation—or, more accurately, an inverse relationship—between Trails A and WCST in sibs, yet these scores were correlated in patients. Keefe et al. ***(62)*** found a correlation of 0.22 between WCST PE and Trails B in sibs. On the other hand, Toomey et al. ***(64)*** found fairly high correlations between attention and verbal memory and between attention and abstraction in a cohort of 54 first-degree relatives, including parents, sibs, and offspring, but no significant correlation between WCST and memory measures (on the WMS-R). Second, using multiple regression, measures of WM, verbal memory, and psychomotor speed shared only modest portions of variance (less than 15% in siblings) ***(66)***. These results are similar even when IQ is used as a

regressor, suggesting that low intelligence does not drive the findings. Third, using factor analysis, several studies ***(66,75,76)*** have shown that WCST and Trails B load on different factors. In nonpatient populations, factor analysis often shows that a broad range of cognitive tests load significantly on the first factor, often referred to as "*g*" (*see* **ref.** ***71*** for review), which accounts for a large portion of total variance. In contrast, these analyses only include tests in which sib groups are impaired. Finally, as described here, Cannon et al., in a critical analysis of MZ and DZ discordant twins, found evidence for four distinct, independent cognitive deficits using canonical discriminant analysis. Thus, these data suggest that poor performance on one test is a relatively poor predictor of performance on different cognitive tests.

The interpretation of the relatively low correlation between tests is unclear. It could mean that the different types of tests (WCST, CPT, WMS-R, and Trails B) tap distinct abnormalities in specific computational networks and have a distinct genetic architectures. On the other hand, the low correlations could simply be an artifact of noisy measurements. In favor of the former, neuroimaging studies suggest that correlations between prefrontal and hippocampal phenotypes, which may be less encumbered by the vicissitudes of neuropsychological measurement error, are low. However, the neural systems related to performance on tests such as WCST and Trails B are very likely to overlap, as based simply on studies from humans with prefrontal brain damage. Dissecting the neurobiological and molecular determinants of performance on these tests may shed light on this problem.

3.4. Are Familial Cognitive Deficits Caused by Genes that Increase Risk for Schizophrenia?

Although schizophrenia-related cognitive deficits appear to be partially familial, how certain is it that they are related to genes that increase the risk for schizophrenia? As noted here, measures such as relative risk only set an upper limit on heritability. Could these familial deficits be caused entirely by shared environmental fac-

tors? Although largely unexplored, the limited evidence relevant to this issue suggests that this is not the case. The most persuasive data comes from the twin study of Cannon et al. described here, in which greater deficits are seen in MZ vs DZ discordant sibs ***(74)***. Consistent with this finding, twin studies of non-schizophrenic cohorts using systematic ascertainment have found fairly high heritability of low cognitive abilities ***(77)***, although the relevance of this to schizophrenia can be questioned. Additional evidence comes from twin and adoption studies of schizophrenia, which suggest that the environmental component of variance related to risk for schizophrenia is unique and thus not shared between sibs ***(78)*** (for review, *see* **ref.** ***71***). This finding implies that relatives of schizophrenics differ from controls in only two respects: 1) having increased genetic risk for schizophrenia, and 2) environmental factors associated with having a relative with schizophrenia. It is theoretically possible but unlikely that the latter could explain familial cognitive deficits, as supported indirectly by two interesting but unreplicated observations. First, children of schizophrenic parents raised by foster parents have impaired performance on a children's version of Trails A and B ***(79)***, indicating that early common environmental factors do not account for cognitive deficits. Second, cognitive deficits seen in psychiatrically healthy siblings are not related to prospectively determined obstetric complications ***(80)***, indicating that in utero or perinatal environmental factors are also not causative. Of course, these studies are by no means conclusive, and further studies of different types of family members along with adoption studies may be needed before more definitive statements can be accepted.

Other explanations may account for the apparent familial nature of these cognitive deficits or otherwise limit the utility of these traits. For example, such deficits could be caused by genetic factors unrelated to the risk for schizophrenia. This is particularly an issue for CPT deficits, which may be found only in siblings of patients who themselves have impaired attention. The same family relationship would also be seen in patients who are short, based simply on the trivial fact that height is partly genetic ***(49)***. Obviously, genes for height would have nothing to do with schizophrenia. On the other

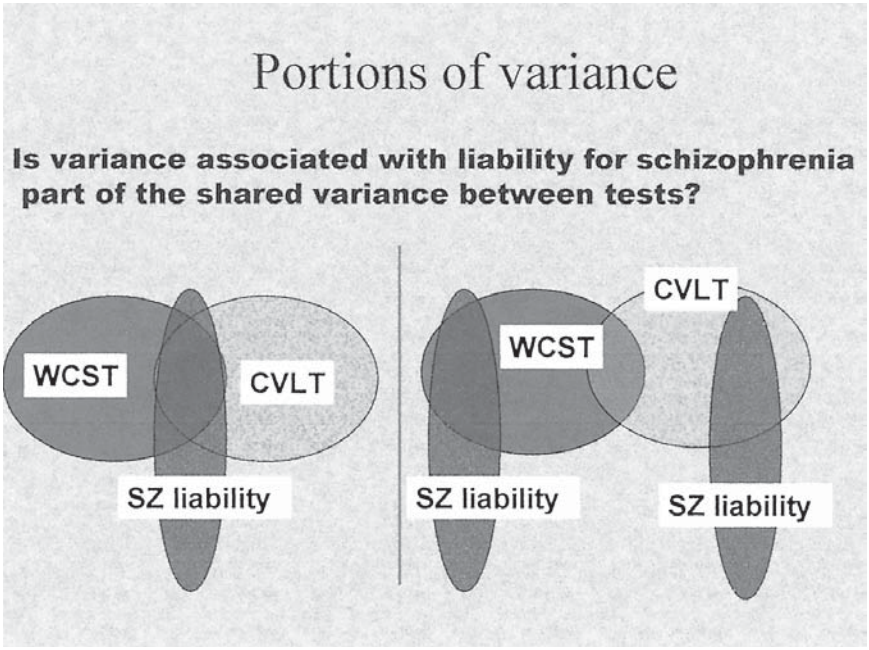

Fig. 1. Models of putative shared variance and liability for schizophrenia.

hand, height is not a significant aspect of brain dysfunction in schizophrenia, as is attention. Another possibility is that the increased rates of impairment in sibs could come from biased ascertainment in sib or control groups, since the former often has many subjects with psychiatric disorders, although the latter may be a "super normal" group. A third potential confounding factor is that, although these traits share little variance (*see* above), this shared variance is the only portion of total variance related to genes that increase the risk for schizophrenia (*see* **Fig. 1**). In this scenario, only a few genes would account for all the cognitive deficits observed in family members, and may also explain only a minor portion of genetic risk for schizophrenia. Unfortunately, data on these areas is limited.

3.5. How Can Cognitive Phenotypes Be Used to Find Schizophrenia Genes?

Is the notion of using cognitive measures to find genes for schizophrenia plausible? Support for this approach comes from a recent finding by Egan et al., using working memory and the WCST as the phenotype ***(81)***. One attractive aspect to using working memory is that its neurobiology is increasingly well understood ***(82–85)***. Working memory appears to be critically dependent on the sustained activity of glutamatergic pyramidal neurons in the prefrontal cortex, which in turn is dependent on optimal D1-mediated dopamine tone ***(85)***. Although there are no reported non-conservative genetic variants that clearly alter function in D1 receptors or other proteins affecting prefrontal dopamine, an important exception is the val158/108met polymorphism in the gene for catechol-o-methyltransferase (COMT) ***(86,87)***. This catabolic enzyme is located on or within postsynaptic neurons and appears to inactivate released dopamine primarily in the prefrontal cortex, in contrast to other regions innervated by dopamine ***(88)***. COMT-knockout mice (males, but not females) show increased dopamine only in the prefrontal cortex ***(89)***. This regional specificity may be a result of the paucity of the dopamine transporter in prefrontal dopaminergic synapses ***(90)***. Also, remarkably, several studies in animals and humans suggest that reduced COMT activity improves some aspects of cognition, although not specifically WM ***(91–93)***. This finding is consistent with other data that increasing synaptic dopamine improves prefrontal cognitive function. The val158/108met polymorphism itself produces a dramatic effect on COMT enzyme activity because of the thermolability of the methionine variant at 37°C ***(87)***. Thus, the valine allele would be expected to reduce WM because of its effects in more rapidly inactivating synaptic dopamine. In a cohort of 175 patients, 219 siblings, and 55 controls, Egan et al. found that the COMT genotype had a significant impact on WM. Specifically, subjects with two valine alleles had worse scores than those with one, who in turn had worse scores than those with none ***(81)*** (*see* **Fig. 2**). Using functional magnetic resonance imaging (fMRI) to examine

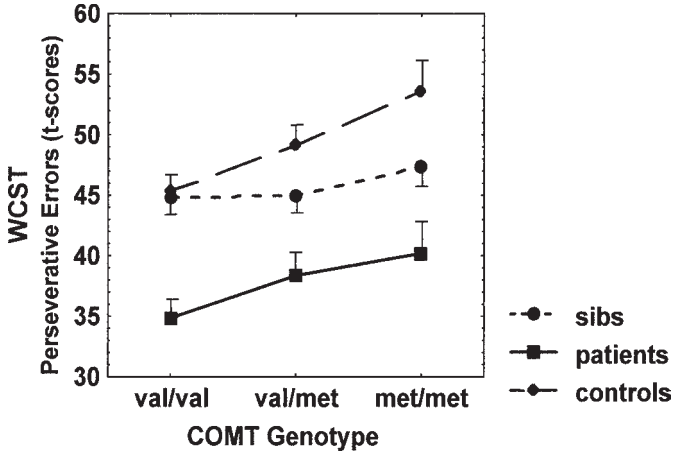

Fig. 2. COMT genotype has a significant effect on executive cognitive function across index cases, their siblings, and healthy controls.

prefrontal physiology, they also found the same allele dosage effect on prefrontal blood-flow alterations, or "efficiency" ***(23,24)***, during a WM task (*see* Chapter 10). Finally, the valine allele was associated with schizophrenia using the transmission/disequilibrium test (TDT) ***(94)***, consistent with several prior TDT-based studies of COMT and schizophrenia ***(95–97)***. Thus, using WM/executive function as an intermediate phenotype pointed to an obvious candidate gene, and subsequently provided a mechanism of action (MOA) by which this gene impaired cognition and a slightly increased risk for schizophrenia. It is also worth noting that the COMT genotype had similar effects in control subjects, supporting the notion that genes for schizophrenia are common alleles that account for a small portion of risk through slight, deleterious effects on cortical function both in patients and unaffected subjects.

3.6. Caveat Emptor

Despite the potential utility of cognitive phenotypes, several questions about this approach persist. First, the magnitude of the

genetic component of phenotypic variance for these traits in families of patients with schizophrenia, although very likely more than zero, remains essentially unknown. Heritability or relative risk estimates are largely unavailable ***(66)***, and the possibility that environmental factors play a substantial role, although unlikely, has not been excluded. Similarly, the magnitude of genetic variance related to risk for schizophrenia contributed by each putative phenotype is also unknown, although it is likely to be greater than zero ***(74)***. Third, cognitive phenotypes may only lead to the identification of a few susceptibility genes. It is possible that most schizophrenia genes are unrelated to cognitive phenotypes, similar to the situation with heart disease and intermediate phenotypes such as hyperlipidemia, smoking, and obesity. Nevertheless, the notion that schizophrenia is fundamentally a disorder of prefrontal and temporal/limbic cortical function ***(98,99)*** substantially increases the likelihood that phenotypic measures of the function of these regions are related to the genetic risk for schizophrenia.

4. Conclusions

Patients with schizophrenia demonstrate a variety of cognitive deficits, which seem most prominent in tasks of prefrontal and temporal/limbic function. These include WM/executive function, attention, and verbal/declarative memory. First-degree relatives have similar yet less severe deficits that are not caused by concomitant psychiatric disorders or substance abuse. These cognitive impairments are only weakly correlated, suggesting that they may represent independent traits related to the genetic risk for schizophrenia. Deficits in WM/executive function, which are attributable to prefrontal cortical processing, may be directly associated with the COMT valine allele, consistent with a number of neurobiological studies of the role of COMT and dopamine on prefrontal function. This allele also appears to slightly increase risk of schizophrenia. Although many uncertainties remain about the use of cognitive phenotypes, they offer important new possibility for methods to pinpoint the genes related to susceptibility for schizophrenia.

References

1. Gold, J., Goldberg, T.E., Kleinman, J.E., and Weinberger, D.W. (1991) The impact of symptomatic state and pharmacological treatment on cognitive functioning of patients with schizophrenia and mood disorders, in *Handbook of Clinical Trials* (P, M. E. B., ed.), Swets and Zeitlinger, Berwyn, PA, pp. 185–214.
2. Weickert, T. and Golberg, T.E. (2000) The course of cognitive impairment in patients with schizophrenia, in *New Research in Schizophrenia* (PD, S. T. H., ed.), Oxford University Press, New York, NY.
3. Goldberg, T. and Green, M. Neuropsychology of schizophrenia, in *Psychopharmacology, Fifth Generation of Progress* (Bloom, F.E., Kupfer, D., Bunney, B. et al., eds.), Lippincott Williams & Wilkins, Philadelphia, PA, in press.
4. Saykin, A.J., Shtasel, D.L., Gur, R.E., Kester, D.B., Mozley, L.H., Stafiniak, P., et al. (1994) Neuropsychological deficits in neuroleptic naive patients with first-episode schizophrenia. *Arch. Gen. Psychiatry* **51(2),** 124–131.
5. Goldstein, G. and Zubin, J. (1990) Neuropsychological differences between young and old schizophrenics with and without associated neurological dysfunction. *Schizophr. Res.* **3,** 117–126.
6. Hyde, T., Nawroz, S., Goldberg, T.E., Bigelow, L.B., Strong, D., Ostrem, J.L., et al. (1994) Is there cognitive decline in schizophrenia? A cross-sectional study. *Br. J. Psychiatry* **164(4),** 494–500.
7. Goldberg, T.E. and Weinberger, D.R. (1996) Effects of neuroleptic medications on the cognition of patients with schizophrenia: a review of recent studies. *J. Clin. Psychiatry* **57(Suppl. 9),** 62–65.
8. Heaton, R.K., Gladsjo, J.A., Palmer, B. W., Kuck, J., Marcotte, T.D., and Jeste, D.V. (2001) Stability and course of neuropsychological deficits in schizophrenia. *Arch. Gen. Psychiatry* **58(1),** 24–32.
9. Blanchard, J. and Neale, J.M. (1994) The neuropsychological signature of schizophrenia: generalized or differential deficit? *Am. J. Psychiatry* **151,** 40–48.
10. Braff, D.L., Heaton, R., Kuck, J., Cullum, M., Moranville, J., Grant, I., et al. (1991) The generalized pattern of neuropsychological deficits in outpatients with chronic schizophrenia with heterogeneous Wisconsin Card Sorting Test results. *Arch. Gen. Psychiatry* **48(10),** 891–898.
11. Shallice, T., Burgess, P.W., and Frith, C.D. (1991) Can the neuropsy-

chological case-study approach be applied to schizophrenia? *Psychol. Med.* **21(3),** 661–673.

12. Gold, J.M., Hermann, B.P., Randolph, C., Wyler, A.R., Goldberg, T.E., and Weinberger, D.R. (1994) Schizophrenia and temporal lobe epilepsy. A neuropsychological analysis. *Arch. Gen. Psychiatry* **51(4),** 265–272.
13. Elvevag, B. and Goldberg, T.E. (2000) Cognitive impairment in schizophrenia is the core of the disorder. *Crit. Rev. Neurobiol.* **14(1),** 1–21.
14. Weickert, T.W., Goldberg, T.E., Bigelow, L.B., Egan, M.F., and Weinberger, D.R. (2000) Differential pattern of cognitive impairment in schizophrenic patients with preserved and compromised intellect. *Arch. Gen. Psychiatry,* in press.
15. Goldberg, T.E., Ragland, J.D., Torrey, E.F., Gold, J.M., Bigelow, L.B., and Weinberger, D.R. (1990) Neuropsychological assessment of monozygotic twins discordant for schizophrenia (see comments). *Arch. Gen. Psychiatry* **47(11),** 1066–1072.
16. Goldberg, T.E., Torrey, E.F., Gold, J.M., Ragland, J.D., Bigelow, L.B., and Weinberger, D.R. (1993) Learning and memory in monozygotic twins discordant for schizophrenia. *Psychol. Med.* **23(1),** 71–85.
17. Goldberg, T.E., Torrey, E.F., Gold, J.M., Bigelow, L.B., Ragland, R.D., Taylor, E., et al. (1995) Genetic risk of neuropsychological impairment in schizophrenia: a study of monozygotic twins discordant and concordant for the disorder. *Schizophr. Res.* **17(1),** 77–84.
18. Green, M.F. (1996) What are the functional consequences of neurocognitive deficits in schizophrenia? *Am. J. Psychiatry* **15(3),** 321–330.
19. Green, M.F., Kern, R.S., Braff, D.L., and Mintz, J. (2000) Neurocognitive deficits and functional outcome in schizophrenia: are we measuring the "right stuff"? *Schizophr. Bull.* **26(1),** 119–136.
20. Berman, K.F., Illowsky, B.P., and Weinberger, D.R. (1988) Physiological dysfunction of dorsolateral prefrontal cortex in schizophrenia. IV. Further evidence for regional and behavioral specificity. *Arch. Gen. Psychiatry* **45(7),** 616–622.
21. Berman, K.F., Ostrem, J.L., Randolph, C., Gold, J., Goldberg, T.E., Coppola, R., et al. (1995) Physiological activation of a cortical network during performance of the Wisconsin Card Sorting Test: a positron emission tomography study. *Neuropsychologia* **33(8),** 1027–1046.
22. Berman, K.F., Zec, R.F., and Weinberger, D.R. (1986) Physiologic dysfunction of dorsolateral prefrontal cortex in schizophrenia. II. Role

of neuroleptic treatment, attention, and mental effort (see comments). *Arch. Gen. Psychiatry* **43(2),** 126–135.

23. Callicott, J.H., Bertolino, A., Mattay, V.S., Langheim, F.J.P., Duyn, J., Coppola, R., et al. (2000) Physiological dysfunction of the dorsolateral prefrontal cortex in schizophrenia revisited. *Cereb. Cortex* **10,** 1078–1092.
24. Callicott, J.H., Mattay, V.S., Bertolino, A., Finn, K., Coppola, R., Frank, J.A., et al. (1999) Physiological characteristics of capacity constraints in working memory as revealed by functional MRI. *Cereb. Cortex* **9(1),** 20–26.
25. Weinberger, D.R. (1987) Implications of normal brain development for the pathogenesis of schizophrenia. *Arch. Gen. Psychiatry* **44(7),** 660–669.
26. Weinberger, D.R. and Berman, K.F. (1988) Speculation on the meaning of cerebral metabolic hypofrontality in schizophrenia. *Schizophr. Bull.* **14(2),** 157–168.
27. Weinberger, D.R., Berman, K.F., and Chase, T.N. (1988a) Mesocortical dopaminergic function and human cognition. *Ann. NY Acad. Sci.* **537,** 330–338.
28. Weinberger, D.R., Berman, K.F., and Illowsky, B.P. (1988b) Physiological dysfunction of dorsolateral prefrontal cortex in schizophrenia. III. A new cohort and evidence for a monoaminergic mechanism. *Arch. Gen. Psychiatry* **45(7),** 609–615.
29. Weinberger, D.R., Berman, K.F., and Zec, R.F. (1986) Physiologic dysfunction of dorsolateral prefrontal cortex in schizophrenia. I. Regional cerebral blood flow evidence. *Arch. Gen. Psychiatry* **43(2),** 114–124.
30. Aleman, A., Hijman, R., de Haan, E.H., and Kahn, R.S. (1999) Memory impairment in schizophrenia: a meta-analysis. *Am. J. Psychiatry* **156(9),** 1358–1366.
31. Pantelis, C., Barber, F.Z., Barnes, T.R., Nelson, H. E., Owen, A.M., and Robbins, T.W. (1999) Comparison of set-shifting ability in patients with chronic schizophrenia and frontal lobe damage. *Schizophr. Res.* **37(3),** 251–270.
32. Elvevag, B., Egan, M.F., and Goldberg, T.E. (2000a) Memory for temporal order in patients with schizophrenia. *Schizophr. Res.* **46(2–3),** 187–193.
33. Elvevag, B., Egan, M.F., and Goldberg, T.E. (2000b) Paired-

associate learning and memory interference in schizophrenia. *Neuropsychologia* **38(12),** 1565–1575.
34. Paul, B., Elvevag, B., Bokat, C., and Goldberg, T.E. (2001) Semantic encoding and episodic memory in schizophrenia.
35. Elvevag, B., Fisher, J.E., Gurd, J.M., and Goldberg, T.E. (2001) Semantic clustering in verbal fluency. (Submitted.)
36. Elvevag, B., Weinberger, D.R., Suter, J.C., and Goldberg, T.E. (2000c) Continuous performance test and schizophrenia: a test of stimulus-response compatibility, working memory, response readiness, or none of the above? *Am. J. Psychiatry* **157(5),** 772–780.
37. Cannon, T.D., Zorrilla, L.E., Shtasel, D., Gur, R.E., Gur, R.C., Marco, E.J., et al. (1994) Neuropsychological functioning in siblings discordant for schizophrenia and healthy volunteers. *Arch. Gen. Psychiatry* **51(8),** 651–661.
38. Kremen, W.S., Seidman, L.J., Pepple, J.R., Lyons, M.J., Tsuang, M.T., and Faraone, S.V. (1994) Neuropsychological risk indicators for schizophrenia: a review of family studies. *Schizophr. Bull.* **20(1),** 103–119.
39. Fish, B. (1977) Neurobiologic antecedents of schizophrenia in children. Evidence for an inherited, congenital neurointegrative defect. *Arch. Gen. Psychiatry* **34(11),** 1297–1313.
40. Parnas, J., Schulsinger, F., Schulsinger, H., Mednick, S.A., and Teasdale, T.W. (1982) Behavioral precursors of schizophrenia spectrum. A prospective study. *Arch. Gen. Psychiatry* **39(6),** 658–664.
41. Cornblatt, B.A. and Keilp, J.G. (1994) Impaired attention, genetics, and the pathophysiology of schizophrenia [published erratum appears in *Schizophr. Bull.* 1994;20(2):248]. *Schizophr. Bull.* **20(1),** 31–46.
42. Nuechterlein, K.H. and Dawson, M.E. (1984) Information processing and attentional functioning in the developmental course of schizophrenic disorders. *Schizophr. Bull.* **10(2),** 160–203.
43. Rutschmann, J., Cornblatt, B., and Erlenmeyer-Kimling, L. (1977) Sustained attention in children at risk for schizophrenia. Report on a continuous performance test. *Arch. Gen. Psychiatry* **34(5),** 571–575.
44. Rutschmann, J., Cornblatt, B., and Erlenmeyer-Kimling, L. (1986) Sustained attention in children at risk for schizophrenia: findings with two visual continuous performance tests in a new sample. *J. Abnorm. Child Psychol.* **14(3),** 365–385.
45. Asarnow, R.F., Steffy, R.A., MacCrimmon, D.J., and Cleghorn, J.M.

(1977) An attentional assessment of foster children at risk for schizophrenia. *J. Abnorm. Psychol.* **86(3),** 267–275.

46. Grunebaum, H., Weiss, J.L., Gallant, D., and Cohler, B.J. (1974) Attention in young children of psychotic mothers. *Am. J. Psychiatry* **131(8),** 887–891.
47. Franke, P., Maier, W., Hardt, J., and Hain, C. (1993) Cognitive functioning and anhedonia in subjects at risk for schizophrenia. *Schizophr. Res.* **10(1),** 77–84.
48. Mirsky, A.F., Lochhead, S.J., Jones, B.P., Kugelmass, S., Walsh, D., and Kendler, K.S. (1992) On familial factors in the attentional deficit in schizophrenia: a review and report of two new subject samples. *J. Psychiatr. Res.* **26(4),** 383–403.
49. Egan, M., Goldberg, T.E., Gscheidle T., Weirick, Bigelow, L., and Weinberger, D.R. (2000) Relative risk of attention deficits in siblings of patients with schizophrenia. *Am. J. Psychiatry* **157(8),** 1309–1316.
50. Keefe, R.S., Silverman, J.M., Mohs, R.C., Siever, L. J., Harvey, P.D., Friedman, L., et al. (1997) Eye tracking, attention, and schizotypal symptoms in nonpsychotic relatives of patients with schizophrenia (see comments). *Arch. Gen. Psychiatry* **54(2),** 169–176.
51. Laurent, A., Biloa-Tang, M., Bougerol, T., Duly, D., Anchisi, A.M., Bosson, J.L., et al. (2000) Executive/attentional performance and measures of schizotypy in patients with schizophrenia and in their nonpsychotic first-degree relatives. *Schizophr. Res.* **46(2–3),** 269–283.
52. Chen, W.J., Liu, S.K., Chang, C.J., Lien, Y.J., Chang, Y.H., and Hwu, H.G. (1998) Sustained attention deficit and schizotypal personality features in nonpsychotic relatives of schizophrenic patients. *Am. J. Psychiatry* **155(9),** 1214–1220.
53. Finkelstein, J.R., Cannon, T.D., Gur, R.E., Gur, R.C., and Moberg, P. (1997) Attentional dysfunctions in neuroleptic-naive and neuroleptic-withdrawn schizophrenic patients and their siblings. *J. Abnorm. Psychol.* **106(2),** 203–212.
54. Maier, W., Franke, P., Hain, C., Kopp, B., and Rist, F. (1992) Neuropsychological indicators of the vulnerability to schizophrenia. *Prog. Neuropsychopharmacol. Biol. Psychiatry* **16(5),** 703–715.
55. Saoud, M., d'Amato, T., Gutknecht, C., Triboulet, P., Bertaud, J.P., Marie-Cardine, M., et al. (2000) Neuropsychological deficit in siblings discordant for schizophrenia. *Schizophr. Bull.* **26(4),** 893–902.
56. Pogue-Geile, M.F., Garrett, A.H., Brunke, J.J., and Hall, J.K. (1991)

Neuropsychological impairments are increased in siblings of schizophrenic patients. *Schizophr. Res.* **4** (abstract), 390.

57. Franke, P., Maier, W., Hain, C., and Klingler, T. (1992) Wisconsin Card Sorting Test: an indicator of vulnerability to schizophrenia? *Schizophr. Res.* **6(3),** 243–249.
58. Scarone, S., Abbruzzese, M., and Gambini, O. (1993) The Wisconsin Card Sorting Test discriminates schizophrenic patients and their siblings. *Schizophr. Res.* **10(2),** 103–107.
59. Yurgelun-Todd, D.A. and Kinney, D.K. (1993) Patterns of neuropsychological deficits that discriminate schizophrenic individuals from siblings and control subjects. *J. Neuropsychiatry Clin. Neurosci.* **5(3),** 294–300.
60. Shedlack, K., Lee, G., Sakuma, M., Xie, S.H., Kushner, M., Pepple, J., et al. (1997) Language processing and memory in ill and well siblings from multiplex families affected with schizophrenia. *Schizophr. Res.* **25(1),** 43–52.
61. Condray, R. and Steinhauer, S.R. (1992) Schizotypal personality disorder in individuals with and without schizophrenic relatives: similarities and contrasts in neurocognitive and clinical functioning. *Schizophr. Res.* **7,** 33–41.
62. Keefe, R.S., Silverman, J.M., Roitman, S.E., Harvey, P.D., Duncan, M.A., Alroy, D., et al. (1994) Performance of nonpsychotic relatives of schizophrenic patients on cognitive tests. *Psychiatry Res.* **53(1),** 1–12.
63. Faraone, S.V., Seidman, L.J., Kremen, W.S., Pepple, J.R., Lyons, M.J., and Tsuang, M.T. (1995) Neuropsychological functioning among the nonpsychotic relatives of schizophrenic patients: a diagnostic efficiency analysis. *J. Abnorm. Psychol.* **104(2),** 286–304.
64. Toomey, R., Faraone, S.V., Seidman, L.J., Kremen, W.S., Pepple, J.R., and Tsuang, M.T. (1998) Association of neuropsychological vulnerability markers in relatives of schizophrenic patients. *Schizophr. Res.* **31(2–3),** 89–98.
65. Staal, W.G., Hijman, R., Hulshoff Pol, H.E., and Kahn, R.S. (2000) Neuropsychological dysfunctions in siblings discordant for schizophrenia. *Psychiatry Res.* **95(3),** 227–235.
66. Egan, M., Goldberg, T., Gscheidle, T., Bigelow, L., Hyde, T., and Weinberger, D. (2001a) Relative risk for cognitive impairments in siblings of patients with schizophrenia. *Biol. Psychiatry* **50(2),** 98–107.

67. Goldberg, T.E., Egan, M.F., Gscheidle, T., Coppola, R., Weickert, T., Kolachana, B.S., et al. (2001) Working memory impairments in schizophrenic sibships: relative risk estimates and effects of COMT Val108/158Met genotype. *submitted*.
68. Byrne, M., Hodges, A., Grant, E., Owens, D.C., and Johnstone, E.C. (1999) Neuropsychological assessment of young people at high genetic risk for developing schizophrenia compared with controls: preliminary findings of the Edinburgh High Risk Study (EHRS). *Psychol. Med.* **29(5),** 1161–1173.
69. Park, S., Holzman, P.S., and Goldman-Rakic, P.S. (1995) Spatial working memory deficits in the relatives of schizophrenic patients. *Arch. Gen. Psychiatry* **52(10),** 821–828.
70. Grove, W.M., Lebow, B.S., Clementz, B.A., Cerri, A., Medus, C., and Iacono, W.G. (1991) Familial prevalence and coaggregation of schizotypy indicators: a multitrait family study. *J. Abnorm. Psychol.* **100(2),** 115–121.
71. Plomin, R., DeFries, J.C., and McClearn, G.E. (1990) *Behavioral Genetics*, 2nd ed., W.H. Freeman and Co., New York, NY.
72. Chen, E.Y., Lam, L.C., Chen, R.Y., Nguyen, D.G., Chan, C. K., and Wilkins, A. J. (1997) Neuropsychological correlates of sustained attention in schizophrenia. *Schizophr. Res.* **24(3),** 299–310.
73. Goldberg, T.E., Torrey, E.F., Berman, K.F., and Weinberger, D.R. (1994) Relations between neuropsychological performance and brain morphological and physiological measures in monozygotic twins discordant for schizophrenia. *Psychiatry Res.* **55(1),** 51–61.
74. Cannon, T.D., Huttunen, M.O., Lonnqvist, J., Tuulio-Henriksson, A., Pirkola, T., Glahn, D., et al. (2000b) The inheritance of neuropsychological dysfunction in twins discordant for schizophrenia. *Am. J. Hum. Genet.* **67(2),** 369–382.
75. Kremen, W.S., Tsuang, M.T., Faraone, S.V., and Lyons, M. J. (1992) Using vulnerability indicators to compare conceptual models of genetic heterogeneity in schizophrenia. *J. Nerv. Ment. Dis.* **180(3),** 141–152.
76. Mirsky, A.F., Anthony, B.J., Duncan, C.C., Ahearn, M.B., and Kellam, S.G. (1991) Analysis of the elements of attention: a neuropsychological approach. *Neuropsychology Rev.* **2(2),** 109–145.
77. Saudino, K.J., Pedersen, N.L., and McClearn, G.E. (1994) The etiology of high and low cognitive ability during the second half of the life span. *Intelligence* **19,** 359–371.

78. Cannon, T.D., Kaprio, J., Lonnqvist, J., Huttunen, M., and Koskenvuo, M. (1998) The genetic epidemiology of schizophrenia in a Finnish twin cohort. A population-based modeling study. *Arch. Gen. Psychiatry* **55(1),** 67–74.
79. Asarnow, R.F., Steffy, R.A., MacCrimmon, D.J., et al. (1978) The McMaster Waterloo project: an attentional and clinical assessment of foster children at risk for schizophrenia, in *The Nature of Schizophrenia: New Approaches to Research and Treatment* (Wynne, L.C., Cromwell, R.L., and Matthysse, S., eds.), Wiley, New York, NY, pp. 339–358.
80. Cannon, T.D., Bearden, C.E., Hollister, J.M., Rosso, I.M., Sanchez, L.E., and Hadley, T. (2000a) Childhood cognitive functioning in schizophrenia patients and their unaffected siblings: a prospective cohort study. *Schizophr. Bull.* **26(2),** 379–393.
81. Egan, M.F., Goldberg, T.E., Kolachana, B.S., Callicott, J.H., Mazzanti, C.M., et al. (2001b) Effect of COMT Val108/158Met genotype on frontal lobe function and risk for schizophrenia. *Proc. Natl. Acad. Sci. USA* **98(12)**, 6917–6922.
82. Sawaguchi, T. and Goldman-Rakic, P.S. (1991) D1 dopamine receptors in prefrontal cortex: involvement in working memory. *Science* **251(4996),** 947–950.
83. Sawaguchi, T. and Goldman-Rakic, P.S. (1994) The role of D1-dopamine receptor in working memory: local injections of dopamine antagonists into the prefrontal cortex of rhesus monkeys performing an oculomotor delayed-response task. *J. Neurophysiol.* **71(2),** 515–528.
84. Seamans, J.K., Floresco, S.B., and Phillips, A.G. (1998) D1 receptor modulation of hippocampal-prefrontal cortical circuits integrating spatial memory with executive functions in the rat. *J. Neurosci.* **18(4),** 1613–1621.
85. Williams, G.V. and Goldman-Rakic, P.S. (1995) Modulation of memory fields by dopamine D1 receptors in prefrontal cortex. *Nature* **376(6541),** 572–575.
86. Lachman, H.M., Papolos, D.F., Saito, T., Yu, Y.M., Szumlanski, C.L., and Weinshilboum, R.M. (1996) Human catechol-O-methyltransferase pharmacogenetics: description of a functional polymorphism and its potential application to neuropsychiatric disorders. *Pharmacogenetics* **6(3),** 243–250.
87. Lotta, T., Vidgren, J., Tilgmann, C., Ulmanen, I., Melen, K., Julkunen, I., et al. (1995) Kinetics of human soluble and membrane-bound catechol O-methyltransferase: a revised mechanism and

description of the thermolabile variant of the enzyme. *Biochemistry* **34(13),** 4202–4210.

88. Karoum, F., Chrapusta, S.J., and Egan, M.F. (1994) 3-methoxytyramine is the major metabolite of released dopamine in the rat frontal cortex: reassessment of the effects of antipsychotics on the dynamics of dopamine release and metabolism in the frontal cortex, nucleus accumbens, and striatum by a simple two pool model. *J. Neurochem.* **63(3),** 972–979.
89. Gogos, J.A., Morgan, M., Luine, V., Santha, M., Ogawa, S., Pfaff, D., et al. (1998) Catechol-O-methyltransferase-deficient mice exhibit sexually dimorphic changes in catecholamine levels and behavior. *Proc. Natl. Acad. Sci. USA* **95(17),** 9991–9996.
90. Sesack, S.R., Hawrylak, V.A., Matus, C., Guido, M.A., and Levey, A.I. (1998) Dopamine axon varicosities in the prelimbic division of the rat prefrontal cortex exhibit sparse immunoreactivity for the dopamine transporter. *J. Neurosci.* **18(7)**, 2697–2708.
91. Gasparini, M., Fabrizio, E., Bonifati, V., and Meco, G. (1997) Cognitive improvement during Tolcapone treatment in Parkinson's disease. *J. Neural. Transm.* **104(8–9),** 887–894.
92. Kneavel, M., Gogos, J., Karayiorgou, K., and Luine, V. (2000) Interaction of COMT gene deletion and environment on cognition. *Society for Neuroscience 30th Annual Meeting*.
93. Liljequist, R., Haapalinna, A., Ahlander, M., Li, Y.H., and Mannisto, P.T. (1997) Catechol O-methyltransferase inhibitor tolcapone has minor influence on performance in experimental memory models in rats. *Behav. Brain Res.* **82(2),** 195–202.
94. Spielman, R.S., McGinnis, R.E., and Ewens, W.J. (1993) Transmission test for linkage disequilibrium: the insulin gene region and insulin-dependent diabetes mellitus (IDDM). *Am. J. Hum. Genet.* **52(3),** 506–516.
95. Kunugi, H., Vallada, H.P., Sham, P.C., Hoda, F., Arranz, M.J., Li, T., et al. (1997) Catechol-O-methyltransferase polymorphisms and schizophrenia: a transmission disequilibrium study in multiply affected families. *Psychiatr. Genet.* **7(3),** 97–101.
96. Li, T., Ball, D., Zhao, J., Murray, R. M., Liu, X., Sham, P.C., et al. (2000) Family-based linkage disequilibrium mapping using SNP marker haplotypes: application to a potential locus for schizophrenia at chromosome 22q11. *Mol. Psychiatry* **5(1),** 77–84.
97. Li, T., Sham, P.C., Vallada, H., Xie, T., Tang, X., Murray, R.M., et

al. (1996) Preferential transmission of the high activity allele of COMT in schizophrenia. *Psychiatr. Genet.* **6(3),** 131–133.

98. Weinberger, D.R., Aloia, M.S., Goldberg, T.E., and Berman, K.F. (1994) The frontal lobes and schizophrenia. *J. Neuropsychiatry Clin. Neurosci.* **6(4),** 419–427.
99. Weinberger, D.R. (1999) Cell biology of the hippocampal formation in schizophrenia. *Biol. Psychiatry* **45(4)**, 395–402.

8

Biochemical Endophenotypes in Personality Disorders

Antonia S. New and Larry J. Siever

1. Introduction

Candidate gene association studies in psychiatric disorders have suffered from difficulties in replication. One strategy for overcoming this difficulty is a focus on endophenotypes. If most psychiatric disorders are polygenic and are caused by numerous genes—each contributing a small effect and conferring susceptibility to the phenotype, then it is difficult to identify the influence of each of these genetic contributions to the phenotype. A recent approach, which is the focus of this volume, is the identification of endophenotypes—components of a syndrome that may be influenced by a smaller number of genes and may be more easily associated with a genotype. Other chapters (7,9,10) cover cognitive, electrophysiologic, and neuroimaging endophenotypes. This chapter reviews promising biochemical endophenotypes, using personality disorders as a model.

There is growing evidence that personality disorders carry a significant degree of biologically based temperament ***(1)***. However, the personality disorders themselves, as defined by DSM-IV, have not been scrutinized with the rigor of field trials; thus it is not surprising that studies of the replicability of diagnoses by different

From: *Methods in Molecular Medicine, vol. 77: Psychiatric Genetics: Methods and Reviews*
Edited by: M. Leboyer and F. Bellivier

diagnostic tools is low *(2)*. More productive neurobiological research has focused on specific dimensions of the personality disorders that may form biologically mediated traits, which predispose to the full-blown disorders, either separately or in synergy *(3)*.

2. Borderline Personality Disorder

Of all the personality disorders, borderline personality disorder carries a high degree of morbidity and even mortality. The temperamental traits believed to predispose to borderline personality disorder are affective instability and impulsive aggression *(3)*. This hypothesis is supported by family and twin studies of borderline personality disorder, which suggest that although the disorder itself may not be heritable, the prominent features of impulsivity and affective instability appear to run in families *(1,4)*. Studies of personality characteristics have shown that suicidality, affective instability, and impulsivity are heritable *(5)*. Impulsive aggression has also been clearly shown to be at least partially heritable, as established by both twin *(6)* and adoption *(7)* studies, with suggested heritability estimates ranging from 20%–62% *(6)*.

2.1. Personality Disorders and Impulsive Aggression

2.1.1. Phenomenology

Impulsive-aggressive behavior in clinical populations of patients with personality disorders is a common clinical phenomenon, and contributes to much of the dysfunction associated with these diagnoses. In particular, a high rate of repeated self-mutilation and episodic dyscontrol has been reported in patients with borderline, antisocial, and histrionic personality disorders, (reflecting impulsive aggression) *(8,9)*. This is not surprising because in DSM-IV, assaultiveness is one of the diagnostic criteria for antisocial personality disorder, and repetitive suicide attempts, self-mutilation, and impulsivity are among the diagnostic criteria for borderline personality disorder. Our data suggest that 51% of patients with personality disorder meet the criteria for intermittent explosive disor-

der-revised research version; of this group meeting IED-R, 64% met borderline or antisocial personality disorder criteria ***(10)***. Suicidal behavior, which is often an impulsive act of self-directed aggression, is also common in patients with personality disorders. In fact, 9% of patients with borderline personality disorder die by suicide ***(11,12)***. As pharmacological and other treatments for mood disorders and psychotic disorders have improved, an increasing proportion of patients who require psychiatric hospitalization carry a primary personality diagnosis ***(13–15)***. A common precipitant for psychiatric hospitalization in patients with personality disorders is an actual or threatened act of aggression directed toward the self or others, and yet the phenomenon of impulsive aggression has not been fully studied.

Further evidence of the significance of impulsive aggressive symptomatology in patients with personality disorders is demonstrated by the frequency of personality disorders in the forensic population. In one study of violent offenders and impulsive firesetters, 47% of the subjects were found to have a personality disorder diagnosis, especially borderline and antisocial personality disorders ***(16)***. In a sample of wife-batterers, higher scores on a measure of borderline personality organization (similar to scores seen in a sample of borderline patients diagnosed by DSM-III criteria) were found compared to controls ***(17)***. Prison inmates in Quebec with antisocial personality disorder were found to have more convictions and an earlier onset of criminal activity than those who did not have antisocial personality disorder ***(18)***. Furthermore, the presence of antisocial personality disorder has been found to increase the likelihood of homicidal violence 10-fold in men and 50-fold in women ***(19)***.

2.1.2. Evidence for Heritability of Impulsive Aggression and Suicidality

Twin and family studies suggest a partially heritable basis for impulsive aggression ***(6,7,20)***. Preliminary data from monozygotic-dizygotic twin studies suggest that although the personality disor-

der diagnoses are not heritable, the traits of impulsive aggression or assertive aggressiveness are significantly heritable ***(6,21,22)***. Studies of adoptees with biological parents with antisocial personality disorder demonstrated a genetic contribution to the development of antisocial personality disorder characterized by antisocial aggressive behavior ***(23,24)***, and this heritable tendency appeared to be brought out by an adverse home environment ***(25)***. Twin studies of suicide also demonstrate a greater concordance for suicidal behavior in monozygotic (MZ) vs dizygotic (DZ) twins, and adoption studies also suggest a heritable component to suicide that is independent of the risk for affective illness itself ***(26–28)***.

2.1.3. Impulsive Aggression and Serotonergic Studies

Abnormalities in central serotonergic activity have been consistently found to be associated with measures of impulsive aggression in patients with personality disorders ***(29–30)***. Studies of cerebrospinal fluid (CSF) have shown that a decrease in 5-hydroxyindolacetic acid (5-HIAA), a metabolite of serotonin, is associated with impulsive aggression in patients with personality disorders, as well as in depressed patients, volunteers, and violent alcoholic offenders ***(31,32)***. The prolactin response to fenfluramine, a serotonin-releasing agent, can be viewed as a measure of net central serotonergic activity ***(29)***, and a blunted response has been associated with impulsive aggression in patients with personality disorders compared to normal controls ***(29,33)***. A blunted response to fenfluramine has also been demonstrated in patients with antisocial personality disorder, a diagnosis that is strongly linked to impulsive aggression ***(30).*** Studies of serotonergic activity in relation to suicide demonstrate low CSF 5-HIAA in patients with a history of violent suicide attempts ***(34,35)***. Postmortem studies have demonstrated decreased imipramine binding in the brains of suicide victims, and more specifically, increases in 5-HT1A autoreceptors in the midbrain of suicide victims ***(36)***, alterations in binding to the serotonin transporter, 5-HT1A and 5-HT2A, in the postmortem brains of suicide victims compared to controls ***(37)***. A blunted neu-

roendocrine response to serotonin-stimulating agents has been found in patients with a history of suicide ***(29,38)***. However, depressed patients with "anger attacks" had a blunted prolactin response to fenfluramine compared to depressed patients with no anger attacks ***(39)***. This suggests that it may be the subset of depressed subjects with irritability and angry outbursts that is most closely associated with decreased serotonergic activity.

Serotonergic agents have been shown to modulate selectively specific regional brain glucose metabolism. In a study comparing glucose metabolism following fenfluramine and placebo in patients with personality disorder and impulsive aggression, normal subjects showed increased metabolism in the orbital frontal and adjacent ventral medial frontal cortex as well as the cingulate and inferior parietal cortex following fenfluramine compared to placebo, and impulsive-aggressive patients appear to show significant increases only in the inferior parietal lobe ***(40)***. Similar results were found in a study of patients with borderline personality disorder ***(41).***

2.2. Affective Instability in Personality Disorders

2.2.1. Phenomenology

Affective instability has been defined as "a predisposition to marked, rapidly reversible shifts in affective state that are extremely sensitive to meaningful environmental events" ***(3)***. Affective instability is one of the defining characteristics of borderline personality disorder in DSM-IV. However, some believe that affective instability is secondary to "affective tempermental dysregulation," and view the borderline diagnosis as representing a part of the bipolar spectrum ***(42)***.

2.2.2. Neuroendocrine

Although the neurochemistry of affective instability has been studied less than that of impulsive aggression, there is some evidence that the cholinergic system plays a role in regulating affect. Depressive symptoms can be elicited with the administration of

cholinomimetics ***(43)***, and patients with depression experience an increase in depression, hostility, and anxiety in response to arecholine ***(44)***. Although cholinergic agents appear to induce a depressed mood in normals and patients with depression, this effect is even more pronounced in patient with borderline personality disorder. Procaine, a pro-cholinergic agent, has been shown to induce a high degree of dysphoria in borderline personality disorder compared to patients with affective disorders and normal controls ***(45)***. The cholinomimetic known as phsyostigmine has been shown to produce significantly more depressive symptoms in borderline patients than placebo, and this effect is greater and more rapid than that seen in healthy controls ***(46)***.

There is also evidence of a disturbance in noradrenergic activity in affectively unstable patients with borderline personality disorder ***(47)***. In healthy subjects, a dysphoric response to dextroamphetamine, a catecholaminergic agent, correlated with measures of affective instability ***(46)***. Furthermore, a heightened growth hormone (GH) response to the alpha-2 agonist, clonidine, has been demonstrated in subjects who are highly reactive to their environment ***(48)***. The GH response to clonidine directly correlated with measures of irritability—which is related to affective instability—in patients with borderline personality disorder but not in depressed subjects ***(49)***. In positron emission tomography (PET) studies of normal subjects, the dysphoric response to procaine correlated with metabolic activation of the left amygdala ***(50)***.

3. Schizotypal Personality Disorder: The Schizophrenia Spectrum

Schizotypal personality disorder (SPD) appears to be related to schizophrenia, phenomenologically as well as genetically and biologically ***(51).*** The disorders are similar in that schizotypal patients display psychotic-like symptoms, including ideas of reference, perceptual distortions, and suspicion, similar to the psychotic symptoms seen in schizophrenia but less severe. In addition, there is evidence that SPD is over-represented in the families of patients

with schizophrenia, and that schizophrenia also occurs more frequently in the families of patients with SPD ***(52,53).***

Research from our laboratory and others suggests that people with SPD demonstrate cognitive impairments that are similar to but less severe than those that occur in schizophrenia. Although schizotypal individuals do not necessarily show the more global cognitive deterioration and broad decreases in intellectual function observed in chronic schizophrenic patients, they do show more selective impairment in working memory (defined as the ability to maintain and manipulate information "on line"), verbal learning (the capacity to learn verbal information over a series of presentation), and sustained attention (the ability to sustain focus on stimuli in the environment).

3.1. Genetic/Family

Family, twin, and adoptive studies have demonstrated a genetic relationship between SPD and schizophrenia, with an increased incidence of SPD in relatives of schizotypal probands ***(51–54)*** and an increase in schizophrenia-related disorders ***(53)*** and schizophrenia itself in the families of patients with SPD ***(55,56)***. The likelihood of having a schizophrenic relative is comparable for probands diagnosed with either SPD or schizophrenia (6.9% vs 6.5%), supporting the common genetic substrates of the two disorders ***(57)***. Both the deficit-like symptoms, which are associated with attentional dysfunction, and psychotic-like symptoms seem to be independently heritable (58). The deficit-like symptoms and cognitive impairment seem to provide core pathology across the spectrum, possibly reflecting a common neurodevelopmental abnormality.

3.2. Structural and Functional Imaging Studies

A growing body of work in this area has begun to identify neuroanatomic similarities and differences between SPD and schizophrenia. Our laboratory and others have shown that lateral ventricle-to-brain ratio (VBR) is increased in SPD patients as well as in schizophrenic patients ***(59–62)***. We have reported ***(60)*** that an

increased VBR involving the lateral ventricle and the frontal horn correlated with increased perseverative errors on the WCST, and frontal-horn VBR correlated with an increased omission error rate in the degraded stimulus continuous performance task (CPT). Magnetic resonance imaging (MRI) studies have identified specific structural abnormali-ties in SPD patients, similar to those seen in schizophrenia ***(63)***, but generally of lesser magnitude and sparing some brain regions.

3.2.1. Functional Imaging

A SPECT study comparing regional cerebral blood flow in SPD subjects and normal controls during performance of the WCST found that the frontal region with the greatest activation in the control subjects was the precentral gyrus. In contrast, SPD subjects showed greatest activation in the middle frontal gyrus ***(61)***. In normal subjects (but not the SPD subjects), task performance was correlated with left PFC activation. The SPD subjects showed a positive correlation between good WCST performance and CBF in the right inferior frontal gyrus, and poor performance correlated with CBF in the right middle frontal gyrus (part of the dorsolateral PFC). This finding suggests that SPD subjects recruit different frontal regions in performance the WCST than control subjects, possibly reflecting dysfunction in the region and/or compensatory strategies to address dysfunction in other areas. PET studies by our group also suggest anomalous lateralization and underactivation of temporal regions as well as altered prefrontal activity in SPD patients during the CVLT. A PET study, also carried out in our laboratory, which examined activation of the thalamic nuclei during a serial word list learning task modeled on the CVLT, found decreased relative metabolism in the mediodorsal nucleus (linked to the frontal cortex), occurring bilaterally in schizophrenic patients but not in SPD patients or normal controls ***(64).*** These findings raise the possibility that cognitive impairment in schizophrenia and SPD may be associated in part with anomalous prefrontal cortical activity, which generally diminished in schizophrenia, but may reflect compensatory function in SPD.

3.2.2. Neurochemistry

Studies of the dopamine metabolite homovanillic acid (HVA) have shown a positive correlation of plasma HVA levels with the psychotic-like symptoms and a negative correlation of HVA levels with deficit-like symptoms and cognitive impairment in patients with SPD and relatives of patients with schizophrenia ***(3,60,65,66)***. Studies of plasma HVA response to a physiologic stressor (a 2-deoxyglucose infusion) ***(67)*** and single-photon emission computed tomography (SPECT) studies of subcortical dopamine release following amphetamine infusions both suggest that SPD patients have a subcortical dopamine responsivity that is decreased compared to schizophrenic patients and more like that of normal controls. These neurochemical findings suggest that *decreased frontal* dopaminergic activity is associated with deficit symptoms and cognitive impairment, and *increased subcortical* dopaminergic activity correlates with psychotic symptoms, and that SPD patients may be less likely to experience psychotic symptoms than schizophrenic patients if they are administered dopaminergic agents.

4. Conclusion

Together, these studies of personality disorders suggest an approach that includes enhancing the phenomenological study of personality diagnoses, with biochemical and neuroimaging characterization of dimensions related to the underlying disorder. This approach will help to identify endophenotypes, which will make genetic studies in this complex area much more powerful.

References

1. Torgersen, S. (2000) Genetics of patients with borderline personality disorder. *Psychiatr. Clin. North Am.* **23(1),** 1–9.
2. Kavoussi, R.J., Coccaro, E.F., Klar, H.M., Bernstein, D., and Siever, L.J. (1990) Structured interviews for borderline personality disorder. *Am. J. Psychiatry* **147(11),** 1522–1525.

3. Siever, L.J. and Davis, K.L. (1991) A psychobiologic perspective on the personality disorders. *Am. J. Psychiatry* **148,** 1647–1658.
4. Silverman, J.M., Pinkham, L., Horvath, T.B., Coccaro, E.F., Howard, K., Schear, S., et al. (1991) Affective and impulsive personality disorder traits in the relatives of patients with borderline personality disorder. *Am. J. Psychiatry* **148(10),** 1378–1385.
5. Bouchard, T.J., Jr. (1994) Genes, environment, and personality. *Science* **264(5166),** 1700–1701.
6. Coccaro, E.F., Bergman, C.S., and McLean, G.E. (1993) Heritability of irritable impulsiveness: a study of twins reared together and apart. *Psychiatry Res.* **48(3),** 229–242.
7. Bohman, M., Cloninger, C.R., von Knorring, A., and Sigvardsson, S. (1984) An adoption study of somatoform disorders: III. Cross-fostering analysis and genetic relationship to alcoholism and criminality. *Arch. Gen. Psychiatry* **41,** 872–878.
8. Virkkunen, M. (1976) Reactive hypoglycemic tendency among habitually violent offenders. *Nutritional Review* **44(Suppl.),** 94–103.
9. Pattison, E. and Kahan, J. (1983) The deliberate self-harm syndrome. *Am. J. Psychiatry* **140(7),** 867–872.
10. New, A.S., Mitropoulou, V., Coccaro, E.F., and Siever, L.J. (1999) The phenomenology of impulsive aggressive symptoms. Abstract from the Annual Meeting of the Society of Biological Psychiatry, Vol. 45, #457.
11. Stone, M.H. (1989) Long-term follow-up of narcissistic/borderline patients. *Psychiatr. Clin. N. Am.* **12(3),** 621–641.
12. Paris, J., Brown, R., and Nowlis, D. (1987) Long-term follow-up of borderline patients in a general hospital. *Compr. Psychiatry* **28(6),** 530–535.
13. Hansson, L. (1989) Utilization of psychiatric inpatient care: a study of changes related to the introduction of a sectorized care organization. *Acta Psych. Scand.* **79(6),** 571–578.
14. Breslow, R., Klinger, B., and Erickson, B. (1993) Crisis hospitalization on a psychiatric emergency service. *Gen. Hosp. Psychiatry* **15(5),** 307–315.
15. Molinari, V., Ames, A., and Essa, M. (1994) Prevalence of personality disorders in two geropsychiatric inpatient units. *J. Geriatr. Psychiatry Neurol.* **7(4),** 295–300.
16. Virkkunen, M., Eggert, M., Rawlings, R., and Linnoila, M. (1996) A prospective follow-up study of alcoholic violent offenders and fire setters. *Arch. Gen. Psychiatry* **53,** 523–529.

17. Dutton, D.G., Starzomski, A., and Ryan, L. (1996) Antecedents of abusive personality and abusive behavior in wife assaulters. *J. Family Violence* **11(2),** 113–132.
18. Hodgin, S. and Cote, G. (1993) Major mental disorder and antisocial personality disorder: a criminal combination. *Bull. Am. Acad. Psychiatry Law* **21(2),** 155–160.
19. Eronen, M., Hakola, P., and Tihonen, J. (1996) Mental disorders and homicidal behavior in Finland. *Arch. Gen. Psychiatry* **53,** 497–501.
20. Coccaro, E., Silverman, J., Klar, H., Horvath, T., and Siever, L.J. (1994) Familial correlates of reduced central serotonergic system function in patients with personality disorders. *Arch. Gen. Psychiatry* **51(4),** 318–324.
21. Alnaes, R. and Torgersen, S. (1989) Clinical differentiation between major depression only, major depression with panic disorder and panic disorder only: childhood, personality and personality disorder. *Acta Psychiatr. Scand.* **79(4),** 370–377.
22. Torgersen, S. (1992) Genetics in borderline conditions. *Acta Psychiatr. Scand. Suppl.* **(379)**, 19–25.
23. Cadoret, R.J., O'Gorman, T.W., Troughton, E., and Heywood, E. (1985) Alcoholism and antisocial personality: interrelationships, genetic and environmental factors. *Arch. Gen. Psychiatry* **42(2),** 161–167.
24. Crowe, R.R. (1974) An adoption study of antisocial personality. *Arch. Gen. Psychiatry* **31(6),** 785–791.
25. Cadoret, R.J., Yates, W.R., Troughton, E., Woodworth. G., and Stewart, M.A. (1995) Genetic-environmental interaction in the genesis of aggressivity and conduct disorders. *Arch. Gen. Psychiatry* **52(11),** 916–924.
26. Roy, A., Rylabder, G., and Sarchiapone, M. (1997) Genetic studies of suicidal behavior. *Psychiatr. Clin. N. Am.* **20(3),** 595–611.
27. Mitterauer, B. (1990) A contribution to the discussion of the role of the genetic factor in suicide, based on five studies in an epidemiologically defined area (Province of Salzburg, Austria). *Compr. Psychiatry* **31(6),** 557–565.
28. Egeland, J.A. and Sussex, J.N. (1985) Suicide and family loadings for affective disorders. *JAMA* **254(7),** 915–918.
29. Coccaro, E.F., Siever, L.J., Klar, H.M., Maurer, G., Cochrane, K., Cooper, T.B., et al. (1989) Serotonergic studies in patients with affective and personality disorders. *Arch. Gen. Psychiatry* **46,** 587–599.

30. O'Keane, V., Moloney, E., O'Neill, H., O'Connor, A., Smith, C., and Dinan, T.C. (1992) Blunted prolactin responses to d-fenfluramine in sociopathy: evidence for subsensitivity of central serotonergic function. *Br. J. Psychiatry* **160,** 643–646.
31. Linnoila, M., DeJong, J., and Virkkunen, M. (1989) Family history of alcoholism in violent offenders and impulsive fire setters. *Arch. Gen. Psychiatry* **46,** 613–616.
32. Virkkunen, M., Rawlings, R., Tokola, R., Poland, R.E., Guidotti, A., Nemeroff, C., et al. (1994) CSF biochemistries, glucose metabolism, and diurnal activity rhythms in alcoholic violent offenders, fire setters, and healthy volunteers. *Arch. Gen. Psychiatry* **51(1),** 20–27.
33. Coccaro, E.F., Berman, M.E., Kavoussi, R.J., and Hager, R.L. (1996) Relationship of prolactin response to d-fenfluramine to behavioral and questionnaire assessments of aggression in personality disordered men. *Soc. Biol. Psychiatry* **37,** 1–8.
34. Asberg, M., Traskman, L., and Thoren, P. (1976) 5-HIAA in the cerebrospinal fluid. A biochemical suicide predictor? *Arch. Gen. Psychiatry* **33(10),** 1193–1197.
35. Asberg, M. (1997) Neurotransmitters and suicidal behavior: the evidence from cerebrospinal fluid studies. *Ann. NY Acad. Sci.* **836,** 158–181.
36. Stockmeier, C.A., Shapiro, L.A., Dilley, G.E., Kolli, T.N., Friedman, L., and Rajkowska, G. (1998) Increase in serotonin-1A autoreceptors in the midbrain of suicide victims with major depression-postmortem evidence for decreased serotonin activity. *J. Neurosci.* **18(18),** 7394–7401.
37. Arango, V., Underwood, M.D., and Mann, J.J. (1997) Postmortem findings in suicide victims. Implications for in vivo imaging studies. *Ann. NY Acad. Sci.* **836,** 269–287.
38. Weiss, D. and Coccaro, E.F. (1997) Neuroendocrine challenge studies of suicidal behavior. *Psychiatr. Clin. N. Am.* **20(3),** 563–579.
39. Fava, M., Vuolo, R.D., Wright, E.C., Nierenberg, A.A., Alpert, J.E., and Rosenbaum, J.F. (2000) Fenfluramine challenge in unipolar depression with and without anger attacks. *Psychiatry Res.* **94(1),** 9–18.
40. Siever, L.J., Buchsbaum, M., New, A.S., Spiegel-Cohen, J., Wei, C.T., Hazlett, E., et al. (1999) d,l-fenfluramine response in impulsive personality disorder assessed with 18F-deoxyglucose positron emission tomography. *Neuropsychopharmacology* **20(5),** 413–423.
41. Soloff, P.H., Meltzer, C.C., Greer, P.J., Constantine, D., and Kelly, T.M. (2000) A fenfluramine-activated FDG-PET study of borderline personality disorder. *Biol. Psychiatry* **47,** 540–547.

42. Akiskal, H.S., Bourgeois, M.L., Angst, J., Post, R., Moller, H., and Hirschfeld, R. (2000) Re-evaluating the prevalence of and diagnostic composition within the broad clinical spectrum of bipolar disorders. *J. Affect. Disord.* **59(Suppl. 1),** S5–S30.
43. Janowsky, D.S., el-Yousef, M.K., and Davis, J.M. (1974) Acetylcholine and depression. *Psychosom. Med.* **36(3),** 248–257.
44. Risch, S.C., Janowsky, D.S., and Gillin, J.C. (1983) Muscarinic supersensitivity of anterior pituitar, ACTH and beta-endophén release in major depressive illness. *Peptides* **4(5),** 789–792.
45. Kellner, C.H., Post, R.M., Putnam, F., Cowdry, R., Gardner, D., Kling, M.A., et al. (1987) Intravenous procaine as a probe of limbic system activity in psychiatric patients and normal controls. *Biol. Psychiatry* **22(9),** 1107–1126.
46. Steinberg, B.J., Trestman, R., Mitropoulou, V., Serby, M., Silverman, J., Coccaro, E., et al. (1997) Depressive response to physostigmine challenge in borderline personality disorder patients. *Neuropsychopharmacology* **17(4),** 264–273.
47. Gurvits, I.G., Koenigsberg, H.W., and Siever, L.J. (2000) Neurotransmitter dysfunction in patients with borderline personality disorder. *Psychiatr. Clin. N. Am.* **23(1),** 27–40.
48. Lane, R.D., Reiman, E.M., Ahern, G.L., Schwartz, G.E., and Davidson, R.J. (1997) Neuroanatomical correlates of happiness, sadness, and disgust. *Am. J. Psychiatry* **154(7),** 926–933.
49. Coccaro, E.F., Harvey, P.H., Kupshaw-Lawrence, E., Herbert, J.L., and Bernstein, D.P. (1991) Development of neuropharmacologically based assessments of impulsive aggressive behavior. *J. Neuropsychiatry* **3(Suppl.),** 44–51.
50. Ketter, T.A., Andreason, P.J., George, M.S., Lee, C., Gill, D.S., Parekh, P.I., et al. (1996) Anterior paralimbic mediation of procaine-induced emotional and psychosensory experiences. *Arch. Gen. Psychiatry* **53,** 59–69.
51. Gunderson, J.G., Siever, L.J., and Spaulding, E. (1983) The search for a schizotype: crossing the border again. *Arch. Gen. Psychiatry* **40,** 15–22.
52. Kendler, K.S., Gruenberg, A.M., and Strauss, J.S. (1981) An independent analysis of the Copenhagen sample of the Danish adoption study of schizophrenia. *Arch. Gen. Psychiatry* **38,** 982–987.
53. Siever, L.J., Silverman, J.M., Horvath, T.B., Klar, H., Coccaro, E., Keefe, R., et al. (1990) Increased morbid risk for schizophrenia-

related disorders in relatives of schizotypal personality disordered patients. *Arch. Gen. Psychiatry* **47,** 634–640.

54. Baron, J.M., Gruen, R., Asnis, L., et al. (1983) Familial relatedness of schizophrenia and schizotypal stats. *Am. J. Psychiatry* **140,** 1437–1442.
55. Schulz, P.M., Schulz, S.C., Goldberg, S.C., et al. (1986) Diagnoses of the relatives of schizotypal outpatients. *J. Nerv. Ment. Dis.* **174,** 457–463.
56. Battaglia, M., et al. (1995) A family study of schizotypal disorder. *Schizophr. Bull.* **21(1),** 33–46.
57. Kendler, K.S., McGuire, M., Gruenberg, A.M., O'Hare, A., Spellman, M., and Walsh, D. (1993) The roscommon family study I. Methods, diagnosis of probands and risk of schizophrenia in relatives. *Arch. Gen. Psychiatry* **50,** 527–540.
58. Kendler, K.S., Ochs, A.L., Gorman, A.M., et al. (1991) The structure of schizotypy: a pilot multitrait twin study. *Psychiatry Res.* **36,** 19–36.
59. Cazzulo, V.A., Giobbio, G.M., et al. (1991) Cerebral structural abnormalities in schizophreniform disorder and in schizophrenia spectrum personality disorders, in *Advances in Neuropsychiatry and Psychopharmacology 1*. (Tamminga, C.A. and Schulz, S.C., eds.). Raven Press, New York, NY, pp. 209–217.
60. Siever, L.J., Rotter, M., Losonczy, M., Guo, S.L., Mitropoulou, V., Trestman, R.L., et al. (1995) Lateral ventricle enlargement in schizotypal personality disorder. *Psychiatry Res.* **57,** 109–118.
61. Buchsbaum, M., Yang, S., Hazlett, E., Siegel, B., Germans, M., Haznedar, M., et al. (1997) Ventricular volume and assymmetry in schizotypal personality disorder and schizophrenia assessed with magnetic resonance imaging. *Schizophr. Res.* **27,** 45–53.
62. Silverman, J.M., Smith, C.J., Guo, S.L., Mohs, R.C., Siever, L.J., and Davis, K.L. (1998) Lateral ventricular enlargement in schizophrenic probands and their siblings with schizophrenia-related disorders. *Biol. Psychiatry* **43,** 97–106.
63. Dickey, C.C., Shenton, M.E., Hirayasu, Y., Fischer, I., Voglmaier, M.M., Niznikiewicz, M.A., et al. (2000) Large CSF volumes not attributable to ventricular volume in schizotypal personality disorder. *Am. J. Psychiatry* **157,** 48–54.
64. Hazlett, E.A., Buchsbaum, M.S., Byne, W., Wei, T.C., Spiegel-Cohen, J., Geneve, C., et al. (1999) Three-dimensional analysis with

MRI and PET of the size, shape and function of the thalamus in the schizophrenia spectrum. *Am. J. Psychiatry* **156,** 1190–1199.

65. Siever, L.J. and Trestman, R.L. (1993) The serotonin system and aggressive personality disorder. International scientific symposium: advances in the pharmacology and clinical applications of serotonin. *Int. Clin. Psychopharmacol.* **8(2),** 33–39.
66. Siever, L.J., Kalus, O., and Keefe, R. (1993) The boundaries of schizophrenia. *Psychiatr. Clin. N. Am.* **16(2),** 217–244.
67. Goodman, M., Mitropoulou, V., New, A.S., Koenigsberg, H.W., and Siever, L.J. (2000) Frontal cortex dysfunction and subcortical dopaminergic activity in schizophrenia spectrum. *Biol. Psychiatry* **47,** 34S.

9

Electrophysiological Phenotypes

Robert Freedman

1. Introduction

The rationale for the use of an endophenotype in genetic studies of psychiatric disorders is that these disorders are likely to be multidetermined, with multiple environmental and genetic origins, so that a discrete biological phenotype is more likely to reflect a single major gene effect. This strategy has been used successfully in the linkage and segregation analysis of several illnesses, including hemochromatosis, for which elevated serum iron was used as the relevant endophenotype, and colon cancer, for which elevated numbers of polyps were used as the phenotype. For psychiatric illnesses, various phenotypic approaches have been suggested. These include the use of variants within the illness, such as differences in age of onset, which may reflect differences in disease (*see* Chapter 6) and the use of personality characteristics as indicators of underlying biological differences. Physiological mechanisms, the focus of this chapter, are favored because of their presumed relationship to synaptic neurotransmission, a basic unit of function in the central nervous system. This chapter describes three examples of physiological endophenotypes that have been used successfully in linkage

From: *Methods in Molecular Medicine, vol. 77: Psychiatric Genetics: Methods and Reviews*
Edited by: M. Leboyer and F. Bellivier

analysis, and discusses general issues in their use and the interpretation of results.

2. Eye-Tracking Dysfunction in Schizophrenia

Although eye-tracking dysfunction in schizophrenia was first discovered in the nineteenth century, its rediscovery in modern times was particularly fortuitous because it came at a time when the genetic basis of schizophrenia was being debated. Eye tracking is recorded in a smooth-pursuit task, in which the subject is asked to follow a slowly moving target, such as a pendulum or its equivalent, which is programmed as a dot moving back and forth on a computer video screen. The task requires activation of the frontal eye fields, which function as a smooth-pursuit system that predicts the movement of the target and maintains the target in the center of the retinal fovea. The frontal eye fields coordinate this pursuit system with visual input from the temporal and occipital cortices, as well as with vestibular input, which accounts for the position of the head itself. Integration of information in the thalamus and cerebellum is performed, before the final output of movement signals through the superior colliculus to neurons in the reticular formation and brainstem motor nuclei. The eye movements, recorded by a variety of non-invasive means, are thus the output of a widely distributed neuronal system. The system can fail in two major ways. The first is that the smooth-pursuit system does not maintain foveal fixation on the target; this function is generally calculated as gain or the velocity of the eyes as a fraction of the velocity of the target. The second is the intrusion of the activity of a second eye movement system, the saccadic system, which is responsible for rapid searches of the visual field. Saccadic movements can override the smooth-pursuit movements, moving the eyes ahead of the target. These anticipatory saccades generally occur in the direction of the target, so that they do not represent a failure of voluntary effort. Rather, they appear to represent a failure to inhibit the activity of the saccadic system at a more involuntary level. After one of these saccades, the smooth-pursuit system slows so that the target reaches the fovea. An exact

opposite sequence can occur: the smooth-pursuit system can lose the target and then a catch-up saccade is generated to refix the target on the fovea. Eye-tracking dysfunction can be measured in several ways. Some researchers measure smooth-pursuit gain and various types of saccadic intrusions using computerized image recognition systems, and others measure the overall performance of the system by considering movements with the same frequency as the target as a signal and all other movements as noise, so that a signal-to-noise ratio can be calculated. Other variants include the use of different systems for measuring eye movement. Generally infrared reflectometry—which triangulates the position of the eye by reflectance off the scleral-ictal junction—is preferred to electro-oculography, which uses skin electrodes near the eyes to measure the movement of the retina, which acts as a 70-mV dipole.

Holzman, who initiated the rediscovery, immediately appreciated the value of eye-tracking dysfunction as a potential endophenotype. He examined monozygotic (MZ) twin and dizygotic (DZ) twin pairs, ascertained by the presence of an affected member. The monozygotic twin pairwise concurrence for schizophrenia and related psychoses was 0.40, consistent with the research findings. Eye tracking was recorded as an electro-oculographic signal while the subject followed a pendulum oscillating at 0.4 Hz. Qualitative ratings of the eye tracking as normal deviant yielded a perfect concordance rate of 1.0 in monozygotic twins. The natural logarithm of the ratio of signal (0.4–0.8 Hz) to the noise (1.2–20.0 Hz) was used as a quantitative variable, with similar results. Thus, the concordance rate for eye-tracking dysfunction exceeded that for schizophrenia, which suggested that the endophenotype of eye tracking was inherited as a major gene effect, whereas schizophrenia itself occurred in a subset of individuals who had inherited this genetic risk ***(1)***.

Subsequently, Holzman showed that the descendants of twins who were discordant for schizophrenia segregated eye-tracking dysfunction in a distribution consistent with single-gene inheritance. Eye movements were recorded by infrared reflectometry and scored qualitatively. The segregation analysis estimated a disease allele frequency of 2.23–3.85% with a penetrance of 8.6–11.4% for

schizophrenia and 62.6–65.5% for eye-tracking dysfunction. The segregation analysis showed evidence for partial pleiotrophy. Certain families in which the schizophrenic proband did not have eye-tracking dysfunction, had clinically unaffected parents and siblings who showed eye-tracking dysfunction. Holzman and his colleagues concluded that a single gene could have pleiotrophic effects that resulted in its expression as eye-tracking dysfunction, schizophrenia, or both conditions ***(2)***.

Arolt and his colleagues used eye-tracking dysfunction as a phenotypic marker in a series of studies of 10 multiply affected families. Four of the 21 schizophrenic individuals, including two with schizoaffective disorder, did not have eye-tracking dysfunction. Nineteen of the 39 individuals who had eye-tracking dysfunction did not have schizophrenia. Arolt and his colleagues examined their infrared reflectometry eye-tracking recordings quantitatively for both types of dysfunction—saccadic intrusion or decreased gain—eye-tracking dysfunction was rated if both parameters were greater than two standard deviations above normal or if one parameter exceeded 2.5 standard deviations above normal ***(3)***.

Genetic analyses were performed independently for the two phenotypes, eye-tracking dysfunction and schizophrenia. Similar disease allele frequencies were assumed: 0.0300 for eye-tracking dysfunction and 0.0354 for schizophrenia. Eye-tracking dysfunction was assumed to be a highly penetrant phenotype with few phenocopies, so that penetrances were 0.98 for homozygotes, 0.90 for heterozygotes, and 0.0060 for normal genotypes. The population prevalence was estimated as 0.059. Schizophrenia was assumed to be less penetrant with more phenocopies, so that penetrances were 0.85 for homozygotes, 0.425 for heterozygotes, and 0.068 for normal genotypes. The population prevalence was estimated to be 0.01. Eye-tracking dysfunction is linked ($Z = 3.70$, theta = 0.0) to D6S271, a marker just centromeric to the HLA site on chromosome 6. The lod score for schizophrenia was positive, but lower ($Z = 1.02$, theta = 0.1). The 6p21 is near, but not identical to, sites found to be linked to schizophrenia by other investigators.

3. P300 and Alcoholism

The P300 is an evoked potential wave that is generally recorded at Pz, an electrode in the center of the scalp near the parietal cortex. It is not evoked by a specific sensory stimulus, but rather by a stimulus that is recognized as a target. Generally, these targets are embedded in a stream of nontargets, so that sometimes the task is called the oddball task, because a rare deviant stimulus (the oddball) is the target. Subjects either count the targets or press a response button. Targets can be presented in any sensory modality. Auditory P300 is reduced in amplitude in schizophrenia, whereas visual P300 amplitude is reduced in alcoholism, as well as in some individuals who are at a genetic risk for alcoholism.

Begleiter and his colleagues used the amplitude of the P300 visually evoked potential, recorded in a target-detection task, to examine the heritability of alcoholism in multiply affected families. Their approach takes advantage of quantitative linkage methods that do not depend upon establishing normal and deviant ranges. A two-stage process was used in this study. First, they scanned the genome using alcoholism alone as a phenotype. A positive, but non-significant lod score was observed at chromosome 4 ($Z = 2.76$). They then employed a bivariate approach, which maximizes the lod score at each locus by estimating pleiotrophy between the clinical diagnosis of alcoholism and the P300 amplitude. Maximum lod scores were obtained near the class I alcohol dehydrogenase locus on chromosome 4 ($Z = 4.75$), using both phenotypes jointly. At this locus, the likelihood analysis found complete pleiotrophy—i.e., the same gene contributed to both the risk for alcoholism and the P300 amplitude. The ADH gene had previously shown evidence for linkage in other studies. Because the Begleiter group used non-parametric linkage methods, they did not estimate disease allele frequency and penetrance ***(4)***.

In an earlier study, Begleiter's group analyzed P300 amplitude alone in the same set of families, but did not consider the clinical diagnosis of alcoholism. Several loci were found to show positive

evidence for linkage, including markers at 6q22.3 (Z = 3.41) and 2q35 (Z = 3.28). The chromosome 4 locus did not show evidence independently for linkage to P300 amplitude ***(5)***.

4. Inhibition of P50 and Schizophrenia

The classic test for inhibitory neuronal function is the conditioning-testing paradigm, in which the response to paired stimuli is measured. The first stimulus excites the principal neurons under study, but also activates or conditions inhibitory mechanisms mediated by a second set of neurons, generally termed interneurons. The effect of these inhibitory mechanisms is not observed until the second or test stimulus is presented. The decrement in the response to the test stimulus, compared to the response to the first or conditioning stimulus, is the test of the strength of the inhibition activated or conditioned during the response to the first stimulus.

The conditioning-testing paradigm has been applied to schizophrenia by recording the cerebral evoked response to pairs of auditory stimuli. The P50 wave of the auditory evoked response showed a significant decrement in normal subjects, but little or no decrement was seen in recordings from patients with schizophrenia. Modeling of the normal response to repeated sounds at the single-neuron level in animals provided the most direct insights into the mechanism of pathophysiology. The pyramidal neurons of the CA3 region of the hippocampus were identified as the source of cerebral evoked responses that show the most marked decrement to repeated stimuli. The decrement depends upon cholinergic stimulation of hippocampal inhibitory interneurons. With such cholinergic stimulation, the interneurons fire prolonged bursts of activity. Such bursts release enough γ-aminobutync acid (GABA) to activate presynaptic $GABA_B$ receptors on the excitatory afferents to the CA3 pyramidal neurons. When the $GABA_B$ receptors have been activated, the release of the excitatory neurotransmitter glutamate is blocked, so that the CA3 pyramidal neurons do not respond to the second stimulus. A key observation is that the receptor for acetylcholine on the hippocampal interneurons is sensitive to the nicotinic antagonist

α-bungarotoxin. Subsequently, this nicotinic receptor was cloned and found to be the product of a specific gene, the α7-nicotinic acetylcholine-receptor subunit gene (CHRNA7). Animals with abnormalities in this gene fail to inhibit the response to repeated auditory stimuli.

In an experiment parallel to Holzman's, we first undertook to demonstrate that abnormal P50 inhibition is a heritable trait, and second, to determine whether the trait is linked to specific genetic loci. We demonstrated autosomal co-dominant inheritance, so that one parent and generally half the siblings of a schizophrenic proband share the abnormality in P50 inhibition. Then, we undertook a screen of all chromosomes to determine if any locus showed significant genetic linkage to the trait. The genetic model assumed an autosomal dominant inheritance with disease allele frequency of 0.05, and the penetrance for abnormal P50 inhibition was set at 0.80; the penetrance for the normal genotype was 0.001. The model resulted in a population frequency of abnormal P50 inhibition of 0.09, with 10% phenocopies. In the nine families used for linkage analysis, 57 of 97 individuals had abnormal P50 inhibition. Of the 36 schizophrenics, 33 had abnormal P50 inhibition. In a scan of over 500 markers at a resolution of 10 cM, a locus on the long arm of chromosome 15 at band 15q14 showed significant indication of genetic linkage to abnormal P50 inhibition, which was analyzed as an independent phenotype. This band was subsequently demonstrated to be the location of CHRNA7, the gene that forms the α7-nicotinic receptor. Further mapping of genetic markers in the region showed that the linkage was maximal at D15S1360, a simple tandem nucleotide repeat within the gene. The logarithm of the odds ratio or lod score was highly significant for a genome-wide scan; $Z = 5.30$, $\Theta = 0.0$ ***(6)***. The lod score for schizophrenia was positive ($Z = 1.33$), but not significant. Molecular investigation of the gene shows that the amino acid coding sequence is normal in most patients, so that the receptor protein itself has a normal biological structure. However, a series of abnormalities in the promoter region of the gene can lead to a reduction in receptors ***(7)***. Thus, converging yet independent neurobiological and genetic evidence identi-

fied the same gene, CHRNA7, as the agent responsible for an alteration in inhibitory neuronal function.

The role of CHRNA7 in the genetic transmission of risk for schizophrenia has been further validated by a series of investigations by other groups that use schizophrenia itself as a phenotype for genetic analysis, so that it is now considered one of the several sites at which there is significant and replicable evidence for contribution to heritability. Using a variety of analytic strategies, other investigators have subsequently shown significant genetic linkage in the 15q14 region. A recent report targeted German patients with Leonhard's periodic catatonia, a form of schizophrenia characterized by chronic illness with facial grimmaces and stereotypies. This form was chosen because it frequently appears in siblings, and thus is likely to be heritable. The maximum lod score is Z = 3.57, and parametric and non-parametric multipoint techniques were used, which identify two nearby loci that closely bracket CHRNA7 ***(8)***. As is true for all loci associated with schizophrenia, there are several reports of positive findings in studies of bipolar disorder. In one study, the 15q14 lod score was 3.46, when families were selected for the presence of psychotic individuals who exhibited a poor response to treatment with lithium ***(9)***.

5. Physiological Phenotypes and the Genetic Analysis of Psychiatric Illnesses

The three examples described here have all used physiological phenotypes to obtain positive evidence for genetic linkage. The first two examples strengthened the evidence of a previously reported linkage that had been found using disease status alone as a phenotype, and the third discovered a new linkage, which has also subsequently been confirmed in studies that used disease status alone. Thus, one could plausibly argue from these three examples that most relevant loci have emerged or subsequently will emerge in studies that use disease status alone as a phenotype, thereby obviating the need for physiological phenotypes. Furthermore, physiological phenotypes are expensive to record and difficult to establish as a reli-

able phenotype. Thus, their use by many investigators in large populations is limited, which lowers the power of genetic studies. Indeed, all three phenotypes came from laboratories that had characterized the measures for more than a decade before applying them to genetic studies. However, the rapid pace of molecular sequencing suggests that coding abnormalities may well be found soon enough for most illnesses, so that a more certain way to discover all the genetic abnormalities in mental illness is simply to sequence a series of affected individuals. What role, then, will these physiological phenotypes play in the future, and should their use be increased, or instead viewed only as a historical feature of human psychiatric genetics?

There are several reasons to believe that—even if all the molecular abnormalities were known—physiological phenotypes would still have considerable value for psychiatric research. Current genetic research suggests that no single genetic locus is causative of any significant number of cases of schizophrenia, because nearly every set of families segregates at least several chromosomal loci. Furthermore, most of these loci are positive in several illnesses—e.g., both schizophrenia and bipolar affective disorder. The interaction between these various loci to produce different clinical syndromes is unlikely to be resolved completely at the genetic level. Although molecular sequences will be profoundly informative, in terms of identifying specific biochemical gene products and predicting their dysfunction, how these biochemical dysfunctions are manifested as physiological dysfunctions and how the dysfunctions interact with each other will likely be more readily appreciated at the phenotypic than the genotypic level. For example, if one genetic abnormality resulted in deficient sensory inhibition, and another decreased the temporal capacity of working memory, then we might expect that these two abnormalities would be synergistic. The brain of an individual with both deficits would be incapable of selecting relevant stimuli—the individual would attempt to process too many stimuli, and, concurrently, the brain would be unable to hold the stimuli in awareness long enough to permit further processing to decide which action to take. The result would be massively disrupted

—even psychotic—behavior. Currently, our phenotypic view is too narrow to permit such analysis. For example, it is unclear how the eye-tracking dysfunction at chromosome 6p interacts with the P50 inhibitory dysfunction as chromosome 15q. However, a complete description of the phenotype at 15q is likely to include some influence on eye tracking, as the α7-nicotinic receptor is expressed in the thalamic lateral geniculate nucleus and the superior colliculus, both neuronal regions known to be involved in the regulation of eye movement. Notably, the physiological dysfunctions described thus far are relatively common in the general population, so that the majority of the individuals who have them, even excluding phenocopies, are not likely to have psychiatric illness. Therefore, combinations of abnormalities are far more likely to be pathogenic than single abnormalities.

6. Conclusion

Physiological phenotypes will be used in the future primarily to explain how genes cause illness, not to find the responsible genes. Intervention into the course of the illness will require understanding how the phenotypic expression of one gene interacts with others to produce illness. Physiological phenotypes have an advantage—they are easily modeled in animals, so that these interactions will likely be studied in parallel in both animal models and humans. From the perspective of the complexity of the brain, it is likely that we will need far more than the three phenotypes presented here to achieve a useful neuronal model of psychiatric illness. Thus, one could argue that the real problem in psychiatric illnesses is not in finding genotypes, but rather in finding phenotypes.

References

1. Holzman, P.S., Kringlen, E., Levy, D.L., and Haberman, S.J. (1980) Deviant eye tracking in twins discordant for psychosis. A replication. *Arch. Gen. Psychiatry* **37,** 627–631.
2. Holzman, P.S., Kringlen, E., Matthysse, S., Flanagan, S.D., Lipton, R.B., Cramer, G., et al. (1988) A single dominant gene can account

for eye tracking dysfunctions and schizophrenia in offspring of discordant twins. *Arch. Gen. Psychiatry* **45,** 641–647**.**

3. Arolt, V., Lencer, R., Nolte, A., Muller-Myhsok, B., Purmann, S., Schurmann, M., et al. (1996) Eye tracking dysfunction is a putative phenotypic susceptibility marker of schizophrenia and maps to a locus on chromosome 6p in families with multiple occurrence of the disease. *Am. J. Med. Genetics* **67,** 564–579.
4. Williams, J.T., Begleiter, H., Porjesz, B. Edenberg, H.J., Foroud, T., Reich, T., et al. s(1999) Joint multipoint linkage analysis of multivariate qualitative and quantitative traits. II. Alcoholism and event-related potentials. *Am. J. Hum. Genetics* **65,** 1148–1160.
5. Almasy, L., Porjesz, B., Blangero, J., Chorlian, D.B., O'Connor, S.J., Kuperman, S., et al. (1998) Quantitative trait loci analysis of human event-related brain potentials: P3 voltage. *Electroencephalogr. Clin. Neurophysiol.* **108,** 244–250.
6. Freedman, R., Coon, H., Myles-Worsley, M., Orr-Urtreger, A., Olincy, A., Davis, A., et al. (1997) Linkage of a neurophysiological deficit in schizophrenia to a chromosome 15 locus. *Proc. Natl. Acad. Sci. USA* **94,** 587–592.
7. Logel, J., Gault, J., Vianzon, R., Hopkins, J., Short, M., Robinson, M., et al. (2000) Mutation screen of the promoter region of the human α7 neuronal nicotinic receptor subunit in normal and schizophrenic individuals. *Soc. Neurosci. Abstr.* **26,** 137.17.
8. Stober, G., Saar, K., and Ruschendorf, F. (2000) Splitting schizophrenia: periodic catatonia-susceptibility locus on chromosome 15q15. *Am. J. Hum. Genet.* **67,** 1201–1207.
9. Turecki, G., Grof, P., Grof, E., et al. (2000) A genome scan using a pharmacogenetic approach indicates a susceptibility locus for bipolar disorder on 15q14. *Biol. Psychiatry* **47,** 69S–70S.

10

Brain Imaging as an Approach to Phenotype Characterization for Genetic Studies of Schizophrenia

Joseph H. Callicott and Daniel R. Weinberger

Introduction

The recent sequencing of the human genome has raised expectations that the identification of susceptibility genes for the major psychiatric disorders should soon follow. However, as detailed in Chapters 6 and 7, the complex genetic architecture of mental illness is likely to pose a continuing challenge to traditional linkage and association approaches. The expansion of functional brain imaging into the realm of psychiatric genetics has come at an opportune time. Two developments in functional brain imaging are likely to transform the role of brain imaging in psychiatric neuroscience and research—the failure of functional brain imaging to produce diagnostically specific measurements of illness vs the technical advances in magnetic resonance imaging (MRI) techniques. However, although the methodological details are growing ever more complex and appear unfamiliar, this dialectic between the psychiatric neuroscientists and the psychiatric clinicians is a familiar one. As was the case two decades ago when the debate centered on the relevance of structural brain imaging findings to the genetic causes of

From: *Methods in Molecular Medicine, vol. 77: Psychiatric Genetics: Methods and Reviews*
Edited by: M. Leboyer and F. Bellivier

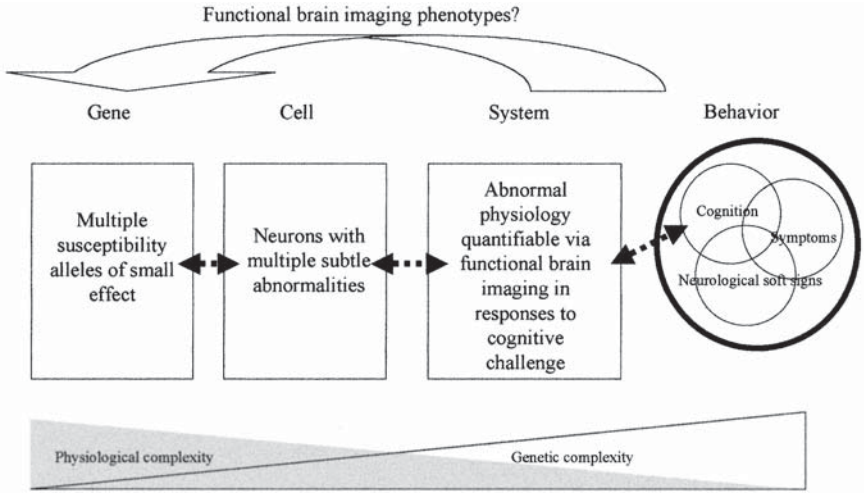

Fig. 1. Functional brain imaging intermediate phenotypes in the search for genetic causation. The central assumptions of the brain imaging intermediate phenotype are illustrated. Schizophrenia is believed to be a complex disorder both in terms of its genetic complexity and its physiological complexity. The functional brain imaging intermediate phenotype attempts to avoid complexity by identifying physiologically "simple" brain imaging findings that arise from "simple" gene interactions. Although the exact inter-relationships between each level (from gene to cell to physiology and finally to behavior) are presently incompletely characterized (illustrated by dashed arrows), the intermediate phenotype can be envisioned as a theoretically based and hypothesis-driven "shortcut" to susceptibility genes.

mental illness, the current quest for candidate functional brain imaging phenotypes should benefit all, since ultimately both clinician and neuroscientist seek to bridge the gap between brain mapping "findings" and neuronal pathology.

The identification of more elemental psychological or physiological characteristics associated with these illnesses and transmitted within families afflicted with these illnesses is the goal of the intermediate phenotype approach to functional brain imaging in schizophrenia (**Fig. 1**). The evolution of the brain imaging intermediate phenotype in schizophrenia is a particularly useful example, illus-

trating both advantages and disadvantages of the intermediate phenotype "mindset." Although functional brain imaging has evolved to become a powerful tool in the search for these intermediate phenotypes, uncertainties regarding the interpretation of particular functional brain imaging findings in patients with schizophrenia (and by inference their ultimate validity as phenotypes) are indicative of the challenges likely to face brain imaging researchers who are interested in a wider variety of psychiatric illnesses ***(11)***.

The union of sophisticated neuropsychological experimental designs and in vivo physiological brain mapping is expected to produce some unexpected and informative findings regarding the wide range of information-processing deficits associated with mental illnesses such as schizophrenia. Although functional brain imaging experiments have thus far failed to find qualitatively distinct brain maps specific to the more complex mental illnesses such as schizophrenia, the ascendancy of functional magnetic resonance imaging (fMRI) and magnetic resonance spectroscopy (MRS) offers two unique potential contributions: quantitative measures of neuronal activity in specific brain regions during a variety of experimental settings via fMRI and quantitative measures of regional neuronal metabolism via MRS. The first section of this chapter summarizes some of the history relevant to brain imaging as one approach to the phenotypic characterization of specific cohorts (i.e., patients with schizophrenia, their relatives, and a healthy comparison population) for genetic studies. Most importantly, two central questions emerge repeatedly in this process of generating putative brain imaging intermediate phenotypes: 1) Does any particular brain imaging abnormality in schizophrenia represent a genetic phenomenon?; and 2) Once a brain imaging abnormality appears to be under significant genetic influence, can this abnormality be defined at the level of a single subject (schizophrenic or otherwise) so that disease specific alleles can be identified?

The second section of this chapter provides a more detailed examination of brain imaging phenotypes currently under study. Functional brain imaging findings in schizophrenia are beginning to accumulate rapidly. Whether any of these findings will lead to the

identification of susceptibility alleles for schizophrenia remains an area of active debate. However, if functional brain imaging researchers are able to address the central questions noted here, then the identification of novel susceptibility alleles should soon follow.

2. The Search for the Intermediate Phenotypes in Schizophrenia

The search for brain imaging phenotypes has been underway since seminal studies in structural brain imaging during the late 1970s and early 1980s (*see* **ref. *56***). These early CT studies showed that lateral ventriculomegaly was not only a replicable and robust group "finding," but also that this trait or phenotype had a direct, "real-world" relationship to individual patients. For example, in one of the early landmark studies, Johnstone et al. ***(28)*** used computerized axial tomography (CT) to examine 17 well-matched healthy volunteers (HV) and 17 institutionalized patients with schizophrenia (SCZ). This study demonstrated that patients as a group had larger ventricles and that (at the level of the individual patient) ventriculomegaly predicted cognitive impairment. The next logical step was to determine whether the lateral ventriculomegaly phenotype was under genetic control (in HV and SCZ) and was also heritable within families afflicted by schizophrenia. Taken together, three seminal studies showed that these CT abnormalities were not simply under genetic influence, but were heritable in HV, SCZ, and the siblings of SCZ. Weinberger et al. ***(60)*** compared the distribution of lateral ventriculomegaly in 7 healthy sibships as compared to 9 sibships in which at least one member had manifest schizophrenia. Ventricle-to-brain ratio (VBR) was examined as a putative heritable phenotype in two ways: quantitatively via intra-class correlation (ICC) ***(2)*** of VBR within sibships; and qualitatively via scatterplots that compared VBR of any given individual to a well-established mean VBR in healthy control populations. Quantitatively and qualitatively, VBR appeared to be heritable in healthy and schizophrenic sibships. Reveley and colleagues ***(43)*** addressed the same questions in monozygotic (MZ) and dizygotic (DZ) twins. Using VBR as the

phenotype, heritability was calculated between 11 pairs of healthy MZ, 8 pairs of healthy DZ twins, and 7 pairs of MZ twins discordant for schizophrenia. VBR was highly heritable in healthy MZ twins ($h^2 = 0.98$) and roughly twice that of healthy DZ twins ($h^2 = 0.45$). VBR was also highly heritable in schizophrenic DZ twins ($h^2 = 0.87$). However, reflecting an apparent genetic loading effect in which strong family history of psychosis (either manifested by frank schizophrenia or simply by family history) predicted greater concordance for the phenotype within sibships as found by Weinberger et al. ***(60)***, the two schizophrenic DZ sibships in the Reveley et al. ***(43)*** study with a strong family history of psychosis showed the least within-pair variance. Finally, DeLisi and colleagues ***(17)*** demonstrated a significant familial component to lateral ventricular enlargement in 11 sibships (including non-psychotic siblings of schizophrenic patients or so-called unaffected siblings [SIBS]) that was not explained by non-genetic or environmental factors such as early head injury or obstetrical complications.

In the decades that followed, MRI largely replaced CT in these investigations ***(38,54,59,61).*** Recently, Lawrie et al. ***(31)*** reported the first MRI volumetric results from the Edinburgh High Risk Project, an ongoing epidemiological monitoring project designed to follow 200 high-risk, but asymptomatic subjects with multiple first- and/or second-degree relatives with manifest schizophrenia in order to determine the factors that are likely to predict which high-risk subjects will eventually develop schizophrenia. As in earlier CT studies, the motivation for examining unaffected members of these families was driven by the assumption that susceptibility alleles (and thus the phenotypic manifestation(s) of these alleles) should be more likely to aggregate within these high-risk subjects. They compared 100 high-risk subjects to 20 first-episode SCZ and 30 matched HV. Of relevance to this discussion of heritability, analyses within the high-risk cohort suggested that a pattern of regional brain volumetric abnormalities (including structures within the anterior mesial temporal lobes and ventricular system) was more frequent in high-risk subjects with the greatest number of affected family members. Although the merits of other MRI volumetric abnormalities contin-

ued to be debated, the consensus remained that lateral ventriculomegaly was both under genetic control and could be used at the individual subject level as an intermediate phenotype ***(16,47,49,52)***. Most importantly, this process eventually led to a specific genetic finding. The first relationship between ventricular enlargement and a specific genetic locus was reported by Shihabuddin et al. ***(48)***, who found a linkage between VBR and the short arm of chromosome 5 (5p14.1-13.1).

The search for intermediate phenotypes in schizophrenia using fMRI and MRS has a much briefer history ***(5,11)***. However, the list of potential candidates for such phenotypes has been steadily growing since the introduction of functional brain imaging to the study of schizophrenia in the mid-1970s. Several of the strengths and weaknesses of this approach are discussed here.

In principle, a phenotype related to genetic risk should be found in ill individuals and in some at-risk individuals who do not manifest illness. Thus, the search for phenotypes at the level of brain imaging begins with patients who manifest schizophrenia. Although functional brain-imaging techniques have been applied to the study of mental illnesses, particularly schizophrenia, for almost 30 yrs, they have failed to generate pathognomonic findings. Pathognomonic brain-imaging findings should be "characteristic" of the illness or disease under study and specific to it. For instance, the most replicated functional brain-imaging finding in patients with schizophrenia remains "hypofrontality." Hypofrontality is widely believed to indicate that patients with schizophrenia have reduced neuronal activity in dorsolateral prefrontal cortex (DLPFC) as assayed by such measures as regional cerebral blood flow (rCBF) or regional cerebral metabolic rate of glucose metabolism (rCMRglu). Fortunately, 30 years of study have added an ever-broadening armamentarium of brain-imaging techniques to assay indirect measures of neuronal activity (such as rCBF or rCMRglu) from single-photon emission tomography (SPET), positron emission tomography (PET), and MRI techniques. Unfortunately, the ongoing and ever-widening application of these functional brain-imaging methodologies to schizophrenia continues to raise as many questions as it answers. This

ongoing controversy may in part explain the lack of candidate brain-imaging phenotypes studies in schizophrenia to date.

However, this apparent conundrum should not prevent the generation of these much-needed phenotypes in schizophrenia. Together with increased genotypic and molecular biological characterization, brain imaging researchers who were interested in identifying a pathognomonic finding in early symptomatic Alzheimer's disease have reported some progress through the use of abnormal rCBF and rCMRglu in the parietal cortex assayed via PET. Ishii et al. ***(26)*** summarized a body of work that *in toto* argued that abnormal brain imaging statistical maps enabled neuroradiologists to more accurately identify mildly symptomatic patients with Alzheimer's disease. More specifically, Ishii et al. ***(26)*** created "cerebral perfusion Z score maps" (Z map) using $H_2{}^{15}O$ PET in 28 patients with probable Alzheimer's disease and 10 matched healthy volunteers. Based on a group comparison of a normative sample of 20 healthy aged volunteers and 20 patients with Alzheimer's disease, Ishii and colleages created a "prototypic" AD brain Z map within the Statistical Parametric Mapping software package—SPM ***(23)***. Four radiologists blinded to diagnosis were then asked to rate the individual conventional rCBF maps and cerebral perfusion "Z maps" of the 28 patients with probable AD and the 10 matched healthy aged volunteers. A receiver operating characteristic (ROC) analysis was then used to compare the two rating systems and found that "Z maps" (mean area under the ROC curve = 0.96) more accurately separated patients with mildly symptomatic Alzheimer's disease from healthy elderly subjects than simple visual inspection of individual rCBF maps (mean area under the ROC curve = 0.584).

Early in the history of the search for candidate brain imaging phenotypes in schizophrenia, it appeared that this statistical mapping approach could be effectively applied to the classification of patients with schizophrenia and healthy controls. However, the relationship between rCBF findings in schizophrenia (such as hypofrontality) at the group level had a complex relationship with the statistical maps at the individual level ***(3,15,35,36,41,42)***. This becomes a critical issue when attempting to apply classification of individual PET

rCBF maps to unaffected siblings of patients with schizophrenia ***(21)***. Fortunately, in parallel with ongoing theoretical debates within the psychiatric neuroscience literature, impressive efforts have been made within the larger brain-imaging literature to formally address the complex relationship between findings at the individual and group level (*see* **refs. *1,8,18,46,62***).

For example, some have argued that functional brain imaging findings in the DLPFC of patients with schizophrenia are either entirely or mostly the result of confounding factors such as performance, medication status, and the effects of chronic illness, although strong opposing evidence has been present in the literature from the very beginning (*see* **ref. *58*** for review). Even if it is accepted that such DLPFC findings in schizophrenia represent physiologically valid measures of neuronal activity in DLPFC (a controversial assumption beyond the scope of this chapter), candidate intermediate phenotypes based on any regional functional brain measure may face additional challenges.

In addition to the two general questions posed at the outset of this chapter, an additional conceptual hurdle must be overcome. Although it is assumed that a candidate intermediate phenotype is heritable, formal estimates of heritability must eventually consider issues of shared genetic susceptibility and shared environment ***(45)***. However, these estimates also include an error term. For example, in structural brain imaging in schizophrenia, although non-genetic factors (such as history of head injury or birth complications) are known to add error to heritability measurements, the very measurements themselves (e.g., manual or automated tracing of ventricles during the quantification of VBR) are known to introduce some error ***(17,43,60)***. As the methodologies have become more complex, estimates of measurement error at the individual brain map level have also become more complicated.

One final source of intense debate regarding the putative validity of abnormal DLPFC rCBF as a candidate brain-imaging phenotype is the issue of performance. Specifically, investigators continue to explore the observation that the most reliable experimental conditions for generating reduced DLPFC rCBF in patients with schizo-

phrenia (and by extension their at-risk siblings) involves cognitive tasks that schizophrenic patients will perform more poorly than a control population—the "performance conundrum." Therefore, even if one succeeded in solving the "statistical mapping problem" in schizophrenia—as illustrated in the work of Ishii and colleages ***(26)*** in suspected Alzheimer's disease—the performance conundrum presents an additional hurdle. For suspected AD, measurements of rCBF at rest were used, and this approach has been criticized as an unreliable measure of abnormal DLPFC brain activity in schizophrenia ***(57)***.

Since there is a growing consensus that performance differences drive DLPFC activation differences between patients and controls, many groups are attempting to map regional responses to cognitive challenge over a range of difficulty ***(7)***. Using a parametric design, the best hope of generating an individual phenotype may be that patients operate on a different "load-response" curve. If patients appear to operate on the same load-response curve, then performance differences will drive brain-map differences in a potentially non-informative fashion. In reference to intermediate phenotypes, the activation differences between patients and controls may reflect a non-genetically driven dissociation that could be found in any two cohorts occupying different areas on the "healthy" load-response curve ***(9,10,14,15)***. However, if the dissociation arises from genetic influences, then individuals (patients, siblings, or controls) could be plotted in reference to the "healthy" curve, and their relative position on this standard curve could represent either a quantitative or qualitative intermediate phenotype ***(9)***.

Circumstantial evidence increasingly indicates that patients with schizophrenia do operate on a distinct load-response curve. For example, the physiological response to working memory (WM) challenge appears to fall on an inverted-U shaped curve for healthy subjects and for patients with schizophrenia. The assumption has been bolstered by the observation of "hypofrontality" in healthy subjects pushed beyond their performance capacity. It is a well-established behavioral observation in humans that WM is capacity-limited ***(37)***. When healthy subjects are pushed beyond their WM

capacity, attenuated PFC activation results. Goldberg et al. ***(24)*** found that healthy subjects became hypofrontal when performing an executive WM task (the WCST) and an auditory shadowing task together vs performing the WCST alone. Using a parametric version of the N-back WM task, Callicott et al. directly demonstrated that healthy subjects become hypofrontal beyond their WM capacity ***(14)***. Other groups have used parametrically manipulated WM tasks that support the notion of an inverted-U shaped response for healthy controls in DLPFC ***(27,34)***. Although far from substantiated, it appears that patients with schizophrenia also show an inverted U-shaped response curve in DLPFC that is a distinct load-response curve from that observed in healthy subjects. Fletcher et al. ***(22)*** were the first group to describe dissociable DLPFC activation curves using a graded semantic memory task. Callicott et al. also tested this hypothesis using a parametric WM task and found support for this dissociation ***(12)***. Furthermore, aberrant DLPFC activation was correlated with in vivo measures of neuronal pathology (^{1}H-MRSI-derived NAA measures) suggesting a pathological neuronal basis for this dissociation in the DLPFC of schizophrenic patients. Although no one experiment has yet captured the full extent of this dissociation, Manoach and colleagues have reported results using a modified Sternberg Item Recognition task that could be interpreted to support distinct load-response curves in DLPFC ***(32,33)***.

3. The Findings

Although structural brain imaging findings have dominated the search for viable intermediate phenotypes, both MRS and fMRI studies have also documented familial findings in patients with schizophrenia and their relatives. Guided by both structural and functional brain imaging findings in schizophrenia, investigators to date have concentrated on the frontal lobes and the hippocampal area. MRS, both proton MRS (^{1}H-MRS) or phosphorous-31 MRS (^{31}P-MRS), remain the only in vivo brain imaging methodologies believed to measure neuronal integrity directly ***(25,55)***. Three studies have used proton MRS (^{1}H-MRS) methodologies to examine

the frontal cortex of first-degree relatives of patients with schizophrenia. The principal metabolite measurements have been n-acetylaspartate (NAA), choline-containing compounds (CHO), and creatine+phosphocreatine (CRE) ***(4)***. Keshavan et al. ***(29)*** used a single-voxel, short echo time (TE) proton MRS (^{1}H-MRS) technique to study medial frontal lobe (anterior cingulate subregion) neuronal integrity in healthy volunteers (HV) (n = 10) and the offspring of schizophrenic patients (n = 9) and found a trend for a reduction in NAA/CHO in the offspring, but failed to identify significant reductions in NAA/CRE. Callicott et al. ***(13)*** used a multiple voxel, long TE proton magnetic resonance spectroscopic imaging (^{1}H-MRSI) technique ***(19)*** in patients with schizophrenia (n = 47), their unaffected siblings (n = 66), and healthy volunteers (n = 66). In DLPFC, ventral lateral prefrontal cortex, and anterior cingulate, SIBS failed to show a significant reduction in NAA measures. Finally, Block et al. ***(6)*** used a combined short and long TE ^{1}H-MRS technique to examine 25 SCZ, 19 NV, 13 subjects with schizophrenia-spectrum disorders, and 35 SIBS. Once again, SIBS failed to show any significant reductions in left dorsal PFC NAA measures. Klemm et al. ***(30)*** used a ^{31}P-MRS technique ***(39)*** to study 14 first-degree relatives of SCZ (children of SCZ plus siblings of SCZ) and 14 matched NV. Based on published reports of abnormal membrane phospholipid turnover in the SCZ ***(40,53)***, Klemm et al. ***(30)*** hypothesized that these first-degree relatives at risk for schizophrenia should also show signs of either decreased membrane phospholipid anabolism or increased membrane phospholipid catabolism in the frontal lobes, as has been seen in adult SCZ. The dependent measurements derived from ^{31}P-MRS were the relative peak areas of "phosphodiesters (%)," "phosphomonoesters (%)," the ratio of phosphomonoesters to phosphodiesters, "total adenosinetriphosphate (%)," "phosphate phosphocreatine (%)," the ratio of phosphate phosphocreatine to total adenosinetriphosphate, "inorganic phosphate (%)," and pH. These dependent measures were taken simultaneously from two regions of interest (ROIs) placed bilaterally within the frontal lobes and group values compared using nonparametric, pairwise Mann-Whitney U tests ($p < 0.05$, uncorrected).

Within these frontal-lobe volumes, first-degree relatives were found to have significantly higher phosphodiester percentages and a lower ratio of total phosphomonoesters to total phosphodiesters. A trend was also found for these high-risk first-degree relatives to have lower phosphate phosphocreatine % ($p = 0.05$). The authors interpreted these results to suggest greater frontal-lobe membrane catabolism in first-degree relatives of SCZ and propose this as a putative intermediate phenotype. On the other hand, there were no significant differences in the other ^{31}P-MRS measurements (e.g., support for the presence of lower-membrane anabolism) and no formal or informal estimates of familiality or heritability performed. In summary, both ^{1}H-MRS and ^{31}P-MRS studies of first-degree relatives of SCZ offer some potential intermediate phenotypes of abnormal PFC neuronal integrity, but the results are not yet conclusive, and require further work to generate formal estimates of familiality and/or heritability.

MRS intermediate phenotypes in the hippocampal area are in a similar state of development. In the same study noted here, Callicott et al. ***(13)*** found that both SCZ and SIBS had reduced NAA measures (NAA/CRE) in the hippocampal area (HIPPO) when compared to NV. Although the number of "sib-pairs" (members of the same family in which a SCZ had reduced HIPPO NAA measures) was too small to perform formal heritability estimates, an intermediate phenotype was defined in the same manner as in prior structural MRI studies ***(17,43,60)***. In essence, any subject (SCZ, SIB, or NV) with HIPPO NAA values below one standard deviation of the NV mean were designated as possessing the HIPPO NAA intermediate phenotype. Familiality within each sib-pair was determined in the following two ways: quantitatively though an ICC ***(2)***, and qualitatively through relative risk (λ_s). Quantitative measures of familiality were not significant, but relative risk calculations ***(44)*** suggested that this putative intermediate phenotype was heritable ($\lambda_s = 3.8 – 8.8$). Recently, Callicott and colleagues (unreported data) have replicated this increased relative risk using the same ^{1}H-MRSI method and a new group of SCZ ($n = 34$), SIBS ($n = 33$), and NV ($n = 38$). In this second experiment, all analyses were directed at

NAA measures (specifically NAA/CRE ratios) in HIPPO. As a group, SIBS and SCZ exhibited a significant reduction in HIPPO NAA/CRE compared with NV. Estimates of heritability were carried out using reduced HIPPO NAA/CRE as the potential intermediate phenotype. Quantitatively, ICC analyses within sib-pairs were again insignificant. Qualitatively, reduced HIPPO NAA/CRE was found in 10–15% of NV, 28–35% of SCZ, and 21–26% of SIBS (λ_s = 4.75–7.60). These findings, although still in a preliminary stage of data analysis, appear to replicate increased relative risk for reduced HIPPO NAA/CRE in SCZ and SIBS, and suggest that decreased HIPPO NAA measures are a viable intermediate phenotype in schizophrenia. It should be noted that there are numerous other brain regions—including the basal ganglia, thalamus, and cerebellum—that are currently under investigation using 1H-MRS and 31P-MRS. The reader is encouraged to read elsewhere for a more detailed discussion of these findings and their potential role as intermediate phenotypes in schizophrenia (*see* **ref. *5***).

fMRI intermediate phenotypes in schizophrenia are more poorly characterized in the published literature. Using a version of the N-back task alluded to here, Egan et al. ***(20)*** were able to show a direct effect of a common functional polymorphism in the catechol-O-methyltransferase gene (Val $^{108/158}$ Met) on PFC fMRI activation. In addition, Callicott et al. (in submission) have found evidence that DLPFC inefficiency during the modified N-back WM task as measured by fMRI may also be a putative brain imaging phenotype. In two separate cohorts of SIBS (n = 23, 25; respectively) and NV (n = 18, 15; respectively), SIBS showed an exaggerated right DLPFC BOLD fMRI response to WM challenge. Although both cohorts were distinct and the BOLD fMRI methodologies were slightly different, second-level analyses of these fMRI data within SPM99 (Wellcome Department of Cognitive Neurology, FIL, UK) have revealed overactivation (or inefficiency) in the same region of right DLPFC in both cohorts. A preliminary analysis of familiality within 7 sib-pairs revealed the following: 1) qualitatively, there appears to be a "susceptibility loading effect" (**Fig. 2A**); 2) quantitatively, there appear to be high correlations between the members of the

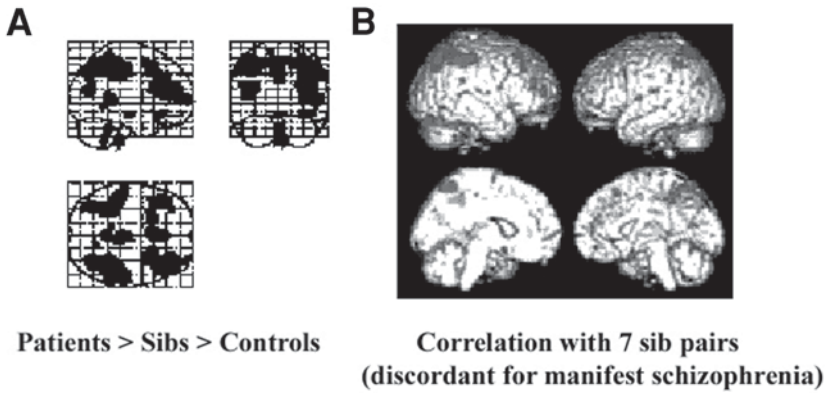

Fig. 2. Greater activation of dorsal PFC during WM challenge in patients and sibs: a statistical exploration of genetic loading. PFC efficiency as an intermediate phenotype. **(A)** Preliminary qualitative analysis: SPM ***(23)*** "glass brain" images on the left represent a preliminary exploration of putative genetic loading for altered cortical responses during a modified version of the N back WM task. This group map identifies areas throughout the WM network in SCZ (n = 13), SIBS (n = 23), and HV (n = 18) across a range of WM load. Putative effects of genetic susceptibility were modeled as a three-step function with the response of sibs intermediate between that of patients and controls (p <0.001, uncorrected). **(B)** Preliminary quantitative analysis: Correlations within sib pairs (n = 7 pairs) of fMRI response across this range of WM load are mapped on the right figure. Correlations were performed within MEDx (Sterling, VA) and significant positive correlations (p < 0.001) are displayed on a standard template (SPM96). These 7 sibling pairs are concordant for the putative fMRI phenotype, but discordant for manifest schizophrenia.

"sib-pair" (**Fig. 2B**). As has been the case for all abnormal group findings in the DLPFC of SCZ that are generated with cognitive challenge studies, the interpretation of any given finding as representing a core feature of the illness (and thus a reasonable candidate intermediate phenotype) is controversial. However, a precedent within the PET functional brain-imaging literature suggests that more complex, network-based analyses may eventually help sort out the apparent confusion regarding regional findings and their putative over-represention in first-degree relatives of SCZ. Spence

et al. ***(51)*** used $H_2{}^{15}O$ PET to study overall brain activation patterns during a verbal fluency cognitive challenge in 10 obligate carriers of SCZ (unaffected subjects with a parent and a child with SCZ), 10 SCZ, and 10 HV. Instead of focusing on brain activation patterns in specific regions, these investigators examined the interrelationship between regional brain activation (i.e., a functional connectivity analysis within SPM) and found "functional disconnection" between activation in DLPFC and anterior cingulate in SCZ, but not obligate carriers. However, these data failed to detect any familial (and by inference potentially heritable) functional disconnection between DLPFC and other regions such as DLPFC and hippocampus (*see* **ref. *57***) in obligate carriers. Given the complex heritability of schizophrenia discussed elsewhere in this book, the use of a small number of obligate carriers does not rule out the possibility that functional disconnection is familial. Regardless, the elegant statistical approach to phenotypic characterization presented in this landmark study is an important contribution. This work ***(50)*** suggests that functional brain imaging can provide rigorous, complex phenotypic characterization of individual SCZ without the potential pitfalls of oversimplification (**Fig. 1**). As is the case for ^{1}H-MRS and ^{31}P-MRS, small sample sizes, technical details, and formal estimates of heritability are significant, yet unsettled impediments to fMRI intermediate phenoytypes in any prospective search for novel susceptibility genes in schizophrenia.

4. Conclusion

Advances in functional brain imaging, coupled with the sequencing of the human genome, are promising developments in schizophrenia research. Armed with as little as a broad knowledge of the brain imaging literature, a standard MRI scanner, and collaborators from a number of other disciplines (notably psychology, radiology, and genetics), investigators across the globe are poised to critically apply this approach. Although minor methodological and technical differences must be clarified, we are likely to unravel some of the genetic and physiological mysteries that have characterized schizophrenia.

References

1. Aguirre, G.K., Zarahn, E., and M, D'Esposito. (1997) Empirical analyses of BOLD fMRI statistics. II. Spatially smoothed data collected under null-hypothesis and experimental conditions. *Neuroimage* **5,** 199–212.
2. Bartko, J.J. and Carpenter, W.T. (1976) On the methods and theory of reliability. *J. Nerv. Ment. Dis.* **163,** 307–317.
3. Berman, K.F., Torrey, E.F., Daniel, D.G., and Weinberger, D.R. (1992) Regional cerebral blood flow in monozygotic twins discordant and concordant for schizophrenia. *Arch. Gen. Psychiatry* **49,** 927–934.
4. Bertolino, A., Nawroz, S., Mattay, V.S., Barnett, A.S., Duyn, J.H., Moonen, C.T., et al. (1996) Regionally specific pattern of neurochemical pathology in schizophrenia as assessed by multislice proton magnetic resonance spectroscopic imaging. *Am. J. Psychiatry* **153,** 1554–1563.
5. Bertolino, A. and Weinberger, D.R. (1999) Proton magnetic resonance spectroscopy in schizophrenia. *Eur. J. Radiology* **30,** 132–141.
6. Block, W., Bayer, T.A., Tepest, R., Traber, F., Rietschel, M., Muller, D.J., et al. (2000) Decreased frontal lobe ratio of N-acetyl aspartate to choline in familial schizophrenia: a proton magnetic resonance spectroscopy study. *Neurosci. Lett.* **289,** 147–151.
7. Braver, T.S., Cohen, J.D., Nystrom, L.E., Jonides, J., Smith, E.E., and Noll, D.C. (1997) A parametric study of prefrontal cortex involvement in human working memory. *Neuroimage* **5,** 49–62.
8. Bullmore, E., Long, C., Suckling, J., Fadili, J., Calvert, G., Zelaya, F., et al. (2001) Colored noise and computational inference in neurophysiological (fMRI) time series analysis: resampling methods in time and wavelet domains. *Hum. Brain Mapp.* **12,** 61–78.
9. Callicott, J. (2001) Functional brain imaging in psychiatry: the next wave, in *Advances in Brain Imaging*, Vol. 4. (Morihisa, J. M., ed.), American Psychiatric Publishing, Inc., Washington, DC, pp. 1–24.
10. Callicott, J. and Weinberger, D. (1999) Functional brain imaging: future prospects for clinical practice, in *Psychiatry in the New Millennium* (Weissman, S., Sashin, M., and Eist, H., eds.), American Psychiatric Press, Inc., Washington, DC, pp. 119–135.
11. Callicott, J. and Weinberger, D. (1999) Functional MRI in psychiatry, in *Medical Radiology: Functional MRI* (Moonen, C. and Bandettini, P., eds.), Springer-Verlag, Berlin, pp. 501–512.

12. Callicott, J.H., Bertolino, A., Mattay, V.S., Langheim, F.J., Duyn, J., Coppola, R., et al. (2000) Physiological dysfunction of the dorsolateral prefrontal cortex in schizophrenia revisited. *Cereb. Cortex* **10,** 1078–1092.
13. Callicott, J.H., Egan, M.F., Bertolino, A., Mattay, V.S., Langheim, F.J., Frank, J.A., et al. (1998) Hippocampal N-acetyl aspartate in unaffected siblings of patients with schizophrenia: a possible intermediate neurobiological phenotype [published erratum appears in *Biol. Psychiatry* (1999) Jan 15;45(2):following 244]. *Biol. Psychiatry* **44,** 941–950.
14. Callicott, J.H., Mattay, V.S., Bertolino, A., Finn, K., Jones, K., Frank, J.A., et al. (1999) Physiological characteristics of capacity constraints in working memory as revealed by functional MRI. *Cereb. Cortex* **9,** 20–26.
15. Callicott, J.H., Ramsey, N.F., Tallent, K., Bertolino, A., Knable, M.B., Coppola, R., et al. (1998) Functional magnetic resonance imaging brain mapping in psychiatry: methodological issues illustrated in a study of working memory in schizophrenia. *Neuropsychopharmacology* **18,** 186–196.
16. Cannon, T.D., van Erp, T.G., Huttunen, M., Lonnqvist, J., Salonen, O., Valanne, L., et al. (1998) Regional gray matter, white matter, and cerebrospinal fluid distributions in schizophrenic patients, their siblings, and controls. *Arch. Gen. Psychiatry* **55,** 1084–1091.
17. DeLisi, L.E., Goldin, L.R., Hamovit, J.R., Maxwell, M.E., Kurtz, D., and Gershon, E.S. (1986) A family study of the association of increased ventricular size with schizophrenia. *Arch. Gen. Psychiatry* **43,** 148–153.
18. DeYoe, E.A., Bandettini, P., Neitz, J., Miller, D., and Winans, P. (1994) Functional magnetic resonance imaging (FMRI) of the human brain. *J. Neurosci. Methods* **54,** 171–187.
19. Duyn, J., Gillen, J., Sobering, G., Van, Z.P., and Moonen, C. (1993) Multisection proton MR spectroscopic imaging of the brain. *Radiology* **188,** 277–282.
20. Egan, M.F., Goldberg, T.E., Kolachana, B.S., Callicott, J.H., Mazzanti, C.M., Straub, R.E., et al. (2001) Effect of COMT Val108/158 Met genotype on frontal lobe function and risk for schizophrenia. *Proc. Natl. Acad. Sci. USA* **98,** 6917–6922.
21. Egan, M.F. and Weinberger, D.R. (1997) Neurobiology of schizophrenia. *Curr. Opin. Neurobiol.* **7,** 701–707.

22. Fletcher, P.C., McKenna, P.J., Frith, C.D., Grasby, P.M., Friston, K.J., and Dolan, R.J. (1998) Brain activations in schizophrenia during a graded memory task studied with functional neuroimaging. *Arch. Gen. Psychiatry* **55,** 1001–1008.
23. Friston, K.J., Holmes, A.P., Worsley, K.J., Poline, J.B., Frith, C.D., and Frackowiak, R.S.J. (1995) Statistical parametric mapping in functional imaging: a general approach. *Hum. Brain Mapp.* **2,** 189–210.
24. Goldberg, T.E., Berman, K.F., Fleming, K., Ostrem, J., Van Horn, J.D., Esposito, G., et al. (1998) Uncoupling cognitive workload and prefrontal cortical physiology: a PET rCBF study. *Neuroimage* **7,** 296–303.
25. Horrobin, D.F. (1998) The membrane phospholipid hypothesis as a biochemical basis for the neurodevelopmental concept of schizophrenia. *Schizophr. Res.* **30,** 193–208.
26. Ishii, K., Sasaki, M., Matsui, M., Sakamoto, S., Yamaji, S., Hayashi, N., et al. (2000) A diagnostic method for suspected Alzheimer's disease using H(2)15O positron emission tomography perfusion Z score. *Neuroradiology* **42,** 787–794.
27. Jansma, J.M., Ramsey, N.F., Coppola, R., and Kahn, R.S. (2000) Specific versus nonspecific brain activity in a parametric N-back task [In Process Citation]. *Neuroimage* **12,** 688–697.
28. Johnstone, E.C., Crow, T.J., Frith, C.D., Husband, J., and Kreel, L. (1976) Cerebral ventricular size and cognitive impairment in chronic schizophrenia. *Lancet* **2,** 924–926.
29. Keshavan, M.S., Montrose, D.M., Pierri, J.N., Dick, E.L., Rosenberg, D., Talagala, L., et al. (1997) Magnetic resonance imaging and spectroscopy in offspring at risk for schizophrenia: preliminary studies. *Prog. Neuropsychopharmacol. Biol. Psychiatry* **21,** 1285–1295.
30. Klemm, S., Rzanny, R., Riehemann, S., Volz, H.P., Schmidt, B., Gerhard, U.J., et al. (2001) Cerebral phosphate metabolism in first-degree relatives of patients with schizophrenia. *Am. J. Psychiatry* **158,** 958–960.
31. Lawrie, S.M., Whalley, H., Kestelman, J.N., Abukmeil, S.S., Byrne, M., Hodges, A., et al. (1999) Magnetic resonance imaging of brain in people at high risk of developing schizophrenia. *Lancet* **353,** 30–33.
32. Manoach, D.S., Gollub, R.L., Benson, E.S., Searl, M.M., Goff, D.C., Halpern, E., et al. (2000) Schizophrenic subjects show aberrant fMRI activation of dorsolateral prefrontal cortex and basal ganglia during working memory performance [In Process Citation]. *Biol. Psychiatry* **48,** 99–109.

33. Manoach, D.S., Press, D.Z., Thangaraj, V., Searl, M.M., Goff, D.C., Halpern, E., et al. (1999) Schizophrenic subjects activate dorsolateral prefrontal cortex during a working memory task, as measured by fMRI. *Biol. Psychiatry* **45,** 1128–1137.
34. Manoach, D.S., Schlaug, G., Siewert, B., Darby, D. G., Bly, B.M., Benfield, A., et al. (1997) Prefrontal cortex fMRI signal changes are correlated with working memory load. *Neuroreport* **8,** 545–549.
35. Marenco, S., Coppola, R., Daniel, D.G., Zigun, J. R., and Weinberger, D.R. (1993) Regional cerebral blood flow during the Wisconsin Card Sorting Test in normal subjects studied by xenon-133 dynamic SPECT: comparison of absolute values, percent distribution values, and covariance analysis. *Psychiatry Res.* **50,** 177–192.
36. Mattay, V.S., Frank, J.A., Santha, A.K.S., Pekar, J.J., Duyn, J.H., McLaughlin, A.C., et al. (1996) Whole brain functional mapping with isotropic MR imaging. *Radiology* **201,** 399–404.
37. Miller, G.A. (1956) The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychol. Rev.* **63,** 81–97.
38. O'Driscoll, G.A., Florencio, P.S., Gagnon, D., Wolff, A.V., Benkelfat, C., Mikula, L., et al. (2001) Amygdala-hippocampal volume and verbal memory in first-degree relatives of schizophrenic patients. *Psychiatry Res.* **107,** 75–85.
39. Ordige, R.J., Connelly, A., and Lohman, J.A. (1986) Image-selected in vivo spectroscopy: a new technique for spatially selective NMR spectroscopy. *J. Magn. Reson.* **66,** 283–294.
40. Pettegrew, J W., Keshavan, M.S., Panchalingam, K., Strycho, S., Kaplan, D.B., Tretta, M.G., et al. (1991) Alterations in brain high-energy phosphate and membrane metabolism in first-episode, drug-naive schizophrenics. *Arch. Gen. Psychiatry* **48,** 563–568.
41. Ramsay, S.C., Murphy, K., Shea, S.A., Friston, K.J., Lammertsma, A.A., Clark, J.C., et al. (1993) Changes in global cerebral blood flow in humans: effect on regional cerebral blood flow during a neural activation task. *J. Physiol.* **471,** 521–534.
42. Ramsey, N.F., Tallent, K., van Gelderen, P., Frank, J.A., Mooner, C.T.W., and Weinberger, D.R. (1996) Reproducibility of human 3D fMRI brain maps acquired during a motor task. *Hum. Brain Mapp.* **4,** 113–121.
43. Reveley, A.M., Reveley, M.A., Clifford, C.A., and Murray, R.M. (1982) Cerebral ventricular size in twins discordant for schizophrenia. *The Lancet*, 540–541.

44. Risch, N. (1983) Estimating morbidity risks in relatives: the effect of reduced fertility. *Behav. Genet.* **13,** 441–451.
45. Risch, N. and Merikangas, K. (1996) The future of genetic studies of complex human diseases [see comments]. *Science* **273,** 1516–1517.
46. Ruttimann, U.E., Unser, M., Rawlings, R.R., Rio, D., Ramsey, N.F., Mattay, V.S., et al. (1998) Statistical analysis of functional MRI data in the wavelet domain. *IEEE Trans. Med. Imaging* **17,** 142–154.
47. Seidman, L.J., Faraone, S.V., Goldstein, J.M., Kremen, W.S., Matsuda, G., Hoge, E.A., et al. (1997) Reduced subcortical brain volumes in nonpyschotic siblings of schizophrenic patients: a pilot magnetic resonance imaging study. *Am. J. Med. Genet.* **74,** 507–514.
48. Shihabuddin, L., Silverman, J.M., Buchsbaum, M.S., Seiver, L.J., Luu, C., Germans, M.K., et al. (1996) Ventricular enlargement associated with linkage marker for schizophrenia-related disorders in one pedigree. *Molecular Psychiatry* **1,** 215–222.
49. Silverman, J.M., Smith, C.J., Guo, S.L., Mohs, R.C., Siever, L.J., and Davis, K.L. (1998) Lateral ventricular enlargement in schizophrenic probands and their siblings with schizophrenia-related disorders. *Biol. Psychiatry* **43,** 97–106.
50. Spence, S.A. (2000) Commenting on neuroimaging. *Br. J. Psychiatry* **176,** 594.
51. Spence, S.A., Liddle, P.F., Stefan, M.D., Hellewell, J.S., Sharma, T., Friston, K.J., et al. (2000) Functional anatomy of verbal fluency in people with schizophrenia and those at genetic risk. Focal dysfunction and distributed disconnectivity reappraised. *Br. J. Psychiatry* **176,** 52–60.
52. Staal, W.G., Hulshoff Pol, H.E., Schnack, H.G., Hoogendoorn, M.L., Jellema, K., and Kahn, R.S. (2000) Structural brain abnormalities in patients with schizophrenia and their healthy siblings. *Am. J. Psychiatry* **157,** 416–421.
53. Stanley, J.A., Williamson, P.C., Drost, D.J., Carr, T.J., Rylett, R.J., Malla, A., et al. (1995) An in vivo study of the prefrontal cortex of schizophrenic patients at different stages of illness via phosphorous magnetic resonance spectroscopy. *Arch. Gen. Psychiatry* **52,** 399–406.
54. Suddath, R.L., Christison, G.W., Torrey, E.F., Casanova, M.F., and Weinberger, D.R. (1990) Anatomical abnormalities in the brains of monozygotic twins discordant for schizophrenia. *N. Engl. J. Med.* **322,** 789–794.

55. Tsai, G. and Coyle, J.T. (1995) N-acetylaspartate in neuropsychiatric disorders. *Prog. Neurobiol.* **46,** 531–540.
56. Weinberger, D.R. (1987) Implications of normal brain development for the pathogenesis of schizophrenia. *Arch. Gen. Psychiatry* **44,** 660–669.
57. Weinberger, D.R. and Berman, K.F. (1996) Prefrontal function in schizophrenia: confounds and controversies. *Phil. Trans. Royal Soc. Med.* **351,** 1495–1503.
58. Weinberger, D.R., Berman, K.F., Iadarola, M., Driesen, N., and Zec, R.F. (1988) Prefrontal cortical blood flow and cognitive function in Huntington's disease. *J. Neurol. Neurosurg. Psychiatry* **51,** 94–104.
59. Weinberger, D.R., Berman, K.F., and Torrey, E.F. (1992) Correlations between abnormal hippocampal morphology and prefrontal physiology in schizophrenia. *Clin. Neuropharmacol.* **15,** 393A–394A.
60. Weinberger, D.R., DeLisi, L.E., Neophytides, A.N., and Wyatt, R.J. (1981) Familial aspects of CT scan abnormalities in chronic schizophrenic patients. *Psychiatry Res.* **4,** 65–71.
61. Wright, I.C., Rabe-Hesketh, S., Woodruff, P.W., David, A.S., Murray, R.M., and Bullmore, E.T. (2000) Meta-analysis of regional brain volumes in schizophrenia. *Am. J. Psychiatry* **157,** 16–25.
62. Zarahn, E., Aguirre, G.K., and M, D'Esposito. (1997) Empirical analyses of BOLD fMRI statistics. I. Spatially unsmoothed data collected under null-hypothesis conditions. *Neuroimage* **5,** 179–197.

11

Psychiatric Genetics

Future and Prospects

Ming T. Tsuang, Levi Taylor, Stephen V. Faraone

1. Introduction

When a completed draft emerged from The Human Genome Project in 2001 ***(1)***, the most ambitious biological enterprise ever undertaken yielded a prize of almost inestimable value: A map of the basepair (bp) sequence of the nearly 6-foot strand of DNA found in almost all of the body's 100 trillion cells ***(2–4)***. As President Clinton has stated, "Without a doubt, this is the most wondrous map ever created by mankind." Data from that project will be generally available to the scientific community, permitting a wide variety of projects in several areas. Medical tests will be created to identify those at risk for a multitude of diseases, so that people will gain time to take preventative measures. Drugs designed on the basis of genetic information will provide more effective treatments, and genetic therapies that work directly at the site of a problematic gene will become increasingly available. Despite the enormous size of the DNA strand, it appears that humans possess the relatively small number of 30,000 genes; the remainder of the material presents the

From: *Methods in Molecular Medicine, vol. 77: Psychiatric Genetics: Methods and Reviews*
Edited by: M. Leboyer and F. Bellivier

opportunity to document our evolutionary heritage with a degree of accuracy that was previously unimaginable.

At a cost of approximately 3 billion dollars, the Human Genome Project required a computing ability millions of times greater than that required by NASA to land a man on the moon. Obviously, the number of statisticians, molecular geneticists, and biomedical specialists required to accomplish the goals of this project was therefore enormous. Moreover, new generations of scientists were trained with project funds in order to meet the enormous demands required for the successful completion of the Human Genome Project. Those scientists developed (and continue to create) new technologies and analytic approaches that have direct implications for the project itself, and can be applied to other areas of the medical, biological, statistical, and computing sciences. Naturally, psychiatric genetics is prominently included among those fields.

2. Analysis in Psychiatric Genetics

Since systematic psychiatric research has provided considerable evidence that many psychiatric disorders have at least some degree of heritability, it seems reasonable to assume that these clinically diagnosed psychiatric disorders would readily lend themselves to genetic analyses. To approach this issue, we have previously noted that the issue of *heterogeneity* is essential ***(5)***. These researchers have emphasized that heterogeneity is of two types, *causal* and *clinical*. Causal heterogeneity applies to conditions in which two or more causes can independently induce the same clinical syndrome. Clinical heterogeneity applies when a single cause leads to more than one clinical syndrome. Both types of heterogeneity are relevant to psychiatric genetics because both illustrate the problems of relying on clinical diagnoses alone in the search for the genes underlying the disorder. Instead, the full spectrum of conditions caused by the relevant genes must be considered, and that spectrum is identified by considering the issue of heterogeneity. The spectrum could be developed from a pool of genetic data derived from a pool of families; in turn, this spectrum could be used to develop a *psychiatric*

genetic nosology. This nosology would enable the formation of categories of patients based on distinct genetic entities.

In order for psychiatric genetic nosology to be effective, it must tackle the problem of *diagnostic accuracy*; viz., the degree to which a diagnosis accurately distinguishes individuals with disorders associated with specific susceptibility genes from those who do not. There are two fundamental types of diagnostic inaccuracy: false-negatives and false-positives. False-negatives can occur for two reasons: The first is found when patients are mistakenly diagnosed as being well or having some other disease when they actually have the relevant disorder. The second cause occurs as a result of *reduced penetrance*—the subject has genes for the disorder, but has not manifested the disease at the time of diagnosis.

False-positive diagnoses are also made for two reasons that are approximately inverse of the causes for false-negatives. The first reason is simply that the patient is misdiagnosed as having the illness. The second reason for a false-positive is that the patient has been accurately diagnosed with the correct general disease, yet does not possess the genetic subform of the illness.

3. Reducing the Rate of False-Negatives and False-Positives

The preceding section suggests that any attempt to develop psychiatric genetic nosology must involve an attempt to reduce the rate of false-positives and false-negatives. It is important to note that it is quite unlikely that genetically influenced brain processes would have a one-to-one correspondence with the complex clinical phenomena that are used to form a positive diagnosis. Instead, psychiatric signs and symptoms are more likely to be more remote effects of genes. Thus, it is almost certain that a promising approach for reducing false-positives and false-negatives in psychiatric genetics would be to directly examine measures of the neurobiology of the brain. Moreover, the utility of that approach is underscored by the fact that very few complex disorders are likely to be caused by a single gene.

In a general way, the strategy described here would be undertaken in the following manner: If a disease features a gene that is also known to relate to some measurable CNS deficit, measurement of that deficit (the presence of phenotypic manifestation of the gene) could be used in linkage analysis. By extension, if the illness is caused by several genes, it would ideally be possible to measure alternate phenotypes, and thereby find neurobiological measures that correspond to each one of those genes. Examination of alternative phenotypes could therefore be highly effective in detecting the gene linkage in a disorder, even if phenotypic presentation with the disease cannot.

Two problems arise in the use of the preceding approach. The first is that, as the number of phenotypic indicators included in a linkage study is increased, the probability also increases that any positive finding will be a result of mere chance. Although this problem can be managed through several statistical procedures, each one entails a loss of power. The second problem with the use of alternative phenotypes is that the procedure's decrease in false-negatives usually involves an increase in false-positives (since the phenotypic indicators are usually more prevalent among controls than the disease under study). For example, the oculomotor dysfunction among relatives of schizophrenic patients occurs at a rate of 14–50% ***(6)***. Because this rate is greater than the rate of schizophrenia and related psychoses among relatives (10%), the inclusion of oculomotor dysfunction decreases the rate of false-negatives ***(6)***. Nevertheless, despite the fact that oculomotor impairment is statistically more common among relatives than for controls, the rate among controls is 2–8× greater than the 1% population risk for schizophrenia. Therefore, the use of oculomotor measures as a phenotypic indicator involves the risk of an increase in false-positives.

One way to decide whether the tradeoff between false-negatives and false-positives warrants the use of a particular phenotypic indicator has been suggested by Risch ***(7,8)***. He showed that, as the ratio of the prevalence among relatives of those with an illness and the prevalence in the general population increases, the power of a linkage study concomitantly increases. Thus, the statistical power

of linkage analysis can be increased by identifying a phenotype that is quite prevalent among relatives but rare in the general population ***(6,9,10)***. An alternative phenotype will thus be most useful if it increases Risch's prevalence ratio.

4. Predictive Genetic Testing

A major goal of psychiatric genetics is to use genetic information to discover preventive or better palliative techniques for treating psychiatric disorders. Unfortunately, predictive presymptomatic or prenatal genetic testing techniques are still in their infancy in psychiatry. That is because, unlike other diseases that have accurate methods for predicting the presence of an illness (*see* **Table 1**), the molecular genetics of psychiatric illnesses are not yet fully understood. Nevertheless, linkage findings in disorders such as Alzheimer's disease and Parkinson's disease suggest that accurate DNA tests to determine at-risk individuals will probably be developed in the near future. However, it is important to consider that for such tests to be clinically useful, there must also be confidence about when the test results should be provided. (For example, the American Society of Human Genetics opposes presymptomatic testing for adult-onset conditions among at-risk children because of the potential for adverse psychological effects on both children and their parents.) Moreover, information derived from the genetic testing should have the potential to improve the lives of those who are tested. (Again, as an example, it has become a common practice in some Orthodox Jewish communities for couples who are contemplating marriage to receive Tay-Sachs testing prior to marriage to enable the couple to make a more informed decision regarding potential children).

In psychiatry, genetic testing may improve the lives of at least some individuals who are found to possess disease genes. For example, if it becomes possible to determine those genes that predispose an individual to schizophrenia, it would become possible to provide proactive interventions for signs of emergent psychosis, and thereby increase the chance of a better prognosis. In fact, if preventive treatments could be developed, early identification may even

Table 1
Currently Available Genetic Tests for Presymptomatic or Prenatal Testing

Disease tested	Reference
Alpha thalassemia	***(11)***
Breast cancer	***(12)***
Cystic fibrosis	***(11)***
Epidermolysis bullosa simplex	***(13)***
Familial Mediterranean fever	***(14,15)***
Fragile X mental retardation	***(16)***
Friedreich's ataxia	***(17)***
Huntington's disease	***(18)***
Gaucher's disease	***(19)***
Kennedy's disease	***(20)***
Myotonic dystrophy	***(21)***
Niemann-Pick disease, type C1	***(22)***
Sickle-cell anemia	***(23)***
Spinocerebellar ataxia type 1	***(24)***
Spinocerebellar ataxia type 3	***(25)***
Tay-Sachs disease	***(26)***

prevent psychosis in some cases. Yet we are still far from reaching that Utopian world.

5. The Future of Psychiatric Treatment

For almost all serious psychiatric illnesses, treatment currently involves both pharmacological and cognitive-behavioral regimens. As noted in the previous section, genetic research promises to supplement that strategy through early detection and prevention, and by developing improved therapeutic techniques targeted at relevant genetic sites.

5.1. Early Identification and Prevention

There are two essential distinctions emphasized by prevention researchers: *Primary* vs *secondary* prevention; and *universal* preventive interventions vs *selective* preventive interventions ***(27–29).***

"Primary prevention" can be any intervention that blocks the onset of a disease. An obvious example is preventing cirrhosis of the liver by limiting alcohol intake. In psychiatry, there is no comparably simple algorithm for preventing mental illnesses, but essentially similar techniques could be yielded by a deeper understanding of the genetic/physiological foundation of any given disorder.

Primary prevention can be implemented on two levels. Prenatal or other early environmental insults can be prevented, thereby preventing the onset of resultant neurodevelopmental abnormalities and ultimately, psychological disorders ***(30,31)***. Later environmental insults can prevent psychopathology, but not initial neurodevelopmental abnormalities ***(32,33)***.

Secondary prevention can only be directed toward alleviating the severity of the effects of a disorder. In this type of prevention, an intervention early in the course of the disease can reduce both the number of active incidences of the disease and the severity of those instances. For this type of intervention, it is essential to identify the presence of the illness as quickly as possible. Clearly, an understanding of the genetic underpinnings of a given illness would further the latter end immensely.

Another model of prevention involves universal, selective, and indicated prevention ***(34)***. Universal preventive programs apply prevention resources to an entire group without regard to the risk levels of particular individuals within that group. As a result, research directed toward distinguishing high risk from low risk is unnecessary.

Selective prevention intervention involves targeting individuals who are more likely to develop a psychiatric illness than most members of the population as a whole. Psychiatric genetic studies provide an ideal method to discover either genes or phenotypic indicators that are predictive of future illness ***(35,36)***.

Indicated prevention is applicable to individuals who display a sufficiently serious risk factor or condition to warrant preventive intervention. Examples of this type of intervention include hypertension control, antituberculosis medication for those who need it as determined by skin test, or dietary modifications for those with abnormally high levels of cholesterol. The basis for deciding whether to implement universal vs indicated prevention protocols is

largely one of cost: Because the cost of indicated preventative measures precludes the implementation among the population at large, there must be a clinical basis for determining that an indicated measure is warranted.

5.2. Medical Interventions for Genetic Defects

In the earliest years of psychopharmacology, drugs were often administered on the basis of the empirical observation that a particular substance exerted a palliative effect, with no real understanding of the psychobiological basis for the observed effect. After substantial advances in the fields of neuroscience and psychopharmacology, a greater understanding of pharmacodynamics emerged, but these findings involved the extant state of the brain as it was when studied; they did not account for *how* the brain came to produce the aberrant proteins requiring intervention. Advances in psychiatric genetics will help to overcome this last causal hurdle. Thus, as stated earlier, as an understanding of psychiatric genetics progresses, medical interventions targeted at specific genetic defects will be improved.

The relatively new field of gene therapy offers real promise in developing medical interventions for genetic disorders in the field of psychiatry ***(37–40)***. Gene therapy is a procedure in which mutant genes are removed from target cells and are replaced with "normal" genes. In other fields, such as hemotology, gene therapy has become an increasingly widespread, successfully implemented technique for curing disease. However, it has not yet been successful in treating diseases with complex genetic foundations. Techniques for doing this are currently being developed, and as they become more refined, gene therapy will become increasingly viable as a strategy for treating mental disorders.

6. Genomics and Schizophrenia

The Human Genome Project has already raised the possibility of a variety of innovations in research. For example, it may soon be

possible to use DNA microarrays to diagnose cancer and infectious disease subtypes, as well as to predict clinical outcomes ***(41)***. The genome could also be used to examine the interactions of the environment, genes, and toxic exposures. It is therefore reasonable to envision that genomics can be applied to gaining a better understanding of schizophrenia, and thereby, developing better clinical methods for attacking the disease.

Niculescu et al. ***(42)*** treated rats with methamphetamine to create an animal model for psychotic mania. Specific brain regions were comprehensively analyzed for changes in gene expression using oligonucleotide GeneChip microarrays. The data were cross-matched against human genomic loci associated with either bipolar disorder or schizophrenia. In this way, they identified several novel candidate genes (such as signal transduction molecules, transcription factors, or metabolic enzymes) that may be involved in the etiology of mood disorders and psychosis. Moreover, their analyses suggested a future classification of candidate genes into two types: psychogenic genes and psychosis-suppressor genes.

Another development that is promising for the future of schizophrenia research and treatment is the quantitative trait loci (QTL) approach (e.g., Williams and Blangero) ***(43)***. This approach involves genetic expressions used in variance component linkage analyses of a quantitative trait in undetermined samples. Comprising effects caused by a single QTL, residual additive genetic factors and individual-specific random environmental variation are considered. The result is the ability to assess more specifically the contribution of specific loci in genomic research. This method offers particular promise to the study of schizophrenia, because of the unique difficulties involved in separating gene-environment interactions. Technological advances in genomics therefore will obviously play a pivotal role in furthering the science.

7. Genomics, Proteomics, and Technolgy

As stated by Rudert ***(44)***, the sequencing of the human genome was only made possible by the massively parallel use of automated

high-throughput technologies that involved the development and interfacing of new hardware and software in a way that would previously have been inconceivable. One of the most relevant advances was the development of the DNA-chip principle, mainly used for RNA expression profiling. This principle has become the basis of new genomics and proteomics technologies, culminating in the lab-on-a-chip concept, which in the next 5–10 yr, could advance at a comparable rate to that of computers over the last 50 yr ***(44)***. The technology has already been applied to the attempt to describe disease and disease risk at the molecular level. Nevertheless, the field is still too new to determine whether it will yield the diagnostic efficacy that some have predicted, particularly as it might be applied to schizophrenia.

The effect of genes on any given biological mechanism is naturally indirect. Rather, genes give rise to proteins that lead to physiological processes. The study of the effect of those proteins is "proteomics." Proteomics refers to the cataloging and analysis of every protein in the human body as it depends on in vivo parameters. Because disease or drugs can alter a protein profile, a description of protein profiles can enhance the understanding of disease as well as the effect of pharmaco-interventions. In the past, since nucleic acids are more chemically homogeneous, genomics was regarded as a more promising approach than proteomics. However, with improvements in proteomic methods, it is likely that proteomics will complement genomics as a tool to study life sciences (*see* **ref. *44***). In turn, an advancement in the understanding of schizophrenia is likely to be facilitated by proteomic as well as genomic research. In fact, as noted by Futterman and Lemberg ***(45)*** our understanding of the human organism and its various ills will, in 50 years, be "transformed beyond recognition through genomics and proteonomics."

8. Ethics and Psychiatric Genetics

The genotyping of human beings has proven to be especially controversial from an ethical standpoint. In turn, the genotyping of human beings in terms of psychological and psychiatric attributes is

especially controversial ***(46–48)***. That is because those fields concern the attributes that most closely correspond to what we understand to constitute "self." Any labeling of the genetic core of psychiatric phenomena therefore stigmatizes some people unfairly, even to the extent of labeling them as "sub-human." Another concern is that health insurance may be denied to people on the basis of their genetic makeup.

Nevertheless, history has shown that any attempt to halt the progress of knowledge is not only fruitless, but dangerous. Moreover, the support of free inquiry is assumed to be the best way to protect humanity's interests, since knowledge *per se* has no ethical valence. Still, it is also true that knowledge in the hands of this unscrupulous or indifferent has often led to disastrous consequences. In light of that last, the Ethical, Legal and Social Initiative (ELSI), the National Institute of Health's Human Genome Project, was created to address that problem. The ELSI "team" consists not only of geneticists, but thinkers from as many relevant areas as possible, including religion, philosophy, law, psychology, and sociology. ELSI researchers continue to address any issues relevant to the ethics of genetics, thereby avoiding many of the disastrous consequences that ensued upon the development of physics a century earlier (e.g., the misuse of atomic energy). In that sense, at least, with the foundation of ELSI, we have not only progressed scientifically or therapeutically in the past century, but (hopefully) ethically as well.

References

1. Science (2001) The human genome map. *Science* **291,** 1218.
2. Collins, F.S. (1999) Shattuck Lecture—medical and societal consequences of the human genome project. *N. Engl. J. Med.* **341,** 28–37.
3. Waterston, R. and Sulston, J.E. (1998) The human genome project: reaching the finish line. *Science* **282**, 53–54.
4. Watson, J.D. (1990) The human genome project: past, present, and future. *Science* **248,** 44–49.
5. Faraone, S.V., Tsuang, D., and Tsuang, M.T. (1999) *Genetics of Mental Disorders: A Guide for Students, Clinicians, and Researchers*. Guilford, New York, NY, pp. 46–89.

6. Faraone, S.V., Kremen, W.S., Lyons, M.J., Pepple, J.R., Seidman, L.J., and Tsuang, M.T. (1995a) Diagnostic accuracy and linkage analysis: how useful are schizophrenia spectrum phenotypes? *Am. J. Psychiatry* **152,** 1286–1290.
7. Risch, N. (1990a) Linkage strategies for genetically complex traits. I. Multilocus models. *Am. J. Hum. Genet.* **46,** 222–228.
8. Risch, N. (1990b) Linkage strategies for genetically complex traits. II. The power of affected relative pairs. *Am. J. Hum. Genet.* **46,** 229–241.
9. Faraone, S.V., Biederman, J., and Monuteaux, M. (2000) Toward guidelines for pedigree selection in genetic studies of attention deficit hyperactivity disorder. *Gen. Epidemiol.* **18,** 1–16.
10. Faraone, S.V., Seidman, L.J., Kremen, W.S., Pepple, J.R., Lyons, M.J., and Tsuang, M.T. (1995b) Neuropsychological functioning among the nonpsychotic relatives of schizophrenic patients: a diagnostic efficiency analysis. *J. Abnorm. Psychol.* **104,** 286–304.
11. Lindeman, R., Wallace, R., Volpato, F., Hu, S.P., and Trent, R.J. (1991) Utility of the polymerase chain reaction (pcr) for prenatal diagnosis of genetic disease. *Pathology* **23,** 158–163.
12. Iacopetta, B., Elsaleh, H., Grieu, F., Joseph, D., Sterrett, G., and Robbins, P. (2000) Routine analysis of p53 mutation in clinical breast tumor specimens using fluorescence-based polymerase chain reaction and single strand conformation polymorphism. *Diagnostic Molecular Pathology* **9,** 20–25.
13. Bonifas, J.M., Rothman, A.L., and Epstein, E.H.J. (1991) Epidermolysis bullosa simplex: evidence in two families for keratin gene abnormalities. *Science* **254,** 1202–1205.
14. Booth, D.R., Gillmore, J.D., Lachmann, H.J., Booth, S.E., Bybee, A., Soyturk, M., et al. (2000) The genetic basis of autosomal dominant familial Mediterranean fever. QJM: *Monthly Journal of the Association of Physicians* **93,** 217–221.
15. Mohey, R., Kelsen, J., Tarp, B., and Obel, N. (2000) Familial mediterranean fever diagnosed by pcr. *Scandinavian J. Infect. Dis.* **32,** 334–335.
16. Duran Dominguez, M., Molina Carrillo, M., Fernandez Toral, J., Martinez Merino, T., Lopez Aristegui, M., Alvarez Retuerto, A., et al. (2001) Molecular diagnosis of fragile x syndrome by polymerase chain reaction: application of a diagnostic protocol in 50 families from northern Spain. *Anales espanoles de pediatria* **54,** 331–339.

17. Campuzano, V., Montermini, L., Molto, M.D., Pianese, L., Cossee, M., Cavalcanti, F., et al. (1996) Friedreich's ataxia: autosomal recessive disease caused by an intronic GAA triplet repeat expansion. *Science* **271,** 1423–1427.
18. Thies, U., Zuhlke, C., Bockel, B., and Schroder, K. (1992) Prenatal diagnosis of Huntington's disease (hd): experiences with six cases and PCR. *Prenatal Diagn.* **12,** 1055–1061.
19. Mistry, P.K., Smith, S.J., Ali, M., Hatton, C.S., McIntyre, N., and Cox, T.M. (1992) Genetic diagnosis of gaucher's disease. *Lancet* **339,** 889–892.
20. Faragher, M.W., Choi, W.T., MacLean, H.E., Warne, G.L., and Zajac, J.D. (1993) Kennedy's disease: clinical presentation and laboratory diagnosis. *Clin. Exp. Neurology* **30,** 61–65.
21. Clark, C., Kelly, K.F., Smith, N., Fairweather, N., Brown, T., Johnston, A., et al. (1991) Prenatal diagnosis for dystrophia myotonica using the polymerase chain reaction. *Prenatal Diagn.* **11,** 467–470.
22. Tsukamoto, H., Yamamoto, T., Nishigaki, T., Sakai, N., Nanba, E., Ninomiya, H., et al. (2001) SSCP analysis by RT-PCR for the prenatal diagnosis of Niemann-Pick disease type c. *Prenatal Diagn.* **21,** 55–57.
23. De Rycke, M., Van De Velde, H., Sermon, K., Lissens, W., De Vos, A., Vandervors, T.M., et al. (2001) Preimplantation genetic diagnosis for sickle-cell anemia and for beta-thalassemia. *Prenatal Diagn.* **21,** 214–222.
24. Sumitani, S., Ishikawa, Y., and Minami, R. (1996) Non-radioactive dna diagnosis of dentatorubral-pallidoluysian atrophy and Machado-Joseph disease using a new primer set. *Clin. Chim. Acta* **254,** 187–195.
25. Hsieh, M., Li, S.Y., Chen, Y.Y., Tsai, C.J., Lin, M.L., Li, C., et al. (1999) A non-radioactive polymerase chain reaction method for diagnosis of Machado-Joseph disease. *J. Formos. Med. Assoc.* **98,** 705–707.
26. Drucker, L., Proia, R.L., and Navon, R. (1992) Identification and rapid detection of three Tay-Sachs mutations in the Moroccan Jewish population. *Am. J. Hum. Genet.* **51,** 371–377.
27. American Academy of Child and Adolescent Psychiatry (1990) Prevention in *Child and Adolescent Psychiatry: The Reduction of Risk For Mental Disorders*. American Academy of Child and Adolescent Psychiatry, Washington, DC.
28. Task Force on Prevention Research (1990) Report of the APA Task Force on Prevention Research. *Am. J. Psychiatry* **147,** 1701–1704.

29. Tsuang, M.T., Stone, W.S., and Faraone, S.V. (2000) Towards the prevention of schizophrenia. *Biol. Psychiatry* **48,** 349–356.
30. Buka, S.L., Tsuang, M.T., and Lipsitt, L.P. (1993) Pregnancy/delivery complications and psychiatric diagnosis. A prospective study. *Arch. Gen. Psychiatry* **50,** 151–156.
31. Zornberg, G.L., Buka, S.L., and Tsuang, M.T. (2000) Hypoxic-ischemia-related fetal/neonatal complications and risk of schizophrenia and other nonaffective psychoses: a 19-year longitudinal study. *Am. J. Psychiatry* **157,** 196–202.
32. Lewis, S.W. and Murray, R.M. (1987) Obstetric complications, neurodevelopmental deviance, and risk of schizophrenia. *J. Psychiatr. Res.* **21,** 413–321.
33. Weinberger, D.R. (1994) Schizophrenia as a neurodevelopmental disorder: A review of the concept, in *Schizophrenia* (Hirsch, S.R. and Weinberger, D.R., eds.), Blackwood Press, London, 293–323.
34. Gordon, R.S. (1983) An operationel classification of disease prevention. *Public Health Rep.* **98(2),** 107–109.
35. Erlenmeyer-Kimling, L. (2000), Neurobehavioral deficits in offspring of schizophrenic parents: liability indicators and predictors of illness. *Am. J. Med. Genet. (Neuropsychiatr. Genet.)* **97,** 65–71.
36. Erlenmeyer-Kimling, L., Hilldoff, Adamo, U., Rock, D., Roberts, S.A., Bassett, A.S., Squires-Wheeler, E., et al. (1997) Original article The New York high-risk project prevalence and comorbidity of axis I disorders in offspring of schizophrenic parents at 25-year follow-up. *Arch. Gen. Psychiatry* **54,** 1096–1102.
37. Friedmann, T. (1994) Gene therapy for neurological disorders. *Trends Genet.* **10,** 210–214.
38. Marshall, E. (1995a) Gene therapy's growing pains. *Science* **269,** 1050–1055.
39. Marshall, E. (1995b) Less hype, more biology needed for gene therapy. *Science* **270,** 1751.
40. Mulligan, R.C. (1993) The basic science of gene therapy. *Science* **260,** 926–932.
41. Schmidt, C.W. (2001) Charting the map of life. *Environ. Health Perspect.* **109,** A24–A29.
42. Niculescu, A.B., Segal, D.S., Kuczenski, R., Barett, T., Hauger, R.L., and Kelsoe, J.R. (2000) Identifying a series of candidate genes for mania and psychosis: a convergent functional genomics approach. *Physiol. Genomics* **4,** 83–91.

43. Williams, J.T. and Blangero, J. (1999) Power of variance component linkage analysis to detect quantitative trait loc. *Ann. Hum. Genet.* **63(Pt 6),** 545–563.
44. Rudert, F. (2000) Genomics and proteomics tools for the clinic. *Current Opinions in Molecular Therapy* **2,** 633–642.
45. Futterman, L.G. and Lemberg, L. (2001) The mysteries of the human genome uncovered—medicine is changed forever. *Am. J. Crit. Care* **10(2),** 125–132.
46. Alexander, J.R., Lerer, B., and Baron, M. (1992) Ethical issues in genetic linkage studies of psychiatric disorders. *Br. J. Psychiatry* **160,** 98–102.
47. Shore, D., Berg, K., Wynne, D., and Folstein, M.F. (1993) Legal and ethical issues in psychiatric genetic research. *Am. J. Med. Genet., Neuropsychiatr. Genet.* **48,** 17–21.
48. Sam, R.J., Bulyzhenkov, V., Prilipko, L., and Christen, Y. (1991) *Ethical Issues of Molecular Genetics in Psychiatry*. Vol. May, (Christen, Y. ed.), Springer Verlag, New York, NY.