Molecular genetics of schizophrenia: a critical review

Neeraj Berry, MD; Vaidehi Jobanputra, PhD; Hemraj Pal, MD

Introduction

Schizophrenia is a chronic disabling disease that affects about 1% of the world’s population. Although the causes of schizophrenia remain unknown, evidence from family, twin and adoption studies clearly demonstrates that it aggregates in families, with the clustering being largely attributable to genetic rather than cultural or environmental factors.1-4

Initial attempts to study the genetic mechanism underlying schizophrenia took the form of segregation analysis or searching for “biologic markers” (such as receptor proteins or neurophysiologic findings from, for example, electroencephalography) that segregate with the disorder.

The advent of molecular genetics was a turning point in schizophrenia research, when it became possible to use both linkage and association methods to study the

Correspondence to: Dr. Neeraj Berry, Department of Psychiatry, All India Institute of Medical Sciences (AIIMS), New Delhi-110029, India; fax 91-11-26588663; berry.neeraj@hotmail.com

Medical subject headings: genetic predisposition to disease; genetics, behavioral; linkage (genetics); linkage disequilibrium; schizophrenia.


Submitted May 17, 2002
Revised May 5, 2003
Accepted May 27, 2003

© 2003 Canadian Medical Association
DNA markers spanning the genome or the candidate gene polymorphism suspected of being related to the disorder. The search began with reports on chromosome 5 by Sherrington et al., though it took almost a decade to reach the pace at which studies are being done today in this field.

Identifying the genes, mode of transmission and responsible chromosomes has proved to be a difficult task in view of the imprecise phenotype, the several phenocopies present and the effect of nongenetic risk factors.

Linkage studies and candidate gene association studies have been used in clinical cohorts collected from a variety of populations, and these studies have given us some hint as to the chromosomes and genes involved. The results from these studies also need careful interpretation regarding their statistical significance and applicability to the whole population.

Nevertheless, with the refinement of phenotypes, the use of endophenotypes, reduction of heterogeneity and extensive genetic mapping, especially after the advent of the Human Genome Project, the chances of discovering mechanisms of inheritance as well the susceptible genes have increased substantially.

In this paper, we review recent and important studies in this field, with particular reference to the various molecular methods used to study the genetic mechanism and their limitations; at the same time, we consider potential improvements that could help us develop a model to predict expression of the disease with more precision.

**Linkage studies**

Linkage analysis seeks to find chromosomal regions within families that tend to be shared among affected relatives but not among unaffected individuals. Conceptually, linkage analysis takes place in 3 steps:

1. A linkage statistic at each of many DNA markers throughout the genome is calculated.
2. The markers that the linkage statistic shows to be transmitted are identified.
3. More markers in that region are tested to locate the gene more precisely.

The statistics are needed in linkage studies to summarize the evidence available from different studies and to help identify the most likely “allele configuration” in case the DNA sample is not available from 1 of the parents. The statistical theory underlying linkage analysis is fairly complex but can be summarized in 2 critical numbers:

1. a statistic whose magnitude increases with the evidence for linkage, and
2. a number between 0 and 1 that indicates the probability of computing the observed statistic in linkage was absent.

The “affected sibling pair” (ASP) method of linkage relies on the ability to identify “identity by descent,” namely, the alleles that were inherited from the same parent. The probability that all ill relative pairs would inherit the same version of the genetic marker is more than that predicted by Mendel’s laws or chance. Similarly, sibling pairs discordant for the disease share alleles at the marker locus more often than predicted by Mendel’s laws. This affected pedigree method can be used without knowing the mode of inheritance, which makes it very appealing in the study of psychiatric disorders.

The second popular method is to compute a logarithm of the odds ratio (LOD) score. The main drawback of this method is the need to specify the mode of transmission. However, to get around this, when the data are analyzed several times under different modes of inheritance, the highest LOD score is close to the true mode of inheritance. The LOD score is computed by first calculating the maximum likelihood estimate — the rate of a marker being linked or not linked to a disease locus under a specific assumption. In calculating these scores, the genotype of entire families is pooled rather than that of the ill relatives alone. A LOD score of 3 is considered to be evidence for linkage, but lately it has been found that this rule works well for single gene diseases rather than for complex diseases such as psychiatric disorders. To conclude whether apparent linkage is “real,” the concept of “genome-wide significance” has been developed — the probability threshold that declares linkage after testing many DNA markers used in a genome scan. Lander and Kruglyak suggest 3 levels of genome-wide significance: suggestive linkage, significant linkage and confirmed linkage, though it is suggested that confirmed linkage only occurs when the results are replicated in an independent study sample.

**Association studies**

Linkage analysis has been extremely successful for finding the genetic basis of diseases with well-defined...
modes of inheritance, with the exception of Alzheimer’s disease; it has not yet identified genes for psychiatric disorders, hence certain researchers have turned to association studies.

The “population-based association study” considers specific genes believed for theoretical reasons to be involved in the pathogenesis of a disorder. Such genes are called “candidate genes.” The gene frequencies between patients and controls are compared. A weak inferential link when reasoning from association studies is the fact that a positive association finding for a specific gene could occur if the gene was very close to the true disease gene. These genes are said to be in “linkage disequilibrium,” that is, rarely separated by crossover.

Because of this phenomenon, we do not necessarily need a candidate gene for association studies. Population-based association studies are limited by potential ethnic differences between patients and controls. Therefore, “family-based association studies” are carried out in families that have at least 1 affected offspring, and a transmission test of linkage disequilibrium (TDT) is applied. This test essentially compares the number of times the heterozygous parents transmit the associated marker to affected offspring as compared with the other marker. If these probabilities differ from those expected by chance, then linkage disequilibrium exists, that is, that gene is associated with the disease. With this technique there is perfect ethnic matching, because the transmitted and nontransmitted alleles are from the same parent.

**Molecular genetic studies**

**Linkage studies**

After the initial enthusiasm regarding a potential link with chromosome 5 suggested by Bassett, subsequent studies have given mixed results. In a genome-wide scan of a nationwide study sample from Finland, the highest LOD scores were found on chromosome arm 5q, whereas another study investigating various regions including 5q in 62 pedigrees from Finland found little support for linkage. 5q has also been implicated in a set of Irish pedigrees and in German and Israeli families. A recent genome-wide scan suggested that 5q33.2 should be intensively investigated by linkage disequilibrium methods. A second region of interest on this chromosome is at the D5S111 locus 5p4.1–p13.1. A significant 2-point LOD score (3.72) and a multipoint score of 4.37 at this locus have been reported in a large Puerto Rico pedigree, using a broad disease definition and dominant inheritance; however, the LOD score has been criticized for not being robust to sensitivity analysis. In contrast, studies of 5 multiplex pedigrees from eastern Canada and certain other studies produced evidence against this region.

Another chromosome to be implicated as being associated with schizophrenia has been chromosome 6, because the 6p22–p24 locus emerged as having one of the strongest associations in 265 Irish families using an additive genetic model and an intermediate phenotypic definition, but this evidence substantially declined when either a narrow or a broad disease definition was used. Certain other studies have also shown a positive association with certain regions of chromosome 6. A study evaluating 28 genetic markers in 10 moderately large Canadian families using ASP analysis found no evidence for linkage, but the positive symptom scale scores on the Positive and Negative Symptoms Scale for schizophrenia (PANSS) were associated with marker D6S1960, suggesting that the locus might be related to the severity of psychotic symptoms. A linkage disequilibrium analysis of 115 ASPs with microsatellite markers also pointed toward a susceptible locus at D6S1960 on 6p. Another linkage analysis performed in 186 multiplex families, assuming locus heterogeneity and moderately broad disease definition, revealed a significant LOD score and genome-wide significance of 5%–8%, again supporting a susceptibility locus on chromosome 6 and a model of locus heterogeneity. Martinez et al. in a follow-up study of a sample that had previously shown suggestive linkage on 6q, found an increase in the LOD score for a 13-cM region after the addition of 43 multiplex pedigrees, and they suggested that a very large sample would be required to narrow down the locus further by ASP linkage methods. There is positive evidence for association with both the long and the short arm of chromosome 6 from genome-wide scans, though there are studies in which linkage could not be replicated; however, interest continues to centre on 6p24–22 and 6q21–22.

Chromosome 6 has several candidate genes, that is, mutations that are hypothesized to predispose individuals to schizophrenia. The spinocerebellar ataxia gene (SCA), LDL-PLA 2, HLA region at 6p21 and, recently, the NOTCH4 gene have been studied; these studies have been primarily negative except for positive link-
age disequilibrium between SCA1 CAG repeats and schizophrenia.28–34

The implicated region in many studies of chromosome 22 is near the site for the velocardiofacial syndrome deletion on 22q, clustering about 4–5 cM around the marker D22S278–9.35 In another study, in 8 Utah multigenerational families, where schizophrenia was phenotyped with P50 auditory evoked potential and ocular motor performance, a genome-wide scan using an autosomal dominant model revealed a significant LOD score (3.55) for marker D22S315, and a nonparametric linkage (NPL) analysis showed evidence for allele sharing over the same broad region (i.e., 3.83).36 There is also evidence from other genome scans, and statistically significant evidence for linkage disequilibrium has been found for polymorphic markers within the catechol-o-methyltransferase (COMT) gene located on this chromosome.37 Several other studies have found negative evidence for linkage in specialized populations such as South African Bantu-speaking families and for allelic association between D22S278 and D22S283 in case–control samples or markers around IL2RB.38–40

In a genome-wide scan of 22 extended Canadian families with high rates of schizophrenia, highly significant evidence of linkage to chromosome 1 in the region 1q21–q22, between the markers D1S1653 and D1S1679 under the recessive model of inheritance with parametric analysis, using narrow disease definition, was reported. The power of the study was unusually high, though the authors claim that this was because of the selection of a sample of dense pedigrees.41 Several other linkage studies including genome-wide scans have also implicated regions on chromosome arm 1q especially involving the DISC1 gene,42,12 though a recent report found no linkage at this locus or any heterogeneity in a large multicentre sample.43

Chromosome 15 became a region of interest after initial reports from US schizophrenia pedigrees, the most promising region being 15q13–q15, which is also host to the α-7 nicotinic receptor gene. A significant linkage was found to a physiologic endophenotype (P50 wave of auditory evoked response) of schizophrenia where D15S1360, containing the gene α-7 nicotinic receptor, was the most positive marker.44 This was further confirmed in the systematic genome scans by the US National Institute of Mental Health (NIMH) Genetics Initiative.45 These scans also found genome-wide significant linkage at 15q14, which maps within 1 cM of the α-7 nicotinic receptor gene.46 Although some studies have failed to confirm linkage at the 15q14 locus, there are significant linkage reports from studies of German, European American, Azorean and Taiwanese families.47–50

Chromosome 11 includes several possible candidate genes for major psychiatric disorders including tryptophan hydroxylase (TPH), the D1 dopamine receptor gene (DRD4) (both at 11p15) and the D2 dopamine receptor gene (DRD2) at 11q23. Therefore, it has been a site of interest to many molecular geneticists. Interest has been maintained in this region because of a suggestive linkage reported by Maziade et al51 and weak evidence from certain genome-wide scans.52,53

In a genome scan of 43 US nuclear families, using a narrow disease definition, significant evidence for linkage was found at 10p for markers D10S1423 and D10S582 on chromosome 10.53 There is also strong evidence from German, Israeli and Irish pedigrees of a vulnerability locus on 10p for markers other than D10S1423 (D10S674, D10S1714).17,43,54,55

The linkage findings for potentially susceptible genes on chromosome 8 and 13 for schizophrenia are at 8p22–2p21 and 13q, respectively, and were first reported in 54 US pedigrees and replicated in 265 multiplex Irish families with a high prevalence of schizophrenia.56–60 In a recent genome scan, strong support of genetic linkage was found for 8p21–p22. Blouin et al56 reported that 13q14–q33 was the only significant region in their genome scan in 54 pedigrees with a maximum NPL score of 4.18 (p < 0.001), satisfying the threshold of a 5% genome-wide significance level.

The hypothesis that a gene in the pseudoautosomal region of the sex chromosomes might confer vulnerability to schizophrenia was based mainly on previous epidemiologic findings of excess anomalous sex chromosomes in psychotic patients, concordance by sex in family studies and sex differences associated with psychosis and underlying brain pathology.61–63 The validity of these findings has, however, been questioned, first, because of the only moderate strength of the association in the reverse direction and, second, because of the incomplete and biased ascertainment of ASPs in samples of limited size. Further, molecular genetic studies have yielded no evidence for the association of schizophrenia with the pseudoautosomal region.64–66 Even DeLisi et al,67 who initially showed moderate association at Xp11, despite listing a new set of markers and a larger cohort, could not find consistent linkage to the X chromosome and proposed that epigenetic modifica-
tion rather than variation in the X chromosome should be considered.

There are several reports of genome-wide scans of linkage studies in schizophrenia. There are several reports of genome-wide scans of linkage studies in schizophrenia. Some of the important features of these scans are listed in Table 1. These reports have again provided statistically significant evidence against a major genetic locus, but weak positive LOD scores were found for as many as 12 chromosomal regions. A meta-analysis of previous samples gave positive results for certain regions, especially on chromosome arms 6p, 6q, 8p and 22q, although drawbacks have been reported in the interpretation of these results.

### Association studies

These studies aim to find the allelic association with a certain disorder. The various mechanisms of this association follow: when a gene is located very close to a true disease gene, for instance at a distance of 1 cM, re-

### Table 1: Genome-wide scans of schizophrenia

<table>
<thead>
<tr>
<th>Study</th>
<th>Sample</th>
<th>Diagnosis</th>
<th>Markers</th>
<th>Analysis</th>
<th>Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>DeLisi et al⁶⁶</td>
<td>382 ASPs</td>
<td>Schizophrenia and SA</td>
<td>396 polymorphic markers</td>
<td>NPL, LOD</td>
<td>No evidence of linkage at 1q, 4p, 5p–q, 6p, 8p, 13q, 15p; significant linkage (LOD &gt; 3) at 10p15–p13 (D10S189)</td>
</tr>
<tr>
<td>DeLisi et al⁶⁹</td>
<td>99 families with at least 1 ASP, cohort of Spanish origin</td>
<td>Schizophrenia or schizophrenia spectrum</td>
<td>404 STR</td>
<td>NPL (Genehunter plus)</td>
<td>No region with significant linkage; little implication (maximum LOD score ~1) on 1p, 2p, 2q, 14p, 8p</td>
</tr>
<tr>
<td>Paunio et al⁷⁸</td>
<td>1200 Finnish individuals, 238 pedigrees</td>
<td>Schizophrenia</td>
<td>315</td>
<td>LOD</td>
<td>Significant LOD scores at 2q (D2S427), 5q (D5S414), 1q</td>
</tr>
<tr>
<td>Garver et al⁷⁴</td>
<td>30 multiplex pedigrees</td>
<td>Schizophrenia, schizophrenia spectrum disorders</td>
<td>Not reported</td>
<td>NPL</td>
<td>Positive linkage at 5p (D5S426), D15S18</td>
</tr>
<tr>
<td>Gurling et al⁷⁵</td>
<td>13 multiplex families</td>
<td>Schizophrenia</td>
<td>365 microsatellite markers</td>
<td>LOD</td>
<td>Linkage at (LOD &gt; 3) 1q33.2, 5q33.2, 8p22.1–p22, 11q21 and evidence for heterogeneity</td>
</tr>
<tr>
<td>Schwab et al¹⁷</td>
<td>305 individuals, 86 SPs</td>
<td>Schizophrenia</td>
<td>463 markers</td>
<td>LOD and NPL</td>
<td>No significant linkage at any region, though suggestive linkage at 6p (D6S2560) MHC, 10p</td>
</tr>
<tr>
<td>Williams et al¹⁷</td>
<td>Two-stage study, 196 ASPs</td>
<td>Schizophrenia (DSM-IV)</td>
<td>229 microsatellite markers</td>
<td>LOD</td>
<td>No region with LOD ≥ 3, suggestive gene of major effect unlikely to exist</td>
</tr>
<tr>
<td>Kaufmann et al¹¹</td>
<td>African American, 30 nuclear families, 98 subjects, 42 SPs</td>
<td>Schizophrenia</td>
<td>459 STR</td>
<td>LOD</td>
<td>No significant linkage as defined by Lander and Kruglyak; evidence suggestive of linkage on 6q16–q24, 8pter–8q12, 9q32–q34, 15p13–15q12, evidence of genetic heterogeneity</td>
</tr>
<tr>
<td>Blouin et al⁹⁶</td>
<td>54 multiplex pedigrees</td>
<td>Schizophrenia</td>
<td>452 microsatellite markers</td>
<td>NPL</td>
<td>Significant linkage at 13q32 and 8p21–p22</td>
</tr>
<tr>
<td>Levinson et al¹¹</td>
<td>269 individuals</td>
<td>Schizophrenia</td>
<td>310 markers</td>
<td>NPL</td>
<td>No significant linkage, suggestive linkage at 2q (D2S410), 10q (D10S1239)</td>
</tr>
<tr>
<td>Hovatta et al¹⁵</td>
<td>Three-stage study, Finnish population</td>
<td>Schizophrenia</td>
<td>Not reported</td>
<td>LOD</td>
<td>Suggestive evidence for linkage at 1q32.2–q41, 4q31, 9q21, Xp11.4–p11.3</td>
</tr>
<tr>
<td>Moises et al¹⁷</td>
<td>Two-stage study, genome scan, stage I pedigrees from Iceland, stage II families from 8 other countries</td>
<td>Schizophrenia</td>
<td>Not reported</td>
<td>LOD</td>
<td>No significant evidence as defined by Lander and Kruglyak; though suggestive linkage at 6p, 9 and 20. Evidence of locus heterogeneity, oligogenic transmission</td>
</tr>
</tbody>
</table>

Note: ASP = affected sibling pair; SA = schizoaffective disorder; NPL = nonparametric linkage; LOD = logarithm of the odds ratio; STR = short tandem repeats; MHC = major histocompatibility complex; DSM = Diagnostic and Statistical Manual of Mental Disorders.
resulting in linkage disequilibrium, or when a gene has a polymorphism within itself that has functional effect resulting in susceptibility to a disease, or when the phenomenon of population stratification occurs, whereby ethnic differences in allele frequencies contribute to observed differences between affected individuals and controls. Most recent studies focus on the polymorphism in or near candidate genes. These studies have also attempted to do a systematic search through the entire genome for linkage disequilibrium.

The dopamine hypothesis of schizophrenia has been accepted for about 3 decades, based on evidence that antipsychotic drugs are D2 receptor antagonists and their efficacy is correlated with affinity for D2 receptors. Defects of the genes in the dopamine system might play an important role in the origin of schizophrenia. Genes for the 5 dopamine receptors in the brain (i.e., D1, D2, D3, D4, D5) have been isolated and cloned. The chromosomal locations of the dopamine receptor genes are as follows: D1 5q35.1; D2 11q22.5; D3 3q13.3; D4 11p15.5; D5 4p15.2.

With regard to the genetics of schizophrenia, the crucial question is whether a linkage or association exists between schizophrenia and the genes for dopamine receptors. Since the advent of newer and atypical antipsychotics, several other neurotransmitters and receptors have also been implicated in the pathogenesis of schizophrenia. Some of the association studies related to these are summarized in Table 2.

Arinami et al found a substitution from serine to cysteine at position 311 of the D2 receptors and reported an increased frequency of the Cys 311 variant in Japanese patients with schizophrenia and earlier onset, good response to medication and a positive family history; this was also reported in another study. Conversely, other case–control studies in different populations reported no association. Arinami et al could not replicate their own studies, and several other researchers have found this substitution to be associated to a greater degree with certain clinical subtypes (disorganized schizophrenia) rather than schizophrenia. A study using a progressive strategy with 3 different approaches (case–control, haplotype relative risk test and TDT) found evidence for linkage between haplotype and schizophrenia, suggesting the need for further investigation at this receptor.

The D4 dopamine receptor gene is an important candidate gene for schizophrenia because of its almost exclusive expression in the limbic system. As well, D4 involvement would combine the dopamine receptor system hypothesis with the limbic system hypothesis. Several studies including a meta-analysis of studies of the Bal1 polymorphism in exon 1 of this gene found excess homozygotes in patients with high familial loading and a good response to neuroleptic treatment (Table 2). There are other studies that have not confirmed this association. However, 2 recent large analyses have provided substantial evidence for an association at the Bal1 site, and it is possible that inconsistent results are the result of a weak association that cannot be detected in every sample.

The D5 receptor may play a major role in schizophrenia given its high density in the frontal cortex and amygdala and its great affinity for an atypical antipsychotic, clozapine. However, several groups have failed to find an association between schizophrenia and the D5 dopamine receptor gene. In fact, Serreti et al could not replicate their own finding of association of DRD4 (exon 1 and exon 3) variants with any major psychosis.

Studies of the D4 dopamine receptor gene have found missense (N351D) and nonsense mutations (C335X). An analysis of combined measures of frontal lobe functions in patients with schizophrenia and controls hints that heterozygotes of C335X might have a vulnerability to mild impairment. Certain other studies have reported significant difference (excess) in allele frequency between patients with schizophrenia and controls, though these results were not replicated in an Italian population (Table 2).

Despite its importance as a candidate gene and a 40-base pair polymorphism of a variable number of tandem repeats in the 3’ untranslated region of the gene coding for the dopamine transporter (DAT), no linkage has been found between schizophrenia and the different alleles or allele combinations of the dopamine transporter gene.

Others have failed to find linkage between schizophrenia and all the dopamine receptor genes. Overall, it seems that molecular genetic studies lend only minor support to the dopamine theory of schizophrenia.

Serotonin-2A (5-HT2A) receptors have received much investigative attention in schizophrenia, because several studies have shown a decrease in the concentration of 5-HT2A receptor in the prefrontal cortex, a response to atypical antipsychotics and a positive association between the A and C polymorphism at position 102 of this receptor gene and schizophrenia. The various mutations screened in 5-HT receptor variants have
been h5-HT₁₃a (Gly-22-Ser, Ile-28-Val, Arg-219-Leu), h5-HT₁₃b (Phe-124-Cys), h5-HT₂₅a (Thr-25-Asn, His-452-Tyr) h5-HT₃c (Cys-23-Ser) and h5-HT₇ (Thr-92-Lys, Pro-279-Leu). The focus of attention has been a silent polymorphism in the 5-HT₁₃a receptor gene, though studies seeking a link with schizophrenia have not yielded positive results consistently (Table 2).

The serotonin transporter gene is another primary candidate for involvement in major psychosis. A functional polymorphism in its upregulatory region (5HT-
TLPR) has been recently reported to be associated with a variety of psychopathologic conditions. Serretti et al\textsuperscript{115} typed patients with schizophrenia for their 5HTTLPR variants and found that the serotonin transporter gene was not a liability factor for symptomatology for schizophrenia; although in another study this region was positively related to individual items on the Brief Psychiatric Rating Scale (BPRS), specifically, intensity of hallucinations, suggesting again a change in presentation according to the allelic variant.\textsuperscript{116}

There is now considerable evidence for glutamatergic system alterations associated with schizophrenia. Reduced expression and regional loss of non-N-methyl-d-aspartate (non-NMDA) glutamate receptors in the temporal lobe of patients with schizophrenia compared with healthy individuals has been reported.\textsuperscript{117} Other studies also report a disproportion in the expression of certain receptor subunits, covariation in several subunits or a long-range position effect to be associated with genetic predisposition to schizophrenia of a mutant allele in the subunit of the NMDA receptor gene.\textsuperscript{118–123}

An alteration in \(\gamma\)-aminobutyric acid (GABA) neurotransmission has been indirectly implicated in the pathogenesis of schizophrenia. GABA receptor subunit genes are plausible candidate genes. However, researchers found weak evidence of linkage between this gene and schizophrenia.\textsuperscript{124} Coon\textsuperscript{125} found a variant (C–G) in the B1 peptide region of this receptor, which cosegregated in small numbers, and they suggested that this variant might be a predisposing allele in rare instances. In an attempt to study the alteration in markers of GABA neurotransmission, including GABA membrane transporter GAT-1, it was seen that GAT-1 mRNA expression is relatively unaltered in most prefrontal cortex GABA neurons in patients with schizophrenia but is reduced below a detectable level in a subset of GABA neurons.\textsuperscript{126}

The tryptophan hydroxylase gene (\(TPH\)), which has been localized to chromosome 11 and codes for the rate-limiting enzymes of serotonin biosynthesis, has also been investigated as a candidate gene along with the \(COMT\) gene and dopamine transporter genes (\(DAT\)). Studies have reported an association between \(TPH\) polymorphism, variation in the region of intron 1 and 7 and negative symptoms of schizophrenia.\textsuperscript{127,128} \(COMT\) inactivates catecholamines, and a common genetic polymorphism in humans is associated with 3- to 4-fold variations in enzyme activity. In a study of 129 Turkish subjects, although no statistically significant difference was found between patients and controls regarding \(COMT\) polymorphism, the BPRS scores were associated with \(COMT\) polymorphism within the schizophrenia group.\textsuperscript{129} Associations between schizophrenia and polymorphisms of the 5-HT\textsubscript{2A} receptor gene, \(TPH\), \(COMT\) and \(DAT\) were tested in an Indian population, but none were found.\textsuperscript{130}

Recent work has identified 4 novel candidate regions by linkage disequilibrium, namely, neuregulin-1 (chromosome 8p21), G72 (chromosome 13q34), dysbindin (chromosome 6p22.3) and proline dehydrogenase (chromosome 22q11), warranting further investigation. These proteins play an important role in cell-to-cell and intracellular transduction.\textsuperscript{131,132}

Epigenetic mechanisms, defined as nonevironmental modifications of gene expression, have gained some attention recently. One such mechanism, “anticipation,” which refers to earlier age of onset and increased severity of disease from parental to offspring generations, has been observed in several neurologic disorders and in schizophrenia.\textsuperscript{133–135} Trinucleotide repeats that cause unstable DNA have been a suggested mechanism for anticipation, and recent studies are searching for trinucleotide repeats in schizophrenia. One such study\textsuperscript{136} reported correlation between the age of onset of spinocerebellar ataxia and CAG repeats at its gene, which is also seen as a candidate gene for schizophrenia.\textsuperscript{29} In many other studies, there is evidence both for and against anticipation in schizophrenia.\textsuperscript{137–140} Several reports implicate nongenetic factors such as “birth cohort effects” or “adjustment of observation time” (for development of the disorder) rather than anticipation in causing early onset and increased severity of the disorder in offspring.\textsuperscript{141,142} Imprinting, another epigenetic phenomenon, has been less well studied in schizophrenia. Imprinting refers to the fact that a mutation in the same gene is expressed in different ways depending on whether the defect is transmitted paternally or maternally. It has been examined for the age of onset where no imprinting effect was generally found.\textsuperscript{143–146}

Discussion

Although most individuals who are first-degree relatives of individuals with schizophrenia do not develop the disease themselves, data from adoption studies are consistent in pointing out that genetic factors rather than intrafamilial culture are responsible for familial
aggregation. Twin studies suggest the same conclusion by observing significantly higher concordance rates in monzygotic than in dizygotic twins. It is also evident that it is not schizophrenia per se that is inherited but a susceptibility to it, and environmental factors appear to be necessary for the disease manifestation in many, if not all, cases. Risk factors include obstetric complications, early developmental delays, life events, immigration and drug abuse. But even in this gene–environment model, the heritability is 66%–85%, again emphasizing the high proportion of liability to genetic influence.147

Molecular studies attempt to identify an alteration in genetic expression that causes susceptibility to a disease. The linkage studies reviewed earlier rule out a single major locus but suggest several regions that may be susceptibility loci and warrant replication and investigation in future studies. These regions (especially those identified in the genome-wide scans) include 5p, 5q, 6p, 6q, 8p, 10p, 13q and 15q.

The issue still remains as to whether these linkages are “real” or false positives. Determining an appropriate statistical threshold at which to accept or reject these findings is important, but to expect a “highly suggestive” leads are vigorously pursued. Moreover, the model proposed by Lander and Kruglyak assumes a single multipoint analysis of a very dense map, whereas most studies compute multipoint statistical analysis, which is likely to give inflated significance levels and make comparison across various studies difficult. The findings for 13q formally meet the criteria for replicated linkage and the other strongest finding is for chromosome 6p,58,24 but it would be discouraging to assume that the other results generated in different genome scans are all false positive. Suarez et al148 suggest that in a disorder caused by a set of interactive loci of small effect, the regions that produce the most positive results in multiple scans should be considered strong “candidate genes,” despite their nonsignificant p values.

The association methods give good results for a gene of modest effect for which a very large sample size is required by the linkage studies. Although the interpretation of these studies is particularly difficult because of problems of repeat analysis, multicentre analyses still support an association between schizophrenia and polymorphism in loci for the D4 dopamine receptor gene and the 5-HT1A receptor gene.

A number of other limitations have been pointed out that lead to a wide discrepancy in results and failure to identify the genes associated with schizophrenia. There is wide variability, from acute psychosis to the schizophrenia spectrum of personality disorders, in what has been included as “cases” in genetic studies. The importance of precise “phenotype selection” is highlighted by reports of overlap of candidate region genes of schizophrenia with those of disorders as distinct as bipolar disorder.190 The current suggestion is that one should classify subjects using affected, unaffected and unknown as categories, which would make the studies statistically more robust.

Considering the multidimensional phenotype of schizophrenia, studies have suggested that multiple symptomatic dimensions such as the different clinical subtypes or biologic endophenotypes may be more closely related to underlying genetic vulnerability. The use of quantitative measures rather than discrete phenotypes may increase the power of a study to detect linkage. The 2 neurobiologic dysfunctions associated with familial schizophrenia: impaired gating of the auditory evoked response (identified by measuring P50 of the auditory evoked potential) and ocular motor dysfunctions (e.g., smooth pursuit eye movement) are the 2 biologic endophenotypes being used for linkage studies.

Another important issue, overlooked in the past, but important in the current context of gene–environment interaction, is ascertainment and sequential extension of pedigrees studied. Initially, pedigrees were ascertained only on the basis of size and number of affected individuals (opportunistic ascertainment) in contrast to “systematic ascertainment.” The assumption then was that the informativeness for linkage was all that mattered, even if there was a nonreplicable sample frame. The most advanced and most demanding designs for combined linkage and association studies initiated to date are those by the NIMH Genetics Initiative on Schizophrenia. This design employs explicit rules for the systematic ascertainment and extension of pedigrees to permit both segregation and linkage analysis, replication, and interpretation of nonlinear genetic and environmental interaction. Future simulation studies are also being carried out to support proper linkage, because present simulation studies have shown that
linkage or association may not be detected because of incomplete penetrance, small samples and allele frequencies at disease and marker loci.

The current evidence is more compatible with “multigenic transmission” in schizophrenia. Certain studies postulate that genetic loci vary across different genetic populations, whereas others point out that in the absence of familial clinical subtypes it is quite possible that there are common multigenic mechanisms.\textsuperscript{190,191}

Risch and Teng\textsuperscript{151} conclude that when considering relative risk of schizophrenia in relatives of affected patients, the “Gottesman and Shields model of multigenic inheritance,” with 3–5 interacting loci, none of which increases the susceptibility by more than 2–3 times, fits best.

Future directions in this area include the following:

- Collection of sufficiently large samples and better methods of pedigree ascertainment to obtain power to detect linkage and then loci to chromosomal regions.
- Application of large-scale family-based association tests across the whole genome.
- Consideration of the interaction between different loci, because we know that linkage and association studies treat these loci independently of each other, and there is a suggestion that these loci might interact to produce other disorders. New methods have been developed that have the power to detect genes of small effect and interaction between loci.
- Potential combination of path, segregation and linkage analysis to increase the power of linkage and association studies to detect genes of small effect.
- Resolution of genetic heterogeneity requires further work on phenotypic classification to identify clinical characteristics that can delineate genetically distinct groups.
- Linkage analysis that is reported, for validity reasons, should be replicable by an independent group of investigators and should hold retain their validity even after sample augmentation (increase in families, subjects).
- Joint application of linkage and association studies as a 2-step strategy. Initial genome-wide scanning using linkage analysis should be followed by high-density association studies.
- Use of state-of-the-art techniques for molecular and quantitative analysis, including nonparametric methods for multipoint linkage, microchip DNA arrays and new genetic markers to identify the source of variance and newer mutations.

Conclusion

Several linkage and association studies have produced positive results, mostly at a suggestive level, for chromosomal regions 1q, 5p, 5q, 6p, 6q, 8p, 10p, 13q, 15q and 22q, but in each case there are reports that were either negative or researchers were unable to replicate the original findings. However, as with other complex diseases, this should not discourage researchers, especially given the novel methods and new statistical tools available. It is apparent that no single locus could cause anything other than a modest effect, though that might be the first to be identified, and that intensive efforts will be required to identify multiple genes of minor effect. At the same time, it is difficult to ignore the possibility that multiple nonlinear interactions between genes and environmental factors might cause such a heterogeneous disorder.

Competing interests: None declared.

References

10. Straub RE, MacLean CJ, O’Neill FA, Walsh D, Kendler KS.


117. Eastwood SL, McDonald B, Burnet PW, Beckwith JP, Kerwin RW, Harrison PJ. Decreased expression of mRNAs encoding non-NMDA glutamate receptors GluR1 and GluR2 in medial temporal lobe neurons in schizophrenia. *Brain Res Mol Brain Res* 1995;29:211-23.


122. Joo A, Shibata H, Ninomiya H, Kawasaki H, Tashiro N, Fuku-


