
Recent studies have supported the hypothesis based upon expectations from population genetics that the high heritability of schizophrenia reflects a combination of relatively common alleles of small effect and rare alleles some with relatively large effects. Genome-wide association studies have identified a number of risk loci at genome-wide levels of significance as well as evidence for a substantial burden of common risk loci. Moreover these recent findings suggest genetic overlap with bipolar disorder which has traditionally been assumed to be genetically distinct from schizophrenia. Genome-wide studies of at least one class of relatively uncommon variant, submicroscopic chromosomal abnormalities often referred to as copy number variations (CNVs), suggest that these confer high risk of schizophrenia. There is evidence both for an increased burden of large, rare CNVs in schizophrenia and that risk is conferred by a number of specific large CNVs as well as by deletions of \textit{NRXN1} which encodes the synaptic scaffolding protein neurexin 1. Many of these CNVs have been implicated in autism, mental retardation, epilepsy and other neurodevelopment disorders. These findings have implications for pathogenesis and nosology of schizophrenia and related disorders, and for future genetic studies.

**Conflict of interest**

All authors declare no conflict of interest.

Schizophrenia is a severe psychiatric disorder with an estimated lifetime prevalence of 1%. Its core features include hallucinations, delusions, altered emotional reactivity, cognitive impairment and disorganized behavior. It often runs a chronic course with many patients responding poorly to medication and suffering from frequent relapses. Consequently it imposes a significant burden on patients, their families and wider society.

Epidemiological, pharmacological and neurobiological studies have advanced our general understanding of the etiology of schizophrenia (1–3); it is now generally accepted that in many cases the origins of the disorder lie early in neurodevelopment and that synaptic dysfunction and altered neural connectivity are likely to be important in its pathogenesis. However this has not implicated a specific pathophysiology or novel treatment targets, and there are no laboratory tests to assist in diagnosis which thus remains clinical and syndromic (Fig. 1). There is therefore a pressing need to understand the pathogenesis of the disorder in order to improve diagnosis and treatment.

**Genetic epidemiology and genetic architecture**

Genetic epidemiological studies have consistently shown that schizophrenia is highly heritable, with heritability estimated at approximately 80% (4). Definitive statements about genetic architecture require the attribution of a high proportion of variance in risk to specific susceptibility alleles, a condition that has not been met for any complex disorder. Nevertheless, the data available from genetic epidemiology and molecular genetics provide the basis for developing rational research strategies and identifying those variants. For some years the dominant hypothesis driving common disease research has been the common-disease-common-variant (CDCV) hypothesis. This argues that genetic susceptibility derives from common alleles (minor allele frequency > 0.01). Justified by a combination of
A Characteristic symptoms
Two or more of the following, each present for a significant portion of time during a one-month period:
- Delusions
- Hallucinations
- Disorganized speech
- Grossly disorganized or catatonic behavior
- Negative symptoms (blunted affect, alogia or avolition).

Only one Criterion A symptom is required if delusions are bizarre or hallucinations consist of a voice keeping up a running commentary on the person's behaviour or thoughts, or two or more voices conversing with each other.

B Social/occupational dysfunction
Since the onset of the disturbance, one or more major areas of functioning, such as work, interpersonal relations, or self-care, are markedly below the level previously achieved.

C Duration
Continuous signs of the disturbance persist for at least six months. This six-month period must include at least one month of symptoms (or less if successfully treated) that meet Criterion A.

D Exclusion
Schizoaffective disorder and mood disorder with psychotic features.

E Substance/general medical condition exclusion
The disturbance is not due to the direct physiological effects of a substance (e.g., a drug of abuse, a medication) or a general medical condition.

F Relationship to a pervasive developmental disorder
If there is a history of autistic disorder or another pervasive development disorder, the diagnosis of schizophrenia is made only if prominent delusions or hallucinations are also present for at least a month (or less if successfully treated).

Fig. 1. Diagnostic and Statistical Manual of Mental Disorders (DSM IV) diagnostic criteria for schizophrenia.

population genetic considerations and pragmatics (5), its key prediction was that genome-wide association studies (GWAS) of common variants might be used to map loci contributing to common diseases. For many diseases, GWAS have been highly successful, but in general the effect sizes of robustly identified common alleles have been smaller than expected (odds ratios < 1.5), and, even for those disorders where a large number of risk alleles have been identified, only a small proportion of variance in risk has been explained. There has been much debate as to the reasons for this so-called missing heritability but essentially this is likely to reflect one or both of the following possibilities (6). Either the effect sizes of typical common risk alleles are too small to have been detected at stringent levels of significance by the studies to date, or the causal variants are not in strong linkage disequilibrium (LD) with the single nucleotide polymorphisms (SNPs) that have been genotyped, the most likely explanation for that being the causal variants are too rare. There is convincing support that for some phenotypes, e.g. height (6), and schizophrenia (7), the former explanation applies to around half of the genetic variance. But it is also clear that low frequency or rare alleles that are not well captured by GWAS also contribute substantially to common phenotypes (8–11).

It has been argued that as a consequence of selection pressures that derive from low fecundity in those affected, the relative contribution of low frequency alleles to common diseases is likely to be the strongest for disorders such as autism and schizophrenia (12, 13). Indeed some proponents of the common disease-rare variant (CDRV) model have suggested that genetic risk to these disorders might entirely be conferred by ultrarare highly penetrant variants, the disease being maintained in the population despite negative selection through new mutation. The debate between CDCV and CDRV hypotheses in regard to schizophrenia has been polarized, but as more empirical data have become available, a more nuanced view is emerging involving a spectrum of risk alleles, common and rare, effects small and large, but each allele contributing only a small fraction to the total population variance (13–17). Such a view is consistent with population genetic theory (5, 18, 19), and the so-called mixed model of schizophrenia (20); it has also received compelling support from recent genomic studies (21).

Early genetic and genomic studies
Work in psychiatric molecular genetics began with linkage studies of small numbers of multiplex families aimed at detecting high penetrance alleles and studies of functional candidate genes. It progressed to larger scale linkage studies and analysis of positional candidate genes as well as studies of microscopic chromosomal abnormalities. Thousands of functional candidate gene studies have been reported, but no strongly replicated findings have emerged (22). This no doubt reflects in part, not only our limited understanding of pathophysiology but also the fact that the great majority of studies have been underpowered to detect either common alleles of small effect or rare alleles of large effect. Positional studies appear to have been more successful; in that, a small number of promising
susceptibility genes has been implicated in multiple genetic studies (22, 23). Those for which the evidence is strongest are NRG1, DTNBPI, DAOA and DISC1. Data consistent with the involvement of these genes in schizophrenia has additionally come from studies of extended and intermediate phenotypes, principally neurocognitive and neuroimaging, and from analyses of gene expression and other aspects of neurobiology. However, in spite of this impressive body of evidence, in no case have specific risk alleles been unambiguously identified in population studies and in no case does the strength and consistency (same alleles or haplotypes across studies) of the genetic evidence equal that for genes now known to be involved in other complex disorders. Detailed discussion of early genetic studies and endophenotype approaches is beyond the scope of this review and readers are directed elsewhere for further background (24, 25).

Recent genomic studies

Recent advances in genomic technology have made possible a new wave of genetic studies of complex diseases. These have allowed researchers to conduct GWAS of large patient and control samples seeking evidence for the role of common risk variants and, genome-wide studies of at least one class of relatively uncommon variant, submicroscopic chromosomal abnormalities often referred to as copy number variations (CNVs).

Genome-wide association studies

To date, a number of loci have been implicated in GWAS of schizophrenia at stringent, genome-wide levels of statistical significance (Table 1). As with other common diseases the odds ratios are small and identified loci only explain a very small proportion of the genetic risk, but an analysis of weakly associated alleles (14) showed that approximately half of the genetic risk is attributable to a very large number of alleles (possibly thousands), many of which are common and of small effect. This finding has also been replicated in non-Europeans (26), and, importantly, in a sample of case-parent trios which excludes population stratification as an explanation (27).

Many of the risk alleles identified in GWAS of schizophrenia such as those at ZNF7804A, MHC and NRGN have also been implicated in bipolar disorder (BD) (31, 33). In addition risk alleles for BD such as those at PBRM1 and CACNA1C have also been implicated in schizophrenia (33, 36, 38). These findings suggest substantial overlap in the common risk alleles conferring susceptibility to the two disorders, a conclusion that has been supported by analysis of a large number of weakly associated alleles (14, 39) and by a recent, large family study (40).

Like schizophrenia, BD is a severe psychiatric disorder. It is characterized by mood instability with patients typically experiencing periods of elevated mood (mania or hypomania) and periods of low mood (depression). During manic episodes patients experience psychotic symptoms similar to those experienced by patients with schizophrenia. The distinction between schizophrenic psychosis and bipolar psychosis was first proposed by the psychiatrist Emil Kraepelin nearly 100 years ago. Clinical work and research in psychiatry has, for the most part, proceeded under the assumption that schizophrenia and BD are distinct clinical entities with separate underlying disease processes and treatments. However, in clinical practice, many patients have prominent mood and psychotic symptoms which do not fit neatly into current diagnostic categories. The recent evidence from GWAS and other studies calls into question the validity of the binary categorical diagnostic classification of these disorders and has implications for how they are approached from a scientific and clinical perspective (41, 42).

There are a number of other important implications of the findings from GWAS in schizophrenia and other disorders (21). First, the combined effect of common risk alleles, while substantial, will be insufficient to account for all genetic risk of schizophrenia and rare alleles are likely to play a role. Second, combined analysis

Table 1. SNPS associated with schizophrenia and BD in GWAS published to date

<table>
<thead>
<tr>
<th>Chromosome</th>
<th>Gene</th>
<th>SNP Location</th>
<th>Odds ratio</th>
<th>Reference(s)</th>
<th>Associated disorder(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>ZNF804A</td>
<td>1.08–1.38</td>
<td>1.14–1.17</td>
<td>(28–31)</td>
<td>Sz, BD</td>
</tr>
<tr>
<td>18</td>
<td>TCF4</td>
<td>1.20–1.30</td>
<td></td>
<td>(32)</td>
<td>Sz</td>
</tr>
<tr>
<td>11</td>
<td>NRGN</td>
<td>1.13–1.19</td>
<td>1.14</td>
<td>(32, 33)</td>
<td>Sz, BD</td>
</tr>
<tr>
<td>6</td>
<td>Extended region MHC</td>
<td>1.13–1.36</td>
<td>1.09–1.14</td>
<td>(14, 32–34)</td>
<td>Sz, BD</td>
</tr>
<tr>
<td>12</td>
<td>CACNA1C</td>
<td>1.13–1.15</td>
<td>1.18</td>
<td>(35, 36)</td>
<td>Sz, BD</td>
</tr>
<tr>
<td>10</td>
<td>ANK3</td>
<td>1.30</td>
<td>1.40</td>
<td>(35)</td>
<td>BD</td>
</tr>
<tr>
<td>3</td>
<td>PBRM1</td>
<td>1.14</td>
<td></td>
<td>(33, 37)</td>
<td>Sz, BD</td>
</tr>
</tbody>
</table>

SNPs, single nucleotide polymorphisms; GWAS, genome-wide association studies; Sz, schizophrenia; BD, bipolar disorder.
of schizophrenia GWAS samples likely to be available in the near future (~30,000 cases and ~40,000 controls) will reveal only a minority of common risk loci at genome-wide levels of significance, and these will explain only a small proportion of risk. Third, evidence from other common traits suggests that the identity of the genes implicated will point to possible disease mechanisms. In schizophrenia and BD, despite the relatively small number of disease loci identified, there are already examples of this. For example, the identification of risk alleles at \textit{CACNA1C} and \textit{NRGN} suggests the involvement of postsynaptic calcium influx, signaling and synaptic sensitivity, while unpublished data showing genome-wide significant association to \textit{mir137} and to four other genes predicted to be silenced by microRNA 137 (encoded by \textit{mir137}) strongly support the involvement of neuronal developmental functions regulated by that molecule (43).

**CNV studies**

Submicroscopic deletions and duplications of segments of DNA, known as CNVs, are important sources of individual genomic variation. CNVs can be detected by various platforms including those used for GWAS. In addition to their role in a number of uncommon syndromes, CNVs have been unequivocally implicated in common disorders and, in particular, neurodevelopmental disorders such as autism and mental retardation (MR) (21, 44). A number of studies have found that the burden of rare (<1%), large (>100 kb) CNVs is increased in schizophrenia (45–48). Moreover, the analysis of large GWAS datasets has allowed individual CNV loci to be identified as conferring risk with high degrees of statistical confidence. These loci include 1q21.1, 15q11.2, 15q13.3, 16p11.2, 16p13.1 and 22q11.2 (44) (Table 2). These are rare relative to associated SNPs (control frequencies typically <0.001) that cumulatively involve 2–3% of cases, and although not fully penetrant confer relatively large effects on an individual's disease risk, with estimated ORs for schizophrenia of 3–20 (49, 50) (Table 2). They also confer risk of a range of neurodevelopmental phenotypes including autism, MR, attention deficit hyperactivity disorder (ADHD) and idiopathic generalised epilepsy (15, 51). This suggests that schizophrenia might best be viewed as one type of phenotypic expression of disorder, with the form of neurodevelopmental phenotype taken dependent upon a combination of genetic (rare, common and polygenic)

In addition to their role in a number of uncommon syndromes, CNVs have been unequivocally implicated in common disorders and, in particular, neurodevelopmental disorders such as autism and mental retardation (MR) (21, 44). A number of studies have found that the burden of rare (<1%), large (>100 kb) CNVs is increased in schizophrenia (45–48). Moreover, the analysis of large GWAS datasets has allowed individual CNV loci to be identified as conferring risk with high degrees of statistical confidence. These loci include 1q21.1, 15q11.2, 15q13.3, 16p11.2, 16p13.1 and 22q11.2 (44) (Table 2). These are rare relative to associated SNPs (control frequencies typically <0.001) that cumulatively involve 2–3% of cases, and although not fully penetrant confer relatively large effects on an individual’s disease risk, with estimated ORs for schizophrenia of 3–20 (49, 50) (Table 2). They also confer risk of a range of neurodevelopmental phenotypes including autism, MR, attention deficit hyperactivity disorder (ADHD) and idiopathic generalised epilepsy (15, 51). This suggests that schizophrenia might best be viewed as one type of phenotypic expression of disorder, with the form of neurodevelopmental phenotype taken dependent upon a combination of genetic (rare, common and polygenic).

<table>
<thead>
<tr>
<th>Locus</th>
<th>Copy number change</th>
<th>Odds ratio</th>
<th>Reference(s)</th>
<th>Associated disorder(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1q21.1</td>
<td>Deletion Duplication</td>
<td>6.6–15.54</td>
<td>(45, 46, 48, 55–58)</td>
<td>Sz, MR, seizures, autism</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3.7–4.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2p16.3 (\textit{NRXN1})</td>
<td>Deletion Duplication</td>
<td>7.5–8.97</td>
<td>(45, 48, 56–63)</td>
<td>Sz, MR, seizures, autism</td>
</tr>
<tr>
<td>3p26.1</td>
<td>Deletion Duplication</td>
<td>NA</td>
<td>(45, 47)</td>
<td>Sz, MR, autism</td>
</tr>
<tr>
<td></td>
<td></td>
<td>NA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3q29</td>
<td>Deletion</td>
<td>17</td>
<td>(45, 57, 62, 64)</td>
<td>Sz, MR, autism</td>
</tr>
<tr>
<td>5p13.2</td>
<td>Deletion Duplication</td>
<td>NA</td>
<td>(45, 58)</td>
<td>Sz</td>
</tr>
<tr>
<td></td>
<td></td>
<td>NA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7q11.2</td>
<td>Duplication</td>
<td>NA</td>
<td>(45, 47)</td>
<td>Sz</td>
</tr>
<tr>
<td>7q22.1</td>
<td>Duplication</td>
<td>NA</td>
<td>(45, 65)</td>
<td>Sz</td>
</tr>
<tr>
<td>7q36.3</td>
<td>Deletion Duplication</td>
<td>NA</td>
<td>(47, 48, 57, 66)</td>
<td>Sz</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4–8.26</td>
<td></td>
<td></td>
</tr>
<tr>
<td>15q11.2</td>
<td>Deletion Duplication</td>
<td>1.94–2.8</td>
<td>(48, 55, 58, 62)</td>
<td>Sz, MR, autism</td>
</tr>
<tr>
<td>15q13.1</td>
<td>Duplication</td>
<td>NA</td>
<td>(59, 65)</td>
<td>Sz</td>
</tr>
<tr>
<td>15q13.3</td>
<td>Deletion Duplication</td>
<td>9.9–12.1</td>
<td>(46, 55, 57, 62)</td>
<td>Sz, MR, seizures</td>
</tr>
<tr>
<td></td>
<td></td>
<td>NA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>16p11.2</td>
<td>Duplication</td>
<td>8.3–26.3</td>
<td>(45, 63, 66–68)</td>
<td>Sz, MR, seizures, autism, ADHD, BD</td>
</tr>
<tr>
<td>16p13.1</td>
<td>Deletion Duplication</td>
<td>NA</td>
<td>(48, 56, 62, 69, 70)</td>
<td>Sz, MR, autism, ADHD</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2.98–3.27</td>
<td></td>
<td></td>
</tr>
<tr>
<td>17p12</td>
<td>Deletion</td>
<td>7.82</td>
<td>(48, 62)</td>
<td>Sz, HNPP</td>
</tr>
<tr>
<td>17q12</td>
<td>Deletion Duplication</td>
<td>∞</td>
<td>(62, 71)</td>
<td>Sz</td>
</tr>
<tr>
<td></td>
<td></td>
<td>NA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>22q11.2</td>
<td>Deletion</td>
<td>21.6–∞</td>
<td>(46–48, 55, 57, 58, 63, 72)</td>
<td>Sz, MR, autism, ADHD, OCD, anxiety, depression</td>
</tr>
</tbody>
</table>

Only replicated loci shown. Odds ratios are not available for all CNVs. In order to construct an odds ratio frequency in controls is required. This is often not available for rare variants.

NA, not available; Sz, schizophrenia; MR, mental retardation; ADHD, attention deficit hyperactivity disorder; BD, bipolar disorder; HNPP, hereditary neuropathy with liability to pressure palsies; OCD, obsessive compulsive disorder.
and environmental factors, and possibly chance (52). Interestingly, and in contrast to the findings with common alleles, there is little evidence that schizophrenia-associated CNVs confer risk for BD. Indeed, as a whole, CNVs are actually less common in BD than in controls, (53) a finding recently replicated by others (54).

To date, CNV studies have only been powered to detect relatively rare events with relatively large effect sizes, and the expectation must be that there will be a spectrum of risk alleles including even rarer CNVs with similar or larger effect sizes, and CNVs with smaller effect sizes that are more common as well as those which are both uncommon and of small effect which will be the most difficult to detect. It is also likely that the number of implicated CNVs will rise with the use of better resolution platforms and larger samples. While it seems likely that only a fraction of CNV susceptibility loci have been identified, it is important to ask to what extent findings till date can illuminate disease biology. Many associated CNVs span multiple genes and it is not immediately clear which are relevant to pathogenesis. An exception is NRXN1 deletion, which is robustly associated with schizophrenia (73, 74) (Table 2) and is found in other neurodevelopmental disorders. NRXN1 encodes the presynaptic neuronal cell adhesion molecule neurexin 1, pointing to the importance of abnormalities of synaptic function in schizophrenia, and to more specific hypotheses concerning the nature of those abnormalities (75). It has also been suggested that there is a general enrichment of genes implicated in synaptic function amongst those disrupted by schizophrenia CNVs (45, 76, 77), but these conclusions remain provisional pending the application of analytic approaches that control for a number of important confounders (78).

Conclusions

Recent advances in genomic technology, particularly the application of GWAS platforms, have led to significant new findings in schizophrenia. Positive results have been obtained from the study of both common and rare variants and indicate a spectrum of risk alleles of variable frequencies and effect sizes, but with each allele contributing only a small fraction to population variance (14, 15, 79).

Overall, GWAS in schizophrenia have lagged behind those in other common diseases, mainly because of the difficulty in amassing sufficiently large samples of adequately phenotyped cases. Studies to date have generally been based on a few thousand cases and controls and will therefore only have identified a small fraction of the loci that could be revealed by larger samples. Indeed comparison with GWAS of other complex diseases and traits suggest that schizophrenia is not atypical in regard to the number of robustly identified loci per size of sample studied and we can expect a greater number of such loci to be identified by larger studies; such studies will also identify the rank order of loci with greater reliability thus benefiting pathway analyses (80).

In addition, although some classes of CNVs are likely to be detected by current GWAS technology, other classes of CNV as well as other types of rare variants are not. As the prospect of affordable genome-wide resequencing becomes more realistic, interest is increasing with the possibility that rare single-base mutations and small insertions and deletions might play a significant role in common diseases. To date, next-generation sequencing studies in schizophrenia have been restricted to those of candidate genes and no such variants have been implicated in the disorder at robust levels of support.

Given the marked reduction in fecundity seen in schizophrenia and the consequent negative selection, mutations that confer a large effect on risk are expected to be rare in the population, and where they occur in multiple unrelated individuals, this will often be so through independent new mutations. A role for de novo CNVs in schizophrenia has been suggested (47), with an increased rate of de novo CNVs seen in sporadic cases. It is likely that other classes of de novo mutation including point mutations and small insertion and deletion mutations will play a significant role in schizophrenia and other neurodevelopmental disorders (5). The study of these highly enriched pathogenic mutations may aid the identification of specific disrupted genes and functionally related proteins that can provide insight into disease pathways. However, the relative contributions of rare inherited mutations, common variation and de novo events remain uncertain and will only be resolved empirically by further adequately powered studies.

From a clinical perspective, perhaps one of the most interesting findings from SNP and CNV studies of schizophrenia is the overlap of risk alleles with other psychiatric and neurodevelopmental disorders. There has been much recent discussion regarding the validity of current categorical classification systems in psychiatry (42) and efforts to improve these systems are in progress (81) (82). The diagnosis of many psychiatric disorders including schizophrenia and BD relies entirely upon clinical features with no validating biomarkers or diagnostic tests. There is therefore no compelling reason to believe that these diagnostic categories reflect discrete pathophysiological processes. GWAS of SNPs in schizophrenia and BD convincingly show overlap in the genetic risk of these disorders, calling into question the dichotomous view of functional psychosis. Furthermore, the finding that CNVs confer susceptibility to a range of neurodevelopmental phenotypes including autism, MR, ADHD and idiopathic generalized epilepsy, suggests that schizophrenia might best be viewed as one type of phenotypic expression of neurodevelopmental disorder, with the form of phenotype dependent upon a combination of genetic (rare, common and polygenic) and environmental factors, and possibly chance (83). It is hoped that future genetic research will translate into improved understanding of pathogenic mechanisms, improved disorder classification and the development of new interventions to benefit patients, their families and wider society.
Doherty et al.

Acknowledgements

This work was supported by an MRC Programme Grant (G0800509) and an MRC Centre Grant (G0801418).

References


40. Walsh T, McClellan JM, McCarthy SE et al. Rare structural variants disrupt multiple genes in neurodevelopmental pathways in schizophrenia. Science 2008: 320: 539–543.


Recent genomic advances in schizophrenia