New discoveries in schizophrenia genetics reveal neurobiological pathways: A review of recent findings

Alex V. Kotlar, Emory University
Kristina B. Mercer, Emory University
Michael Zwick, Emory University
Jennifer Mulle, Emory University

Journal Title: European Journal of Medical Genetics
Volume: Volume 58, Number 12
Publisher: Elsevier | 2015-12-01, Pages 704-714
Type of Work: Article | Post-print: After Peer Review
Publisher DOI: 10.1016/j.ejmg.2015.10.008
Permanent URL: https://pid.emory.edu/ark:/25593/rwj2g

Final published version: http://dx.doi.org/10.1016/j.ejmg.2015.10.008

Copyright information:
© 2015 Elsevier Masson SAS. All rights reserved.
This is an Open Access work distributed under the terms of the Creative Commons Attribution 4.0 International License (https://creativecommons.org/licenses/by/4.0/).
New discoveries in schizophrenia genetics reveal neurobiological pathways: a review of recent findings

Alex V. Kotlar\textsuperscript{a}, Kristina B. Mercer\textsuperscript{b}, Michael E. Zwick\textsuperscript{c,d}, and Jennifer G. Mulle\textsuperscript{b,c}

\textsuperscript{a}Emory graduate program in Genetics and Molecular Biology, Emory University, Atlanta GA
\textsuperscript{b}Department of Epidemiology, Rollins School of Public Health, Emory University, Atlanta GA
\textsuperscript{c}Department of Human Genetics, Emory University School of Medicine, Atlanta GA
\textsuperscript{d}Department of Pediatrics, Emory University School of Medicine, Atlanta GA

Abstract

Schizophrenia research has undergone a recent transformation. By leveraging large sample sizes, genome-wide association studies of common genetic variants have approximately tripled the number of candidate genetic loci. Rare variant studies have identified copy number variants that are schizophrenia risk loci. Among these, the 3q29 microdeletion is now known to be the single largest schizophrenia risk factor. Next-generation sequencing studies are increasingly used for rare variant association testing, and have already facilitated identification of large effect alleles. Collectively, recent findings implicate voltage-gated calcium channel and cytoskeletal pathways in the pathogenesis of schizophrenia. Taken together, these results suggest the possibility of imminent breakthroughs in the molecular understanding of schizophrenia.

Keywords

Schizophrenia Genetics; Psychiatric Genetics; copy number variation; GWAS

Introduction

Schizophrenia (SZ) is a severe psychiatric disorder with a prevalence of approximately 0.5–1\% [1]. Symptoms of SZ are classified into positive and negative categories. Positive symptoms include delusions, hallucinations, disorganized speech, and disorganized or catatonic behavior, while alogia (lack of speech), avolition (lowered motivation), and emotional blunting encompass negative symptoms [2]. Schizophrenia is diagnosed based on the observation of at least one core symptom (delusions, hallucinations, or disorganized speech), and at least one additional symptom (a second core symptom, grossly disorganized behavior, or negative symptoms) over a 6-month period in which the individual is disturbed...
for the majority of a 1-month block. Males have typical onset in their early 20s and females in their late 20s [3].

Established SZ risk factors include paternal age [4], maternal parity [5], obstetric complications [6] including low birth weight for gestational age [5, 6], season of birth [7–10], cannabis use [11], infection by *Toxoplasma gondii* [12], and urban birth [13]. Evidence of a genetic contribution to risk has accumulated since the 1930's [14]. Initial efforts to establish the relative genetic contribution relied on adoption and fostering study designs, which analyzed phenotypic concordance of parents with offspring that are reared apart. In 1966, Heston reported that children born to affected mothers, when adopted and fostered by different parents, were at increased risk of schizophrenia [15]. Further evidence of a genetic contribution to SZ risk was provided by Kety et al's elegant adoption studies in 1968 and 1976 [16, 17]. Data from twin studies, which analyze the phenotypic concordance between monozygotic and dizygotic twins, suggested that genetic factors account for approximately 80% of the total variance (additive genetic variance: $a^2 = 0.82$, C.I=0.71–0.90) [18, 19] although estimates have varied widely [19]. For instance, a 2009 analysis of 9,009,202 individuals across approximately 2 million families concluded that the variance due to additive genetic effects was approximately 64% (C.I. 61.5% – 72.2%) [20]. Thus, it was revealed that genetic factors play a significant role in SZ susceptibility. Furthermore, these findings suggested that genetic studies could help obtain a greater mechanistic understanding of SZ.

While family and twin studies led to the proposition that schizophrenia was under the influence of a “major gene” [14, 21, 22], observations of departures from simple transmission models led to the consideration of polygenic inheritance models [16, 23]. Initial attempts to map SZ loci using linkage analysis had mixed results, and while there were notable successes, including the identification of DISC1 [24] and NRG1 [25], many linkage peaks failed to replicate [26]. These findings led to the conclusion that the genetic architecture of SZ is incredibly heterogeneous. Thus, in order for gene discovery to be a viable strategy for understanding SZ, both larger sample sizes and alternative experimental strategies were needed.

In the last several years, researchers have combined larger sample sizes with methods including whole-exome sequencing (in the analysis of rare variants), copy number variant analysis strategies, and genome wide association studies (GWAS). This work has led to unprecedented success in expanding the understanding of SZ genetic risk factors, and for the first time in the history of SZ genetics, there are now genetic variants that are considered true risk loci. These established alleles include more than 128 common single nucleotide polymorphisms (SNPs), and at least fifteen rare copy number variants (CNVs) (Figure 1). Together, recent studies illuminate several plausible genetic pathways, and support a polygenic model of SZ, in which genetic risk is largely due to either the presence of a very rare, large effect copy number variant (with odds ratios as high as of 41 as seen in the case of the 3q29 deletion [27]), or by the coincident inheritance of many small effect alleles (Figure 2). This review will focus on this very recent history of schizophrenia genetics, including new findings and the biological hypotheses that these findings have inspired.
Copy number variants

In the past decade the success of research focused on the genetics of schizophrenia has advanced remarkably, providing evidence for the association of copy number variants (CNVs) in risk for disease [28, 29]. CNVs, which are defined as gains or losses of genomic material of at least 1 kilobase (kb) in size, can encompass a single exon of a gene, an entire gene, or even multiple genes. Very large CNVs are typically rare in the genome [30]. When comparing SZ cases and unaffected controls, it has been repeatedly demonstrated that case samples have an increased genomic burden of duplications and deletions sized 100kb and larger [31, 32] and this enrichment increases as CNV size increases. Large studies combining case/control cohorts have identified several recurrent loci that are associated with SZ. Replication of the findings across several populations adds credence to the validity of these data and strongly suggests a functional role for these CNVs in the pathophysiology of schizophrenia. There are now at least 15 distinct genomic regions where dosage changes of varying lengths (120kb to 9Mb) are associated with risk for SZ; the increased odds for disease range from 2 to > 40 [27, 33] (Figure 2). Many of these same CNVs have also been implicated in risk for autism, suggesting shared neurological pathways between the psychiatric disorders [34, 35]. In fact, Walsh et al report an enrichment of genes involved in brain development within SZ-associated CNV intervals [32]. Other work has identified abnormalities in signaling complexes, brain structure, and conductivity among carriers of CNV risk loci [36–38]. Further research into the biological pathways disrupted by these CNVs will allow us to better understand the mechanisms responsible for the progression of neuropsychiatric disease such as SZ.

1q21 Deletion and Duplication

The 1.35Mb 1q21.1 deletion was among the first CNV to be associated with SZ, with an odds ratio of 14.83 [39]. It was identified among a combined cohort of 4,718 cases and 41,201 controls from Iceland, United Kingdom, Germany, Finland, Italy, Denmark, Norway, The Netherlands and China. The association at 1q21.1 was supported by a second study published by the International Schizophrenia Consortium (ISC) among a sample comprised of similar ancestry [31]. The reciprocal duplication has also been associated with SZ (odds ratio (OR) = 3.7–4.5) [40], suggesting that any dosage imbalance of these genes confers risk for SZ. Among the many genes within this copy number variable region (approx. 8), the gene for connexin 50, GJA8, has been reported to associate with SZ [41]. Gene ontology and network analysis also supports this association [42].

2p16.3/NRXN1 Deletion

Deletions at 2p16.3 (variable sizes of <1 Mb) that encompass a partial or full-length region of the neurodevelopment and synaptic adhesion gene NRXN1 are associated with SZ [43, 44]. In particular, deletions that span exons of NRXN1 are thought to be responsible for the increase in risk (OR= 8.97, p=0.0027) [45]. The deletion breakpoints are highly variable, and have enabled further investigation into the specific gene regions associated with SZ outcome. Clinical outcome appears to be related to the deletion of specific exons in NRXN1, which have been observed among SZ cases across various studies [46].
3q29 Deletion

The 3q29 deletion (0.84–1.6 kb) has one of the largest effect sizes of any SZ risk factor [27, 47]. This 1.6 Mb CNV contains 22 protein-coding genes, among which PAK2, DLG1 and FBXO45 have been proposed as candidate schizophrenia susceptibility loci [27]. The initial association was detected in 2010, with a 16.98 odds ratio [47]. This finding has subsequently been replicated in two independent studies [40, 48]. As additional cases are ascertained, the strength of association continues to be revised upward. The most recent meta-analysis of 25,904 SZ cases and 62,871 controls found an odds ratio of 41.1 (p= 5.8 × 10−8, 95% C.I 5.6–1953.6) [27].

It is yet to be determined which genes in the 3q29 interval are causative. The largest GWAS to date did not find individually associated 3q29 single nucleotide variants [49]. In order to detect association with such infrequent alleles it is useful to consider mutation groups, rather than individual nucleotide associations, and the first study that used this strategy found that 3q29 genes as a group represented the only significant CNV gene set [50]. Although these mutations seemed to be concentrated in the DLG1 gene, with 5 out of 2,536 cases carrying DLG1 variants compared to none among 2,543 controls, this result was not significant after multiple hypothesis correction. Implicating individual genes in the 3q29 interval will require larger samples sizes. In order to understand the mechanisms by which the 3q29 deletion increases SZ risk, biological studies are also necessary. Such investigations are already underway, facilitated by the use of an online registry which has ascertained over 58 carriers (3q29deletion.org).

7q36.3/VIPR2 Duplication

Variably sized microduplications at 7q36.3, spanning a region of 362 kb located upstream and across the coding region of the vasoactive intestinal peptide receptor gene (VIPR2), are associated with SZ [51]. Many of these duplications result in an increase in both VIPR2 transcript and intracellular cAMP within lymphocytes harboring the 7q36.3 variant. VIPR2 encodes for the VPAC2 protein, which is a G-protein coupled receptor that, when activated by its receptor ligand, initiates neuronal signaling cascades [52]. This finding, originally identified using a study cohort of Caucasian SZ cases (OR= 14.1), was replicated among a study cohort of Chinese cases and controls (OR= infinity, 95% CI = 1.327–infinity) [53]. Levinson et al also replicated the association, specifically for the VIPR2 exonic duplications (OR=4.0, p=0.002) [40]. Despite the strong evidence for 7q36.3 duplications in increased susceptibility for SZ, data from a large case/control provided by Rees et al do not support the original findings [33]. Unlike many of the other CNVs associated with SZ, the 7q36.3 duplications are highly variable in breakpoints and consequently, size. Thus replication may depend on which specific duplications are used in association tests.

7q11.23 Duplication

Mulle et al were the first to report an association between a 7q11.23 duplication and SZ [54]. A meta-analysis of 14,387 SZ cases and 28,139 control subjects resulted in a Mantel-Haenszel corrected OR of 10.78. This duplication has also been associated with autism [55, 56]. The reciprocal 1.4–1.5 Mb deletion causes Williams–Beuren Syndrome (WBS) which has been clinically well characterized and results in symptoms that include facial
dysmorphisms, short stature, cardiac defects, and cognitive impairments [57]. Individuals with WBS are described as having a “cocktail party personality,” where they are highly social and their verbal abilities are preserved despite cognitive impairments and developmental delay. This highly social phenotype is contrasted to the social withdrawal seen in ASD and SZ [54]. This recent finding awaits replication in another large cohort.

15q11.2 and 15q13.3 Deletions and 15q11-q13 Duplication

Deletions at 15q11.2 (470 kb; OR= 2.73) and 15q13.3 (1.58 Mb; OR= 11.54), like the 1q21.1 deletion, were among the first CNVs identified as SZ risk factors in 2008 [31, 39]. Both these findings have since been replicated and are now well-established SZ susceptibility loci [33, 40, 58–60]. An atypical 129kb duplication at 15q11.2, containing a single gene (UBE3A) involved in synapse development has also been identified in a single SZ case [61, 62].

A maternally derived duplication (> 5 Mb) at 15q11-q13, associated with Prader-Willi syndrome, has been reported with nominal significance (p=0.01; OR = 7.3) in several schizophrenic and schizoaffective individuals [63]. More SZ cases (N=2) harboring the 15q11-q13 duplication were reported in a second study, although no measure of statistical significance was reported [64]. An odds ratio of 13.2 has been predicted for this CNV duplication based on combined data from approximately 15,000 cases [33].

16p11.2 and 16p12.1 Deletion, 16p13.1 Deletion and Duplication

SZ associated CNVs along three regions of chromosome 16 have been identified: 16p11.2, 16p12.1, and 16p13.1. Microduplications (but not deletions) at 16p11.2, spanning a region of 600kb and containing 28 genes (29.56–20.11 Mb, hg19) associate with SZ among several thousand cases and controls (OR = 8.4) [65]. At this exact locus, both deletions and duplications are also associated with ASD [66], highlighting the overlapping loci associated with SZ and ASD also observed for at least six other CNVs [34].

A different deletion within the 16p11.2 chromosomal region (28.73 – 28.95 Mb, hg19), located distal to the previously described CNV, has also been identified in two independent cohorts of Ashkenazi Jewish and Bulgarian descent, and has been replicated in a large combined cohort from the US, Europe and Japan (all samples, OR = 6.25) [67]. Rees et al also report a significant finding for this CNV in their large, combined cohort revealing an odds ratio of 3.39 [33].

Among the newest of the SZ associated CNVs, a 480 kb deletion at 16p12.1 was identified in 2014 in a study cohort of approximately 7,000 total SZ cases [68]. This finding was replicated using a separate cohort of nearly 15,000 cases (reported in the same study). This deletion, comprised of 7 genes, is associated with a 2.72-increased risk of developing SZ. Although the data are convincing, this is the only study to date that has reported this finding.

Deletions and duplications within a genomic region at 16p13.1, which had previously been associated with autism and mental retardation [69, 70], were examined for association with SZ among 4,345 affected individuals and 35,079 controls from eight European populations [71]. Several distinct copy number variants exist within this region, and each was tested for
association with SZ. Carriers of duplications located in a 1.5 Mb region of chromosome 16 (approximate region = chr16: 15.1–16.6 Mb, hg19) strongly associate with risk for SZ with an OR of 7.27 [71]. This region contains the genes *NTAN1* (involved in memory) [72] and *NDE1* (involved in neurogenesis) [73]. Notably, for the latter, a rare single nucleotide variant (chr16: 15785118) associates with SZ (p=0.039), strengthening a role for *NDE1* in the development of the disease [74]. Support for the association of a 16p13.1 CNV duplication and SZ is provided by two other studies [58, 75]. Most recently, Rees et al report an association with an odds ratio of 2.3 among a total of 12,029 cases and 69,289 [33].

### 17p12 Deletion, 17q12 Deletion and Duplication

The genomic region 17p12 was first identified in a linkage study of families with SZ affected relatives, but only provided results suggestive of a relationship with SZ [76]. Additional evidence was provided several years later by Kirov et al, where the data showed a 1.3 Mb deletion among several SZ cases; one of which had a parent and three siblings with SZ. This CNV was found to be statistically significant, with an odds ratio of 10, in 5,300 cases and 39,000 controls across two independent cohorts [58]. Another chromosome 17 deletion at 17q12, which is associated with intellectual disability and autism, was observed among schizophrenic study participants from the GAIN and the SGENE Consortium. This 1.4 deletion was not observed among 52,448 healthy controls and has a predicted OR of 4.49 [77]. Rees et al support both deletion findings [33]. A recent study by Szatkiewicz et al also found evidence for an association between SZ and a duplication at 17q12 (OR=4.16, P=0.018) [48].

### 22q11.2 Deletion

The 22q11.2 deletion, which is an established cause of DiGeorge/VCFS syndrome, was known to be associated with SZ as early as 1994 [78]. The deletion is 3Mb (~65% of carriers) or 1.6 Mb (~35% of carriers) and has one of the largest effect sizes for schizophrenia, with an odds ratio of 28 or greater [33]. The frequency of the deletion in the general population is estimated to be 1 in 4,000, but in SZ it is 1 in 100 [79]. Intriguingly, recent evidence suggests that a duplication within the same general region (22q11.2) provides protection against the development of schizophrenia [80]. It has also been noted that among 22q11.2 deletion carriers, cognitive decline in childhood may signal the beginning of the prodromal period [81].

### Rare variant and genome wide association studies

While copy number variant analyses have provided tremendous insight into the genetic architecture of schizophrenia, the role of other classes of genetic variation in SZ risk were also studied. In recent years, two strategies have been used to substantially increase the understanding of the impact that single nucleotide variants have on SZ risk. These strategies differ fundamentally in their approaches to detecting association. Rare variant methods, like Sequence Kernel Association Test (SKAT) [82], associate groups of variants at a genetic locus with a phenotypic outcome. The quality of the underlying sequencing annotation data, generated by tools like PLINK and Seqant [83, 84], may increase the effectiveness of rare variant methods. In contrast, GWAS determine the strength of independent associations
between common variants and disease, and is most powerful when common variants have large effect sizes. For heterogeneous disorders like schizophrenia, both study designs require thousands of samples to detect robust genetic signals.

**Next generation sequencing for rare variants**

In 2014, two complimentary whole-exome sequencing studies analyzed rare variants in SZ cases and controls. Fromer et al looked at the distribution of non-synonymous variants in 623 affected trios (SZ & schizoaffective). In single gene analysis, only the gene TAF13 (encoding transcription initiation factor TFIID subunit 13) contained an exome-wide significant recurrence of loss of function mutations ($p = 10^{-16}$) [85]. A gene set analysis found significant enrichment in the n-methyl d-aspartate receptor (NMDAR) complex, the actin-mediated cytoskeleton protein (ARC) complex, and Fragile X mental retardation protein (FMRP) target groups [85]. In a second exome sequencing study, Purcell et al reported similar findings among 2,536 SZ case and 2,543 controls. No individual genes were significant under burden or SKAT analysis [50]. However, gene set analysis (conducted by summing individual burden scores) yielded significant enrichments of disruptive rare variants in ARC, postsynaptic density protein 95 (NLG4/PSD-95) complex, NMDAR complex, and voltage gated calcium channel (VGCC) groups [50]. The common findings of rare deleterious variants in both the ARC and NMDAR gene sets between these two studies is highly encouraging. However, these results also suggest that future rare variant analysis will need to use even larger sample sizes to identify individually associated genes.

**Genome wide association studies**

While rare variant analyses have made important contributions, in recent years the bulk of SZ genetic variant discovery has been driven by genome wide association studies. However, the utility of this design has only become apparent as sample sizes have increased. In 2008, the first GWAS of SZ, conducted across 1,471 samples, reported no genome-wide significant loci [86]. This study was powered to detect 80% of true associations, for single nucleotide polymorphisms (SNPs) of at least 10% minor allele frequency (MAF) and 1.83 effect size (genotype relative risk). Their negative results suggested that common variants in SZ did not have particularly large effect sizes, at least relative to copy number variants (Figure 2). A somewhat larger study in the same year identified a putative association with single nucleotide polymorphism (SNP) rs1344706, located in the zinc finger protein gene ZNF804A [87], which has since been independently replicated [88]. The first major results came in 2009, when the International Schizophrenia Consortium (ISC) combined three cohorts to analyze 8,014 cases and 19,080 controls of European descent. The ISC found robust support for a polygenic model of schizophrenia, suggesting that at least 1/3rd of SZ genetic variance could be explained by common variants of individually small impact [89]. They also found suggestive evidence of an association with VGCC genes, which were previously reported in connection with bipolar disorder (CACNA1C, $p = 7.7 \times 10^{-6}$) [90]. VGCC genes, such as CACNA1C, are now considered leading candidate SZ risk loci [49, 50]. Two companion studies reported 7 SNPs in the MHC (major histocompatability complex), TCF4 (Transcription Factor 4) and NRGN (Neurogranin) [91], as well as common variants located at 6p22.1[92].
These early successes drove an explosion in SZ GWAS sample sizes. In 2011, a GWAS of 21,856 discovery samples, and 29,839 replication samples were utilized to uncover 5 new loci [93]. Among the genome-wide significant loci were CACNA1C, previously identified by the ISC, and MIR137 (microRNA 137), which is a regulator of neuronal development. A 2013 GWAS conducted across a similar number of samples reported 13 new loci, as well as the replication of several others, including CACNA1C, and MIR137 [94]. Finally, in 2014, a tour-de-force GWAS of 36,989 cases and 113,075 controls reported 128 genome-wide significant SNPs in 108 loci, 83 of which were new [49]. Their polygenic risk analysis suggests that these 128 SNPs explain approximately 3.4% of SZ genetic variance [49]. Many more loci may await discovery in GWAS that interrogate larger numbers of samples. However, by now it is clear that the individual effect sizes of these undiscovered loci are unlikely to approach those seen in copy number analysis.

**Biological pathways**

Findings from genome wide association studies and exome sequencing studies are converging on genes in the ARC signaling complex, NMDAR complex, VGCC, and FMRP target pathways. Considered a potential “master regulator of synaptic plasticity” [95], ARC is a highly-conserved gene that links neuronal stimulation to synaptic remodeling. ARC is transcribed within neuronal soma in response to synaptic activation [96], and the resulting mRNA is rapidly trafficked to activated dendritic spines, where it undergoes local translation [97]. ARC is critical in the maintenance, but not induction, of long-term potentiation (LTP), which is the persistent strengthening of synapses, long-term depression (LTD), which is the persistent weakening of synapses, and long-term (but not short term) memory consolidation [98]. The stimulus-dependent elevation of ARC expression is in part mediated by the activity of NMDAR and VGCC [99]. At least 28 genes are known to interact with ARC, and form the basis of the “ARC complex” groups tested in the recent exome-sequencing experiments of Fromer et al and Purcell et al [36, 50, 85].

NMDAR are voltage-dependent ligand-binding receptors that mediate excitatory signaling, which plays a crucial role in LTP [100], LTD [101], and both short and long term memory consolidation [102–104]. These receptors allow for the influx of calcium ions in response to binding of glutamate and n-methyl d-aspartate (NMDA), as well as the co-agonist glycine [105]. They are found in a wide range of cell types, including, but not limited to, neurons, astrocytes, and oligodendrocytes [106–109]. In the last 40 years, the hypothesis of NMDAR hypofunction in SZ has gained significant attention. NMDAR hypofunction is seen in the dorsolateral prefrontal cortex and hippocampus of post-mortem brains collected from SZ-diagnosed individuals [110–113]. Unfortunately, drugs targeting NMDAR have shown inconclusive efficacy in clinical trials [114, 115], which may reflect the underlying genetic heterogeneity of the disorder.

Evidence surrounding VGCC involvement in SZ is relatively new: several individual VGCC genes, such as CACNA1C have been replicated in independent GWAS [49, 93]. Furthermore, a VGCC gene set was implicated in a recent whole-exome sequencing experiment [50]. Like NMDAR, VGCC are also critical to cell excitability across a range of cell types, including neurons [116]. They play key roles in LTP and long-term (but not
short-term) memory consolidation [104]. Functional studies of rare variants may clarify the role that VGCC plays in SZ occurrence.

Fragile X mental retardation protein (FMRP) is an mRNA-targeting protein, encoded by FMR1 [117]. FMRP impacts LTP and LTD, by regulating translation of a key set of mRNAs. These mRNAs have been identified and make up the gene set of “FMRP targets” [118]. FMRP may also interact directly with some neurologically active proteins. For instance, knockdown experiments have suggested that FMRP may interact with N-type calcium channel Cav2.2 [119] and repress NMDAR subunit translation [120]. Post-mortem ex-vivo studies have reported significantly lower FMRP levels in the cerebella of schizophrenic patients when compared with controls [121]. However, more research is needed to conclusively establish the link.

### Overlap with other psychiatric disorders

Targeted sequencing, GWAS, and copy number variant analyses support the possibility that genetic risk factors for schizophrenia overlap with those associated with autism spectrum disorders (ASD) and intellectual disability [122]. Fromer et al, in their 2014 whole exome sequencing experiment, found statistically significant support for concordance with ASD and intellectual disability loss-of-function de-novo gene sets. Genes in the intersection include SCN2A, which encodes a voltage gated sodium channel subunit that is critical in the generation of action potentials (a key step in neuronal signal propagation) and which has been implicated in both ASD and epilepsy [123–126]. Enrichment was also found in NMDAR complex genes, which have been implicated in ASD [127] and ID [128]. In the complimentary case/control whole-exome sequencing study, Purcell et al reported an enrichment of rare mutations in a 738 member FMRP target gene set [50]. Loss of function mutations in FMR1 cause intellectual disability (Fragile X), and are a leading monogenic cause of autism [129]. Furthermore, several FMRP targets overlap with candidate autism risk loci [118]. The VGCC complex gene set was also significantly enriched. Interestingly, both of these rare-variants studies found significant enrichment in ARC complex genes, which interact with NMDAR and VGCC in learning and memory processes, and are implicated in ASD [99, 130].

The largest GWAS of SZ to date reported similar associations. Of note, VGCC genes, including CACNA1C, have now been replicated in several independent GWAS experiments, were enriched in the rare-variant analyses conducted by Purcell et al [49, 93] and represent a leading autism susceptibility pathway [131]. Furthermore, mutations in calcium channel genes have recently been reported in connection with ASD, SZ, attention deficit-hyperactivity disorder (ADD), bipolar disorder (BP), and major depressive disorder (MDD) [132]. Other notable associations include the NMDAR gene, and GRIN2A.

Several SZ-associated copy number variants are also implicated in ASD and ID. Exonic deletions of NRXN1 have been associated with risk for speech delay, mild dysmorphic features, epilepsy and autism spectrum disorders [46]. The 3q29 deletion is associated with mild to moderate intellectual disability and ASD [133]. The 7q11.23 duplication is also associated with ASD [55], while the reciprocal deletion gives rise to Williams-Beuren
Syndrome, where the phenotype includes intellectual disability [134]. Several other CNVs have been connected with ASD and/or ID, including deletions in 15q11.2 [135], 15q13.3 [136], 22q11.2 [137], and duplications of 1q21.1 [138]. Future studies, conducted across a larger number of samples, will be needed to delineate the genetic boundaries between SZ, ASD, and intellectual disability.

Future Directions

The estimated heritability of SZ ranges widely, with some estimates as high as 80–88% [18, 19]. At least fifteen rare CNVs, with effect sizes ranging from approximately 2–40, and more than 128 common, small effect-size single nucleotide polymorphisms have been identified as genetic risk factors in schizophrenia (Figure 1). The common variants uncovered by the largest SZ GWAS to date are estimated to explain 3.4% of SZ variance, a fraction of the estimated SZ heritability. Although this may seem like a paucity of the total variance, it is important to note how rapidly the field is advancing: almost all of the common variants now known were discovered in the last four years. This accelerating progress has been driven by the use of larger sample sizes, and suggests that even larger cohorts will lead to a windfall of discovery. Researchers are now well positioned to lead a breakthrough in the understanding of SZ biology, through the combination of genetic variant statistical analysis and functional investigation.

The exciting results discussed here must be translated into biologically meaningful information. This will be accomplished by functional studies of SZ-associated variants, particularly those rare variants with the largest effect sizes. We now have early evidence that many variants affect shared pathways, including voltage gated calcium channel (VGCC), n-methyl d-aspartate receptor complex, activity-regulated cytoskeleton protein (ARC) complex, and FMRP target genes. The study of biological perturbations caused by severe, gene-disrupting mutation may help clarify the means by which SZ presents in the clinic, and in combination with phenotypic characterization of patients carrying specific rare variants, we may be able to transform our understanding of SZ, from a highly heterogeneous disorder, to one that impacts a tractable number of “master systems”. These findings may ultimately be leveraged to offer personalized medicine, refine treatment options, and improve the outcome of individuals with SZ.

Acknowledgements

Thanks to David Cutler for manuscript suggestions. This work was supported by NIH grant MH100917(JGM).

References


[14]. Kallmann, FJ. DOI. 1938. The genetics of schizophrenia.


[24]. Blackwood D, Fordyce A, Walker M, Clair DS, Porteous D, Muir W. Schizophrenia and affective disorders— cosegregation with a translocation at chromosome 4q2 that directly


[27]. Mulle, JG. DOI. 2015. The 3q29 deletion confers greater than 40-fold increase in risk for schizophrenia, Molecular Psychiatry.


[83]. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de Bakker PI, Daly MJ, Sham PC. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet. 2007; 81:559–575. [PubMed: 17701901]


Figure 1. Locations of SZ-associated alleles
Chromosomal locations of SZ-associated variants: copy number variants [33, 51] (red squares), recurrent de-novo mutations and rare variants identified by NGS [50, 85] (blue circles), and GWAS loci [49] (green triangles).
Figure 2. A comparison of significant allele odds ratios and population frequencies
Shown is a comparison of odds ratio (y-axis) and population frequency (log scale; x-axis) for selected GWAS loci (left side of center y-axis) and CNV (right side of center y-axis) associated with SZ. GWAS SNPs from [49] were ranked by odds ratio, and a representative SNP from each quartile was selected for display here.