Unraveling the genetic architecture of copy number variants associated with schizophrenia and other neuropsychiatric disorders

Roman Rutkowski, Emory University
Jason Schroeder, Emory University
Georgette Gafford, Emory University
Stephen Warren, Emory University
David Weinshenker, Emory University
Tamara Caspary, Emory University
Jennifer Mulle, Emory University

Journal Title: Journal of Neuroscience Research
Volume: Volume 95, Number 5
Publisher: Wiley: 12 months | 2017-05-01, Pages 1144-1160
Type of Work: Article | Post-print: After Peer Review
Publisher DOI: 10.1002/jnr.23970
Permanent URL: https://pid.emory.edu/ark:/25593/s6973

Final published version: http://dx.doi.org/10.1002/jnr.23970

Copyright information:
© 2016 Wiley Periodicals, Inc.

Accessed April 26, 2019 12:16 PM EDT
Unraveling the genetic architecture of copy number variants associated with schizophrenia and other neuropsychiatric disorders

Timothy P. Rutkowski¹, Jason P. Schroeder¹, Georgette M. Gafford¹, Stephen T. Warren¹, David Weinshenker¹, Tamara Caspary¹, and Jennifer G. Mulle¹,²

¹ Department of Human Genetics, Emory University School of Medicine, Atlanta, GA 30322, USA.
² Department of Epidemiology, Rollins School of Public Health, Emory University, Atlanta, GA 30322, USA.

Abstract

Recent studies show that the complex genetic architecture of schizophrenia (SZ) is driven in part by polygenic components, or the cumulative effect of variants of small effect in many genes, as well as rare single-locus variants with large effect sizes. Here we discuss genetic aberrations known as copy number variants (CNVs), which fall in the latter category and are associated with a high risk for SZ and other neuropsychiatric disorders. We briefly review recurrent CNVs associated with SZ, and then highlight one CNV in particular, a recurrent 1.6-Mb deletion on chromosome 3q29, which is estimated to confer a 40-fold increased risk for SZ. Additionally, we describe the use of genetic mouse models, behavioral tools, and patient-derived induced pluripotent stem cells as a means to study CNVs in the hope of gaining mechanistic insight into their respective disorders. Taken together, the genomic data connecting CNVs with a multitude of human neuropsychiatric disease, our current technical ability to model such chromosomal anomalies in mouse, and the existence of precise behavioral measures of endophenotypes argue that the time is ripe for systematic dissection of the genetic mechanisms underlying such disease.

Keywords

CNVs; genetics; schizophrenia; CRISPR; behavioral assays
I. Introduction

Overview of Schizophrenia

Approximately seven out of 1000 individuals (0.7%) are afflicted with schizophrenia (SZ) (McGrath et al., 2008), with symptoms being grouped into three categories: positive, negative, and cognitive. Onset is typically between the ages of 18-25 years, although presentation is about five years earlier for males than females (American Psychiatric Association, 2013). Positive symptoms, sometimes collectively referred to as “psychosis,” are characterized by hallucinations, delusions, conceptual disorganization, and thought disorder (Association, 2013). Negative symptoms are characterized by social withdrawal and lack of motivation (Andreasen, 1995, Lewis and Lieberman, 2000). Cognitive deficits are perhaps the most detrimental symptoms and include impaired attention, executive function, declarative memory (ability to recall and use known information), and working memory (ability to process and manipulate information) (Lewis and Lieberman, 2000, Perry et al., 2000, Lett et al., 2014). A vast array of pharmaceuticals is used to treat SZ, but the effectiveness of each drug varies, as shown by a meta-analysis that revealed heterogeneity in both drug response and side effects (Miyamoto et al., 2012, Leucht et al., 2013). While drug treatment can ameliorate the positive symptoms of SZ, the negative and cognitive symptoms remain recalcitrant (Makinen et al., 2008, Fumagalli et al., 2009, Hill et al., 2010, Miyamoto et al., 2012, Tcheremissine et al., 2012, Millan et al., 2016). Furthermore, a recent study suggests that low doses of medication combined with talk therapy and family support are more effective at improving outcome than drug treatment alone (Kane et al., 2015). These studies highlight the complications involved in treating SZ but hold out the hope that a better understanding of the disease’s molecular pathology could improve treatment strategies.

Risk factors for SZ include obstetric complications, urban birth, immigrant status, young maternal age, advanced paternal age, and family history (Preti et al., 2000, Jablensky et al., 2005, Sorensen et al., 2014). Family history has a large effect size, with a ~10-fold increase in risk for those with an affected first-degree relative, suggesting a strong role for genetics in the etiology of SZ. Elegant studies of adoptive families in Denmark first confirmed the genetic risk (Kety et al., 1976, Kety et al., 1994). Of 14,000 children who were adopted away from their biological families between the years 1924-1947, 47 were diagnosed with SZ. Forty-seven control adoptees were also identified, and the rate of SZ was examined in both biological and adoptive relatives of SZ cases and control adoptees. The researchers found a 12.5-fold excess of SZ in the biological relatives of those with SZ, supporting genetics as a strong risk factor (Heston, 1966, Kety et al., 1971, Wender et al., 1974, Kety et al., 1976, Kety et al., 1994, Mulle, 2012).

Because genetic factors are a significant contributor to the development of SZ, collaborative efforts have focused on identifying genes disrupted in patients. The history of schizophrenia genetics is long, voluminous, and even infamous (Owen, 1992), with excellent reviews available elsewhere (Sullivan et al., 2012). Here we focus on the most recent findings, including recent genome-wide association studies (GWAS). Results from a 2008 GWAS used 738 cases and 733 controls but could uncover no loci significantly associated with SZ,
suggesting this study was underpowered to detect common variants (Sullivan et al., 2008, Mulle, 2012). A larger GWAS conducted in 2009 by the International Schizophrenia Consortium unearthed the first evidence in support of SZ as a polygenic disorder (International Schizophrenia et al., 2009). This discovery prompted a wave of SZ GWAS with larger sample sizes. A 2013 study used a sample size of 21,246 cases and 38,072 controls and identified 13 new loci of interest (Ripke et al., 2013). A follow-up study in 2014 with a sample size of 36,989 cases and 113,075 controls identified 108 schizophrenia-associated single-nucleotide polymorphisms (SNPs) and pointed to multiple biological pathways involved in the onset of SZ (Schizophrenia Working Group of the Psychiatric Genomics, 2014). These data provided further support for polygenic influences in SZ.

In contrast to polygenic influences, copy number variants (CNVs, reviewed below) associated with SZ are rare, often contain multiple genes, and are associated with large effect sizes. While we discuss other CNVs associated with SZ, we focus more in depth on the 3q29 deletion due to the recent discovery of the significant risk it confers for SZ. In addition, we examine the possibility that SZ exists on a continuum with other psychiatric disorders. Finally, we go on to discuss both the molecular and behavioral tools that can be used to study CNVs, allowing us to learn more about the underlying biology of associated conditions.

II. Copy Number Variants

What Are CNVs?

CNVs are defined as a gain or loss of a segment of DNA greater than one kilobase (> 1 kb). Some CNVs may be larger than 100 kb, although CNVs of this size typically occur at low frequency in the population. Multiple genes and regulatory regions can be affected in CNVs of all sizes (Itsara et al., 2009, Dasouki et al., 2011, Torres et al., 2016). In 2008, a study revealed the increased presence of large, rare CNVs in schizophrenic patients versus controls (Walsh et al., 2008). These findings are consistent with the hypothesis that CNVs may be significant genetic risk factors for SZ. Additional studies with larger sample sizes led to the discovery of several recurrent CNVs (Table I) associated with SZ and other psychiatric disorders (International Schizophrenia, 2008, Stefansson et al., 2008, Moreno-De-Luca et al., 2010, Mulle et al., 2010, Mulle, 2012, Rees et al., 2014b). These SZ-associated CNVs share many characteristics. First, these CNVs result in multiple symptoms apart from neuropsychiatric disorders, among them congenital defects and craniofacial abnormalities (Hastings et al., 2009, Torres et al., 2016). Second, they cover large regions of the genome that contain many genes affecting multiple signaling pathways, cell types, and tissues. SZ-associated CNVs are also rare (<0.1% in the general population) and confer high risk for the onset of SZ; the increased risk for SZ associated with CNVs is many times the minimal risk associated with SNPs (Bodmer and Bonilla, 2008, Xu et al., 2008, Mulle, 2012). It is important to emphasize that these associations have been replicated by multiple studies, suggesting they are molecular targets worth dissecting to understand the biology of this disease.

Recurrent CNVs, including those associated with SZ, arise due to the presence of repetitive sequences within the genome, known as low copy repeats (LCRs) (Sasaki et al., 2010).
LCRs are stretches of DNA that are typically 10-500 kb (though their size can vary), with greater than 90% sequence identity to another place in the genome (Bailey et al., 2001). The presence of LCRs puts the genome at risk for chromosomal rearrangements that can cause CNVs, such as deletions and duplications (Stankiewicz and Lupski, 2002). These rearrangements occur because LCRs serve as substrates for non-allelic homologous recombination (NAHR), which is recombination between two highly similar DNA sequences that are not alleles (Shaw and Lupski, 2004). The result of NAHR depends on the orientation of the LCRs: if the repetitive sequences are in the same direction, NAHR causes a deletion or duplication, although not necessarily in similar proportions (Turner et al., 2008). Conversely, if the repetitive sequences are oriented in opposite directions, the result is an inversion (Shaffer and Lupski, 2000). A recent review provides further information regarding the mechanism of NAHR (Sasaki et al., 2010).

**What Have CNVs Told Us About SZ?**

Given the large size of many CNVs, it is difficult to pinpoint what in the region leads to the phenotype. One way to gain some insight into how CNVs confer high risk for SZ is to perform Gene Set Enrichment Analysis (GSEA). Briefly, GSEA is a statistical tool that allows one to determine the biological pathways associated with a phenotype based on the family of genes grouped with the disorder of interest (Subramanian et al., 2005). Recently, Kirov et al. used GSEA on gene sets assembled from proteomic analyses of the synapse and discovered significant enrichment of genes that code for proteins in the post-synaptic density (PSD) in de novo CNVs associated with SZ (Kirov et al., 2012). Specifically, genes encoding elements of the N-methyl-D-aspartic acid receptors (NMDAR) network (proteins that interact with NMDAR (NMDAR complex)) and proteins that interact with the cytoskeleton-associated protein (ARC) complex were significantly enriched. A replication study was conducted on 7,907 patients with SZ and 10,585 controls, and the results showed significant enrichment of genes of the NMDAR complex, but not the ARC complex, in the CNVs of schizophrenic patients (Kirov et al., 2012). A more recent study analyzed 623 schizophrenia trios (individuals with SZ and their unaffected parents) via exome sequencing and found that small, de novo mutations in NMDAR and ARC complexes were enriched in SZ patients (Fromer et al., 2014). This overlap in proteomic and genomic approaches implicates the disruption of networks associated with NMDARs or ARC as relevant to the disease mechanism behind SZ.

**CNVs Associated with SZ**

There are multiple recurrent CNVs associated with SZ and other psychiatric disorders, as summarized in Table 1. Here, we discuss three CNVs in particular, the 22q11 deletion, the 16p11.2 duplication, and the 3q29 deletion, due to the high risk they confer for SZ (Rees et al., 2014b). We go on to discuss the 3q29 deletion in more detail because recent data indicate it confers the greatest genetic risk for SZ (Mulle, 2015).

The most common of the recurring CNVs, the 22q11 deletion, occurs in approximately 1 in 4,000 live births. Most of these reported deletions are 3 Mb in size, which consists of ~60 genes; there is also a smaller recurrent deletion, 1.5 Mb consisting of 28 genes, that comprises the remainder of reported deletions. Both deletion sizes are associated with...
multiple symptoms, such as cardiac defects, learning difficulties, and facial dysmorphisms (Bassett and Chow, 2008, Drew et al., 2011, Torres et al., 2016). Additionally, this deletion is associated with multiple neuropsychiatric disorders, among them autism spectrum disorder (ASD) and intellectual disability (ID). Patients with this deletion are also at risk for SZ (Grozeva et al., 2012, Monks et al., 2014, Rees et al., 2014a, Rees et al., 2014b); notably, this CNV carries the second highest risk for developing the disease (Rees et al., 2014b, Kotlar et al., 2015).

Among the genes located in this region, two candidate genes are proposed to contribute to the SZ phenotype: catechol-o-methyl transferase (COMT) and proline dehydrogenase 1 (PRODH). COMT encodes the postsynaptic enzyme known to regulate the degradation of dopamine (Yavich et al., 2007); the SZ phenotype has been associated previously with dopamine deficits (Yamaguchi et al., 2015). PRODH encodes an enzyme responsible for glutamate production in the mitochondria. Loss of normal activity of PRODH has been linked to psychiatric disorders, such asSZ (Jacquet et al., 2002, Torres et al., 2016).

The 16p11.2 duplication spans a region of 600 kb and has a population prevalence of 1 in 1,000 (Blumenthal et al., 2014). This CNV is associated with a plethora of symptoms, including metabolic dysregulation and microcephaly (Maillard et al., 2015, Arbogast et al., 2016). In addition to SZ, this duplication is associated with neuropsychiatric disorders, such as ASD and ID (Arbogast et al., 2016).

Among the 28 genes in the region, there are multiple candidates that might contribute to the neuropsychiatric disorder (Arbogast et al., 2016). One of these genes is double C2 domain alpha (DOC2A). DOC2A regulates the release of Ca2+-dependent neurotransmitters, and mouse studies indicate that DOC2A functions in memory processes (Mochida et al., 1998, Sakaguchi et al., 1999, Arbogast et al., 2016). A second candidate gene in the 16p11.2 region is TAO kinase 2 (TAOK2). This gene is a kinase that has multiple roles, including the regulation of dendrite formation on cortical neurons (Arbogast et al., 2016); it is also an ASD susceptibility gene (de Anda et al., 2012).

A recently discovered CNV, the 1.6-Mb 3q29 deletion, was first reported in 2005 in six patients with ID and musculoskeletal abnormalities (Willatt et al., 2005). There are now more than 40 reported cases of individuals with this deletion. Most of the reported deletions are de novo, with an incidence rate of one in 30-40,000 births; however, inheritance has been documented (Digilio et al., 2009, Stefansson et al., 2014). While the symptoms of the deletion are heterogeneous, most patients experience varying degrees of ID and speech delay (Cox and Butler, 2015b); the deletion is also associated with anxiety, feeding problems during the first year of life, decreased weight at birth, and dental abnormalities (Glassford et al., 2016). The deletion was first reported to be associated with schizophrenia in 2010, and this association has since been replicated (Mulle et al., 2010, Levinson et al., 2011, Rees et al., 2014b, Szatkiewicz et al., 2014). A recent meta-analysis revealed that the 3q29 deletion increases the risk of developing schizophrenia by 41.1-fold (Mulle, 2015). This deletion is also linked to bipolar disorder (Clayton-Smith et al., 2010, Glassford et al., 2016, Green et al., 2016). Case reports connected the 3q29 deletion to ASD (Cobb et al., 2010, Quintero-
Rivera et al., 2010, Sagar et al., 2013), and two recent reports confirm this association (Sanders et al., 2015, Glassford et al., 2016).

Within the 1.6-Mb deletion interval, there are 22 protein-coding genes, but none have been identified definitively as key to the neuropsychiatric sequelae (Mulle, 2015). The three genes in the interval nominated as candidates are: discs large 1 (DLG1), p21 protein (Cdc42/Rac)-activated kinase 2 (PAK2), and F-box protein 45 (FBXO45). Of these, DLG1 has generated the most interest. DLG1 is a scaffold protein (meaning its role is to assist in the localization and organization of multiple proteins) capable of interacting with subunits of NMDARs. Briefly, DLG1 interacts with GluN2 (a subunit of NMDARs) and is thought to assist in the maturation of NMDARs by localizing them to the Golgi complex, suggesting a function in synaptic signaling (Walch, 2013). In addition, there is clinical evidence implicating DLG1 as a candidate gene for SZ. In one study, multiple SNPs within DLG1 were identified in a cohort of Japanese males with SZ. Further statistical analyses indicated that one of those SNPs showed significant genotypic association within this cohort of patients (Uezato et al., 2012). In addition, Purcell et al. found an enrichment of DLG1 mutations in SZ patients using exome sequencing on a cohort of 2,536 cases and 2,543 controls (Purcell et al., 2014). With mounting evidence of its role in NMDAR function and the risk it confers in the onset of SZ, larger cohorts of SZ patients should be assessed for mutations in DLG1.

Until rigorous studies are done, it is strictly speculative to nominate candidate genes within 22q11, 16p11.2, and 3q29 loci (and those mentioned in Table 1) as associated with SZ. History is littered with failures of the candidate gene approach to unravel complex neuropsychiatric disorders (Sullivan et al., 2012, Farrell et al., 2015). To avoid this, both in vivo and in vitro tools, discussed later in this review, are required to dissect which gene(s) contribute to the phenotype. Once accomplished, this work may lead to an understanding of the pathophysiology of these CNVs, which would give us powerful mechanistic insight into the etiology of SZ.

Neural Disorders on a Spectrum

One observation from CNV studies is that a single CNV may confer risk for multiple neuropsychiatric disorders—ASD, ID, and SZ (Figure 1)—suggesting the possibility of a shared etiology among these conditions (Guilmatre et al., 2009, Crespi and Crofts, 2012, Torres et al., 2016). The idea that these conditions may exist on a continuum is controversial, particularly among clinicians, but genetic evidence to support this hypothesis continues to mount. Here we briefly discuss the evidence that these disorders are not separate entities, but linked.

Per the DSM-V, ASD is characterized by deficits in social communication and interaction, the presence of repetitive actions, and is often first observed during childhood (Association, 2013). The notion that ASD and SZ are linked is not new, as they were even believed to be one and the same prior to the 1970s (Chisholm et al., 2015). These disorders share several symptoms, including deficits in social reciprocity (reciprocal response in a conversation between two people). One striking difference between these disorders was thought to be psychosis, a specific hallmark of SZ but not ASD (Barneveld et al., 2011, Chisholm et al., 2015); however, multiple studies have shown that individuals diagnosed with ASD during
childhood are at a higher risk of also developing psychosis later in life (Joshi et al., 2013, Sullivan et al., 2013). Indeed, in some studies up to 35% of ASD patients experienced psychosis (Konstantareas and Hewitt, 2001, Chisholm et al., 2015). There are also shared molecular pathways implicated separately in the cause of both ASD and SZ (Gao and Penzes, 2015). One such pathway involves the aforementioned NMDARs; genes encoding various NMDAR subunits are disrupted in several CNVs associated with both ASD and SZ. Genetically engineered mice with perturbations of NMDAR subunits show impaired social interaction and phenotypes corresponding to both ASD and SZ, respectively (Belforte et al., 2010, Won et al., 2012, Gao and Penzes, 2015). NMDARs, as well as other pathways, including mammalian target of rapamycin pathway (mTOR) signaling (Gao and Penzes, 2015), hint at molecular similarities between these disorders.

ID is another condition associated with both ASD and SZ. ID is a heterogeneous disorder defined as the inability of an individual to learn new skills, to comprehend new information, and to thrive independently (Ropers, 2008). Diagnosis typically occurs before age 18 years, and the severity of ID depends on a test of one’s intelligence quotient (IQ). An IQ between 50-70 is considered mild ID, while an IQ less than 50 indicates severe ID. Studies have shown that roughly 40% of patients with ID also have ASD, and an estimated 50-80% of patients with ASD have ID (Brereton et al., 2006, Matson and Shoemaker, 2009). Furthermore, a separate study found that ID was present in approximately 4% of a cohort of SZ patients (Morgan et al., 2008). The common molecular pathway among the three diagnosed disorders is unclear, but many of the characterized CNVs are known to underlie ASD, ID, and SZ (Torres et al., 2016). Mutations in NRXN1, which maintains proper synaptic activity by modulating calcium, have been found in individual patients who have all three disorders (Vinas-Jornet et al., 2014). Another recent review also highlights the link between calcium signaling and neuropsychiatric disorders, such as ASD and SZ (Kotlar et al., 2015).

If these disorders are on a continuum, suffering from one disorder would suggest a person is more likely to suffer from the others, which seems to be supported by the current data. Confirming the co-occurrence of these three disorders will require large cohorts that can be followed longitudinally. Studying how this occurs in patients is necessary, but will be slow and expensive; however, modeling CNVs in mouse combined with behavioral studies and the use of patient-derived induced pluripotent stem cells (iPSCs) may allow us to dissect the mechanism of this continuum.

### III. Tools to Study CNVs

The growing number of CNVs associated with neurodevelopmental and psychiatric disorders demands that we better understand how and why they confer risk, putting our efforts into modeling these CNVs in vivo. The ability to precisely manipulate the mouse genome to recapitulate specific chromosomal lesions enables straightforward generation of useful mouse models. Combined with careful behavioral phenotyping, the consequence of such lesions can be measured. In this section we review the power of using large deletions and duplications to understand disease etiology and focus on how mouse models are generated and behaviorally phenotyped. We also introduce the use of (iPSCs) as another
means to understand genomic and molecular changes caused by a CNV, and to improve drug efficacy.

**Chromosome Engineering Using Cre-loxP**

Manipulation of the mouse genome to generate chromosomal deletions, duplications, inversions, and translocations has a rich history. Starting with George Snell’s demonstration that the deleterious effects of X-rays could be inherited (Russell, 1954), classic mouse geneticists realized the power of using irradiation to induce deletions and study the resulting phenotypes. The approach evolved to include additional mutagens and to identify inversions and chromosomal translocations (Russell, 1951, Russell et al., 1989, Davisson et al., 1990, Stubbs et al., 1997, Yu and Bradley, 2001). Indeed, gamma irradiation induced the chromosome 16 duplication (syntetic to part of human chromosome 21) that first modeled Down syndrome (Davisson et al., 1990, Reeves et al., 1995). However, the challenge of using irradiation or chemical mutagens was the inability to control the breakpoints, making it hard to genetically mimic human disease. This challenge was met partly through studies that induced nested deletions, whose phenotypes could be compared to define critical regions (You et al., 1997, Schimenti et al., 2000, Naf et al., 2001). The ultimate solution came with the advent of Cre-loxP technology, which enabled targeting of a loxP site precisely at the desired chromosomal breakpoint (Lewandoski and Martin, 1997).

Cre-lox technology grew popular initially as a way to generate conditional alleles of single genes, enabling spatial or temporal control of deletion (Sauer and Henderson, 1988, Orban et al., 1992). Allan Bradley’s group exploited the system for chromosome engineering. By sequentially targeting distinct loxP sites to two points on a chromosome, Cre-mediated recombination results in a deletion or inversion, depending on whether the loxP sites are in the same or opposite orientation. If the two loxP sites are targeted to different chromosomes, chromosomal translocations are possible. In reality, the efficiency of Cre recombining two loxP sites decreases as the distance between the two loxP sites grows. Thus, the Bradley group included a selectable marker whose gene was split between the two loxP sites requiring recombination for its intact expression. They chose to use the hypoxanthine phosphoribosyl transferase (Hprt) minigene split into its 5’- Hp and 3’-rt halves. This approach enabled generation of mice with specific chromosomal anomalies using ES cells through six steps: 1) targeting of a 5’Hp-loxP construct, 2) targeting of a loxP-3’rt construct, 3) Cre transfection, 4) HAT selection (for recombined Hprt expression), 5) Southern blot and fluorescence in situ hybridization (FISH) analysis to confirm the desired chromosomal event occurred, and 6) blastocyst injection and chimera production (Yu and Bradley, 2001, Zheng et al., 2001), making for a labor-intensive and expensive, albeit feasible, approach. The Bradley lab greatly improved this approach by generating a resource of premade targeting vectors and mapping them throughout the genome so that researchers could order up the clones necessary for chromosomal engineering (Adams et al., 2004).

Chromosome engineering led to a series of mouse models that precisely mimicked the genetic architecture of human patients. In many cases the approach enabled the etiology of the disease to be linked to individual or small groups of genes. For example, the mouse model of the recurrent 22q11 microdeletion (DiGeorge syndrome) displayed congenital
heart disease similar to patients, which was subsequently proven to be due to haploinsufficiency of the \textit{Tbx1} gene (Lindsay et al., 1999) (Lindsay et al., 2001). To dissect constitutional chromosome 17 deletions, three distinct mouse lines were made: a 6.9-Mb deletion, the reciprocal duplication, and a smaller 1.8-Mb deletion. Analysis of each line alone and subsequently in combination with each of the other lines made it clear that a small portion of the region led to the heart defects observed in patients, likely due to haploinsufficiency (Yu et al., 2006). Regulatory elements can be uncovered via chromosomal engineering, as well. By generating a large deletion covering most of the \textit{HoxD} cluster and investigating what aspects of gene expression could be rescued with large human transgenes, Spitz and colleagues identified the location of enhancer sequences (Spitz et al., 2005). Taken together, chromosome engineering revolutionized our ability to molecularly dissect chromosomal regions and anomalies.

Chromosomal engineering has proved useful for generating mouse models of neurodevelopmental and neuropsychiatric disease. The 22q11 deletion mice revealed that \textit{Dgcr8}, a gene that encodes a component of the microRNA biogenesis pathway, is a significant contributor to the SZ phenotype and not to the cardiac deficits. This highlights the power of deletion analysis to tease apart multiple mechanisms of disease. In addition, disruption of miRNA biogenesis affects genes throughout the genome, emphasizing the complexity underlying CNV-associated psychiatric disorders (Forstner et al., 2013). Patients with recurrent 15q11-13 duplication syndrome display symptoms that include autism, intellectual disability, seizures, and developmental delay. To genetically recapitulate the disease, researchers targeted the loxP sites in \textit{trans} to generate a mouse carrying a 6.3-Mb duplication of the syntenic region on mouse chromosome 7. These mice exhibit an autistic endophenotype (aspects of a disorder that are quantitative and can be modeled in a rodent; see the SZ-Related Rodent Behaviors section of this review), providing an animal model to study this recurrent CNV (Nakatani et al., 2009). Down syndrome (DS) and monosomy 21, on the other hand, affect chromosomal regions with multiple syntenic regions in the mouse genome. Using Cre-lox, researchers have generated deletions on both mouse chromosome 10 and 17 to model the human CNV, 21q22.3 (monosomy 21). The mouse model displays reduced learning and memory function, consistent with the intellectual deficits in patients (Yu et al., 2010). Similarly, researchers generated duplications on mouse chromosomes 17, 16, and 10 to model the trisomic regions of DS. These mice are allowing distinct DS phenotypes, such as heart abnormalities, to be connected to discrete genomic regions (Liu et al., 2011). In the course of chromosome engineering, both deletions and duplications can be generated concordantly, depending on whether the loxP sites are in \textit{cis} or \textit{trans}. Indeed, researchers have created both 16p11.2 deletion and duplication mice and observed aspects of the disorder the same as in patients. The deletion mouse model, but not the reciprocal duplication mouse model, recapitulated the human disorder with an autistic endophenotype. This work underscores the challenge of genetic heterogeneity when using mouse models to study human disorders arising from recurrent CNVs (Horev et al., 2011).

\textbf{Modulating Double-Strand Breaks to Generate CNVs}

While the power of chromosomal engineering is undeniable, so is the technical élan it requires. Recent advances using nucleases appear to be lowering the technical barrier for
chromosomal engineering. This transformation began with including transcription activator-like effector nucleases (TALENs) and zinc-finger nucleases (ZFNs), and these in turn are now being superseded by clustered regularly interspaced short palindromic repeats (CRISPRs). All three technologies depend on the same principle George Snell used: induction of double-strand breaks (DSBs) to induce rearrangements (Xiao et al., 2013, Wijshake et al., 2014). But, unlike irradiation and similar to Cre-loxP, all three permit specification of the breakpoint.

TALENs and ZFNs are similar genetic engineering tools in that both consist of a nuclease, which is fused to a protein that binds to the genetic locus of interest. Once bound to DNA, a DSB is generated. Repair of DSBs is often mediated via nonhomologous end joining (NHEJ), which can occur perfectly or imperfectly, resulting in the insertion or deletion of genetic material (indel)(Lee et al., 2012, Sander and Joung, 2014). Depending on the indel, there may be a functional consequence. When TALENs or ZFNs are targeted to two positions on a chromosome, a deletion or (less frequently) an inversion can occur. Indeed, both TALENs and ZFNs are sufficient to generate CNVs. In zebrafish, researchers used TALENs to generate both deletions and inversions up to 1 Mb in size (Xiao et al., 2013). In a human cell line, ZFNs were able to create a deletion and duplication of a 230-kb region, as well as a 15-kb inversion (Lee et al., 2012). Despite such proofs of principle, TALENs and ZFNs never became popular. While the exact reasons are unclear, here at the dawn of the CRISPR age, the issue is now moot.

CRISPR is popular because the system is straightforward to design and optimize. CRISPR relies on a short sequence of RNA, known as the guide RNA (gRNA), to target the nuclease (originally Cas9) to the site of DNA cleavage to generate a DSB. The repair mechanisms are identical to those described earlier for TALENs and ZFNs: to create a deletion or inversion using CRISPR, two CRISPRs to the desired breakpoints are needed (Sander and Joung, 2014). If those two CRISPRs target the chromosomes in trans, a duplication is possible. Chromosomal translocations can be made by designing CRISPRs to appropriate breakpoints on distinct chromosomes. The design for creating a variety of CNVs using CRISPR is depicted in Figure 2. The same logic of the original geneticists who exploited deletion analysis can again be used to design nested deletions and define critical regions. Such subdeletions can share a breakpoint with a larger deletion or occur within the larger deletion (Figure 3).

In terms of chromosome engineering in mice, CRISPR provides an unparalleled advantage over previous methods: CRISPRs and Cas9 nucleases can be directly injected in the one-cell zygote, eliminating the need for ES cell targeting or blastocyst injection. Within a month, the offspring can be screened directly, and even though some may be mosaic, the fact that editing occurs during the first few cleavages of the developing embryo increases the likelihood of germline transmission. Of critical note, the fact that a DSB is induced increases the risk of unwanted rearrangement, making Southern analysis necessary to ensure that only the desired manipulation occurred (Wang et al., 2013). In addition, FISH is well suited to confirm that the chromosomal rearrangement is correct (Zheng et al., 2001). Such controls will always be needed when manipulating the genome. The advantage of being able to
directly edit the genome is that it allows chromosomal engineering to be performed more quickly and cheaply than ever before.

While most CRISPR-generated mouse alleles are for single genes, using two CRISPRs at a time can generate CNVs \textit{in vivo} (Singh et al., 2015), including the generation of a 30-Mb deletion and the creation of a chromosomal translocation (Essletzbichler et al., 2014, Torres et al., 2014). This also includes the generation of an 11-Mb inversion that modeled a form of lung cancer driven by the \textit{EML4-ALK} oncogene, recapitulating the human disorder and showing the use of ALK inhibitors blocked tumor formation (Maddalo et al., 2014). Similar manipulations can be created \textit{in vitro}, one report used CRISPR to model the recurrent CNVs 16p11.2 and 15q13.3 in iPSCs. This study is proof of principle that the breakpoints of a recurrent CNV can be easily controlled using CRISPR and should be possible in mouse (Tai et al., 2016).

One of the initial concerns about CRISPR was off-target cleavage creating second-site mutations (Wang et al., 2013), a problem that has been well addressed with better algorithms for CRISPR target design with the advent of the new generation of nucleases (Kleinsteve et al., 2016, Slaymaker et al., 2016, Tsai and Joung, 2016). While there may still be some concerns in cell culture, when working \textit{in vivo}, even if second-site mutations are induced, the founder mouse must be bred such that mutations will recombine out of the experimental cohort over time. Thus, researchers can be confident that the phenotype should be attributable to the CNV being interrogated.

There will surely be challenges using CRISPR for chromosomal engineering. In the past two decades, gene targeting demonstrated that some genomic regions are more difficult to manipulate, so it is possible that some CNVs may prove harder to model. In addition to NHEJ, DSBs can also be repaired through homology-directed repair (HDR) from a DNA template donor (Singh et al., 2015). Providing a donor of the desired outcome may make it possible to increase the frequency of a desired manipulation (Chen et al., 2011). Alternatively, one could imagine a hybrid approach using CRISPRs in ES cells and adapting the \textit{Hprt} minigene selection, or finding an equivalent gene that could be split and then selected for \textit{in vivo}. Such problems, as well as unforeseen ones, will surely be solved in time; the enormous amount of biology that the approach can reveal will drive solutions.

In terms of studying neuropsychiatric disease, CNVs can now be created in mouse with great precision for less money and in a shorter time than ever before. As explained in the next section, this is especially important when performing behavioral assays, which can be highly strain specific. Due to the relatively low cost and short timeframe of chromosomal engineering using CRISPR, it is reasonable to generate the same deletion (or other manipulation) on multiple strain backgrounds to dissect the mechanism of such complex diseases. CRISPR is now poised to revolutionize our ability to engineer chromosomes, and thus to understand CNV in human variation and in disease states.

\textbf{SZ-Related Rodent Behaviors}

Once a mouse model of SZ is created, it is imperative to choose behavioral tests that reflect this psychiatric disorder. While it is naïve to suppose that a “schizophrenic” mouse could be
engineered to encompass all these symptoms, the SZ endophenotype can be modeled and tested individually in rodents. Endophenotypes are not themselves symptoms or signs of the disease, but are rather components of the disease that are quantitative and amenable to assessment in rodents in the laboratory. This reductionist approach has vastly improved our understanding of the mechanisms behind SZ and its potential treatment (McGrath et al., 2004). In the following section, we will address how behavioral tests are chosen that reflect SZ endophenotypes, and how these tests can act as correlates to SZ behavior.

One important aspect of modeling disease in animals is construct validity, which denotes a common underlying neurobiology in both the animal model and the human psychiatric disorder (Nestler and Hyman, 2010). Given the 40-fold increased incidence of SZ in patients with the 3q29 deletion, mice generated to carry the deletion would seem an ideal rodent model capable of providing powerful insight into the underlying genetics and neurobiology of SZ. In addition to construct validity, rodent models should strive to achieve face and predictive validity. Face validity is the extent to which a model “looks like the disease”; if the mouse 3q29 deletion recapitulates the behavioral features of SZ, it will have face validity. If treatments (e.g., antipsychotic drugs) that provide therapeutic benefit in humans with disease also ameliorate the endophenotypes observed in mice carrying the 3q29 deletion, the model then possesses predictive validity and can be used as a preclinical screen for novel compounds.

Mouse behaviors relevant to SZ endophenotypes can be categorized into three main symptom clusters: positive, negative, and cognitive (Powell and Miyakawa, 2006, Nestler and Hyman, 2010). The most common measure thought to reflect positive schizophrenic symptoms in rodents is hyperactivity. Increased locomotor activity at baseline or in a novel environment is seen in multiple proposed SZ rodent models (Lipska and Weinberger, 2000). This parallels the psychomotor agitation observed in some schizophrenic patients. In addition, psychotomimetic drugs, such as ketamine, phencyclidine, and amphetamine, which induce psychosis in human controls and exacerbate symptoms in SZ patients, induce hyperactivity and stereotypic behaviors in rodents, while rodent models of SZ have shown increased sensitivity to the locomotor-activating consequences of psychotomimetic drugs. The finding that SZ patients exhibit enhanced dopamine release following amphetamine injection (Kegeles et al. 2010) validates this assay as a preclinical assessment for rodent models of SZ. These results should be interpreted with caution, however, because hyperactivity is not a defining characteristic of SZ (Kegeles et al., 2010) and is not present in all patients.

Social withdrawal is one of the more debilitating aspects of SZ. This endophenotype is quantified in rats and mice by exposing them to a “stranger” rodent of the same species. There are multiple ways to assess social interaction in rodents, from allowing the stimulus rodent free movement throughout an area to confinement to a specific location. Whatever particular procedure is employed, the amount of time investigating the stimulus rodent(s) serves as the dependent variable. As a note of caution, if experimental subjects are tested in an apparatus other than their home cage, these sociability measures can be confounded by anxiety induced by a novel environment. In situations when this is a concern, subjects should either be tested in their home cage or undergo extensive habituation to the test.
enclosure. In addition to asociality, a subset of SZ patients suffers from anhedonia, or blunted affect. These negative symptoms can also be measured in rodent behavioral tests. Sucrose preference testing, in which subjects are allowed access to a sweetened solution and the amount consumed is recorded, as well as detailed examination of mouth and tongue movements following sucrose consumption, allow for quantification of motivation and reward. Intracranial self-stimulation, which uses voluntary operant responses to determine the sensitivity of the reward system, can also be employed to assess hedonic state. As is the case with many endophenotypes examined in rodent models of schizophrenia, these symptoms can occur in other psychiatric conditions and thus are not exclusive to SZ.

Cognitive symptoms of schizophrenia have also been assessed in rodent models of the disease. These cognitive endophenotypes can be subcategorized into deficits in working memory, declarative memory, and attention. Working memory can be measured in the eight-arm radial maze, delayed alternation in the T-maze, or spontaneous alternation in the Y-maze. These tasks require subjects to integrate their recent previous behavior into current choice selection. For example, in the eight-arm radial maze, re-entries into a food-baited arm after the reward has been found and consumed are identified as working memory errors. Working memory errors can also result from perseveration and a lack of behavioral flexibility. For this reason, it is difficult to determine whether performance deficits in these tasks are the result of impaired working memory or poor impulse control. This concern is somewhat offset by the knowledge that SZ patients also exhibit reduced behavioral flexibility and perseveration (executive function) (Crider, 1997). In addition to working memory, SZ is also associated with impairments in declarative memory (Saykin et al., 1991, Saykin et al., 1994, Heinrichs and Zakzanis, 1998). The eight-arm radial maze, Morris water maze, and Barnes maze all measure declarative spatial memory in rodents. These tasks require subjects to use the interrelationship between extra-maze cues to locate food reinforcement (radial maze) or escape from an aversive stimulus (Morris water maze and Barnes maze). For example, in the Morris water maze, subjects are trained to locate an invisible escape platform in a pool of water using spatial visual cues in the room. The latency and distance traveled to locate the platform over trials are the primary dependent variables for task acquisition. Recall memory can be assessed in a probe trial, during which the escape platform is removed and the amount of time spent in and around its former location are recorded. Adding an inter-trial interval and concurrent exposure to a previously encountered “stimulus” rodent, paired with a novel conspecific in the social interaction assays mentioned above, allow for assessment of social memory. Mice have an innate preference for novelty, and will normally spend more time interacting with a novel subject compared to a familiar subject, so an equal time spent investigating a novel and familiar animal can reveal social memory impairment. Of particular relevance to the current review, the DF(16)A+/− mouse model of 22q11.2 deletion exhibits strongly impaired social memory (Piskorowski et al., 2016), likely suggesting a similar deficit in mice carrying the 3q29 deletion.

The rodent behavioral task with the best-regarded face validity for assessing SZ endophenotypes is prepulse inhibition (PPI), a measure of sensorimotor gating, which is associated with impaired selective attention in SZ patients (Scholes and Martin-Iverson, 2010). This phenomenon is measured in humans and rodents by exposing subjects to an
auditory startle stimulus. This startle stimulus is assessed in the presence or absence of a less intense non-noxious stimulus. In normal humans and control mice, the pre-pulse decreases the response to the startle stimulus. Importantly, SZ patients and some rodent models of SZ, including 22q11.2 microdeletion, demonstrate a deficit in PPI, signifying preattention processing deficits (Prasad et al., 2008). Latent inhibition is a related process in which exposure to a stimulus decreases future associations formed with that stimulus. SZ patients are generally resistant to latent inhibition, particularly when acutely symptomatic (Weiner, 2003, Lubow, 2005), as are some mutant mice models of the disorder (Zeng et al., 2001, Rimer et al., 2005). Additional cognitive deficits commonly seen in SZ patients include attention and executive function deficiencies. These processes can also be modeled by rodent behavioral tasks. The Wisconsin Card Sorting (WCS) task reveals attentional set shifting problems in schizophrenic patients (Owen et al., 1993). Rats and mice can be trained to use odor and texture cues to discriminate for food reinforcement. In a manner similar to the WCS, the rules of the discrimination can be changed during testing, such that subjects must shift their attention amongst the cues to continue receiving reward. The five-choice serial reaction time task (5-CSRTT) requires subjects to attend to five separate locations for presentation of a visual stimulus. They can then report detection of the stimulus by performing an operant response for food reinforcement. The 5-CSRTT measures spatial attention but can also detect impulsivity and perseverative behavior (Robbins, 2002, Chudasama and Robbins, 2004). Very recently, the Novel Methods leading to New Medications in Depression and Schizophrenia (NEWMEDS) initiative developed and validated a battery of tasks intended to examine cognitive SZ endophenotypes (Hvoslef-Eide et al., 2015). The battery includes the 5-CSRTT and an additional test of attention, as well measures of cognitive flexibility, spatial working memory, and long-term memory. The entire test battery can be performed in the same touchscreen operant chambers, allowing for high-throughput characterization of SZ-associated endophenotypes, and represents the vanguard of behavioral phenotyping. Importantly, the NEWMEDS battery is fully automated, eliminating or reducing confounds produced by experimenter bias or subject anxiety as a result of extensive investigator handling and intervention. Task automation is recommended whenever possible, and fortunately most if not all of the SZ-related tasks described earlier are capable of automation with a video tracking system, and software-controlled and -monitored operant and startle chambers. These tools allow experimenters to analyze multiple different maze and conditioning environments, providing for automated scoring of cognitive tasks, such as the Morris water maze and PPI. Even more advanced tracking systems allow assessment of multiple subjects in the same apparatus, fully automating measurement of social behavior, as well.

When testing mice for behavioral abnormalities, challenges are always likely. One potential challenge is that some phenotypes will be more robust on different strain backgrounds. To overcome this, it is beneficial to study a CNV on more than one background strain. As examples, mice with poor eyesight (e.g., adult FVB mice have retinal degeneration) or coordination have trouble performing in the Morris water maze task, mice with high levels of anxiety typically display low levels of social interaction, and mice with auditory deficiencies can appear deficient in PPI; using mice with overt alterations in brain morphology (e.g., 129 mice have a partial deficit in corpus callosum formation) may
introduce more variability in some tests (Wahlsten, 1982, Chang et al., 2002). Thus, it is critical to have proper controls and alternatives for each behavioral test. Many of these controls are built into the test, such as eyesight (visible platform in Morris water maze) and hearing (baseline startle in PPI). It is also important for the studies to be powered adequately to account for intra-strain variability in behavioral response. Thus, testing 8-16 mice per genotype is recommended (depending on whether or not males and females are combined in the analysis), which in the literature is more than sufficient to uncover statistically significant differences between genotypes, even in mixed genetic background populations (Weinshenker et al., 2002, Marino et al., 2005, Grubb et al., 2009). It is also important to note that not all symptoms of SZ are present to the same degree in all patients, indicating that genetic modifiers regulate the expression of these symptoms. Because baseline performance in cognitive and behavioral tasks is highly variable between different strains of mice, the choice of strain background is a critical factor for experimental success. For example, it might be difficult to observe a deficit in PPI caused by the deletion on a strain background that would have difficulty performing PPI due to hearing difficulties. As such, one might choose to investigate the phenotypic consequences of the deletion on four strains of mice initially, and then pare down to one or two strain backgrounds that give the most robust phenotypes for subsequent analysis of subdeletions of the interval and single-gene knockouts. A good resource for understanding which strain to choose for behavioral assays is the Mouse Phenome Database (http://phenome.jax.org), which provides a compilation of mouse behavioral data based on strain background.

As discussed, manipulation of the mouse genome in combination with behavioral assays provides a powerful means to study CNVs associated not only with SZ, but other psychiatric disorders, as well. Combining the ease of using the mouse model in behavioral studies with the availability of tools to investigate the mouse genome will move the field towards a better understanding of causes of the CNV. Coupled with pharmacological studies that can be done on mice carrying these recurrent CNVs, we are more likely to find better ways of alleviating some of the symptoms of SZ.

**Using iPSCs to Study CNVs**

Another tool to understand how CNVs confer risk for a psychiatric disorder is induced pluripotent stem cells (iPSCs). Briefly, fibroblasts or PBMCs from a patient are isolated and treated with a cocktail of transcription factors (SOX2, OCT4, KLF4, and c-MYC). Treatment with these factors reprograms the cells into iPSCs, which can then be differentiated into the cell type of interest, such as neurons (Takahashi et al., 2007). The neurons are then characterized using a series of analyses, among them cellular morphology, electrophysiology, and molecular characterization (Wen et al., 2014). Thus, the use of iPSCs provides a means to study potential neuronal dysfunction caused from CNVs. For example, iPSCs have been derived from a series of patients with either the deletion or duplication at the 7q11.23 locus. The deletion is the cause of the well-known Williams-Beuren syndrome, and the duplication is also associated with both schizophrenia (Mulle et al., 2014) and autism (Sanders et al., 2015). Study of iPSCs resulted in the identification of additional candidate genes outside the breakpoints that may contribute to the phenotype (Adamo et al., 2015, Urban and Purmann, 2015). In another study using iPSCs from patients with SZ-
associated DISC1 mutations, abnormal synaptic signaling was observed, providing molecular insight into the cause of SZ (Wen et al., 2014).

The future is promising for the use of iPSCs in the study of CNVs and psychiatric disorders. As techniques continue to improve, the next steps will include comprehensive, coordinated analysis of iPSCs from multiple patients containing different SZ-associated CNVs for cross-comparison and potential identification of common molecular mechanisms. As mentioned in the CNV section of the review, NMDARs and ARC are two protein families found enriched in CNVs associated with SZ. The use of iPSCs is an ideal system to determine whether the expression of these protein families (or others) is altered in the respective CNV. The discovery of altered biological pathways in the various CNVs could then lead to personalized treatment for the patient groups.

V. Conclusion

Extensive collaboration is necessary for the next phase of SZ research. Collective efforts are what led to the discovery of SZ-associated CNVs, and it will take a comparable effort to uncover the disease mechanism, with the ultimate goal of effective treatments. The first step in this challenging endeavor is to understand how each of these CNVs results in SZ on a molecular level. Modeling the effect of CNVs in animals and iPSCs can be invaluable tools to identify the key genes within CNV intervals that contribute to neurodevelopmental phenotypes and further correlate genotype and phenotype. With the development of CRISPR, not only is generating these CNVs conceivable, but so is generating a series of smaller deletions/duplications for each CNV to narrow the interval and specify which molecular targets contribute to the phenotypes. One resource that may help us find the gene that correlates to the respective phenotype is mousephenotype.org, a database of information on knockout mouse models, created in the hope of one day assigning a biological function to each gene in the mouse genome.

The database also includes behavioral assay results from knockout mice, providing yet another tool for unlocking the molecular complexity of a CNV. Finally, cross-CNV comparisons will be necessary to unravel relevant biological pathways. This is especially true for the 3q29 deletion and 22q11 microdeletion. These two CNVs account for only 0.082% and 0.29%, respectively, of all SZ cases (Rees et al., 2014b); however, of all the CNVs associated with SZ, they confer the highest risk (Grozeva et al., 2012, Mulle, 2015). It is possible that an understanding of how they confer such a high risk for SZ at a molecular level could lead to understanding of the mechanisms of SZ. While the 22q11 deletion has been dissected extensively, there is still more to learn (Forstner et al., 2013); the molecular mechanism of the 3q29 deletion is not known. The better we understand SZ at the molecular level, the more likely we are to find effective treatments.

The discovery of CNVs that cause neuropsychiatric disorders has led to the notion of an etiologic spectrum among ASD, ID, and SZ. Other genetic factors, or modifier loci, for pathogenesis are likely to be important. In the most severe case, some individuals have a “second hit” variant that influences penetrance at the primary CNV (Veltman and Brunner, 2010, Girirajan et al., 2012). With the advent of routine whole-genome sequencing for
clinical indications, future studies in humans will investigate genomic background when assessing risk. Mouse models will also help us understand the contribution of background genomic influences associated with a given CNV. This will require the creation of multiple genetic mouse models, generating an identical CNV on different genetic strains. Although this is no small task, the use of CRISPR and standardized behavioral assays put answers within reach. This type of work, albeit involved, might also provide in vivo evidence to support the continuum model. With the tools outlined in this review, the future promises exciting new research that will help us understand the molecular mechanisms behind CNV-associated neuropsychiatric phenotypes, with the ultimate goal of effective treatment for patients’ symptoms.

Acknowledgements

The authors would like to thank Sarah Bay and Cheryl Timms Strauss for editing and the NIGMS for funding (R01 GM097331).

References


J Neurosci Res. Author manuscript; available in PMC 2017 November 01.


Kegeles LS, Abi-Dargham A, Franke WG, Gil R, Cooper TB, Slifstein M, Hwang DR, Huang Y, Haber SN, Laruelle M. Increased synaptic dopamine function in associative regions of the striatum in schizophrenia. Arch Gen Psychiatry. 2010; 67:231–239. [PubMed: 20194823]


Deletion 17q12 is a recurrent copy number variant that confers high risk of autism and schizophrenia. Am J Hum Genet. 2010; 87:618–630. [PubMed: 21055719]


Mulle JG. The 3q29 deletion confers >40-fold increase in risk for schizophrenia. Mol Psychiatry. 2015; 20:1028–1029. [PubMed: 26055425]


Rutkowski et al. Page 23

J Neurosci Res. Author manuscript; available in PMC 2017 November 01.


Significance

Copy number variants (CNVs) are large genomic aberrations that confer risk for multiple psychiatric disorders, including SZ. Most CNVs affect the function of multiple genes, making it difficult to understand the molecular mechanism of these CNVs. Cre-loxP is one technique that has proved successful to model CNVs and identify individual genes that contribute to a phenotype; however, this genetic tool, while powerful, is both time-consuming and expensive. Fortunately, with recent advances in genomic engineering via the use of CRISPR technology, it is now feasible to dissect CNVs and identify individual genes contributing to the phenotypes of psychiatric disorders.
Figure 1. Risk for SZ, ID, and ASD associated with recurrent CNVs
Odds ratios (ORs) for SZ were obtained from (Rees et al., 2014b, Mulle, 2015). ORs for ID were estimated using ID data from (Kaminsky et al., 2011), when zero CNV were observed in controls the OR was estimated by Haldane's correction (adding 0.5 to each cell in the 2×2 matrix). ORs for ASD were estimated from ASD case data (Sanders et al., 2015), n = 4687 cases) and controls (Cooper et al., 2011), n = 8329 controls) using Fisher's exact test in R; when zero CNV were observed in controls the OR was estimated by Haldane's correction.
Figure 2. Generating CNVs using CRISPR

A. To generate an inversion or deletion, CRISPRs are designed at the desired breakpoints to create DSBs. The resulting inversion or deletion is generated only if the DSBs are created on the same homologous chromosome (cis).

B. If the DSBs are generated on different homologous chromosomes (trans), this may result in a duplication.

Generation of DSBs in cis or trans is random and cannot be controlled.

C. As an example of how to generate a translocation, two CRISPRs are needed. The first CRISPR is designed to create a DSB outside the first gene of interest on chromosome 3, and the second CRISPR is designed to create a DSB outside the other gene of interest on chromosome 10.
Designing subdeletions of a CNV is relatively simple. Briefly, the recurrent deletion is divided into three (or more) segments (A, B, C). Each of these segments is separately deleted using the corresponding CRISPR. Functional studies can then be done on each subdeletion to determine which segment contains the gene(s) that confer risk for a disorder, such as schizophrenia (SZ).

Other rearrangements are possible with this design, as outlined in Figure 2.
<table>
<thead>
<tr>
<th>Location</th>
<th>Type of CNV</th>
<th>Psychiatric Disorder(s)</th>
<th>Additional Symptoms</th>
<th>Gene(s) of Interest</th>
<th>Refs</th>
</tr>
</thead>
<tbody>
<tr>
<td>1q21.1</td>
<td>Deletion and duplication</td>
<td>ASD, ID, SZ</td>
<td>Microcephaly (deletion) macrocephaly (duplication)</td>
<td>PRKAB2</td>
<td>(Torres et al., 2016)</td>
</tr>
<tr>
<td>3q29</td>
<td>Deletion</td>
<td>ASD, ID, SZ</td>
<td>Speech delay, facial dysmorphisms</td>
<td>DLG1, PAK2</td>
<td>(Cox and Butler, 2015b, Glassford et al., 2016)</td>
</tr>
<tr>
<td>7q11.2</td>
<td>Duplication</td>
<td>ASD, ID, SZ</td>
<td>Speech delay</td>
<td>GTF2I</td>
<td>(Somerville et al., 2005, Vellem and Mervis, 2011, Escudero and Johnstone, 2014, Adamo et al., 2015)</td>
</tr>
<tr>
<td>15q11.2</td>
<td>Deletion</td>
<td>ASD, ID, SZ</td>
<td>Motor and speech delay</td>
<td>CYFIP1</td>
<td>(Cox and Butler, 2015a)</td>
</tr>
<tr>
<td>15q11.2-13.1</td>
<td>Duplication</td>
<td>ASD, ID, SZ</td>
<td>Language impairment, facial dysmorphisms</td>
<td>HBI152 (a small nucleolar RNA)</td>
<td>(Nakatani et al., 2009, Stewart et al., 2011, Coppola et al., 2013)</td>
</tr>
<tr>
<td>15q13.3</td>
<td>Deletion</td>
<td>ASD, ID, SZ</td>
<td>Hypotonia, cardiac defect</td>
<td>CHRNA7</td>
<td>(Ziats et al., 2016)</td>
</tr>
<tr>
<td>16p11.2</td>
<td>Duplication</td>
<td>ASD, ID, SZ</td>
<td>Decreased body mass index</td>
<td>DOC2A, TAOK2</td>
<td>(Arbogast et al., 2016)</td>
</tr>
<tr>
<td>16p13</td>
<td>Duplication</td>
<td>ASD, ID, SZ</td>
<td>Heart defects</td>
<td>NDE1</td>
<td>(Kuang et al., 2011, Ramalingam et al., 2011, Kimura et al., 2015, Torres et al., 2016)</td>
</tr>
<tr>
<td>17p12</td>
<td>Deletion</td>
<td>SZ</td>
<td>Neuropathy</td>
<td>PMP22</td>
<td>(Kirov et al., 2009, Luigetti et al., 2014)</td>
</tr>
<tr>
<td>17q12</td>
<td>Deletion</td>
<td>ASD, ID, SZ</td>
<td>Kidney abnormalities</td>
<td>LHX1</td>
<td>(Moreno-De-Luca et al., 2010, Palumbo et al., 2014, Clissold et al., 2016)</td>
</tr>
<tr>
<td>22q11</td>
<td>Deletion</td>
<td>ASD, ID, SZ</td>
<td>Congenital defects, facial dysmorphisms</td>
<td>COMT, PRODH</td>
<td>(Bassett and Chow, 2008, Biswas and Furniss, 2016)</td>
</tr>
</tbody>
</table>

Multiple CNVs with their respective symptoms, characteristics, and corresponding references.