Review

Endophenotypes in Schizophrenia: Digging Deeper to Identify Genetic Mechanisms

Tiffany A. Greenwood 1, Andrew Shutes-David 2,3, Debby W. Tsuang 2,4,*

1 Department of Psychiatry, University of California San Diego, La Jolla, CA 92093, USA
2 Geriatric Research, Education, and Clinical Center, VA Puget Sound Health Care System, Seattle, WA 98108, USA
3 Mental Illness Research, Education, and Clinical Center, VA Puget Sound Health Care System, Seattle, WA 98108, USA
4 Department of Psychiatry and Behavioral Sciences, University of Washington, Seattle, WA 98195, USA

* Correspondence: Debby W. Tsuang, Email: dwt1@uw.edu.

ABSTRACT

Schizophrenia (SZ) is a severe psychotic disorder that is highly heritable and common in the general population. The genetic heterogeneity of SZ is substantial, with contributions from common, rare, and de novo variants, in addition to environmental factors. Large genome-wide association studies have detected many variants that are associated with SZ, yet the pathways by which these variants influence risk remain largely unknown. SZ is also clinically heterogeneous, with patients exhibiting a broad range of deficits and symptom severity that vary over the course of illness and treatment, which has complicated efforts to identify risk variants. However, the underlying brain dysfunction forms a more stable trait marker that quantitative neurocognitive and neurophysiological endophenotypes may be able to objectively measure. These endophenotypes are less likely to be heterogeneous than the disorder and provide a neurobiological context to detect risk variants and underlying pathways among genes associated with SZ diagnosis. Furthermore, many endophenotypes are translational into animal model systems, allowing for direct evaluation of the neural circuit dysfunctions and neurobiological substrates. We review a selection of the most promising SZ endophenotypes, including prepulse inhibition, mismatch negativity, oculomotor antisaccade, letter-number sequencing, and continuous performance tests. We also highlight recent findings from large consortia that suggest the potential role of genes, particularly in the neuregulin and glutamate pathways, in several of these endophenotypes. Although endophenotypes require additional time and effort to assess, the insight into the underlying neurobiology that they provide may ultimately reveal the underlying genetic architecture for SZ and suggest novel treatment targets.

KEYWORDS: endophenotype; schizophrenia; genetics; neurocognition; neurophysiology
SCHIZOPHRENIA: A GENETICALLY AND PHENOTYPICALLY COMPLEX DISORDER

Clinical Course and Presentation

Schizophrenia (SZ) is a severe and persistent psychotic disorder with a population prevalence of nearly 1% [1]. The profound clinical heterogeneity of SZ is reflected in its broad range of deficits and symptom severity. The positive symptoms associated with SZ include hallucinations, delusions, disorganized speech, and disorganized behavior, whereas the negative symptoms include diminished facial and emotional expression, poverty of speech, lack of motivation, an inability to experience pleasure, and social withdrawal. To receive a diagnosis of SZ, two or more of these symptoms must be consistently present for at least 1 month and must be severe enough to cause social and/or occupational dysfunction and neglect of hygiene. Due to the extreme heterogeneity in individual symptom profiles, Eugen Bleuler aptly described SZ as the “Group of Schizophrenias” in 1911, only 3 years after he first coined the term “schizophrenia” [2].

The onset of SZ most commonly occurs in late adolescence or early adulthood [3] and is generally chronic and unremitting, often requiring multiple hospitalizations over the course of a lifetime [4]. Although about 25% of patients with SZ will experience a full recovery with medication [5], most patients require long-term care and never return to their previous level of functioning. Given the severity of symptoms, about 40% of SZ patients attempt suicide and 5% succeed [6].

The treatment of SZ is complex. Antipsychotic medications, which target the positive symptoms, are expensive, are only partially effective, and have troublesome side effects that, when paired with the lack of insight that is characteristic of SZ, often result in non-compliance [7,8]. In fact, the Clinical Antipsychotic Trials of Intervention Effectiveness (CATIE) study, which investigated the effectiveness of five commonly used antipsychotics in a sample of 1460 patients with SZ, found that about 50% of patients discontinue their medications by 6 months and that 74% of patients discontinue their medications by 18 months [9]. Thus, it is imperative that we increase our understanding of the underlying etiology of SZ to facilitate the development of novel pharmacological agents that can target specific symptoms in this heterogeneous disorder.

The Genetics of Schizophrenia

Twin and multiplex family studies dating back to the 1940s have produced a range of heritability estimates for SZ. These family-based estimates of what one might describe as broad-sense heritability take into account the role of rare variants and were summarized in a widely quoted 2003 meta-analysis of SZ as a heritability estimate of 81% [10,11]. In contrast, single-nucleotide polymorphism (SNP)–estimated heritability is calculated by way of population-based genome-wide association
studies (GWAS) of common, additive genetic variation, and such studies, which do not take rare variants into account, have suggested that SZ heritability may be closer to 60–70% [10,12,13]. Note that when discussing the heritability of specific SZ endophenotypes later in this review, we are generally using broad-sense heritability.

Although sporadic, non-familial cases of SZ are often observed, family history remains the strongest and best replicated risk factor for developing SZ [14]. Compared with the ~1% prevalence in the general population, relatives of an individual with SZ have a significantly increased risk of developing the disorder—that is, the closer the relationship between a relative and an individual with SZ, the more genes that are shared and the higher the risk, which ranges from 2 to 4% for the second-degree relatives of individuals with SZ to 10 to 15% for the first-degree relatives of individuals with SZ [10]. Still, the concordance rate for monozygotic (MZ) twins, who share their entire genome, is only 45%. These studies support a strong role of genetics in SZ susceptibility but also suggest that environmental and other risk factors are involved.

Identifying the genetic variants captured by these heritability estimates has been challenging. Early efforts to map the genes contributing to SZ risk through family-based linkage and association studies implicated many chromosomal regions and suggested plausible candidate genes [15–19]. However, there have been difficulties in narrowing the source of the signal within linkage regions and in identifying functional variants within the implicated genes, and these difficulties have been compounded by inconsistent replication of findings from different samples. These issues are largely due to the fact that SZ is a complex disorder that results from the interaction of many genetic and non-genetic factors [1,20]. As such, there is no clear one-to-one relationship between genotype and phenotype. This is evident, for example, in family studies, wherein unaffected relatives often share risk variants with their affected family members. This reduced penetrance also affects case-control studies, as some portion of clinically unaffected individuals are likely carriers of SZ risk genes without expressing the phenotype. Heterogeneity also complicates gene identification, as different genes (i.e., genetic heterogeneity) or different variants of the same gene (i.e., allelic heterogeneity) may contribute to disease risk. Finally, sporadic, or non-familial, cases of SZ likely have different underlying genetic architecture than familial cases, yet both sporadic and familial SZ cases are combined in case-control studies of SZ. Such factors cause a significant reduction in the power of genetic studies, which has, in part, been addressed through more recent technological advances and the use of increasingly larger samples.
THE MODERN ERA OF GENOMICS IN SCHIZOPHRENIA

Genome-Wide Association Studies (GWAS)

The field of psychiatric genetics shifted to GWAS in 2007 with the introduction of microarray-based genotyping and the culmination of efforts to map the structure of the human genome. Such studies rely on allele frequency differences for SNPs in large samples of unrelated SZ cases and controls to implicate genomic regions containing risk genes. Although early SZ GWAS suffered from relatively small samples, combining results across studies has begun to produce replicable findings. For example, in a GWAS of 479 SZ cases and 2937 controls, followed by replication testing of 12 suggestive loci in 16,726 subjects, O’Donovan and colleagues identified \textit{ZNF804A} as a possible susceptibility gene for SZ [21]. An initial GWAS of 2681 cases and 2653 controls from the Molecular Genetics of Schizophrenia (MGS) sample by Shi and colleagues found no loci reaching genome-wide significance, yet the meta-analysis of 8008 cases and 19,077 controls identified 7 SZ-associated SNPs on chromosome 6p22.1, which represents the extended major histocompatibility complex (MHC) region [22]. Likewise, an initial GWAS by Stefansson and colleagues in 2663 cases and 13,498 controls produced no significant findings, yet follow-up genotyping and meta-analysis implicated the MHC and variants in the neurogranin (\textit{NRGN}) and transcription factor 4 (\textit{TCF4}) genes [23]. Another study by the International Schizophrenia Consortium confirmed the association to the MHC region and demonstrated that common genetic variants in aggregate may account for >30% of the risk for SZ [24].

More recently, the formation of the Psychiatric Genomics Consortium (PGC) as the largest consortium in the history of psychiatry has enabled mega-analyses of SZ. In a landmark publication of ~37,000 cases and ~113,000 controls, the PGC identified 108 distinct regions that met genome-wide significance, 83 of which had not been previously reported [25]. These findings included many genes that are biologically plausible candidates, such as \textit{DRD2}; several genes related to glutamate signaling; and the MHC region, yet the associated variants collectively only account for 3.4% of the risk for SZ. A considerable proportion of the observed heritability is thus not detectable in these very large studies of unrelated SZ cases and controls [26,27].

Next-Generation Sequencing (NGS) Studies

Another strategy in the genetic investigation of SZ has been the application of next-generation sequencing (NGS) methods that were pioneered in cancer research. Contrary to the GWAS focus on common variants, these methods seek to identify copy number variants (CNVs) and single-nucleotide variants (SNVs) that occur in less than 0.1% of the general population but that have elevated frequencies and functional roles in SZ. CNVs associated with SZ are generally large, recurrent
deletions or deletions encompassing regions that are highly damaging and that consequently have large effects on risk, with odds ratios ranging from 2 to 60. Studies have collectively identified SZ-associated recurrent CNVs in several chromosomal regions (e.g., 1q21.1, 3q29, 7q11.2, 15q13.3, 16p11.2, and 22q11), all of which are also associated with autism and intellectual disability, suggesting a shared genetic architecture between these disorders [28–32]. These CNVs tend to involve neurodevelopmental pathways, and genes in the neuregulin and glutamate pathways are disproportionately disrupted in SZ [29,32,33]. Although a recent study by the PGC, which represents the largest genome-wide analysis of CNVs for any psychiatric disorder, confirmed the presence of SZ-associated CNVs in these regions, the identified loci were present in only 1.4% of the 21,094 SZ cases and account for only ~0.85% of the heritability of SZ [33].

Whole-exome sequencing (WES) studies attempt to identify SNVs in the coding sequence of genes that may contribute to risk for SZ by examining frequency differences between cases and controls through mutational burden analyses. The extremely high rate of background mutation complicates such studies, which to date have been underpowered and have generally failed to provide insight into the mechanism of illness. For example, in a WES study of 166 cases and 307 controls with follow-up genotyping of novel, potentially damaging variants in an independent sample of 2617 cases and 1800 controls, no variant reached study-wide significance [34]. Another WES of 2536 cases and 2543 controls with a focus on gene-based tests of disruptive variants similarly failed to find significance, although this study importantly confirmed a polygenic contribution of very rare (<1/10,000), damaging mutations to SZ risk [35]. Likewise, exome-sequencing studies in subsets of SZ subjects have been somewhat informative. For example, in an exome-sequencing study that investigated subjects with sporadic SZ, Gulsuner and colleagues found that de novo loss-of-function mutations occur more frequently in cases with sporadic SZ than in controls [36]; more recently, the findings in this study were supported by Yang and colleagues [37].

**Gene-Expression Studies**

The human genome is made up of DNA that is organized into 20,000 to 25,000 genes. Nearly every human cell contains the same genes, but different cells and tissue have different patterns of gene expression. These differences in gene expression may contribute to diseases like SZ.

Before each gene is expressed, the DNA must first be copied, or transcribed, into messenger RNA (mRNA). These transcriptions, which are collectively referred to as a transcriptome, control the function of cells. Several groups have explored the use of peripheral blood gene-expression signatures [38] and used transcriptomes to investigate the biological consequences of SZ-associated genes; these studies have
suggested a potential role for ARC, NMDA receptors, synapse and immunity dysfunction, and CNS development in SZ [39].

**Epigenetic Studies**

Epigenetic “marks” in DNA change how genes are expressed in the process of DNA transcription to RNA and subsequent translation to protein. Environmental factors such as smoking, maternal starvation, and in utero viral infection can interact with genes via processes such as methylation and histone modification to change gene expression in the developing brain. These functional changes can be stable and have long-lasting chromatin modifications that regulate both local and genome-wide gene activity.

DNA methylation is a normal biochemical process whereby atoms are added to DNA as part of gene expression, but environmental factors can lead to the formation of differentially methylated genetic regions. Some studies have demonstrated an enrichment of such changes in SZ-related gene networks [40]. For example, one study showed that SZ-associated methylation differences were observed at 923 cytosine-phosphate-guanine (CpG) islands and that some of the replicated differentially methylated positions overlapped with top-ranked SZ regions from GWAS [41].

Nucleosomal histone modifications contribute to genome organization and function, with various histone methylation and acetylation marks (including H3-trimethyl-lysine 4 [H3K4me3] and H3-acetyl-lysine-27 [H3K27ac]) serving as key regulators for active gene promoters and enhancers [42]. Studies of histone modification are ongoing in SZ [43], and publicly accessible resources are available, including results from the dorsolateral prefrontal cortex and anterior cingulate cortex of SZ brains and UCSC browser visualizations for cell-specific maps (see [http://www.psychencode.org/](http://www.psychencode.org/) or [https://www.synapse.org/#!Synapse:syn4566010](https://www.synapse.org/#!Synapse:syn4566010)). Given the relative ease that histone posttranslational modifications can be observed in nearly all somatic tissues, this type of epigenetic information could serve as a useful marker if an association is found between histone modifications and SZ.

**Summary**

Although SZ is highly heritable, identifying risk genes has proven difficult due to the complex nature of the disorder. Large-scale GWAS employing a case-control design have detected many common variants associated with SZ diagnosis. However, genome-wide significant variants account for only a small portion of the overall risk, the effect size of each individual allele is very small, and the pathways by which these variants impact risk remain largely unknown. Studies of rare variation have shed light on possible underlying pathways, yet CNVs and SNVs account for only a small percentage of SZ cases. Likewise, although emerging technologies in gene-expression and epigenetic studies may ultimately be
fruitful, the challenge of translating genetic findings to precision medicine in SZ still persists.

THE VALUE OF DEEP PHENOTYPING IN SCHIZOPHRENIA RESEARCH

A key limitation faced by researchers in SZ is that the heterogeneous clinical phenotype complicates the diagnosis, therefore diluting the genetic signal. To address this issue, Gottesman and Shields introduced the concept of “endophenotypes” in SZ [44], and the application of endophenotypes to the study of SZ has become known as “deep phenotyping.” These state-independent traits are associated with illness, are observed in unaffected family members at a higher rate than in the general population, and co-segregate with illness in families [44]. To be useful in genetic studies, an endophenotype must also be significantly heritable, reliably measured, and stable over time [45,46]. Given that endophenotypes fall within the genotype-to-phenotype pathway of the disease process, they form a critical link between genetic risk variants and the more direct lower-level biological processes that contribute to observable symptoms [47,48]. In other polygenic, complex conditions, like hypercholesterolemia, endophenotypes (i.e., cholesterol) have been particularly useful in helping to characterize the underlying biological dysfunction and to develop targetable treatments [49].

Endophenotypes are quantitative in nature and reflect a continuous distribution of trait values in the general population, consistent with polygenic theories of SZ [50–52]. Under this scenario, as detailed in Figure 1, individuals carrying non-penetrant risk variants do not diminish a linkage signal, as with an analysis of diagnosis. Indeed, SZ patients, unaffected relatives, and controls provide a range of scores and are thus informative regarding the underlying liability. Therefore, by using endophenotypes to measure the underlying liability to SZ as a quantitative trait, we gain significant power and increased resolution compared to analyses of diagnosis, which use a somewhat subjective threshold to dichotomize a continuous trait [53,54]. For a disorder like SZ, which has a prevalence of ~1% and heritability of 60–80%, the analysis of a quantitative endophenotype is expected to be approximately 100-fold more efficient than the analysis of the categorical diagnosis, which translates to a 10-fold increase in power [53]. The analysis of an endophenotype thus requires a substantially smaller sample size to achieve comparable power to an analysis of SZ diagnosis. By enriching for individuals with higher liability to SZ, who presumably carry more risk alleles, family-based designs provide a further increase in power [55].
Figure 1. The liability threshold model. Endophenotypes capture the full spectrum of information in liability distribution, whereas diagnosis dichotomizes the distribution into “affected” and “unaffected” according to a threshold defined by subjective symptom profiles.

NEUROPHYSIOLOGICAL AND NEUROCOGNITIVE ENDOPHENOTYPES IN SCHIZOPHRENIA

Neurocognitive dysfunction is a hallmark of SZ, leading to impaired quality of life and poor functional outcome [56–58]. As such, neurocognitive deficits are endorsed by the Measurement and Treatment Research to Improve Cognition in Schizophrenia (MATRICS) and the Food and Drug Administration (FDA) as targets for the development of new treatments for SZ [59,60]. Data from the North American Prodrome Longitudinal Study (NAPLS) has further demonstrated that measures of neurocognitive functioning in SZ may have predictive utility in identifying clinically high-risk subjects who will convert to psychosis [61,62]. Some researchers have suggested that the neurocognitive deficits in attention, learning, and memory that characterize SZ may arise from a primary impairment in the filtering or “gating” of sensory information [63]. According to this hypothesis, an inability to filter out irrelevant external stimuli leads to the psychotic symptoms (e.g., hallucinations, delusions) and cognitive disorganization observed in SZ. Therefore, studying an array of endophenotypes that range from largely automatic (i.e., preattentive) neurophysiological measures to highly volitional (i.e., attentive) neurocognitive measures may provide a window into the biological processes underlying SZ. Examples of several promising endophenotype paradigms are discussed below and shown in Figure 2.

Many neurophysiological and neurocognitive endophenotype measures exhibit medium-to-large deficits in SZ, exhibit moderate-to-substantial stability, and are independent of clinical state [64–67]. Here, we focus on three widely used neurophysiological, inhibitory endophenotypes that have been shown to rank highly across such
parameters: prepulse inhibition, mismatch negativity, and oculomotor antisaccade [64]. Given the large body of literature describing neurocognitive measures and SZ, we describe a few representative tests with a focus on working memory and attention. Reviews of additional endophenotype paradigms that have been used in genomic and clinical-outcome studies of SZ [64–67], including neuroimaging studies [68,69], can be found elsewhere. Figure 3 demonstrates the large effect sizes of some potentially promising endophenotypes, which are described below.

Figure 2. The range of SZ-related endophenotypes from automatic to volitional. PPI and MMN are neurophysiological endophenotypes representing automatic, or preattentive, processes that can be elicited with no active participation from the test subject. PPI is measured as the percent inhibition (i.e., gating) of the startle reflex in response to a weak prestimulus. MMN is measured as the difference between the waveforms for a standard stimulus ERP and a deviant stimulus ERP. Neurocognitive endophenotypes like CPT and LNS are volitional, or attentive, and require active participation from the subject. CPT measures sustained attention as the signal/noise discrimination ratio \( d' \) of blurred target stimuli (in the DS-CPT) or pairs of identically presented stimuli (in the CPT-IP). LNS measures working memory as the correct reordering of a verbally presented list of intermixed numbers and letters. AS is measured as the ratio of correct antisaccades (i.e., looking in the opposite direction of the presented stimulus) to total interpretable saccades (i.e., eye movements in any direction). Although AS is a neurophysiological measure, subjects are required to actively participate and control their response. AS: antisaccade; CPT: continuous performance test; DS: degraded stimulus; ERP: event-related potential; IP: identical pairs; LNS: letter-number sequencing; MMN: mismatch negativity; PPI: prepulse inhibition; SZ: schizophrenia.

Prepulse Inhibition (PPI)

PPI of the acoustic startle reflex is a measure of sensorimotor gating that may reflect the biological processes underlying gating impairments in SZ. Compared to healthy controls, patients with SZ fail to inhibit their startle response (i.e., eye blinking) when they hear a startle-eliciting stimulus (i.e., a pulse) after hearing a weaker or softer stimulus that precedes the pulse by 30-to-120 ms (i.e., the prepulse). There is a consistent body of literature demonstrating robust PPI deficits in SZ that are relatively stable across time [64,66,70,71], despite the emergence and nearly ubiquitous use of atypical antipsychotics, which enhance PPI and make deficits more difficult to detect and quantify [72,73]. PPI deficits are also observed in unaffected relatives of SZ patients and are significantly heritable in SZ families (45–50%), suggesting that PPI may be a marker of risk for SZ rather than a consequence of illness [66,74–76]. Furthermore, while all SZ patients demonstrate comparable deficits in PPI regardless of family history of illness, the broad-sense heritability of PPI deficits is substantially increased in families with a higher genetic vulnerability to SZ (47%) compared to that observed when family history for illness is not considered (29%) [77].

Mismatch Negativity (MMN)

MMN is a component of the event-related potential (ERP) that is elicited when an infrequently presented “deviant” stimulus occurs in a sequence of repetitive “standard” auditory stimuli [78]. The deviant sound can differ in one or more perceptual features, such as duration, frequency, or intensity. MMN refers to the difference between the waveforms for the standard stimulus ERP and the deviant stimulus ERP and usually peaks between 150 and 250 ms after the presentation of the
deviant stimulus [79]. Like PPI, MMN can be elicited regardless of whether a subject is paying attention to the sequence, and it is thus considered an index of preattentive sensory processing. Reduced MMN amplitude is one of the most robust findings in SZ [80], with a large effect size of about 1.0 that is relatively unaffected by medication and remains stable across time and fluctuations in clinical symptoms [64,81–84]. MMN is also 68% heritable and is present in individuals at genetic risk for developing SZ [85–90]. MMN is also associated with psychosocial functioning and has shown promise for predicting the onset of psychosis in high-risk populations [91–95]. Finally, we note that the P3a, another measure of early auditory information processing, is a promising endophenotype that is often measured in tandem with MMN [64,84,96].

**Oculomotor Antisaccade (AS)**

The AS task is a widely used electrophysiological measure of inhibitory failure, which is an important cause of the cognitive and clinical abnormalities found in SZ [97]. This task requires participants to fixate on a central target and respond to a peripheral cue by looking in the opposite direction at the same distance; it is measured as the ratio of correct to total interpretable saccades (0 to 1). Compared to healthy controls, SZ patients display significant deficits on an overlap version of the AS task, with a large effect size of about 1.0 [64,98,99]. Studies have also found that AS performance, latency, and gain are stable measures of pathophysiological disturbances in SZ [100,101]. In addition, large family and twin studies have shown that the AS task has a high heritability of 42–57% [74,102] with documented deficits in first-degree relatives of SZ patients [103,104]. These data suggest that AS performance represents a useful endophenotype for SZ.

**Letter-Number Sequencing (LNS)**

Working memory is the ability to maintain and manipulate the internal representation of a stimulus in transient memory [105]. Deficits in working memory have been extensively documented and are profoundly affected and persistent in SZ [106] and in relatives of patients with SZ [107]. These deficits are thought to reflect the fundamental cognitive disturbances that are associated with SZ [108]. A subtest of the well-standardized and normed Wechsler Memory Scale III [109], the LNS is a prototypical measurement of verbal working memory information storage with manipulation. Administration of this task involves the verbal presentation of a series of intermixed letters and numbers, increasing in difficulty from 2 stimuli to 8 stimuli as subjects advance in the task. In the reorder condition of the LNS, subjects repeat the numbers in ascending order followed by the letters in alphabetical order, which provides an assessment of executive functioning and working memory. The LNS can also be administered in the forward condition, wherein subjects simply repeat the numbers and letters in the order presented,
which provides an assessment of transient online storage and retrieval. SZ patients typically perform worse (i.e., obtain lower scores) than controls for both conditions, but larger differences are usually observed for the reorder condition. The LNS ranks highly among neurocognitive endophenotypes, as it demonstrates large SZ-related deficits and longitudinal stability [64]. The LNS is also highly heritable, both individually (39%) and as part of a composite measure of working memory (27%) [74,110].

**Continuous Performance Test (CPT)**

Deficits in attention were included in Bleuler’s first descriptions of SZ, and such deficits remain a key nosological feature of the disorder, as well as a primary area of emphasis in SZ research [2,65]. More recent studies of endophenotypes in SZ have narrowed in on deficits in sustained focused attention, which are most commonly measured using CPTs. These tests typically take 15 min to administer and require subjects to monitor a series of target stimuli (e.g., letters, digits, shapes) for 30–100 ms between periods of vigilance during which no stimuli are presented. Several different versions of the CPT are available to assess various aspects of sustained focused attention and vigilance. For example, the Degraded Stimulus CPT (DS-CPT) is associated with a high perceptual load and involves the identification and response to highly blurred, black and white target stimuli [111,112], whereas the Identical Pairs version of the CPT (CPT-IP) is associated with a working memory load and requires the subject to respond each time a string of 3 or 4 digits is presented twice in a row [113,114]. SZ patients consistently show large performance deficits on CPTs, with meta-analyses reporting a mean effect size of 1.18 [115]. These deficits are typically not affected by psychotropics [116,117] and are relatively persistent across time and clinical status [113,118]. Deficits in CPTs are also reliably detected in the unaffected relatives of SZ patients, with moderate heritability in healthy families (39–49%) [113] and higher heritability observed in SZ families (38–79%) [74,119,120].

**GENETIC STUDIES OF ENDOPHENOTYPE DEFICITS IN SCHIZOPHRENIA**

Despite the increasing use of endophenotypes in the study of SZ, few large-scale genetic studies of endophenotype deficits in SZ have been published. In the largest genome-wide study of PPI to date, investigators from the Learning on Genetics of SZ Spectrum (LOGOS) identified and replicated common variants in the *NGF* and *CALN1* genes that reached genome-wide significance [121]. Although this study only included healthy males, polygenic risk scores calculated based on the PGC SZ cohort showed that an increased risk for SZ was associated with PPI deficits. Subsequent meta-analyses have offered additional evidence to support an abnormal neurotrophin profile in SZ, with the observance of decreased peripheral blood *NGF* levels in SZ patients [122], and have
highlighted the polygenicity of PPI, suggesting the relevance of specific variants in \textit{COMT}, \textit{GRIK3}, \textit{PRODH}, and \textit{TCF4} [123]. A recent GWAS of pursuit initiation and AS error rate conducted by the Bipolar-Schizophrenia Network on Intermediate Phenotypes (B-SNIP) consortium implicated genes involved in nuclear trafficking and gene silencing (\textit{IP08}), axonal guidance and synaptic specificity (\textit{PCDH12}), transduction of nerve signals (\textit{NRSN1}), retinal degeneration (\textit{LMO7}), and synaptic glutamate release (\textit{SH3GL2}) in the regulation of oculomotor inhibition [124].

Other studies have examined the relationship between neurocognitive functioning and polygenic risk for SZ in large, community-based samples. For example, one study found that higher risk for SZ was associated with lower neurocognitive functioning in a very large sample of community-dwelling adults from the UK Biobank (\(N = 112,151\)) [125]. These findings are consistent with those of the Avon Longitudinal Study of Parents and Children (ALSPAC), who found that higher risk for SZ was associated with lower IQ [126]. Similarly, the Cognitive Genomics Consortium (COGENT) reported a genetic correlation of \(-0.17\) between general cognitive function and SZ in their population-based sample of >35,000 individuals [127]. Although these studies suggest that a portion of the genetic risk for SZ may be mediated through genetic effects on neurocognitive functioning, the reliance on assessments of IQ and global cognitive summaries in population-representative samples provides little insight into the neurobiological mechanisms that may mediate the genetic overlap between cognitive functioning and SZ. Large genetic studies of neurocognitive deficits in the context of SZ are thus needed to identify the genes contributing to this shared genetic architecture.

The Consortium on the Genetics of Schizophrenia (COGS) has conducted candidate gene and linkage analyses of many SZ-related endophenotypes, including those reviewed here. The COGS results suggest that the genes contributing to the specific endophenotypes reviewed here converge on a functional gene network, which is summarized in Figure 4. In association studies that genotyped 130 families using a custom genotyping array that evaluated 94 candidate genes for SZ, the COGS identified 71 SNP-endophenotype associations that involved 28 genes and 1 or more of the endophenotypes discussed in this review and that were strong enough to satisfy an experiment-wide significance level of 0.05, after accounting for linkage disequilibrium, endophenotype correlations, family structure, gene size, and multiple testing. Furthermore, each of the associated genes shown in Figure 4 contained at least one SNP with an \textit{a posteriori} chance of \(\geq 84\%\) of being a true finding of association, indicating that the observed associations most likely represent true positive results. Additionally, these studies identified 8 genes (\textit{CTNNA2}, \textit{ERBB4}, \textit{GRID2}, \textit{GRIK3}, \textit{GRIK4}, \textit{NOS1AP}, \textit{NRG1}, and \textit{RELN}) with pleiotropic associations across multiple endophenotypes in the COGS sample, as well as in an independent case-control study [128–130]. These 8 genes are also associated with at least two of the
endophenotypes that are the focus of this review and span neurophysiological and neurocognitive domains. Linkage analyses of a larger sample of 296 COGS families using a genome-wide SNP array identified many regions of at least suggestive evidence for linkage, with several candidate and other genes of interest located beneath the linkage peaks, including ADORA2A, ATXN7, CSMD1, GRIK2, GRIN2B, GRIN3A, HTR2A, and SLC6A3 [131].

![Ingenious Pathway Analysis](image-url)

**Figure 4.** Ingenuity Pathway Analysis of the genes identified through association and linkage analyses of the endophenotypes discussed in this review. Genes are represented as nodes with solid lines representing direct protein-protein interactions and dashed arrows representing more indirect effects on expression, activation, or inhibition. Genes and interactions in black show association (blue box) or linkage (red box) with the endophenotypes listed in this review in the COGS study, whereas those in gray represent other interacting genes. Genes in a bold font are associated with at least two endophenotypes from this review and span neurophysiological and neurocognitive domains. Genes identified as associated with SZ diagnosis by GWAS of common variants are indicated with a “C,” whereas those identified by studies of rare and de novo variation are indicated with an “R.” *GRIN1* showed evidence of both association and linkage with other neurocognitive endophenotypes in the COGS study, and both DISC1 and GRID2 showed evidence of linkage with neurophysiological endophenotypes in the COGS study. Note that CSMD1 and CTNNA2 were prominent genes in these analyses, with external support from studies of SZ diagnosis, but do not directly connect into the defined network. COGS: Consortium on the Genetics of Schizophrenia; GWAS: genome-wide association study; SZ: schizophrenia.

Studies of common, rare, and *de novo* variation in SZ offer independent support for many of the genes implicated through linkage or association studies of SZ-related endophenotypes. For example, a recent large GWAS of SZ diagnosis conducted by the PGC and a GWAS of neurocognitive endophenotypes for SZ identified associations with
ATXN7, CSMD1, DRD2, GRIN2A, GRIN3A, and GRM3 [132,133]. Other studies have shown a significant burden of rare variants in GRIN1, GRIN3A, and SLC1A2 in SZ patients [35,134], as well as a disproportionate disruption of genes in the neuregulin and glutamate pathways [29]. Studies of de novo variation have implicated CSMD1, CTNNA2, DBH, DISC1, GRID2, GRIN2B, HTR2A, RELN, and SLC6A3 in SZ and related disorders [135–143].

As Figure 4 demonstrates, genetic analyses of both SZ-related endophenotypes in family-based samples and large SZ case-control studies implicate glutamatergic neurotransmission and synaptic plasticity pathways in mediating SZ susceptibility, which is consistent with the glutamate hypothesis and an abundance of evidence implicating NRG1-ErbB4 signaling in SZ [144–146]. For a comprehensive list of genes that have been associated with SZ, see the summary of genetic findings compiled by Ren and colleagues [147].

TRANSLATIONAL STUDIES OF SCHIZOPHRENIA-ASSOCIATED DEFICITS

Cross-species investigations are important for the development of novel therapeutic interventions in SZ. For this reason, the endophenotype strategy has been advocated by the National Institute of Mental Health as part of the Research Domain Criteria (RDoC) framework [148]. Several endophenotypes are amenable to translational studies, which use intrinsic or induced deficits in animal models to mimic those observed in humans and thereby provide important insight into the neurobiological processes underlying SZ.

PPI is one example of a SZ-related endophenotype that can be reliably measured across species to understand the biology of brain-based inhibitory mechanisms [149]. PPI deficits in rats can be induced through the administration of dopamine (DA) agonists like apomorphine (APO). In one study of two rat strains displaying varying sensitivity to APO, 104 genes were found to significantly differ in their expression of the nucleus accumbens, which represents a control center for PPI. Many of these genes correlated with PPI deficits and were clustered within the DA receptor-signaling, synaptic long-term potentiation (involving glutamate transmission), and inositol phosphate metabolism pathways, suggesting that these pathways may substantially influence PPI and may therefore represent therapeutic targets [150]. By examining these findings in the context of human genetic studies, we can begin to piece together how various pathways may be involved in each specific endophenotype. For example, of the 12 human genes associated with PPI-related deficits in SZ patients and their families, 8 show significant differences in gene expression related to PPI in rats [128–130,150,151].

Some neurocognitive endophenotypes have also been investigated in animal models. Given the hypothesis that neurocognitive dysfunction in SZ results from impaired stimulus filtering and that this attentional
impairment may relate to reduced expression of muscarinic acetylcholine receptors in SZ, some investigators have suggested that mice treated with scopolamine, a muscarinic receptor antagonist, may provide a model of attentional dysfunction in SZ. Using a five-choice version of the CPT (5C-CPT) developed specifically for rodents, the investigators found that mice treated with scopolamine exhibited significantly impaired attention that was characterized by reduced target detection and increased impulsive responding [152]. These findings are comparable to those observed in humans with SZ using a reverse-translated version of the 5C-CPT, which was designed to minimize demands on perceptual, visual learning, processing speed, or working memory functions and which has been validated against the CPT-IP [152]. Given that the FDA has endorsed neurocognitive deficits as a key treatment target, these and other related findings hold promise to facilitate the development of pro-cognitive therapeutics that may improve functioning in SZ patients.

In the future, investigations of the transcriptome, DNA methylation, and histone modification may also have translatable potential. These kinds of studies represent a promising area for SZ research, but such methods have not yet been entirely incorporated into SZ endophenotype studies. It is likely that the integration of these gene expression and epigenetic resources with endophenotype-related animal models may be informative in future studies.

CONCLUSIONS AND FUTURE DIRECTIONS

SZ is a clinically heterogeneous disorder in which patients exhibit a broad range of neurobiological deficits and symptom severity, and this profound heterogeneity has complicated efforts to identify genetic risk variants. The use of SZ-related endophenotypes, rather than diagnosis, in genetic studies offers several advantages that may help address these complications and facilitate the identification of genetic risk variants and aberrant molecular pathways [153–155]. Traditional case-control studies collapse the substantial phenotypic variation in SZ into a single, categorical diagnosis. However, recent studies have shown that phenotypic specificity is actually more important than a large sample size for detecting true genetic associations in genetic studies of complex diseases, including SZ [156,157]. Although the use of large samples remains important in SZ research, endophenotypes can provide the needed phenotypic specificity, as well as increased biological relevance. For a complex disorder like SZ, the genetic signal from studies that use endophenotypes rather than diagnosis should thus be stronger, more direct, and associated with fewer variants, each of larger individual effect. As each endophenotype reflects a discrete neurobiological function or pathway, they also may be useful in parsing the complex etiology of SZ and reducing the associated heterogeneity. Finally, by capitalizing on the inherent power advantages of a quantitative trait, the analysis of an endophenotype requires a substantially smaller sample

size, nearly 10-fold, to achieve comparable power to an analysis of SZ diagnosis. Although assessing endophenotypes may incur an additional burden with regard to recruitment and testing, the potential payoff suggests that a deep phenotyping strategy is well worth the effort.

Genetic analyses of endophenotypes can also supplement traditional analyses of SZ diagnosis, providing additional biological insight into these results. One simple way to utilize endophenotypes is to inform existing large-scale GWAS of SZ diagnosis [132]. Using enrichment and polygenic scoring methods, we can assess how much of the common genetic risk is shared between SZ and a particular endophenotype [24,158]. In addition to estimating the overall genetic correlation between a given endophenotype and SZ, we can also investigate whether SNPs that increase the risk for SZ overlap more than we would expect by chance with SNPs that result in deficits in, for example, PPI or neurocognition.

Endophenotypes are also ideally suited to multivariate phenotype methods, which can be used to construct individual phenotypic profiles and to identify subgroups of similar SZ patients [159]. Methods that can simultaneously incorporate both phenotype and genomic information may prove particularly useful in this regard [160]. The identified subgroups can then be correlated with clinical factors, such as illness severity, clinical course, level of functioning, and medication response, to establish individualized treatment protocols. By investigating the relationship of clinical, genetic, and environmental factors to subgroups defined by endophenotypes, which reflect the underlying neurobiological dysfunctions, we may gain a better understanding of Bleuler’s “Group of Schizophrenias.”

The endophenotype strategy thus complements large-scale case-control efforts, providing the necessary biological insight to facilitate precision medicine in SZ. Ultimately, we may be able to distill the number of endophenotypes to a smaller robust subset that accounts for a majority of the variance between cases and controls [161]. Using endophenotypes to dig deeper, going beyond the broad clinical diagnostic category of SZ, will facilitate the identification of pathways leading from genetic variation to brain dysfunction. A better understanding of the genetic substrates underlying deficits in key domains defined by SZ-related endophenotypes will provide a molecular roadmap to guide the development of novel pharmacological interventions aimed at improving real-world functioning, as well as information to guide precision-based treatment decisions.

AUTHOR CONTRIBUTIONS

TAG and DWT conceptualized the paper. All authors performed key literature reviews to describe the relationships between endophenotypes, genetics, and SZ. All authors contributed to the organization and the writing of the paper.
CONFLICTS OF INTEREST

The authors declare that there are no conflicts of interest.

FUNDING

This material is based upon work supported in part by Merit Review Award I01 BX002241 from the US Department of Veterans Affairs Biomedical Laboratory Research and Development Service at the VA Puget Sound Health Care System. The authors were previously supported by R01 MH065571 and R01 MH065558 from the National Institute of Mental Health.

REFERENCES


29. Walsh T, McClellan JM, McCarthy SE, Addington AM, Pierce SB, Cooper GM, et al. Rare structural variants disrupt multiple genes in neurodevelopmental pathways in schizophrenia. Science. 2008 Apr 25;320(5875):539-43. doi: 10.1126/science.1155174


94. Perez VB, Woods SW, Roach BJ, Ford JM, McGlashan TH, Srihari VH, et al. Automatic auditory processing deficits in schizophrenia and clinical high-


118. Nuechterlein K, Dawson M, Ventura J, Yee-Bradbury C. Longitudinal stability of vigilance and span of apprehension deficits in the early phase of schizophrenia. In Sixth Annual Meeting of the Society for Research in Psychopathology; 1991 Dec; Cambridge, USA.


134. Fiorentino A, Sharp SI, McQuillin A. Association of rare variation in the glutamate receptor gene SLC1A2 with susceptibility to bipolar disorder and


156. Liang SG, Greenwood TA. The impact of clinical heterogeneity in schizophrenia on genomic analyses. Schizophr Res. 2015 Feb;161(2-3):490-5. doi: 10.1016/j.schres.2014.11.019


How to cite this article: