Genetic architecture of schizophrenia: a review of major advancements

Sophie E Legge[1*], Marcos L Santoro[2,3*], Sathish Periyasamy[4,5*], Adeniran Okewole[6*], Arsalan Arsalan[7*], Kaarina Kowalec[8,9*]

[1] MRC Centre for Neuropsychiatric Genetics and Genomics, Division of Psychological Medicine and Clinical Neurosciences, Cardiff University, Cardiff, United Kingdom

[2] Departamento de Morfologia e Genética, Universidade Federal de Sao Paulo, Sao Paulo, Brazil

[3] Laboratory of Integrative Neuroscience, Departamento de Psiquiatria, Universidade Federal de Sao Paulo, Sao Paulo, Brazil

[4] Queensland Brain Institute, The University of Queensland, Brisbane, QLD 4072, Australia

[5] Queensland Centre for Mental Health Research, The University of Queensland, Brisbane, QLD, Australia

[6] Neuropsychiatric Hospital Aro, Abeokuta, Nigeria

[7] Department of Pharmacy, University of Peshawar, Peshawar, Pakistan

[8] Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden

[9] College of Pharmacy, University of Manitoba, Winnipeg, Canada

[*] All authors made an equal contribution

Corresponding author: Kaarina Kowalec, PhD, University of Manitoba, Rady Faculty of Health Sciences, 750 McDermot Avenue, University of Manitoba, Winnipeg, Manitoba, Canada R3E 0T5 (204) 272-3104, kaarina.kowalec@ki.se.

Manuscript word count: 4,438/5,000

Abstract word count: 230/250

Abstract

Schizophrenia is a severe psychiatric disorder with high heritability. Consortia efforts and technological advancements have led to a substantial increase in knowledge of the genetic architecture of schizophrenia over the past decade. In this article, we provide an overview of the current understanding of the genetics of schizophrenia, outline remaining challenges, and summarise future directions of research. World-wide collaborations have resulted in genome-wide association studies (GWAS) in over 56,000 schizophrenia cases and 78,000 controls, which identified 176 distinct genetic loci. The latest GWAS from the Psychiatric Genetics Consortium, available as a pre-print, indicates that 270 distinct common genetic loci have now been associated with schizophrenia. Polygenic risk scores can currently explain around 7.7% of the variance in schizophrenia case-control status. Rare variant studies have implicated 8 rare copy-number variants, and an increased burden of loss-of-function variants in SETD1A, as increasing the risk of schizophrenia. The latest exome sequencing study, available as a pre-print, implicates a burden of rare coding variants in a further 9 genes. Gene-set analyses have demonstrated significant enrichment of both comman and rare genetic variants associated with schizophrenia in synaptic pathways. To address current challenges, future genetic studies of schizophrenia need increased sample sizes from more diverse populations. Continued expansion of international collaboration will likely identify new genetic regions, improve fine mapping to identify causal variants, and increase our understanding of the biology and mechanisms of schizophrenia.

Introduction

Schizophrenia is a severe and often chronic psychiatric disorder causing substantial personal and societal burden from severe and long-term disability. Schizophrenia is characterised by positive (e.g. hallucinations, delusions), negative (e.g. alogia, avolition, anhedonia), and disorganized symptoms (e.g. speech, behaviour) (Dollfus & Lyne, 2017), and is also associated with cognitive impairment. The typical onset of schizophrenia occurs in late adolescence to early adulthood, with a peak in prevalence around 40 years of age (Charlson et al., 2018). Individuals with schizophrenia have a 15-20 years shorter life expectancy in comparison to the general population (Tanskanen, Tiihonen, & Taipale, 2018).

Schizophrenia is highly heritable and our understanding of the genetic architecture of schizophrenia has greatly increased over the past decade. This progress has arisen primarily through advances in molecular genetics technology, making feasible and affordable large scale genotyping and sequencing, and through world-wide collaboration amalgamating research samples, exemplified by the Psychiatric Genetic Consortium (PGC). This collaborative effort has led to the identification of hundreds of common risk variants, rare variants and copy number variants (CNVs), providing key insights into the biological basis of schizophrenia (*Figure 1*).

Prevalence and incidence

The global age-standardised point prevalence of schizophrenia is estimated to be 0.28% (95% uncertainty interval: 0.24-0.31) (Charlson et al., 2018). Globally, the number of prevalent schizophrenia cases increased from 13.1 million in 1990 to 20.9 million cases in 2016, with the largest increase occurring in Eastern sub-Saharan Africa and North Africa/Middle East triggered by high population growth in those areas (Charlson et al., 2018). The incidence of schizophrenia has been reported to be 15.2/100,000 persons, with a median male: female rate ratio of 1.4 (McGrath, Saha, Chant, & Welham, 2008).

Although a low prevalence disorder, schizophrenia plays a major role in the global burden of disease, contributing approximately 13 million years of life lived with disability (Charlson et al., 2018). Approximately 1 in 7 schizophrenia cases recover, based on clinical and social functioning indicators (Jaaskelainen et al., 2013). Poor outcomes are common and include premature mortality, long-term hospitalisation, treatment-resistance, and poor quality of life. The risk of suicide in those with schizophrenia is very high, with estimates of ~1 in 3 people with schizophrenia attempting suicide during their lifetime (Pompili et al., 2007). Individuals with schizophrenia are also subject to social disability (i.e. relationships with others, self-care), with research evidence including over 15 years of follow up time found that 25% have severe social disability (Wiersma et al., 2000).

Individuals affected by schizophrenia experience a two- to threefold increased risk of death (median standardised mortality ratio for all-cause mortality = 2.6) (McGrath et al., 2008). Reasons for the excess deaths in schizophrenia include a high suicide and accident rate, poor health behaviors (e.g. smoking, substance abuse, obesity), adverse effects of antipsychotics, and under-diagnosis or under-treatment of general medical comorbidities (Bitter et al., 2017; Jerrell, McIntyre, & Tripathi, 2010; Limosin, Loze, Philippe, Casadebaig, & Rouillon, 2007; Olfson, Gerhard, Huang, Crystal, & Stroup, 2015).

Environmental factors

Environmental factors act throughout life to influence the likelihood of disease, are not necessarily causative factors, but can be largely grouped by timing into three categories: early development, proximal, and onset factors (Stilo & Murray, 2019). Early development factors include obstetric complications (Cannon, Jones, & Murray, 2002), and advanced paternal age (Sipos et al., 2004). Proximal factors include social adversity, migration, and urbanicity. Meta-analyses have demonstrated migration as an important risk factor for schizophrenia in first- and second-generation migrants (Bourque, Van Der

Ven, & Malla, 2011; Cantor-Graae & Selten, 2005), with a greater risk in migrants from developing countries (Cantor-Graae & Selten, 2005); pointing towards a role for psychosocial adversity in schizophrenia aetiology. Lastly, factors occurring around the time of schizophrenia onset include primarily drug abuse, trauma, and social adversity (Stilo & Murray, 2019).

Heritability

Heritability is a measure of the proportion of variation of a trait that is attributable to genetic inheritance (Young et al., 2018). Heritability estimates of schizophrenia are based on family studies (e.g. familial aggregation and twin studies) and vary based on the study methodology. While estimates have ranged from 41-87% (Chou et al., 2017), the current estimate of heritability of schizophrenia is approximately 80% (Owen, Sawa, & Mortensen, 2016).

Common genetic variation

Genome-wide association studies (GWAS) investigate millions of common genetic variants (or single nucleotide polymorphism, SNP) simultaneously to determine their association with a trait. Many of the early GWAS of schizophrenia were underpowered to identify common SNPs, which have typically been shown to have small effects in schizophrenia. Including the largest GWAS to date from the PGC, which is currently available as a pre-print, over 300 independent genome-wide significant variants ($p \leq 5 \times 10^{-8}$) have been associated with schizophrenia. *Figure 2* visualizes the history of schizophrenia GWAS reported by key publications. In European populations, common variant associations identified from GWAS explain around one third (24%) of genetic liability to schizophrenia (Lee et al., 2012; Ripke et al., 2013; Schizophrenia Working Group of the Psychiatric Genomics Consortium, 2014).

The first reported genome-wide significant loci associated with schizophrenia came in 2009 (International Schizophrenia Consortium et al., 2009; Shi et al., 2009; Stefansson et al., 2009). In the

subsequent decade, efforts by the PGC have spearheaded much of the common variant discoveries in schizophrenia genetics. The first GWAS by the PGC identified significant associations in seven loci, five of which were novel and two of which had been previously implicated (Schizophrenia Psychiatric Genome-Wide Association Study (GWAS) Consortium, 2011). The major breakthrough for schizophrenia GWAS came from the second GWAS by the PGC by identifying 128 independent associations in 108 loci, 83 of which had not been reported previously (Schizophrenia Working Group of the Psychiatric Genomics Consortium, 2014). A subsequent meta-analysis identified 179 independent associations mapping to 145 independent loci, 52 of which were not previously reported by the PGC (Pardiñas et al., 2018). Utilizing GWAS fine mapping, chromosome conformation capture, and summary-data-based Mendelian randomisation analyses, potential causal genes were mapped to 33 loci (Pardiñas et al., 2018). Among these genes that are potentially causally related to schizophrenia, *SLC39A8* deficiency has also been associated with severe neurodevelopmental disorders supposed via impaired manganese transport and glycosylation. As oral galactose supplementation is a treatment option for complete normalization of glycosylation (Park et al., 2015), it emphasizes a therapeutic potential for schizophrenia. The largest GWAS to date is the third effort from the PGC, currently available as a pre-print, in 69,369 people with schizophrenia and 236,642 controls, identifying a total of 329 linkage disequilibrium-independent significant SNPs mapping to 270 distinct loci (Schizophrenia Working Group of the Psychiatric Genomics Consortium et al., 2020). Fine-mapping analyses refined these signals to 130 genes that are likely to explain these associations.

One of the most robust common variant associations with schizophrenia, among European populations, has been the 6p22.1 locus. This locus was first reported in 2009 in three independent studies as a genome-wide significant association (Purcell et al., 2009; Shi et al., 2009; Stefansson et al., 2009), and has been repeatedly confirmed since then in the major studies reported in *Figure 2*. The association at this locus, in an extended region around the Major Histocompatibility Complex, is likely to reflect causal

varaints located in at least three independent regions, one of which involves alleles of complement component 4 (C4) (Sekar et al., 2016). The associated alleles promote greater expression of *C4A* in the brain. This locus was also reported in studies in Chinese populations (Z. Li et al., 2017; Yue et al., 2011), however, a subsequent meta-analysis of the largest East Asian schizophrenia GWAS (Lam et al., 2019) did not replicate this finding.

Many genetic studies have demonstrated population-specific characteristics due to differences in underlying genetic architecture and environmental exposures, highlighting the importance of investigating trait/disease heterogeneity in population genetic studies. There have been GWAS studies reported in Indian (Periyasamy et al., 2019), African American (Fiorica & Wheeler, 2019) and Latin American (Bigdeli et al., 2019) populations but, to date, large scale schizophrenia GWAS has primarily been investigated in people of European ancestry and more recently East Asian ancestries. The largest schizophrenia GWAS in individuals of East Asian ancestries reported 21 independent associations in 19 loci, which included the top three associations shared with the European studies, and an additional 14 associations compared to the previous study of Chinese ancestry (Lam et al., 2019). The subsequent meta-analysis of East Asian and European ancestries reported 208 independent associations in 176 independent genetic loci, 53 of which were novel. These findings suggest that although the common genetic basis for schizophrenia is largely shared between populations, there are also likely to be population-specific risk variants driven by underlying differences in linkage disequilibrium and/or allele frequency.

Polygenic risk score prediction of schizophrenia: Polygenic risk scores (PRS) have emerged as an informative tool for studying the effects of genetic liability and may potentially be a clinically useful application of GWAS results. PRS is a single measure of the cumulative effects of common variants associated with a disorder, with higher scores indicating higher genetic liability (Lewis & Vassos, 2020). The variance explained varies depending on the GWAS p-value threshold used for calculating the PRS

(from all SNPs to only genome-wide significant SNPs); a p-value threshold of <0.05 from currently powered GWAS explains the greatest amount of variance in schizophrenia case-control status (Schizophrenia Working Group of the Psychiatric Genomics Consortium, 2014; Schizophrenia Working Group of the Psychiatric Genomics Consortium et al., 2020). First demonstrated in 2009 (Purcell et al., 2009), PRS can currently explain around 7.7% of the variance in schizophrenia case-control status (Schizophrenia Working Group of the Psychiatric Genomics Consortium *et al.*, 2020). The ability of PRS to explain schizophrenia case-control status is reduced in samples from a health care setting compared with the original GWAS research cohorts (Zheutlin et al., 2019). Although the ability of PRS to predict schizophrenia is currently insufficient for diagnostic purposes, there may be greater promise for sampling individuals at the extreme ends of the PRS distribution. How schizophrenia PRS could be applied clinically remains unclear given many individuals with a high PRS remain unaffected and in contrast to other common diseases that have benefited from PRS, such as coronary artery disease and type 2 diabetes, there is currently no preventative strategy in place for schizophrenia (Lewis & Vassos, 2020).

A key challenge in PRS analysis is the application across the major different ancestral populations. PRS derived from alleles discovered in GWAS of Europeans ancestries explain less variance when applied to African and Asian populations than in European ancestry samples, likely due to differences in allele frequencies and linkage disequilibrium structures (Curtis, 2018; Martin et al., 2019). PRS prediction of schizophrenia computed from European ancestry populations has been reported to be only 45% as accurate in East Asian compared to European individuals, despite a broadly shared genetic aetiology (Lam et al., 2019). The performance of PRS derived from a European ancestry GWAS applied to individuals with an African ancestry has been demonstrated to be particularly poor (Curtis, 2018; Vassos et al., 2017). Given the vast majority of genetic studies are conducted in populations of European

ancestry, greater diversity in GWAS must be prioritised to realise the potential of PRS (Martin et al., 2019).

Dissecting the clinical heterogeneity of schizophrenia using PRS: PRS has been used to disentangle the clinical heterogeneity of schizophrenia by investigating the phenotypic markers of genetic liability, particularly in relation to treatment outcomes, symptom severity, and cognitive ability (Dennison, Legge, Pardiñas, & Walters, 2020; Rees & Owen, 2020). A higher PRS for schizophrenia has been associated with a more chronic illness course, indexed by the number and length of hospital admissions (Meier et al., 2016). Results from studies investigating an association between schizophrenia PRS and treatment-resistance to antipsychotics are conflicting, indicating that outcome specific PRSs are likely to be required (Frank et al., 2015; Horsdal et al., 2017; Kowalec et al., 2019; Legge et al., 2020; Zhang et al., 2019).

Schizophrenia PRS has been associated with negative and disorganised symptom dimensions in individuals with schizophrenia, although the reported variance explained is small (Bipolar Disorder and Schizophrenia Working Group of the Psychiatric Genomics Consortium, 2018; Fanous et al., 2012; Jonas et al., 2019). There is no evidence to suggest that schizophrenia PRS is associated with the severity of positive symptoms in individuals with schizophrenia, but it has been associated with the presence of psychotic symptoms in first-episode samples and bipolar disorder (Allardyce et al., 2018; Ruderfer et al., 2014). There has been conflicting findings from studies investigating the association between schizophrenia PRS and the cognitive deficits observed in individuals with schizophrenia (Dickinson et al., 2020; Richards et al., 2020).

A further area of interest has been the relationship between PRS and intermediate phenotypes relevant to schizophrenia such as neuroimaging measures. A higher schizophrenia PRS has been associated with lower connectivity in areas disrupted in individuals with schizophrenia (Cao, Zhou, & Cannon, 2020).

There does not currently appear to be any strong evidence for an association between PRS and brain structural changes relevant to schizophrenia (van der Merwe et al., 2019).

Shared common genetic heritability with other psychiatric disorders: Common genetic liability for schizophrenia has been robustly found to have pleiotropic effects on related psychiatric disorders (***Figure 3***). Schizophrenia has significant genetic correlations with bipolar disorder (r_g = 0.68), major depressive disorder (r_g = 0.34), obsessive compulsive disorder (r_g = 0.33), ADHD (r_g = 0.22), anorexia nervosa (r_g = 0.22), and autism spectrum disorder (r_g = 0.21) (Brainstorm Consortium et al., 2018). In addition, schizophrenia had a positive genetic correlation with neuroticism (r_g = 0.19), negative correlations with subjective well-being (r_g = -0.30), body mass index (r_g = -0.10), and intelligence (r_g = -0.20) (Brainstorm Consortium et al., 2018). These results are also consistent in non-European populations (Lam et al., 2019), and in a PRS analysis of a large real-world sample of patients from four US-based healthcare systems (Zheutlin et al., 2019). These findings indicate that a substantial proportion of the common genetic architecture among psychiatric disorders and cognition is shared and has implications for the validity of current clinical diagnostic boundaries. These findings are supported by the observation that schizophrenia is often comorbid with other psychiatric disorders, with the highest rates for substance use disorders (Regier et al., 1990) and mood disorders (Buckley, Miller, Lehrer, & Castle, 2009). Possible explanations for this pleiotropy include the presence of a general genetic psychopathology factor that increases risk for multiple psychiatric disorders or that additional genetic and environmental factors influence the eventual clinical presentation (Dennison et al., 2020).

SNP-based heritability estimates: Current estimates place the SNP-based heritability, or the phenotypic variance due to genetic variation tagged by polymorphisms derived from array genotyping, at approximately 24% (Pardiñas et al., 2018; Schizophrenia Working Group of the Psychiatric Genomics Consortium, Ripke, Walters, & O'Donovan, 2020). The disparity in heritability estimates from family

studies and genome-wide association studies indicates there are susceptibility genes yet to be identified and has led to a search for the "missing" heritability. There are several potential sources of missing heritability including common variants that current studies are not yet powered to detect, but also gene-environment interactions, epigenetic variation, and rare genetic variation.

Rare genetic variation

Rare variants are typically defined as those with a minor allele frequency <1% and include single nucleotide variants (SNVs), altering one or a small number of bases, and insertion-deletion variants, which can vary in size from those affecting single nucleotides to those classified as CNVs that affect thousands or millions of bases. CNV studies to date are typically based on the same genotyping platforms used to conduct GWAS, while other rare variants have largely been studied using whole exome sequencing (WES). Whole genome sequencing can capture both types of variants as well as other forms of structural changes, but to date, such studies in schizophrenia are small in scale. Rare variant studies are poised to be highly informative by pinpointing causal genes, and understanding molecular and cellular aspects of the disease process (Sullivan & Geschwind, 2019). Including the most recent exome sequencing effort reported in pre-peer reviewed format, 10 genes have been implicated through exome sequencing studies to date, and eight rare CNVs have been associated with schizophrenia.

Single nucleotide variants (SNVs) and small insertion deletion (indel) variants: The current sample sizes included in WES or whole genome sequencing cohorts are insufficiently powered to detect significant associations at the single variant level. However, an alternative statistical strategy exists, whereby variants are pooled together at the gene or genome level and then the burden of those variants can be compared between cases and controls. A WES study demonstrated an enrichment of damaging ultra-rare variants in 2536 schizophrenia cases, compared with 2543 controls (Purcell et al., 2014). Similar findings were reported from an extension of the same Swedish cohort of 12332 individuals (4946 cases,

6242 controls) analysed along with 45376 whole exomes from the Exome Aggregation Consortium (ExAC), whereby an excess burden of disruptive and damaging (as predicted by bioinformatics) ultra-rare variants was found in cases (odds ratio=1.07, 95% confidence interval [CI]=1.05-1.09) (Genovese et al., 2016). However, no individual gene was found to have an significant excess of damaging ultra-rare variants, demonstrating the highly polygenic nature of schizophrenia.

Another WES study identified a significant association between *SETD1A* loss of function (LoF: variants predicted to result in the loss of protein-coding function) variants and schizophrenia by combining a whole-exome case control sequencing study of 4264 schizophrenia cases, and 9343 controls, with *de novo* mutation analysis in 1077 parent-proband trios (Singh et al., 2016). A subsequent follow-up study found the burden of LoF variants was enriched in LoF intolerant gene sets in schizophrenia cases compared to controls (odds ratio=1.24, 95% CI=1.16-1.31), although no individual gene was implicated (Singh et al., 2017). Similarly, the recent population-specific WES studies of South African (Gulsuner et al., 2020) and Taiwanese (Howrigan et al., 2020) ancestry did not identify any implicated gene but found the burden of LoF variants to be enriched in highly brain-expressed and evolutionarily constrained genes. Currently, the largest rare exome sequencing effort is from the Schizophrenia Exome Sequencing Meta-Analysis (SCHEMA) Consortium, which pooled data from 24,248 schizophrenia cases and 97,322 controls. As reported in pre-peer reviewed format, 10 genes (including *SETD1A*), reached genome wide significance for an excess of ultra-rare variants predicted to be damaging in cases compared to controls ($p<2x10^{-6}$) (Singh, Neale, & Daly, 2020), and a further 22 reached suggestive levels of significance as defined by a false discovery rate of 0.05.

De novo variants (DNVs): DNVs are variants that are present in offspring as a result of a new mutation event, and are therefore absent in the parents. Schizophrenia is associated with a reduction in reproductive fecundity, leading to the hypothesis that DNVs with large effect sizes play a role in the genetic aetiology. In a recent WES study of 3444 schizophrenia parent-proband trios, LoF DNVs were

found to be significantly enriched in LoF intolerant genes, but no gene individually achieved exome wide significance for enrichment of LoF DNVs (Rees et al., 2020). However, the burden of DNVs was significantly higher in genes previously associated with neurodevelopmental disorders, among which *SLC6A1* encoding a gamma-aminobutyric acid (GABA) transporter, was significantly enriched with missense variants. In gene-set analysis, DNVs were enriched in evolutionary constrained genes and in those implicated in multiple neurodevelopmental disorders. A separate study also noted that the DNV burden in schizophrenia is smaller as compared to early onset neurodevelopmental disorders (Howrigan et al., 2020).

Copy number variations (CNVs): CNVs are either duplications or deletions, ranging from 50 base pairs to mega bases in the genome and can span a whole gene or multiple genes in a region. CNVs have been consistently implicated in the aetiology of schizophrenia, with the first associated CNV for schizophrenia being a large deletion on chromosome 22q11.2, which confers a 20-fold increased risk, with approximately 25% of carriers developing schizophrenia. The largest CNV study to date in 21,094 cases and 20,227 controls found 8 CNVs (6 deletions and 2 duplications) to be significantly associated with schizophrenia (***Figure 1***) (CNV and Schizophrenia Working Groups of the Psychiatric Genomics Consortium, 2017). These 8 loci collectively explain 0.85% of variance in liability to schizophrenia, with 1.4% of the cases carrying these risk CNVs. This is consistent with other studies demonstrating large effect sizes of CNVs, even though the absolute number of cases with CNVs is small (Rees et al., 2014). Studies have shown that CNV penetrance (i.e., the proportion of CNV carriers demonstrating the phenotype of interest) in schizophrenia is lower compared with other neurodevelopmental disorders ((Kirov et al., 2014; Vassos et al., 2010). The penetrance of CNVs in schizophrenia seems to be additionally influenced by the burden of commmon risk alelles (Tansey et al., 2016), with evidence supporting an additive joint effect between the two classes of risk variants (Bergen et al., 2019).

Functional annotation of genetic variants

The polygenic nature of schizophrenia imposes a challenge to understand how, where, and when genetic variation acts to increase vulnerability of developing the disorder. Following up on the findings from genome-wide association approaches, the aim is to then gain biological insights by identifying genes, tissues, cells, and biological pathways associated with schizophrenia.

Enrichment of gene-sets: Gene-set analysis investigates whether sets of genes, grouped by biological pathway or expression in particular tissues, are enriched for variants associated with schizophrenia. Consistent with the hypothesis that schizophrenia is primarily a disorder of neuronal dysfunction, genes highly expressed in the brain, mainly in cortical inhibitory interneurons and excitatory neurons from cerebral cortex and hippocampus, are strongly enriched for SNPs associated with schizophrenia (Pardiñas et al., 2018; Schizophrenia Working Group of the Psychiatric Genomics Consortium et al., 2020). Other notable gene-set findings included an association within a target gene of antipsychotic drugs (*DRD2*), genes involved in glutamatergic neurotransmission and synaptic plasticity (e.g. *GRM3, GRIN2A, SRR, GRIA1*) and additional genes encoding voltage-gated calcium channel subunits (*CACNA1C, CACNB2* and *CACNA1I*) (Schizophrenia Working Group of the Psychiatric Genomics Consortium, 2014).

Likewise, analyses of gene ontology classifications for schizophrenia have demonstrated an enrichment in histone methylation and synaptic pathways, in particular postsynaptic proteins and structures have been enriched for all classes of risk variants (G. Kirov et al., 2012; Schizophrenia Working Group of the Psychiatric Genomics Consortium et al., 2020; Singh et al., 2020). Notably, as the sample sizes have increased in genetic studies, enrichment analyses have shown a convergence of rare and common variants findings, pointing to concordant genes (e.g. *GRIN2A*, *SP4*, *STAG1*, and *FAM120A*) and pathways, which strongly supports their relevance in schizophrenia.

Transcriptome-wide association studies (TWAS): TWAS use expression quantitative trait loci (eQTLs –

variants that affect expression) of genes in a specific tissue to predict expression of that gene. Upon

merging this information with GWAS summary statistics, we can infer which genes might be causally

related to schizophrenia and whether they are up or downregulated. A TWAS using genetic data from

the second PGC GWAS and expression data from brain, blood, and adipose tissues identified 157 genes

to be associated with schizophrenia, with 42 associated with chromatin organization, highlighting that

regions responsible for gene expression regulation can be potential targets (Gusev et al., 2018). This was

subsequently followed by another study reporting 107 genes, of which 11 were also differentially

expressed and in the same direction as found previously in schizophrenia brain samples (Gandal et al.,

2018). A more recent TWAS evaluated 5301 genes with *cis*-heritable expression (variance explained by

SNPs close to the gene) in the dorsolateral prefrontal cortex and identified 89 genic associations, 20 of

which were novel (Hall et al., 2020).

Transcriptomic and single-cell studies: Although TWAS are able to indicate possible causal genes from

GWAS results, TWAS cannot directly extract time specific information of disease nor which kind of cell

type is involved. As an initial step to rectify this gap, the Common Mind Consortium generated

transcriptomic data from dorsolateral prefrontal cortex (258 schizophrenia cases, 279 controls) and

intersected with 142 GWAS associations, to demonstrate an overlap of 20 variants potentially

influencing gene expression of one or more genes (Wang et al., 2018).

Recently, single cell RNAseq data has offered new interpretation of GWAS results. For schizophrenia,

PsychENCODE-generated single cell RNAseq data identified spatiotemporal loci (M. Li et al., 2018) and

mapped the associated genomic loci to pyramidal cells, medium spiny neurons, and interneurons in

adult cortical cells (Skene et al., 2018) and to neural progenitors (M. Li et al., 2018), oligodendrocyte

precursors and fetal microglia (Polioudakis et al., 2019).

Future perspectives and increasing diversity

In the next few years, we should expect that larger sample size GWAS and rare variant studies will identify more genes and refine the biological processes implicated in schizophrenia. Similarly, efforts to improve functional annotations provided by newer iterations of consortia such as PsychENCODE, GTEx, CommonMind Consortium and others will aid further interpretation of GWAS results. It is hoped that these insights could contribute to drug discovery and personalised medicine efforts. Future research is also likely to focus on investigating how the different elements of genetic risk combine and how genetic risk interacts with environmental factors to ultimately cause schizophrenia.

As mentioned previously, it is of paramount importance to improve diversity of schizophrenia genetic studies, by including non-European populations. By 2019, 79% of all GWAS samples (including non-schizophrenia studies) are from individuals of European descent (Martin et al., 2019). For schizophrenia, the latest GWAS pre-print included ~20% non-European samples, the majority of which were from East Asia (Schizophrenia Working Group of the Psychiatric Genomics Consortium et al., 2020). Increasing genetic diversity will ensure emergent clinical applications are applicable to all human populations, and at the same time, will increase the power to identify genetic associations and aid fine mapping (Y. R. Li & Keating, 2014). As new ongoing consortia and biobanks are established, it is essential that ancestry disparities are addressed.

Conclusions

In the last two decades, the sample size for genetic studies of schizophrenia has increased from hundreds to hundreds of thousands of individuals, resulting in the identification of >300 common variants, 10 genes with a burden of rare coding variants, and at least 8 CNVs. Genetic studies of schizophrenia provided the first compelling evidence of the value of GWAS in investigating common

genetic risk for psychiatric disorders, which has led to an expansion of sample sizes and subsequently the identification of many genetic variants for other psychiatric disorders (Schizophrenia Working Group of the Psychiatric Genomics Consortium, 2014). Moreover, GWAS demonstrated the polygenic nature of psychiatric disorders and how hundreds or thousands of variants contribute to liability to schizophrenia. Rare variant studies in schizophrenia have also made important achivements, including identification of overlap of implicated genes and gene sets with other neurodevelopmental disorders. This pleiotropy further adds to the existing complex picture emerging out of these studies (Sullivan & Geschwind, 2019).

For the coming years, we should expect not only an increase in the sample sizes but also an increase in diversity of genetic studies, which should identify new regions, improve fine mapping and increase our understanding of the biology and mechanisms of schizophrenia. Likewise, as sequencing prices decrease, larger whole genome sequencing studies should identify other rare variants that play a role in the disease. Now that many genes have been implicated with schizophrenia, functional studies will be critical to better understand how and when schizophrenia vulnerability acts during the neurodevelopment.

Acknowledgements

Funding Statement

Conflicts of Interest

References

Allardyce, J., Leonenko, G., Hamshere, M., Pardiñas, A. F., Forty, L., Knott, S., … Escott-Price, V. (2018). Association Between Schizophrenia-Related Polygenic Liability and the Occurrence and Level of Mood-Incongruent Psychotic Symptoms in Bipolar Disorder. *JAMA Psychiatry*, *75*(1), 28–35. https://doi.org/10.1001/jamapsychiatry.2017.3485

Bergen, S. E., Ploner, A., Howrigan, D., O'Donovan, M. C., Smoller, J. W., Sullivan, P. F., … Kendler, K. S. (2019). Joint contributions of rare copy number variants and common SNPs to risk for schizophrenia. *American Journal of Psychiatry*, *176*(1), 29–35. https://doi.org/10.1176/appi.ajp.2018.17040467

Bipolar Disorder and Schizophrenia Working Group of the Psychiatric Genomics Consortium. (2018). Genomic Dissection of Bipolar Disorder and Schizophrenia, Including 28 Subphenotypes. *Cell*, *173*(7), 1705-1715.e16. https://doi.org/10.1016/j.cell.2018.05.046

Bitter, I., Czobor, P., Borsi, A., Fehér, L., Nagy, B. Z., Bacskai, M., … Takács, P. (2017). Mortality and the relationship of somatic comorbidities to mortality in schizophrenia. A nationwide matched-cohort study. *European Psychiatry*, *45*, 97–103. https://doi.org/10.1016/j.eurpsy.2017.05.022

Bourque, F., Van Der Ven, E., & Malla, A. (2011). A meta-analysis of the risk for psychotic disorders among first- and second-generation immigrants. *Psychological Medicine*, *41*(5), 897–910. https://doi.org/10.1017/S0033291710001406

Brainstorm Consortium, Anttila, V., Bulik-Sullivan, B., Finucane, H. K., Walters, R. K., Bras, J., … Murray, R. (2018). Analysis of shared heritability in common disorders of the brain. *Science (New York, N.Y.)*, *360*(6395), eaap8757. https://doi.org/10.1126/science.aap8757

Buckley, P. F., Miller, B. J., Lehrer, D. S., & Castle, D. J. (2009). Psychiatric comorbidities and schizophrenia. *Schizophrenia Bulletin*, *35*(2), 383–402. https://doi.org/10.1093/schbul/sbn135

Cannon, M., Jones, P. B., & Murray, R. M. (2002). Obstetric complications and schizophrenia: Historical and meta-analytic review. *American Journal of Psychiatry*. Am J Psychiatry. https://doi.org/10.1176/appi.ajp.159.7.1080

Cantor-Graae, E., & Selten, J. P. (2005, January). Schizophrenia and migration: A meta-analysis and review. *American Journal of Psychiatry*. Am J Psychiatry. https://doi.org/10.1176/appi.ajp.162.1.12

Cao, H., Zhou, H., & Cannon, T. D. (2020). Functional connectome-wide associations of schizophrenia polygenic risk. *Molecular Psychiatry*. https://doi.org/10.1038/s41380-020-0699-3

Charlson, F. J., Ferrari, A. J., Santomauro, D. F., Diminic, S., Stockings, E., Scott, J. G., … Whiteford, H. A. (2018). Global epidemiology and burden of schizophrenia: Findings from the global burden of disease study 2016. *Schizophrenia Bulletin*, *44*(6), 1195–1203. https://doi.org/10.1093/schbul/sby058

Chou, I. J., Kuo, C. F., Huang, Y. S., Grainge, M. J., Valdes, A. M., See, L. C., … Doherty, M. (2017). Familial aggregation and heritability of schizophrenia and co-aggregation of psychiatric illnesses in affected families. *Schizophrenia Bulletin*, *43*(5), 1070–1078. https://doi.org/10.1093/schbul/sbw159

CNV and Schizophrenia Working Groups of the Psychiatric Genomics Consortium. (2017). Contribution of copy number variants to schizophrenia from a genome-wide study of 41,321 subjects. *Nature Genetics*, *49*(1), 27–35. https://doi.org/10.1038/ng.3725

Curtis, D. (2018). Polygenic risk score for schizophrenia is more strongly associated with ancestry than with schizophrenia. *Psychiatric Genetics*, *28*(5), 85–89. https://doi.org/10.1097/YPG.0000000000000206

Dennison, C. A., Legge, S. E., Pardiñas, A. F., & Walters, J. T. R. (2020). Genome-wide association studies in schizophrenia: Recent advances, challenges and future perspective. *Schizophrenia Research*, *217*, 4–12. https://doi.org/10.1016/j.schres.2019.10.048

Dickinson, D., Zaidman, S. R., Giangrande, E. J., Eisenberg, D. P., Gregory, M. D., & Berman, K. F. (2020). Distinct Polygenic Score Profiles in Schizophrenia Subgroups With Different Trajectories of Cognitive Development. *The American Journal of Psychiatry*, *177*(4), 298–307. https://doi.org/10.1176/appi.ajp.2019.19050527

Dollfus, S., & Lyne, J. (2017, August 1). Negative symptoms: History of the concept and their position in diagnosis of schizophrenia. *Schizophrenia Research*. Elsevier B.V. https://doi.org/10.1016/j.schres.2016.06.024

Fanous, A. H., Zhou, B., Aggen, S. H., Bergen, S. E., Amdur, R. L., Duan, J., … Schizophrenia Psychiatric Genome-Wide Association Study (GWAS) Consortium. (2012). Genome-wide association study of clinical dimensions of schizophrenia: polygenic effect on disorganized symptoms. *The American Journal of Psychiatry*, *169*(12), 1309–1317. https://doi.org/10.1176/appi.ajp.2012.12020218

Frank, J., Lang, M., Witt, S. H., Strohmaier, J., Rujescu, D., Cichon, S., … Rietschel, M. (2015). Identification of increased genetic risk scores for schizophrenia in treatment-resistant patients. *Molecular Psychiatry*, *20*(2), 150–151. https://doi.org/10.1038/mp.2014.56

Gandal, M. J., Zhang, P., Hadjimichael, E., Walker, R. L., Chen, C., Liu, S., … Geschwind, D. H. (2018). Transcriptome-wide isoform-level dysregulation in ASD, schizophrenia, and bipolar disorder. *Science (New York, N.Y.)*, *362*(6420). https://doi.org/10.1126/science.aat8127

Genovese, G., Fromer, M., Stahl, E. A., Ruderfer, D. M., Chambert, K., Landen, M., … McCarroll, S. A. (2016). Increased burden of ultra-rare protein-altering variants among 4,877 individuals with schizophrenia. *Nat Neurosci*, *19*(11), 1433–1441. https://doi.org/10.1038/nn.4402

Gusev, A., Mancuso, N., Won, H., Kousi, M., Finucane, H. K., Reshef, Y., … Price, A. L. (2018). Transcriptome-wide association study of schizophrenia and chromatin activity yields mechanistic disease insights. *Nature Genetics*, *50*(4), 538–548. https://doi.org/10.1038/s41588-018-0092-1

Hall, L. S., Medway, C. W., Pain, O., Pardiñas, A. F., Rees, E. G., Escott-Price, V., … O'Donovan, M. C. (2020). A transcriptome-wide association study implicates specific pre- and post-synaptic abnormalities in schizophrenia. *Human Molecular Genetics*, *29*(1), 159–167. https://doi.org/10.1093/hmg/ddz253

Horsdal, H. T., Meier, S. M., Wimberley, T., Agerbo, E., Gasse, C., & MacCabe, J. H. (2017). Polygenic Risk Score for Schizophrenia and Treatment-Resistant Schizophrenia. *Schizophrenia Bulletin*, *43*(5), 1064–1069. https://doi.org/10.1093/schbul/sbx007

Howrigan, D. P., Rose, S. A., Samocha, K. E., Fromer, M., Cerrato, F., Chen, W. J., … Neale, B. M. (2020). Exome sequencing in schizophrenia-affected parent-offspring trios reveals risk conferred by protein-coding de novo mutations. *Nature Neuroscience*, *23*(2), 185–193. https://doi.org/10.1038/s41593-019-0564-3

International Schizophrenia Consortium, Purcell, S. M., Wray, N. R., Stone, J. L., Visscher, P. M., O'Donovan, M. C., … Sklar, P. (2009). Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature*, *460*(7256), 748–752. https://doi.org/10.1038/nature08185

Jaaskelainen, E., Juola, P., Hirvonen, N., McGrath, J. J., Saha, S., Isohanni, M., … Miettunen, J. (2013). A Systematic Review and Meta-Analysis of Recovery in Schizophrenia. *Schizophrenia Bulletin*, *39*(6), 1296–1306. https://doi.org/10.1093/schbul/sbs130

Jerrell, J. M., McIntyre, R. S., & Tripathi, A. (2010). Incidence and costs of cardiometabolic conditions in patients with schizophrenia treated with antipsychotic medications. *Clinical Schizophrenia and Related Psychoses*, *4*(3), 161–168. https://doi.org/10.3371/CSRP.4.3.2

Jonas, K. G., Lencz, T., Li, K., Malhotra, A. K., Perlman, G., Fochtmann, L. J., … Kotov, R. (2019). Schizophrenia polygenic risk score and 20-year course of illness in psychotic disorders. *Translational Psychiatry*, *9*(1), 300. https://doi.org/10.1038/s41398-019-0612-5

Kirov, G., Pocklington, A. J., Holmans, P., Ivanov, D., Ikeda, M., Ruderfer, D., … Owen, M. J. (2012). De

novo CNV analysis implicates specific abnormalities of postsynaptic signalling complexes in the pathogenesis of schizophrenia. *Molecular Psychiatry*, *17*(2), 142–153. https://doi.org/10.1038/mp.2011.154

Kirov, George, Rees, E., Walters, J. T. R., Escott-Price, V., Georgieva, L., Richards, A. L., … Owen, M. J. (2014). The penetrance of copy number variations for schizophrenia and developmental delay. *Biological Psychiatry*, *75*(5), 378–385. https://doi.org/10.1016/j.biopsych.2013.07.022

Kowalec, K., Lu, Y., Sariaslan, A., Song, J., Ploner, A., Dalman, C., … Sullivan, P. F. (2019). Increased schizophrenia family history burden and reduced premorbid IQ in treatment-resistant schizophrenia: a Swedish National Register and Genomic Study. *Molecular Psychiatry*. https://doi.org/10.1038/s41380-019-0575-1

Lam, M., Chen, C.-Y., Li, Z., Martin, A. R., Bryois, J., Ma, X., … Huang, H. (2019). Comparative genetic architectures of schizophrenia in East Asian and European populations. *Nature Genetics*, *51*(12), 1670–1678. https://doi.org/10.1038/s41588-019-0512-x

Lee, S. H., DeCandia, T. R., Ripke, S., Yang, J., Schizophrenia Psychiatric Genome-Wide Association Study Consortium (PGC-SCZ), P. F., International Schizophrenia Consortium (ISC), M. E., … Wray, N. R. (2012). Estimating the proportion of variation in susceptibility to schizophrenia captured by common SNPs. *Nature Genetics*, *44*(3), 247–250. https://doi.org/10.1038/ng.1108

Legge, S. E., Dennison, C. A., Pardiñas, A. F., Rees, E., Lynham, A. J., Hopkins, L., … Walters, J. T. R. (2020). Clinical indicators of treatment-resistant psychosis. *The British Journal of Psychiatry : The Journal of Mental Science*, *216*(5), 259–266. https://doi.org/10.1192/bjp.2019.120

Lewis, C. M., & Vassos, E. (2020). Polygenic risk scores: from research tools to clinical instruments. *Genome Medicine*, *12*(1), 44. https://doi.org/10.1186/s13073-020-00742-5

Li, M., Santpere, G., Imamura Kawasawa, Y., Evgrafov, O. V, Gulden, F. O., Pochareddy, S., … Sestan, N. (2018). Integrative functional genomic analysis of human brain development and neuropsychiatric risks. *Science (New York, N.Y.)*, *362*(6420). https://doi.org/10.1126/science.aat7615

Li, Y. R., & Keating, B. J. (2014). Trans-ethnic genome-wide association studies: advantages and challenges of mapping in diverse populations. *Genome Medicine*, *6*(10), 91. https://doi.org/10.1186/s13073-014-0091-5

Li, Z., Chen, J., Yu, H., He, L., Xu, Y., Zhang, D., … Shi, Y. (2017). Genome-wide association analysis identifies 30 new susceptibility loci for schizophrenia. *Nature Genetics*, *49*(11), 1576–1583. https://doi.org/10.1038/ng.3973

Limosin, F., Loze, J. Y., Philippe, A., Casadebaig, F., & Rouillon, F. (2007). Ten-year prospective follow-up study of the mortality by suicide in schizophrenic patients. *Schizophrenia Research*, *94*(1–3), 23–28. https://doi.org/10.1016/j.schres.2007.04.031

Martin, A. R., Kanai, M., Kamatani, Y., Okada, Y., Neale, B. M., & Daly, M. J. (2019). Clinical use of current polygenic risk scores may exacerbate health disparities. *Nature Genetics*, *51*(4), 584–591. https://doi.org/10.1038/s41588-019-0379-x

McGrath, J., Saha, S., Chant, D., & Welham, J. (2008). Schizophrenia: a concise overview of incidence, prevalence, and mortality. *Epidemiol Rev*, *30*, 67–76. https://doi.org/10.1093/epirev/mxn001

Meier, S. M., Agerbo, E., Maier, R., Pedersen, C. B., Lang, M., Grove, J., … Mattheisen, M. (2016). High loading of polygenic risk in cases with chronic schizophrenia. *Molecular Psychiatry*, *21*(7), 969–974. https://doi.org/10.1038/mp.2015.130

Olfson, M., Gerhard, T., Huang, C., Crystal, S., & Stroup, T. S. (2015). Premature mortality among adults with schizophrenia in the United States. *JAMA Psychiatry*, *72*(12), 1172–1181. https://doi.org/10.1001/jamapsychiatry.2015.1737

Owen, M. J., Sawa, A., & Mortensen, P. B. (2016, July 2). Schizophrenia. *The Lancet*. Lancet Publishing Group. https://doi.org/10.1016/S0140-6736(15)01121-6

Pardiñas, A. F., Holmans, P., Pocklington, A. J., Escott-Price, V., Ripke, S., Carrera, N., … Walters, J. T. R.

(2018). Common schizophrenia alleles are enriched in mutation-intolerant genes and in regions under strong background selection. *Nature Genetics*, *50*(3), 381–389. https://doi.org/10.1038/s41588-018-0059-2

Polioudakis, D., de la Torre-Ubieta, L., Langerman, J., Elkins, A. G., Shi, X., Stein, J. L., … Geschwind, D. H. (2019). A Single-Cell Transcriptomic Atlas of Human Neocortical Development during Mid-gestation. *Neuron*, *103*(5), 785-801.e8. https://doi.org/10.1016/j.neuron.2019.06.011

Pompili, M., Amador, X. F., Girardi, P., Harkavy-Friedman, J., Harrow, M., Kaplan, K., … Tatarelli, R. (2007). Suicide risk in schizophrenia: learning from the past to change the future. *Annals of General Psychiatry*, *6*(1), 10. https://doi.org/10.1186/1744-859X-6-10

Purcell, S. M., Moran, J. L., Fromer, M., Ruderfer, D., Solovieff, N., Roussos, P., … Sklar, P. (2014). A polygenic burden of rare disruptive mutations in schizophrenia. *Nature*, *506*(7487), 185–190. https://doi.org/10.1038/nature12975

Purcell, S. M., Wray, N. R., Stone, J. L., Visscher, P. M., O'Donovan, M. C., Sullivan, P. F., … Moran, J. L. (2009). Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature*, *10*(AuGuST), 8192–8192. https://doi.org/10.1038/nature08185

Rees, E., Han, J., Morgan, J., Carrera, N., Escott-Price, V., Pocklington, A. J., … Owen, M. J. (2020). De novo mutations identified by exome sequencing implicate rare missense variants in SLC6A1 in schizophrenia. *Nature Neuroscience*, *23*(2), 179–184. https://doi.org/10.1038/s41593-019-0565-2

Rees, E., & Owen, M. J. (2020). Translating insights from neuropsychiatric genetics and genomics for precision psychiatry. *Genome Medicine*, *12*(1), 43. https://doi.org/10.1186/s13073-020-00734-5

Rees, E., Walters, J. T. R., Georgieva, L., Isles, A. R., Chambert, K. D., Richards, A. L., … Kirov, G. (2014). Analysis of copy number variations at 15 schizophrenia-associated loci. *The British Journal of Psychiatry : The Journal of Mental Science*, *204*(2), 108–114. https://doi.org/10.1192/bjp.bp.113.131052

Regier, D. A., Farmer, M. E., Rae, D. S., Locke, B. Z., Keith, S. J., Judd, L. L., & Goodwin, F. K. (1990). Comorbidity of mental disorders with alcohol and other drug abuse. Results from the Epidemiologic Catchment Area (ECA) Study. *JAMA*, *264*(19), 2511–2518. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/2232018

Richards, A. L., Pardiñas, A. F., Frizzati, A., Tansey, K. E., Lynham, A. J., Holmans, P., … Walters, J. T. R. (2020). The Relationship Between Polygenic Risk Scores and Cognition in Schizophrenia. *Schizophrenia Bulletin*, *46*(2), 336–344. https://doi.org/10.1093/schbul/sbz061

Ripke, S., O'Dushlaine, C., Chambert, K., Moran, J. L., Kähler, A. K., Akterin, S., … Sullivan, P. F. (2013). Genome-wide association analysis identifies 13 new risk loci for schizophrenia. *Nature Genetics*, *45*(10), 1150–1159. https://doi.org/10.1038/ng.2742

Ruderfer, D. M., Fanous, A. H., Ripke, S., McQuillin, A., Amdur, R. L., Gejman, P. V., … Kendler, K. S. (2014). Polygenic dissection of diagnosis and clinical dimensions of bipolar disorder and schizophrenia. *Molecular Psychiatry*, *19*(9), 1017–1024. https://doi.org/10.1038/mp.2013.138

Schizophrenia Psychiatric Genome-Wide Association Study (GWAS) Consortium. (2011). Genome-wide association study identifies five new schizophrenia loci. *Nature Genetics*, *43*(10), 969–976. https://doi.org/10.1038/ng.940

Schizophrenia Working Group of the Psychiatric Genomics Consortium. (2014). Biological insights from 108 schizophrenia-associated genetic loci. *Nature*, *511*(7510), 421–427. https://doi.org/10.1038/nature13595

Schizophrenia Working Group of the Psychiatric Genomics Consortium, Ripke, S., Walters, J. T., & O'Donovan, M. C. (2020). Mapping genomic loci prioritises genes and implicates synaptic biology in schizophrenia. *MedRxiv*, 2020.09.12.20192922. https://doi.org/10.1101/2020.09.12.20192922

Sekar, A., Bialas, A. R., De Rivera, H., Davis, A., Hammond, T. R., Kamitaki, N., … McCarroll, S. A. (2016). Schizophrenia risk from complex variation of complement component 4. *Nature*, *530*(7589), 177–

183. https://doi.org/10.1038/nature16549

Shi, J., Levinson, D. F., Duan, J., Sanders, A. R., Zheng, Y., Péer, I., … Gejman, P. V. (2009). Common variants on chromosome 6p22.1 are associated with schizophrenia. *Nature*, *460*(7256), 753–757. https://doi.org/10.1038/nature08192

Singh, T., Kurki, M. I., Curtis, D., Purcell, S. M., Crooks, L., McRae, J., … Barrett, J. C. (2016). Rare loss-of-function variants in SETD1A are associated with schizophrenia and developmental disorders. *Nature Neuroscience*, *19*(4), 571–577. https://doi.org/10.1038/nn.4267

Singh, T., Neale, B. M., & Daly, M. J. (2020). Exome sequencing identifies rare coding variants in 10 genes which confer substantial risk for schizophrenia on behalf of the Schizophrenia Exome Meta-Analysis (SCHEMA) Consortium*. *MedRxiv*, 2020.09.18.20192815. https://doi.org/10.1101/2020.09.18.20192815

Singh, T., Walters, J. T. R., Johnstone, M., Curtis, D., Suvisaari, J., Torniainen, M., … Barrett, J. C. (2017). The contribution of rare variants to risk of schizophrenia in individuals with and without intellectual disability. *Nature Genetics*, *49*(8), 1167–1173. https://doi.org/10.1038/ng.3903

Sipos, A., Rasmussen, F., Harrison, G., Tynelius, P., Lewis, G., Leon, D. A., & Gunnell, D. (2004). Paternal age and schizophrenia: A population based cohort study. *British Medical Journal*, *329*(7474), 1070–1073. https://doi.org/10.1136/bmj.38243.672396.55

Skene, N. G., Bryois, J., Bakken, T. E., Breen, G., Crowley, J. J., Gaspar, H. A., … Hjerling-Leffler, J. (2018). Genetic identification of brain cell types underlying schizophrenia. *Nature Genetics*, *50*(6), 825–833. https://doi.org/10.1038/s41588-018-0129-5

Stefansson, H., Ophoff, R. A., Steinberg, S., Andreassen, O. A., Cichon, S., Rujescu, D., … Collier, D. A. (2009). Common variants conferring risk of schizophrenia. *Nature*, *460*(7256), 744–747. https://doi.org/10.1038/nature08186

Stilo, S. A., & Murray, R. M. (2019, October 1). Non-Genetic Factors in Schizophrenia. *Current Psychiatry Reports*. Current Medicine Group LLC 1. https://doi.org/10.1007/s11920-019-1091-3

Sullivan, P. F., & Geschwind, D. H. (2019, March 21). Defining the Genetic, Genomic, Cellular, and Diagnostic Architectures of Psychiatric Disorders. *Cell*. Cell Press. https://doi.org/10.1016/j.cell.2019.01.015

Tansey, K. E., Rees, E., Linden, D. E., Ripke, S., Chambert, K. D., Moran, J. L., … O'Donovan, M. C. (2016). Common alleles contribute to schizophrenia in CNV carriers. *Molecular Psychiatry*, *21*(8), 1153. https://doi.org/10.1038/mp.2015.170

Tanskanen, A., Tiihonen, J., & Taipale, H. (2018). Mortality in schizophrenia: 30-year nationwide follow-up study. *Acta Psychiatrica Scandinavica*, *138*(6), 492–499. https://doi.org/10.1111/acps.12913

van der Merwe, C., Passschier, R., Mufford, M., Ramesar, R., Dalvie, S., & Stein, D. J. (2019). Polygenic risk for schizophrenia and associated brain structural changes: A systematic review. *Comprehensive Psychiatry*, *88*, 77–82. https://doi.org/10.1016/j.comppsych.2018.11.014

Vassos, E., Collier, D. A., Holden, S., Patch, C., Rujescu, D., St Clair, D., & Lewis, C. M. (2010). Penetrance for copy number variants associated with schizophrenia. *Human Molecular Genetics*, *19*(17), 3477–3481. https://doi.org/10.1093/hmg/ddq259

Vassos, E., Di Forti, M., Coleman, J., Iyegbe, C., Prata, D., Euesden, J., … Breen, G. (2017). An Examination of Polygenic Score Risk Prediction in Individuals With First-Episode Psychosis. *Biological Psychiatry*, *81*(6), 470–477. https://doi.org/10.1016/j.biopsych.2016.06.028

Wang, D., Liu, S., Warrell, J., Won, H., Shi, X., Navarro, F. C. P., … Gerstein, M. B. (2018). Comprehensive functional genomic resource and integrative model for the human brain. *Science*, *362*(6420). https://doi.org/10.1126/science.aat8464

Wiersma, D., Wanderling, J., Dragomirecka, E., Ganev, K., Harrison, G., An Der Heiden, W., … Walsh, D. (2000). Social disability in schizophrenia: its development and prediction over 15 years in incidence cohorts in six European centres. *Psychological Medicine*, *30*(5), 1155–1167. Retrieved from

http://www.ncbi.nlm.nih.gov/pubmed/12027051

Young, A. I., Frigge, M. L., Gudbjartsson, D. F., Thorleifsson, G., Bjornsdottir, G., Sulem, P., … Kong, A. (2018). Relatedness disequilibrium regression estimates heritability without environmental bias. *Nature Genetics*, *50*(9), 1304–1310. https://doi.org/10.1038/s41588-018-0178-9

Yue, W.-H., Wang, H.-F., Sun, L.-D., Tang, F.-L., Liu, Z.-H., Zhang, H.-X., … Zhang, D. (2011). Genome-wide association study identifies a susceptibility locus for schizophrenia in Han Chinese at 11p11.2. *Nature Genetics*, *43*(12), 1228–1231. https://doi.org/10.1038/ng.979

Zhang, J. P., Robinson, D., Yu, J., Gallego, J., Fleischhacker, W. W., Kahn, R. S., … Lencz, T. (2019). Schizophrenia polygenic risk score as a predictor of antipsychotic efficacy in first-episode psychosis. *American Journal of Psychiatry*, *176*(1), 21–28. https://doi.org/10.1176/appi.ajp.2018.17121363

Zheutlin, A. B., Dennis, J., Linnér, R. K., Moscati, A., Restrepo, N., Straub, P., … Smoller, J. W. (2019). Penetrance and pleiotropy of polygenic risk scores for schizophrenia in 106,160 patients across four health care systems. *American Journal of Psychiatry*, *176*(10), 846–855. https://doi.org/10.1176/appi.ajp.2019.18091085

Figure 1

Genetic architecture of schizophrenia. The risk allele frequency of SNPs and CNVs identified in Psychiatric Genomics Consortium (PGC) schizophrenia datasets are shown on the x axis. The effect sizes of risk allele are shown on the y axis. Many common alleles with small effect sizes have been identified (shown by diamonds), consistent with the common disease common variant hypothesis. Using genomewide approaches, some rare copy number variants have also been identified (shown by squares), consistent with common disease rare variant hypothesis. Only a few less-common variants (0.01<MAF≤0.05) have been identified due to large sample sizes. The effect size of identified risk allele is approximately inversely proportional to allele frequency.
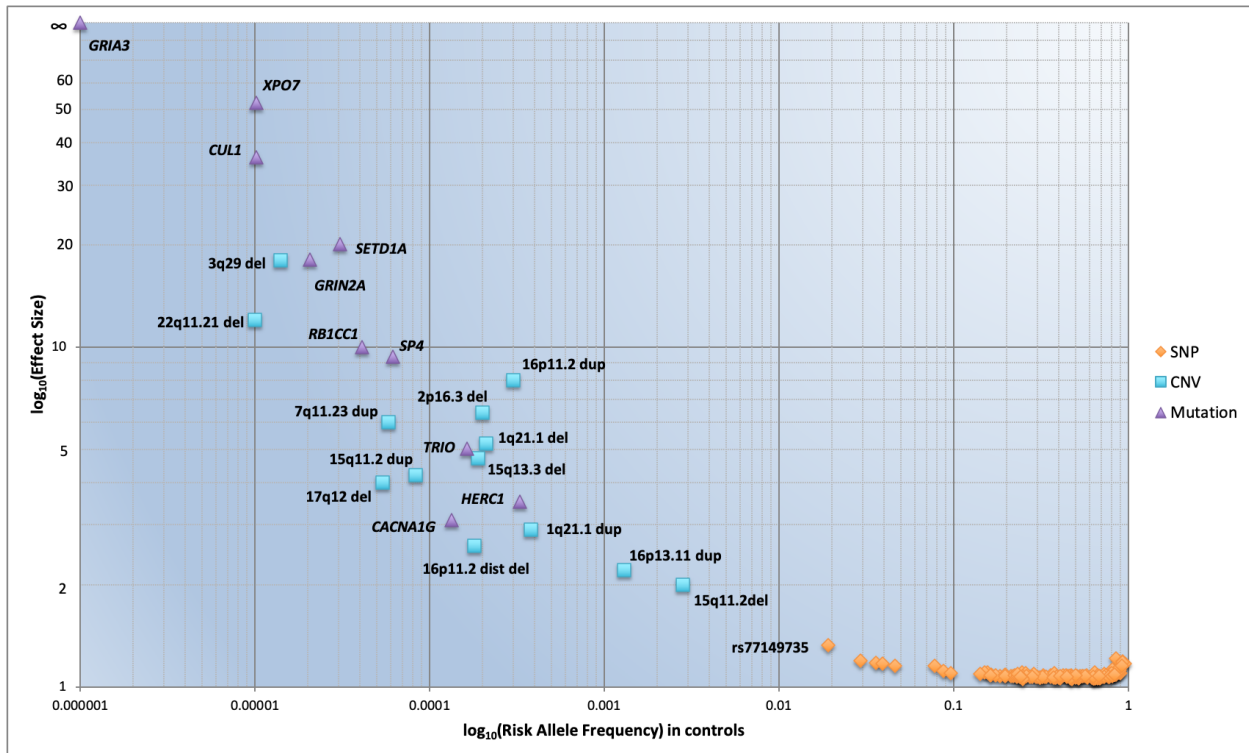
Figure 2

The history of reported schizophrenia GWAS: The key publications [PMID, year of publication] since the initial discovery of schizophrenia genome-wide significant loci are shown on the x-axis. The sample size of reported schizophrenia GWASs are shown on the left-hand y axis. The number of independent loci discovered in reported GWASs are shown on the right-hand y axis. SWG-PGC, schizophrenia working group of the psychiatric genomics consortium.

	[19571808], 09	[19571811], 09	[19571809], 09	[21926974], 11	[22037552], 11	[22037555],11	[23974872], 13	[25056061], 14	[28991256], 17	[29483656], 18	[31740837], 19	SWG-PGC medRxiv, 20
Family-Trios							581	1235				
Replication-Control	21093	15490	16424	21397	5603	4539	19762	66236	5770	154224		142627
Replication-Case	10282	4686	4327	8442	4027	4383	7413	1513	4384	5662		1979
Discovery-Control	13498	3587	2653	12462	1599	6468	18310	45604	83493	64643	78818	94015
Discovery-Case	2663	3322	2681	9394	746	3750	13833	34241	50874	40675	56418	67390
Loci	3	1	1	7	2	2	22	108	113	145	176	270

Figure 3

Genetic correlations (rg) of psychiatric disorders and related phenotypes with schizophrenia (Brainstorm Consortium et al., 2018). Error bars represent 95% confidence intervals.
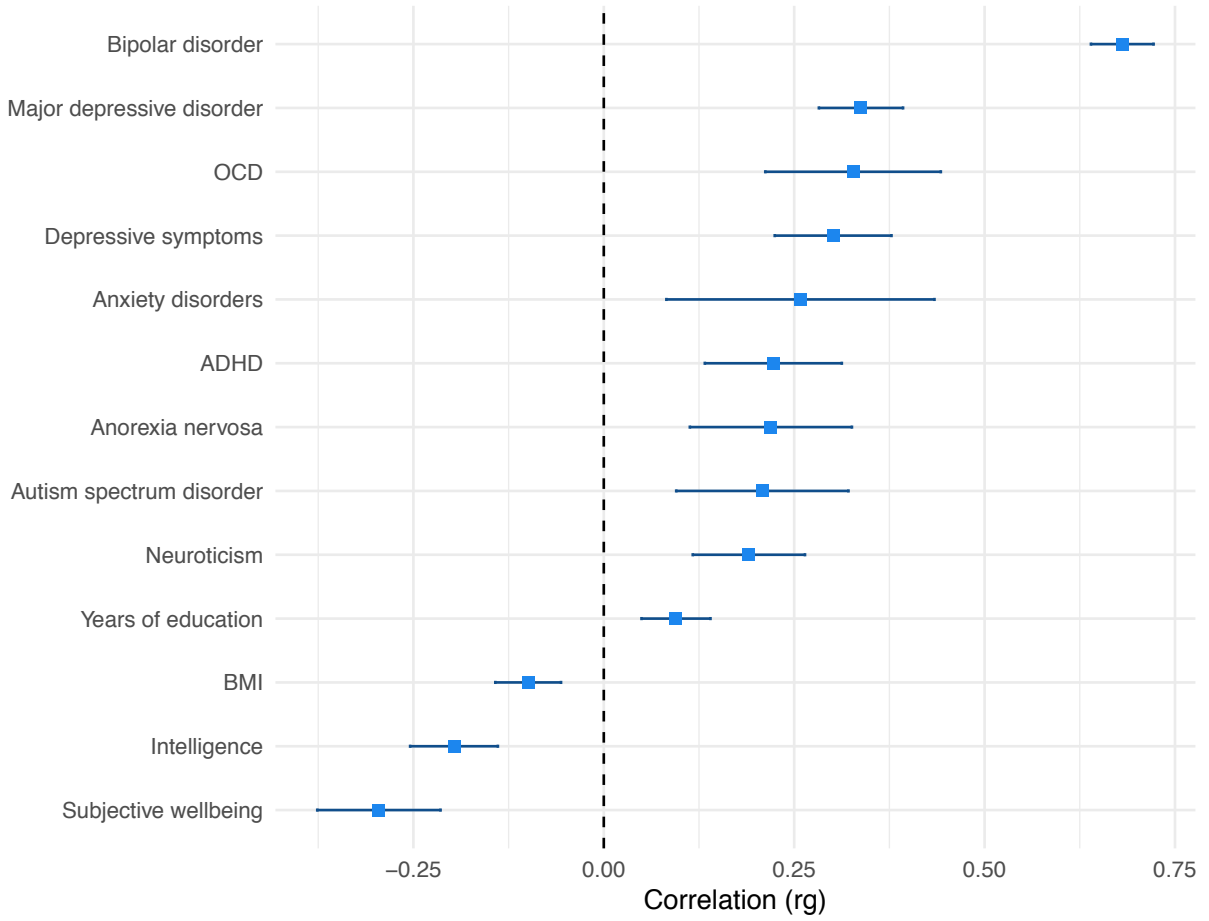
Figure 4

Overview of schizophrenia genetic findings.

SUMMARY OF SCHIZOPHRENIA GENETICS

HERITABILITY
Schizophrenia has a strong genetic component. Twin and family studies estimate heritability at ~80%. SNP based estimates currently explain ~24% of the heritability.

Common variants (>1%)

SNPs
The largest GWAS to date by the PGC (69,369 cases, 236,642 controls) identified 329 independent genome-wide significant SNPs mapping to 270 distinct loci. Fine-mapping refined association to 130 genes most likely to explain these associations.

CNVs
The largest study (21,094 cases, 20,227 controls) in schizophrenia identified 8 CNV loci to be associated, collectively explaining 0.85% of variance in liability. 1.4% of cases carried risk CNVs. The most strongly associated CNV for schizophrenia is 22q11.2 deletion which confers a 20-fold risk.

Rare variants (<1%)

SNVS
The latest SCHEMA Consortium analysis (24,248 cases, 97,322 controls) identified 10 genes with an excess of ultra-rare variants in schizophrenia cases. Other studies have found an increased burden of disruptive and damaging ultra-rare variants, and loss-of-function variants enriched in loss-of-function intolerant genes.

DNVs
In an analysis of 3444 trios, loss-of-function DNVs were found to be significantly enriched in loss-of-function intolerant genes in schizophrenia cases and this burden was higher for genes previously associated with neurodevelopmental disorders.

POLYGENIC RISK SCORES
PRS is a single measure of common genetic liability. It can explain ~7.7% of the variance in schizophrenia case/control status with an AUC ~0.71. PRS calculated in one population explains less variance when applied to another, highlighting need for greater diversity in GWAS.

PLEIOTROPY
Risk variants for schizophrenia are also associated with other psychiatric and somatic disorders. Genetic correlations of common variants show significant overlap with most psychiatric conditions, cognitive ability, and BMI. Rare variants associated with schizophrenia are all also associated with intellectual disability and developmental disorders.

BIOLOGICAL INSIGHTS
Gene set analyses provide evidence that risk variants for schizophrenia act by disrupting the functioning of the synapse through multiple inter-related mechanisms. GWAS and rare variant studies have identified associations with targets of FMRP, voltage-gated calcium ion channel complexes, synaptic plasticity, and abnormal long-term potentiation. Rare variant studies have also found associations with ARC and NMDAR gene sets.