Current progress in understanding Schizophrenia using genomics and pluripotent stem cells: A Meta-analytical overview

Ashwani Choudhary 1, Ritu Nayak 1, David Peles 1, Liron Mizrahi 1 and Shani Stern 1*
1 Sagol Department of Neurobiology, University of Haifa, Haifa 3498838, Israel
*correspondence at ssstern@univ.haifa.ac.il

ABSTRACT

Schizophrenia (SCZ) is a highly heritable, polygenic neuropsychiatric disease, which disables the patients as well as decreases their life expectancy and quality of life. Common and Rare variants studies on SCZ subjects have provided more than 100 genomic loci that hold importance in the context of SCZ pathophysiology. Transcriptomic studies from clinical samples have informed about the differentially expressed genes (DEGs) and non-coding RNAs in SCZ patients. Despite these advancements, no causative genes for SCZ were found and hence SCZ is difficult to recapitulate in animal models. In the last decade, induced Pluripotent Stem Cells (iPSCs)-based models have helped in understanding the neural phenotypes of SCZ by studying patient iPSC-derived 2D neuronal cultures and 3D brain organoids. Here, we have aimed to provide a simplistic overview of the current progress and advancements after synthesizing the enormous literature on SCZ genetics and SCZ iPSC-based models. Although further understanding of SCZ genetics and mechanisms using these technological advancements is required, the recent approaches have allowed to delineate important cellular mechanisms and biological pathways affected in SCZ.

Keywords: Schizophrenia, GWAS, Transcriptomics, Non-coding RNAs, iPSC, Patient-iPSC-derived neurons
1. INTRODUCTION

Schizophrenia (SCZ) is a complex neurological disease that manifests mainly in adulthood although reported to have a neurodevelopment origin (Bloom 1993; Fatemi and Folsom 2009; Ahmad et al. 2018; Schultz, North, and Shields 2007). According to the WHO report, globally approximately 24 million people are affected with SCZ with a worldwide prevalence of around 1% (Stilo and Murray 2010; Rhoades, Jackson, and Teng 2019). SCZ patients are more likely to die younger than the general population mainly due to co-morbid conditions like cardiovascular, diabetes and communicable diseases (Wildgust, Hodgson, and Beary 2010). With the rise in global life expectancy and aging populations, psychiatric disorders like SCZ will further increase the socio-economic liability (Lee et al. 2018), particularly in low and middle-income countries. A few reports have also highlighted the worldwide increase in the incidence of mental disorders like anxiety and depression including psychosis with the COVID-19 pandemic (Xie, Xu, and Al-Aly 2022; Brown et al. 2020; Chacko et al. 2020).

SCZ is characterized by positive symptoms like hallucinations and delusions, negative symptoms like social and emotional withdrawal, and cognitive symptoms like confusion and lack of decision-making (Tandon et al. 2013; Kahn et al. 2015; Page et al. 2022). SCZ is clinically diagnosed according to the criteria mentioned in the DSM V manual based on the characteristic symptoms of hallucinations and delusions; personal and work-related relationships; duration of symptoms and a pre-existing psychiatric condition etc. (Tandon et al. 2013; Kahn et al. 2015). SCZ was observed in relatives and families of patients and is believed to be inherited even before the advent of newer technologies like Genome-Wide association studies (GWAS) and Next-Generation sequencing (NGS) (Henriksen, Nordgaard, and Jansson 2017). Meta-analytical studies of monozygotic and dizygotic twins have suggested an 80% heritability factor for SCZ (Cardno and Gottesman 2000; Sullivan, Kendler, and Neale 2003). Although both candidate gene approaches and high-throughput genomics approaches have
identified approximately 170 genomic loci associated with SCZ but the majority are not useful for clinical applications (Henriksen, Nordgaard, and Jansson 2017; Lam et al. 2019; Trubetskoy et al. 2022). An increasing number of reports have formidably associated environmental risk factors in the development of SCZ. Environmental risk factors such as maternal infections and stress as well as maternal malnutrition, which leads to pregnancy and birth complications, also support the neuro-developmental hypothesis (Stilo and Murray 2010; Fatemi and Folsom 2009). Other risk factors for SCZ include socioeconomic status, urban living, increased paternal age, migration status, head injury, autoimmune diseases, epilepsy, and drug abuse (Cannon, Jones, and Murray 2002; Malaspina et al. 2001; Lederbogen et al. 2011).

SCZ is primarily managed clinically with first and second-generation antipsychotics that are known to act on the dopamine D2 receptor and are more effective for positive symptoms (Kane and Correll 2010; Tandon, Nasrallah, and Keshavan 2010). Recent reports show that second-generation antipsychotics, despite being thought to have better efficacy and fewer neurological adverse effects, show comparable efficacy to first-generation antipsychotics and equally increase the incidences of metabolic diseases in SCZ patients (De Hert et al. 2011; Nielsen et al. 2015; Leucht et al. 2013). Treatment-resistant schizophrenia (TRS), which accounts for 30% of all SCZ cases, is treated with Clozapine, which is successful in reducing antipsychotic symptoms, especially in TRS patients, even though the mechanism of action of Clozapine is not fully understood (Potkin et al. 2020; Meltzer et al. 2003). The current gap in the treatment of SCZ and other psychiatric diseases (Kohn et al. 2004) is mainly in the development of drugs that could treat the negative and cognitive symptoms effectively with minimal side effects (Kane and Correll 2010).

Neuroimaging techniques have highlighted structural abnormalities in the brain of SCZ patients. Enlargement of lateral ventricles, thinner cortical areas and reduced hippocampal volumes are established changes found in SCZ patients (van Haren et al. 2011; Haijma et al. 2011).
Combining neuroimaging tools with the genetic variants data has been a new field of research to further validate and understand the function of the genetic variants in the SCZ patient subgroups (Callicott et al. 2005). Neuroimaging tools have also been used to validate two environmental risk factors – urban life and migration status (Akdeniz et al. 2014).

Animal models have significantly aided in the understanding of brain development, particularly in informing about evolutionarily conserved cellular and molecular pathways (Pinnapureddy et al. 2015). Although transgenic models have helped in unraveling some aspects of monogenic diseases, the complexity and heterogeneity of polygenic diseases are difficult to replicate in such model systems (Kaiser, Zhou, and Feng 2017). Because of the interactions of multiple genetic and environmental risk factors, neuropsychiatric diseases such as SCZ, Bipolar disorder (BD), and Autism spectrum disorders (ASD) are difficult to study entirely in an animal model system (Nestler and Hyman 2010). At the cellular and molecular level, the use of iPSC technology in neuropsychiatric diseases has provided an alternative strategy to understand the complexity of patient-specific cell lines albeit with limitations as reviewed previously (Soliman et al. 2017; De Los Angeles et al. 2021; Nayak et al. 2021).

In this paper, we present a collective update of the findings from SCZ-iPSC model systems and knowledge gained through GWAS & transcriptomic studies on SCZ clinical samples. Although previous reports and reviews have discussed SCZ iPSC models, GWAS and transcriptomic studies in detail, our goal is to present a quantitative meta-analytic overview of the overall prevailing knowledge to understand the current advancements in the field of SCZ disease pathogenesis and progression.

2. DISCUSSION

2.1. What insights do genomic studies in SCZ provide?
Recent studies have established the existence of interaction between multiple genetic and environmental factors in the SCZ disease epidemiology (Kahn et al. 2015). Nonetheless, the contribution of genetic factors is large as SCZ is among the highly heritable neuropsychiatric disorder (more than BD and ASD) although without definite causative genes (Sullivan, Daly, and O'Donovan 2012; Schizophrenia Working Group of the Psychiatric Genomics 2014). Using classical genetics techniques some genes were found to be associated with SCZ but in very small sample sizes and which do not follow the Mendelian form of inheritance (Farrell et al. 2015; Sullivan, Daly, and O'Donovan 2012; Trifu et al. 2020). Over the last decade, GWAS has revolutionized the field of complex genetic diseases such as SCZ, by massively genotyping large sample sizes of patients to find genotype-phenotype relationships (Schizophrenia Working Group of the Psychiatric Genomics 2014; Uffelmann et al. 2021). GWAS and NGS have helped in capturing genetic variants, which constitute to partially reveal SCZ genetics (Rhoades, Jackson, and Teng 2019).

2.1.1. Common Variants

Psychiatric genomics consortium (PGC) published studies in large sample sizes to provide common gene variants associated with SCZ. Surprisingly, miR 137, a micro-RNA that regulates neural development, was shown to have the highest association with SCZ in the analysis of populations with European ancestry in 2011 (Schizophrenia Psychiatric Genome-Wide Association Study 2011a). Their subsequent paper in 2014 with samples from both European and Asian ancestry reported SNPs linked to genes associated with glutamatergic neurotransmission and synaptic plasticity (Schizophrenia Working Group of the Psychiatric Genomics 2014). DRD2 gene, which encodes for the dopamine D2 receptor that is also the known SCZ drug target, was also reported to be significantly associated with SCZ (Schizophrenia Working Group of the Psychiatric Genomics 2014). MHC (Major histocompatibility Complex) region, another important genetic variant, was first found to be
associated with SCZ by the Stefansson group (Purcell et al. 2009) and was also reported by subsequent studies (Irish Schizophrenia Genomics and the Wellcome Trust Case Control 2012; Ripke et al. 2013). A recent paper by the PGC consortium has found 120 genes to be involved in fundamental processes such as synaptic organization, neuronal differentiation, and neuronal transmission. Among these genes were the glutamate receptor subunit GRIN2A and transcription factor SP4 and other genes that were linked to rare disruptive coding variants in patients with SCZ (Trubetskoy et al. 2022).

To search for consensus genetic variants across different studies, we have performed a meta-analysis of SCZ GWAS as mentioned in the methods section. The majority of the GWAS on SCZ populations to date has been performed in European and East Asian populations (Fig1.a.). A few studies have also been done on African, Latin American, and South Asian populations among others. Our analysis resulted in a list of highly significant genes that were repeatedly reported as significant across different GWAS publications (p-value <0.01, see Methods) (Fig1.b). CACNA1C was found to be reported in most of the studies along with the HLA locus of the MHC region. CACNA1C is a calcium channel known to have a role in various neurodegenerative diseases (Indelicato and Boesch 2021). It has also been found to be associated with BD in GWAS analysis of common variants (Schizophrenia Psychiatric Genome-Wide Association Study 2011a). NOTCH 4, which has a role in inflammatory pathways, was also found to be highly replicated in multiple studies (Harb et al. 2021). Together with the MHC region, it highlights the important role that the immune system plays in SCZ. Some long intergenic non-coding RNAs were also reported by several studies, further reinforcing the role of non-coding RNAs in regulating gene expression as well as mechanisms that have not yet been fully understood (Fig1.b). TRIM 26 is another important gene that is known to regulate interferon-gamma signaling pathways and is associated with neural tube defects (Zhang et al. 2015; Zhao et al. 2021). However, the discovery of common genetic
variants in SCZ has only provided insights into a fraction of SCZ genetics and fuelled a debate about the missing heritability in complex diseases (van Dongen and Boomsma 2013; Zuk et al. 2012; Maher 2008). This has also led to the hypothesis of the presence of undiscovered rare variants with high penetrance (Rhoades, Jackson, and Teng 2019).

2.1.2. Rare variants

Over the years, genetic studies in SCZ patients have been able to identify eleven effective rare structural variants in the genomic region, which have been associated with SCZ and reported to increase the risk of SCZ (Rees et al. 2014). These rare structural variants have neither very high penetrance, nor specificity, as they are also associated with other neurodevelopmental disorders. (Rees et al. 2014; Sullivan, Daly, and O'Donovan 2012) (Table 3). Among these rare variants, the copy number variation of 22q11.21 is well studied and increases the SCZ risk by 20 times but is also associated with Di-George syndrome, developmental delays, and intellectual disability (Cleynen et al. 2021). Most of these structural variants span a large genomic region and are difficult to study in model systems (Table 3). iPSC models of some of these variations have been studied and discussed below.

In addition to the rare structural variants, other rare variants have been identified recently using exome sequencing studies and Whole-genome sequencing approaches. RELN mutations were found to be associated with SCZ in Whole-genome sequencing of a Chinese family with SCZ-affected members. SETD1A loss of function gene variant was found to be significantly associated with SCZ and other developmental disorders in an exome study (Singh et al. 2016). A recent large study on exome sequencing by SCHEMA has found 10 ultra-rare variants in 10 genes, which included representative populations from all seven continents (Singh et al. 2022). Hence, more research using exome sequencing and NGS is required for the exploration of further important genes. Despite these findings, the major portion of SCZ genetics remains undiscovered (Singh et al. 2022). Moreover, the discovered genetic variants have not been
helpful in clinical management although recent development of Polygenic risk scores (PRS) profiling for estimating the cumulative effect of small effect-common genetic variants along with rare variants could be helpful in the prediction and intervention of SCZ in clinics in the future (Iyegbe et al. 2014).

2.1.3. Gene expression studies

The lack of discovery of causal genes by GWAS in SCZ makes the study of gene expression pertinent (Rhoades, Jackson, and Teng 2019; Stern, Zhang, et al. 2022). Technological advancements in RNA sequencing have replaced the incomprehensive microarrays-based gene expression assays and allowed for the study of the whole transcriptome of the cells and tissues. A meta-analytical review (Merikangas et al. 2022) has recently reported the list of shortlisted differentially expressed genes in SCZ after analyzing the transcriptomic studies performed from SCZ samples to date. These studies were performed mostly on post-mortem brain tissues and blood of SCZ patients (Fig.1.c). A few studies were also done on LCLs, skin fibroblasts, and iPSCs from SCZ clinical samples (Fig.1.c). The differentially expressed genes were shortlisted by Merikangas et al. (Merikangas et al. 2022) based on their three or more reports of the genes in different studies. We analyzed to find the overlapping shared genes between the highly reported genetic variants as determined by us in Fig 1.b. and the list of 160 differentially expressed genes (DEGs) in SCZ as mentioned by Merikangas et al (Merikangas et al. 2022). We found a list of 14 genes that were common between the highly reported common variant genes and the list of DEGs (Fig1.d). Further analysis of such genes could help in understanding their biological role in SCZ pathophysiology.

Noncoding RNAs have emerged as important biological regulators in brain development and disorders (Qureshi and Mehler 2012) (Salta and De Strooper 2012). As mentioned earlier, SNP associated with miR-137 has been shown to have a strong statistical association with SCZ in GWAS (Schizophrenia Psychiatric Genome-Wide Association Study 2011b) (Fig1.b). In our
analysis, we found miR-137 to be differentially expressed in PBMCs and blood of SCZ patients (Fig1.e). Furthermore, miR-7, miR-9, miR-26b, miR-34a, miR-132, miR-181b, and miR-212 were observed to be differentially expressed in a variety of tissues, including PBMCs, blood, and postmortem-brain samples from SCZ patients (Fig1.e). Some iPSC studies have also suggested the role of miRNAs in SCZ iPSC-derived NPCs and Neurons (Fig 1.e). GWAS datasets have also shown an association with long intergenic non-coding RNAs in several studies (Fig1.b). Recently, several IncRNAs have been reported to be differentially expressed in the SCZ clinical samples (Barry et al. 2014; Tian et al. 2018; Liu et al. 2018). Some of the IncRNAs reported in multiple publications have also been listed in Fig 1e. Circular RNAs, a new class of non-coding RNAs, also were shown to be involved in SCZ (Mahmoudi et al. 2019) but not enough studies have been published for meta-analysis according to our criteria (see Methods).

2.2. iPSC-based models and how they have helped in understanding SCZ pathophysiology

Since its discovery in the previous decade (Takahashi and Yamanaka 2006), iPSC technology has been utilized to model a variety of diseases, including neurodegenerative and neuropsychiatric disorders(Stern, Sarkar, Galor, et al. 2020; Stern, Sarkar, Stern, et al. 2020; Stern, Zhang, et al. 2022; Brant et al. 2021; Quraishi et al. 2019; Stern, Lau, et al. 2022; Rowe and Daley 2019). It has a considerable advantage in the research of fundamental mechanisms and pathways affecting neuropsychiatric illnesses such as SCZ, BD, and ASD that are otherwise difficult to model due to a lack of causative genes (Soliman et al. 2017; Stern, Sarkar, Stern, et al. 2020; De Los Angeles et al. 2021). Chiang et al. first reported the generation of iPSCs from an SCZ patient harboring a DISC1 mutation (Chiang et al. 2011). We have performed a quantitative meta-analysis for iPSC-based models of SCZ in Fig 2a (i-iii) and Fig 2b. as described in the Methods section by analyzing the findings of more than 50 papers.
published since the first report in 2011. The relevant phenotypes observed among the different SCZ-iPSC models are plotted in Fig. 2. b. iPSCs can be derived from a variety of somatic tissue but the majority of the research groups have utilized fibroblasts for reprogramming control and SCZ patient samples (Fig.2a.i.). The seminal articles used both lentiviral and episomal vectors to derive iPSCs for SCZ-iPSC models (Chiang et al. 2011; Brennand et al. 2011). Integration-free reprogramming approaches such as Sendai virus vectors and mRNA reprogramming methods have also been successfully employed for reprogramming into iPSCs (Fig.2.a.ii.).

Different techniques related to transcriptomic analysis (RNA-seq, scRNA-seq), gene editing (CRISPR/Cas9, TALENs), high-resolution microscopy, and functional analysis (Multielectrode arrays, Calcium imaging, and Patch-clamp) have been used with iPSCs, neural progenitor cells (NPCs), and neurons derived from SCZ patients to study the mechanisms underlying SCZ pathology. iPSCs are typically differentiated into neural cells through Embryoid Bodies (EBs), neuronal rosettes, and NPCs using specific chemicals and small molecules, and thereafter terminally differentiated into neuronal subtypes using neurotrophic and specific growth factors to obtain neural cell types. For NPC/ Neural Stem Cell (NSC) induction from iPSCs, dual SMAD inhibition has become the preferred method (Chambers et al. 2009). The majority of the studies were conducted on cortical glutamatergic neurons or mixed populations containing both excitatory and inhibitory neurons (Fig.2.a.iii.). A few reports also used overexpression of specific transcription factors to directly induce excitatory or inhibitory neurons from iPSCs to study the SCZ disease mechanisms (Li et al. 2019; Sellgren et al. 2019; Pak et al. 2021). Some studies generated cerebral brain organoids to further understand the SCZ disease pathology in a model, which better mimics brain structure (Fig.2.a.iii) (Kathuria et al. 2020; Stachowiak et al. 2017; Sawada et al. 2020). Here, we briefly discuss the disease relevant findings from SCZ-iPSC-based models, which are also presented in Fig.2.b in the form of a heatmap (described in the Methods section). The heatmap
summarizes SCZ phenotypes, which were found to be either upregulated or downregulated among various neural cell types studied in SCZ-iPSC-based model studies.

### 2.2.1. Disease-relevant phenotypes observed in SCZ iPSC models

*Abnormal NPCs, Neurites, and Synaptic connections*

Brennand et al. conducted the first detailed study using an iPSC-based model to investigate SCZ in 2011, reporting reduced connectivity and neurite extension as well as lower expression of post-synaptic density protein 95 (PSD95) in iPSC-derived cortical neurons from patients with a family history of SCZ (Brennand et al. 2011). Transcriptomic analysis pointed towards significant changes in gene expression related to glutamate, Wnt and cAMP signaling pathways (Brennand et al. 2011). Furthermore, treatment with the antipsychotic drug loxapine improved neuronal connectivity and restored altered gene expression to some extent (Brennand et al. 2011). *In vitro* studies using iPSC-derived neurons and monocyte-derived microglia-like cells from SCZ patients have shown defective synaptic pruning in addition to abnormal synaptogenesis (Sellgren et al. 2019). Synaptic deficits, as well as change in redox state, were also observed in iPSC-derived cortical interneurons of idiopathic SCZ patients (Kathuria et al. 2019a). Treatment with N-acetyl serine, an antioxidant, was found to restore the observed synaptic deficits in SCZ patient-specific cortical interneurons (Kathuria et al. 2019a). In a recent study of discordant monozygotic twin pairs with SCZ, abnormal dendritic arborization resulted in hypoexcitable neurons and synaptic deficits in iPSC-derived dentate gyrus (DG) neurons of co-twin affected with SCZ compared to unaffected co-twin siblings. Moreover, gene expression analysis showed alteration of the DG development pathway, Wnt signaling pathway, and synapse-related pathways in the iPSC-derived DG neurons of SCZ-affected co-twins (Stern, Zhang, et al. 2022).

SCZ-iPSCs-based models with highly penetrant genetic risk factors have shed some light on the cellular phenotypes and genetic pathways. As previously stated, a chromosome 22q11.2
deletion is a known high-risk factor for SCZ. iPSCs derived from the 22q11.2 deletion possessed decreased differentiation ability (Pedrosa et al. 2011; Toyoshima et al. 2016). In addition, smaller neurosphere size and relatively poor neurite extension were also observed after differentiation of the patient iPSCs with 22q11.2 deletion compared to controls (Pedrosa et al. 2011; Toyoshima et al. 2016). In the 15q11.2 deletion iPSC model, another known SCZ risk factor, NPCs lacked apical polarity and adherent junctions indicating a disruption in neuronal development (Yoon et al. 2014). Additionally, iPSC-based cell models were used to study DISC1 mutation as it also carries an increased risk of SCZ (Wen et al. 2014). Synaptic vesicle release was observed to be dysfunctional in iPSC patient-derived neurons with mutant DISC1 compared to controls, and this was rescued when DISC1 was selectively edited in the patient cell line using TALENs, indicating that this mutation affects synaptic activity (Wen et al. 2014). In addition, TALEN and CRISPR-Cas9 editing were employed to introduce the DISC1 mutation into iPSCs resulting in an aberrant neuronal differentiation and a dysregulation of the Wnt signaling pathway as reported earlier by other groups (Srikanth et al. 2015). Exonic deletions of Reelin (RELN) have been linked to SCZ (Costain et al. 2013). In a recent study, glutamatergic and GABAergic iPSC-derived neurons from an SCZ patient with a RELN deletion were reported to have dendritic shortening and fewer synapse counts (Ishii et al. 2019). Those patient-derived neurons were compared to neurons differentiated from isogenic iPSCs specifically edited to have RELN deletion using CRISPR/Cas9 technology (Ishii et al. 2019). The cellular phenotypes of neurons differentiated from genetically edited iPSCs with RELN deletion were comparable with iPSCs of SCZ patients with RELN deletion, confirming the involvement of this gene in SCZ (Ishii et al. 2019). Another study investigating the role of SCZ risk factor, NRXN1 deletion in SCZ-related phenotypes, found altered neurotransmitter release in neurons directly induced from SCZ patient iPSC. They also found neurotransmitter release impairment in induced neurons when iPSCs were edited to have a
heterozygous deletion of NRXN1 using TALENs. Importantly, no such defects were observed in mouse neurons deficient of NRXN1 gene reinforcing the importance of human iPSC model for SCZ disease phenotypes (Pak et al. 2021).

**Mitochondrial dysfunction**

An increasing number of reports have linked mitochondrial dysfunction to SCZ (Dror et al. 2002; Prabakaran et al. 2004; Iwamoto, Bundo, and Kato 2004; Rosenfeld et al. 2011). Mitochondrial protein expression and a reduction in membrane potential were reported in SCZ iPSC NPCs, which suggested a compromised mitochondrial function (Robicsek et al. 2013). Altered mitochondrial morphology in SCZ iPSC-NPCs such as smaller and less connected mitochondria and a reduction in mitochondrial density around the nucleus has also been observed (Brennand et al. 2015). Other researchers reported a higher mitochondrial oxygen consumption (Da Silveira Paulsen et al. 2012) and increased reactive oxygen species (ROS) levels (Robicsek et al. 2013; Brennand et al. 2015) in SCZ iPSC-derived NPCs. Robicsek et al. in a subsequent study used isolated active normal mitochondria (IAN-MIT) technology to transfer healthy mitochondria to SCZ iPSCs which rescued the impaired differentiation of SCZ-iPSCs into glutamatergic neurons (Robicsek et al. 2017). A study found a reduction in ATP levels due to reduced oxidative phosphorylation at mitochondrial complexes I and IV in the forebrain excitatory neurons directly induced from iPSCs from 22q11.2 SCZ subjects (Li et al. 2019). Furthermore, mitochondrial protein expression was specifically altered and not the nuclear-encoded protein in those neurons (Li et al. 2019). Ni et al. found dysregulation of oxidative phosphorylation genes and altered mitochondrial function in cortical interneurons differentiated from SCZ patient-derived iPSCs compared to iPSCs derived from the control group (Ni et al. 2020). Surprisingly, they observed no such significant changes in glutamatergic neurons derived from the same SCZ patient iPSCs. Furthermore, Alpha Lipoic Acid/Acetyl-L-Carnitine (ALA/ALC) was found to restore the observed dysregulation of oxidative
phosphorylation genes in cortical interneurons (Ni et al. 2020). These observations suggest that alterations in Mitochondria structure and function may play an important role in SCZ pathology.

Altered Neuronal functions

Evidence for aberrant glutamate and GABA signaling in neurons derived from SCZ iPSC have been provided by several studies. Both transcriptomic and proteomic analyses have found dysregulation in receptor proteins as well as differential expression of glutamate and GABA signaling pathway genes (Brennand et al. 2011; Brennand et al. 2015; Kathuria et al. 2019b; Kathuria et al. 2020; Narla et al. 2017; Tiihonen et al. 2019). Calcium imaging and microelectrode array experiments have been used to measure the altered excitatory and inhibitory responses of SCZ patient-specific neurons in several studies (Tiihonen et al. 2019; Kathuria et al. 2019b; Kathuria et al. 2020; Hathy et al. 2020; Stern, Zhang, et al. 2022; Sarkar et al. 2018). Electrophysiological recordings have found reduced activity in iPSC-derived neurons from SCZ patients as well as changes in excitatory and inhibitory synaptic activity (Fig 2b). Clozapine was found to affect the differential calcium response to glutamate and GABA observed in SCZ iPSC-derived neurons (Tiihonen et al. 2019). Recent studies have found associations between electrophysiological measures and SCZ clinical status in iPSC-based models (Page et al. 2022). Cortical neurons derived from SCZ patients had altered sodium channel function and firing potential and increased GABA transmission as compared to controls (Page et al. 2022).

Findings from 3D models of SCZ brain organoid

Cortical malformation was observed in brain organoids of three SCZ patients and linked to FGFR1 signaling (Stachowiak et al. 2017). In addition, expression of the Reelin protein, known to have an important role in corticogenesis and earlier reported to be deficient in SCZ patients
and 2D SCZ-based models, was found to be decreased in the SCZ patient-derived organoids (Stachowiak et al. 2017). The brain organoid studies from monozygotic twins with SCZ revealed changes in the excitatory and inhibitory balance with an increase in GABAergic synaptic genes compared to healthy controls (Sawada et al. 2020). NPC proliferation and changes in the ventricles/rosettes per organoid were observed in DISC1 deficient SCZ brain organoids as compared to controls (Srikanth et al. 2018). Also, WNT signaling genes were found to be dysregulated in the transcriptomic analysis of these SCZ-iPSC brain organoids providing mechanistic clues to these observed cellular phenotypes (Srikanth et al. 2018). A recent transcriptomic study of cerebral organoids from SCZ patients found dysregulation of a few SCZ-associated genes that were previously found in related GWAS as well as the alteration in the genes related to mitochondrial function (Kathuria et al. 2020). The mitochondrial functional assay revealed alteration in Mitochondria-oxygen consumption rate in SCZ brain organoids along with changes in excitatory and inhibitory synaptic function measured via microelectrode arrays (Kathuria et al. 2020). Another recent proteomic study using brain organoids derived from SCZ patients found proteins involved in axonal pathways and morphogenesis to be depleted in SCZ. Two proteins (Pleiotrophin and Podocalyxin) reported earlier in GWAS studies were also found to be altered (Notaras et al. 2021). Together, these studies suggest alterations in NPC proliferation/differentiation, alterations of neurites and synaptic connections, and dysregulation of mitochondrial function in SCZ.

**Insights from Environmental risk factor studies**

Prenatal stress is known to be an important environmental risk factor for SCZ. Maternal immune activation (MIA) is a form of prenatal stress that can result from microbe infection during pregnancy (Brown and Derkits 2010; Estes and McAllister 2016). MIA has been associated with an increase in the copy number of LINE-1 retrotransposons in animal models as well as SCZ iPSC-derived neurons from 22q.11 deletion patients (Bundo et al. 2014). LINE-
1 retrotransposons are known to create somatic heterogeneity in neurons and were found to be dysregulated in SCZ postmortem brain samples (Doyle et al. 2017). Herpes simplex virus type 1 (HSV-1) infection of SCZ iPSC-derived neurons altered the expression of genes related to glutamatergic signaling and mitochondrial function compared to control iPSC-derived neurons which discloses the aggravating role of viral infection on SCZ genetic background (D’Aiuto et al. 2014). Rodent models have shown induction of several cytokines in the fetal brain after MIA indicating a neuro-inflammatory response (Estes and McAllister 2016). In experiments mimicking neuro-inflammation, IFN gamma treatment to the human iPSC-derived NPCs reflected the similar transcriptomic response reported in SCZ and Autism (Warre-Cornish et al. 2020). Another report studying the effect of activated microglia conditioned media on SCZ-iPSC-derived neurons found differential responses in cortical interneurons and glutamatergic neurons (Park et al. 2020). Cortical interneurons showed changes in gene expression related to metabolic pathways as well as mitochondrial dysfunction, reduced arborization, and synapse deficits when exposed to activated microglia conditioned media while the glutamatergic neurons derived from the SCZ–iPSCs did not change significantly in the study (Park et al. 2020). Moreover, in the presence of mitochondrial function- enhancer compounds (Alpha Lipoic Acid/Acetyl-L-Carnitine), these deficits in cortical interneurons were found to be rescued (Park et al. 2020). But besides these preliminary findings, more exhaustive iPSC-based research into the impact of environmental risk factors on SCZ-relevant biological pathways is required. In addition, the 2D iPSC-based models do not recapitulate the in vivo scenario of immunological and inflammatory response that necessitates a coordinated response from neuronal, immunological, and glial cells (Balan, Toyoshima, and Yoshikawa 2019).

2.2.2. Key cellular and molecular pathways identified

As discussed above, several studies have reported phenotypes as well as molecular signatures relevant to SCZ pathology in NPCs and differentiated neural cells from SCZ iPSC-based
A WNT signaling pathway dysregulation has been consistently reported by various groups employing both 2D and 3D iPSC-derived culture and organoids as well as postmortem clinical samples through transcriptomic and proteomic studies (Sawada et al. 2020; Topol et al. 2015; Brennand et al. 2011; Srikanth et al. 2015; Srikanth et al. 2018; Panaccione et al. 2013). Prior cellular and animal model studies have established the role of WNT signaling in cortical neurogenesis (Wrobel et al. 2007; Mutch et al. 2009). Hence, the dysregulation of the WNT signaling pathway may contribute to the observations of aberrant NPC proliferation/migration and synaptic alterations.

Another consistent observation in many studies has been mitochondrial dysfunction and oxidative stress (Brennand et al. 2015; Kathuria et al. 2020; Robicsek et al. 2013; Da Silveira Paulsen et al. 2012; Ni et al. 2020). Since normal mitochondrial function is important for neuritogenesis and axonogenesis, the alterations of the mitochondrial oxidative pathway may contribute to abnormal neurite and dysregulation of axonal pathways that were observed in SCZ-iPSC-based models (Leon et al. 2016). A few SCZ iPSC studies discussed above have also tested this proof of concept in rescue experiments by enhancing the mitochondrial function using compounds or transferring healthy mitochondria into iPSC-derived neurons (Park et al. 2020; Robicsek et al. 2017; Ni et al. 2020). Moreover, oxidative stress was found to regulate the WNT signaling pathway in neuronal and non-neuronal cells (Zhang, Tannous, and Zheng 2019; Almeida et al. 2007). In addition, oxidative stress is also known to regulate synaptic plasticity (Massaad and Klann 2011) and affect synaptic transmission in Alzheimer’s disease (Kamat et al. 2016). N-acetyl serine, an antioxidant, was found to rescue morphological alterations of synapses in SCZ patient-derived cortical interneurons indicating the role that oxidative stress plays in synaptic impairments in SCZ (Kathuria et al. 2019a). We, therefore, speculate that oxidative stress also plays an important role in the observed phenotypes in SCZ. However, these events may be interconnected and influence one another. The contribution of
glial cells to SCZ pathology was not studied extensively but in a relatively small number of studies (Akkouh et al. 2020; Akkouh et al. 2021; Windrem et al. 2017; Szabo et al. 2021). Mice implanted with human iPSC-derived glial cells from SCZ patients were observed to have abnormal behavior including anxiety and reduced social interaction (Windrem et al. 2017). The transcriptome analysis of differentiated glial progenitor cells of SCZ patients grown in vitro also showed dysregulation of genes related to glial differentiation and synaptic pathways in the same study (Windrem et al. 2017).

Although these lines of evidence point toward the underlying pathophysiology of SCZ, it is important to note that SCZ-based iPSC models have been derived from SCZ patients with different subtypes (Idiopathic or genetic) and have varied clinical phenotypes. Hence, it is important to further stratify the SCZ clinical samples and delineate the important pathways in each clinical subgroup for the designing and discovery of better therapeutic alternatives.

3. CONCLUSIONS

Advancements in genomics have increased the understanding of complex polygenic diseases like SCZ. GWAS and transcriptomic studies have significantly helped in the discovery of genetic variants and molecular pathways important in the pathophysiology of SCZ. While the causative genes of SCZ are still obscure, the discovery of common genetic variants from GWAS can be utilized for calculating PRS for further stratifying the SCZ patients’ samples based on their association with clinical symptoms (Zheutlin et al. 2019; Page et al. 2022). Apart from patient-specific cell lines from clinically diagnosed individuals, rare but highly penetrant mutations discovered via genetic studies & implicated in SCZ have been modeled using iPSCs to study the SCZ disease-specific neural cells in vitro. Rare genetic variants have aided in understanding the crucial genetic and molecular pathways but some of them are known to have a pleiotropic effect and are also associated with other neuropsychiatric diseases. Genome editing tools are still technically challenging for large genomic deletions and allow currently
for the editing of small genomic regions that are relevant for SCZ; some of these small genomic regions mutations have already been modeled but much more work remains. 2D cultures derived from iPSCs can be used for differentiating into SCZ patient-specific neuronal subtypes but the lack of coherence between the different methods of iPSC reprogramming and differentiation can add confounding factors to the conclusions (Nayak et al. 2021). In addition, further methodological advancements in iPSC disease modeling are required to fully recapitulate the in vivo brain development system (Noh et al. 2017; Balan, Toyoshima, and Yoshikawa 2019).

SCZ management necessitates a better understanding of the pathophysiology and progression of the disease, as well as improved diagnostic methods for prevention and treatment (Chien et al. 2013). Transcriptomic studies in response to drug treatment in SCZ cohorts may enable finding associations with genomic signatures and biological markers that can eventually help in SCZ patient stratification and the development of precision medicine (Stern et al. 2018). Future iPSC studies should further utilize the knowledge that was gained by the genetic studies to classify SCZ patient samples into subtypes that are based on the clinical phenotypes to better understand the underlying pathophysiology and mechanisms that is specific to the SCZ disease spectrum.
4. METHODS

4.1 SCZ GWAS Genes analysis

To collect the comprehensive data of GWAS on SCZ subjects, we used GWAS Catalog (https://www.ebi.ac.uk/gwas/) (Buniello et al. 2018) and searched the database with the term “Schizophrenia” and downloaded the resulting association data table (.tsv file). This data was imported to Matlab for analysis. We used the column SNP_GENE_IDS which contains the Ensembl ID of the genes that contain SCZ-related rs (SNP-risk allele) within their sequence. Ensembl ID was matched with HGNC approved gene symbol using the method described below.

To count the number of publications for each ethnic group, we unified multiple studies from the same publication. The African sub-categories and the “Hispanic” category were unified into the “Hispanic or Latin American” category. For each gene, we determined the frequency by counting the number of times a gene variant was reported in a study with a p-value of <0.01. Note that one research article may include several studies (see the (Buniello et al. 2018) for details). This frequency was used to determine the font size in the word clouds in Figure 1 (b) and Figure 1 (d). Word-cloud figures were generated by Matlab’s word cloud function. Genes that are marked in red also appear in the list of 160 differentially expressed genes (DEGs) related to SCZ which was reported in the recent meta-analysis review (Merikangas et al. 2022).

4.2. Common genes between SCZ GWAS and SCZ DEGs

To intersect these two lists of genes, gene symbols were first standardized using HUGO Gene Nomenclature (https://www.genenames.org/) as described below and both the lists were searched for common genes among them and the word cloud was plotted using Matlab. The font size of the gene name is proportional to the number of times it has been reported as DEGs across different transcriptomic studies in SCZ clinical samples.
4.3. **Standardization of Gene nomenclature**

We downloaded the list of Ensembl ID and their matching symbol from biomart: https://www.ensembl.org/biomart/martview/. We chose the “Ensembl Genes 106“ database and the “Human Genes (GRCh38.p13)” dataset. We chose the attributes “Gene Stable ID” and “Gene name” and downloaded the result CSV file. We also downloaded the database of HGNC Gene symbols from http://ftp.ebi.ac.uk/pub/databases/genenames/hgnc/tsv/hgnc_complete_set.txt which includes for each gene: approved gene symbol, previous gene symbols, alias gene symbols, and the Ensembl ID.

For each Ensembl ID in the GWAS data, we did the following process to obtain its gene symbol. We searched for the ensemble ID both in the biomart results and in the HGNC database. If the Ensembl ID was found in the HGNC database we used the matching approved symbol from the HGNC database. If it didn’t match any ensemble ID in the HGNC database, we used the biomart symbol and compared it to the list of approved HGNC symbols. If the symbol was an approved HGNC symbol then we used it. If it was not an approved symbol, we searched it in the list of previous gene symbols in the HGNC database. If the symbol matched a previous symbol of a gene then we used the currently approved HGNC symbol of that gene (and not the previous symbol). This allowed matching HGNC approved nomenclature for all of the Ensembl IDs without ambiguity. However, three Ensembl IDs (ENSG00000145075, ENSG00000284548, ENSG00000284548) did not have a matching symbol either in biomart or in HGNC and were not used for the intersection. These three IDs are labeled as novel genes on the Ensembl website (using https://www.ensembl.org search). This process resulted in three gene symbols that were changed from their biomart symbol to the matching HGNC:
Table 1. Gene variants from the GWAS database standardized according to HGNC approved nomenclature

<table>
<thead>
<tr>
<th>Previous symbol (biomart)</th>
<th>HGNC approved nomenclature (used for intersection)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CTB-178M22.2</td>
<td>TENM2-AS1</td>
</tr>
<tr>
<td>GPR1-AS</td>
<td>CMKLR2-AS</td>
</tr>
<tr>
<td>LINC02163</td>
<td>NIHCOLE</td>
</tr>
</tbody>
</table>

For the DEGs we did a similar process; if a gene symbol matched an approved HGNC gene symbol then we used it. If it did not match an approved HGNC gene, then we searched the previous gene names in the HGNC database and used the current symbol name. This process resulted in updating the following gene symbols:

Table 2. DEGs are standardized according to the HGNC approved nomenclature

<table>
<thead>
<tr>
<th>Previous symbol</th>
<th>HGNC approved symbol (used for intersection)</th>
</tr>
</thead>
<tbody>
<tr>
<td>C2orf82</td>
<td>SNORC</td>
</tr>
<tr>
<td>MARCH2</td>
<td>MARCHF2</td>
</tr>
<tr>
<td>SMEK2</td>
<td>PPP4R3B</td>
</tr>
<tr>
<td>APOPT1</td>
<td>COA8</td>
</tr>
<tr>
<td>C10orf54</td>
<td>VSIR</td>
</tr>
<tr>
<td>C9orf16</td>
<td>BBLN</td>
</tr>
</tbody>
</table>
4.4. Pie-chart distribution

All the pie charts were generated using Matlab from the collected data for respective figures. For Fig 1a and 1c, as mentioned above the respective data for ethnic groups in GWAS and the tissues used for transcriptomic studies were collected from the GWAS catalog and Merikangas et al.,2022. Fig1. a shows the number of publications that included each of the ethnic groups. Note that some papers included more than one ethnic group. For Fig 2a (i-iii), the pie charts show the distribution of data collected from approximately 52 papers, which used the iPSC-based model of SCZ. Note that for Fig 2a (iii), a single paper may include several cell types. The article list was generated from Pubmed using the search term "Schizophrenia iPSCs" from the timeline of 2011 to 2022. The relevant original research articles were referred and the data was collected for the "Type of tissue used" for reprogramming and "reprogramming methods" used for derivation of iPSCs as well as the "neural cell type" studied. We did not count approximately 8-10 articles for our analysis that had worked with previously derived iPSCs.

4.5. Heat maps
For Fig 1e, the differentially expressed miRNAs and lncRNAs from SCZ clinical samples including SCZ-iPSC-based models were listed. We analyzed for the presence of at least 2 reports for the same miRNAs and lncRNAs in any SCZ clinical sample. The list was converted to a heatmap using Matlab. The analysis did not include non-coding RNAs reported from animal models of SCZ. Also, not enough reports of circular RNAs in SCZ for metanalysis were available. Dark shade indicates the presence of differentially expressed non-coding RNAs in particular tissue and white indicates an absence of differentially expressed noncoding RNAs. Darker shades indicate more than one report. The color bar represents the frequency of differentially expressed non-coding RNAs across different publications.

For Fig 2b. The phenotypes reported from the SCZ-iPSC-related papers were listed and plotted into heat maps using Matlab. We also referred to data from the tables mentioned in (Räsänen et al. 2022). The blue color denotes the downregulation (↓) and the green upregulation (↑) of disease phenotypes. We have used arrows for simple representation. Both cortical glutamatergic neurons and induced glutamatergic neuron phenotypes have been unified in one cell type. A similar analysis was performed for GABAergic neurons. The color bar represents the frequency of phenotypes observed across different SCZ-iPSC-based models.

4.6. Tables

Table 1 and Table 2 are the list of genes for which the nomenclature was standardized according to HGNC approved symbols. Table 3 is based on previous literature reports on SCZ rare structural variants. We specifically also referred to the detailed meta-analysis report in (Rees et al. 2014) for rare structural variants reported for SCZ.
Table 3. A list of eleven rare structural variants implicated in SCZ and reported to be significant in previous genetic meta-analysis studies (Rees et al. 2014). See methods for further details. ASD-Autism spectrum disease; ADHD-Attention deficit-hyperactivity disorder

<table>
<thead>
<tr>
<th>Variation</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>1q21.1</td>
<td>Deletion</td>
</tr>
<tr>
<td></td>
<td>Duplication</td>
</tr>
<tr>
<td>NRXN1</td>
<td>Deletion</td>
</tr>
<tr>
<td>3q29</td>
<td>Deletion</td>
</tr>
<tr>
<td>WBS</td>
<td>Duplication</td>
</tr>
<tr>
<td>15q11.2</td>
<td>Deletion</td>
</tr>
<tr>
<td>AS/PWS</td>
<td>Duplication</td>
</tr>
<tr>
<td>15q13.3</td>
<td>Deletion</td>
</tr>
<tr>
<td>16p13.11</td>
<td>Duplication</td>
</tr>
<tr>
<td>16p11.2</td>
<td>Duplication</td>
</tr>
<tr>
<td>22q11.2</td>
<td>Deletion</td>
</tr>
</tbody>
</table>
APPENDIX 1

FIGURE LEGENDS

Figure 1.

(a) A pie chart distribution of GWAS performed from different populations/ethnic groups across different GWAS publications. The plotted GWAS datasets were extracted from the GWAS catalog as mentioned in the methods. The pie chart includes only the discovery cohorts and no replication cohorts as described in the methods.

(b) Genes significantly associated with SCZ according to the GWAS catalog. The font size is proportional to the frequency of the reported gene variant across different studies. Only studies which reported p-value < 0.01 for a specific variant were counted. Gene variants that were also reported as DEGs are colored in red.

(c) A pie chart distribution of tissue types used in different transcriptomic studies on SCZ clinical samples. The list of studies was chosen based on the analysis done by Merikangas et al., 2022. For details, see the Methods section. LCL-lymphoblasts cell lines; PFC- Prefrontal cortex; DLPFC-Dorsolateral prefrontal cortex; STG- Superior temporal gyrus regions; OFC: orbitofrontal cortex.

(d) A list of Common genes between Fig.1 (b) and highly reported DEGs across multiple transcriptomic studies on SCZ clinical samples (Merikangas et al., 2022). The font size of the gene is proportional to the frequency of the reported genes across different GWAS studies.

(e) A heatmap showing shortlisted differentially expressed noncoding RNAs (miRNAs and lncRNAs) in different SCZ clinical tissues and patient-specific iPSCs. The list of miRNAs and lncRNAs was shortlisted based on the report of miRNAs and lncRNAs in at least 2 studies done on SCZ clinical samples including SCZ patient iPSCs. Dark shade indicates the presence of noncoding RNAs in particular tissue and white indicates the absence of noncoding RNAs. Darker shades indicate more than one report. PFC- Prefrontal cortex; DLPFC-Dorsolateral prefrontal cortex; STG- Superior temporal gyrus regions of postmortem brain. The color bar represents the frequency of non-coding RNAs across different publications. See the Methods section for more detail.

Figure 2.

(a) A pie chart showing the distribution of i. Different clinical tissues are used for deriving iPSCs in different SCZ-iPSC-based publications. ii. The type of reprogramming
method used for deriving SCZ patient-specific iPSCs in different publications. iii.
Neural cell types studied in different SCZ-iPSC-based publications

(b) A heatmap for the phenotypes observed during differentiation of SCZ-iPSCs into
different cell types. The blue color denotes the downregulation (↓) and the green
denotes an upregulation (↑) of disease phenotypes. No information or report is denoted
by a black background. Differential gene expression observed in neural cell type has
been denoted with both upregulation and downregulation (↑↓) symbol for DEGs and
both blue and green colors has been used for DEGs. Both cortical glutamatergic neurons
and induced glutamatergic neuron phenotypes have been unified in one cell type. A
similar analysis has been done for GABAergic neurons. Proliferation-n stands for NPC
proliferation while differentiation-N stands for neuronal subtype differentiation.

Table 1. Gene variants from the GWAS database standardized according to HGNC approved
nomenclature

Table 2. DEGs standardized according to the HGNC approved nomenclature

Table 3: A list of eleven rare structural variants implicated in SCZ and reported to be
significant in previous genetic meta-analysis studies (Rees et al. 2014). See methods for further
details. ASD-Autism spectrum disease; ADHD-Attention deficit-hyperactivity disorder
REFERENCES

Ahmad, Ruhel, Vincenza Sportelli, Michael Ziller, Dietmar Spengler, and Anke %J Cells Hoffmann. 2018. 'Tracing early neurodevelopment in schizophrenia with induced pluripotent stem cells', 7: 140.


Almeida, Maria, Li Han, Marta Martin-Millan, Charles A. O'Brien, and Stavros C. Manolagas. 2007. 'Oxidative Stress Antagonizes Wnt Signaling in Osteoblast Precursors by Diverting β-Catenin from T Cell Factor- to Forkhead Box O-mediated Transcription*', Journal of Biological Chemistry, 282: 27298-305.

Balan, Shabeesh, Manabu Toyoshima, and Takeo %J Neurobiology of disease Yoshikawa. 2019. 'Contribution of induced pluripotent stem cell technologies to the understanding of cellular phenotypes in schizophrenia', 131: 104162.


Irish Schizophrenia Genomics, Consortium, and Consortium the Wellcome Trust Case Control. 2012. 'Genome-wide association study implicates HLA-C*01:02 as a risk factor at the major histocompatibility complex locus in schizophrenia', *Biological Psychiatry*, 72: 620-28.


Iwamoto, Kazuya, Miki Bundo, and Tadafumi Kato. 2004. 'Altered expression of mitochondria-related genes in postmortem brains of patients with bipolar disorder or schizophrenia, as revealed by large-scale DNA microarray analysis', *Human Molecular Genetics*, 14: 241-53.


Kathuria, Annie, Kara Lopez-Lengowski, Bradley Watmuff, Donna McPhie, Bruce M. Cohen, and Rakesh Karmacharya. 2019b. 'Synaptic deficits in iPSC-derived cortical interneurons in schizophrenia are mediated by NLGN2 and rescued by N-acetylcysteine', *Transl Psychiatry*, 9: 321.


schizophreniaNetwork-China Genetic, and Netherlands the. 2019. 'Comparative genetic architectures of schizophrenia in East Asian and European populations', Nature Genetics, 51: 1670-78.


Mahmoudi, Ebrahim, Chantel Fitzsimmons, Michael P. Geaghan, Cynthia Shannon Weickert, Joshua R. Atkins, Xi Wang, and Murray J. Cairns. 2019. 'Circular RNA biogenesis is decreased in postmortem cortical gray matter in schizophrenia and may alter the bioavailability of associated miRNA', Neuropsychopharmacology : official publication of the American College of Neuropsychopharmacology, 44: 1043-54.


derived homogeneous populations of developing schizophrenia cortical interneurons have compromised mitochondrial function', 25: 2873-88.


Potkin, Steven G., John M. Kane, Christoph U. Correll, Jean-Pierre Lindenmayer, Ofer Agid, Stephen R. Marder, Mark Olfson, and Oliver D. Howes. 2020. 'The neurobiology of treatment-resistant schizophrenia: paths to antipsychotic resistance and a roadmap for future research', npj Schizophrenia, 6: 1.


Purcell, Shaun M., Naomi R. Wray, Jennifer L. Stone, Peter M. Visscher, Michael C. O’Donovan, Patrick F. Sullivan, Pamela Sklar, Shaun M. Purcell, Jennifer L. Stone, Patrick F. Sullivan,


Schizophrenia Psychiatric Genome-Wide Association Study, Consortium. 2011a. 'Genome-wide association study identifies five new schizophrenia loci', Nature Genetics, 43: 969-76.

———. 2011b. 'Genome-wide association study identifies five new schizophrenia loci', Nat Genet, 43: 969-76.

Srikanth, Priya, Karam Han, Dana G Callahan, Eugenia Makovkina, Christina R Muratore, Matthew A Lalli, Honglin Zhou, Justin D Boyd, Kenneth S Kosik, Dennis J Selkoe, and Tracy L Young-Pearse. 2015. 'Genomic DISC1 Disruption in hiPSCs Alters Wnt Signaling and Neural Cell Fate', Cell Reports, 12: 1414-29.


Topol, Aaron, Shijia Zhu, Ngoc Tran, Anthony Simone, Gang Fang, and Kristen J. Brennand. 2015. 'Altered WNT Signaling in Human Induced Pluripotent Stem Cell Neural Progenitor Cells Derived from Four Schizophrenia Patients', Biological Psychiatry, 78: e29-e34.


Trifu, Simona Corina, Bianca Kohn, Andrei Vlasie, and Bogdan-Eduard Patrici. 2020. 'Genetics of schizophrenia (Review)', Experimental and therapeutic medicine, 20: 3462-68.


Zhang, Xiaojuan, Lijun Pei, Runting Li, Wei Zhang, Hua Yang, Yongchao Li, Yu Guo, Pingping Tan, Jingdong J. Han, Xiaoying Zheng, and Runlin Z. Ma. 2015. 'Spina bifida in fetus is associated with an altered pattern of DNA methylation in placenta', Journal of Human Genetics, 60: 605-11.


Fig.1.

(a) 

(b) 

(c) 

(d) 

(e) 

Blood Brain iPSCs

miRNAs lncRNAs

Frequency of noncoding RNAs in publications

CC-BY-ND 4.0 International license

made available under a which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is

The copyright holder for this preprint this version posted August 22, 2022. doi: bioRxiv preprint
Figure 1.

(a) A pie chart distribution of GWAS performed from different populations/ethnic groups across different GWAS publications. The plotted GWAS datasets were extracted from the GWAS catalog as mentioned in the methods. The pie chart includes only the discovery cohorts and no replication cohorts as described in the methods.

(b) Genes significantly associated with SCZ according to the GWAS catalog. The font size is proportional to the frequency of the reported gene variant across different studies. Only studies which reported p-value < 0.01 for a specific variant were counted. Gene variants that were also reported as DEGs are colored in red.

(c) A pie chart distribution of tissue types used in different transcriptomic studies on SCZ clinical samples. The list of studies was chosen based on the analysis done by Merikangas et al., 2022. For details, see the Methods section. LCL- lymphoblasts cell lines; PFC- Prefrontal cortex; DLPFC-Dorsolateral prefrontal cortex; STG- Superior temporal gyrus regions; OFC: orbitofrontal cortex.

(d) A list of Common genes between Fig.1 (b) and highly reported DEGs across multiple transcriptomic studies on SCZ clinical samples (Merikangas et al., 2022). The font size of the gene is proportional to the frequency of the reported genes across different GWAS studies.

(e) A heatmap showing shortlisted differentially expressed noncoding RNAs (miRNAs and IncRNAs) in different SCZ clinical tissues and patient-specific iPSCs. The list of miRNAs and IncRNAs was shortlisted based on the report of miRNAs and IncRNAs in at least 2 studies done on SCZ clinical samples including SCZ patient iPSCs. Dark shade indicates the presence of noncoding RNAs in particular tissue and white indicates the absence of noncoding RNAs. Darker shades indicate more than one report. PFC- Prefrontal cortex; DLPFC- Dorsolateral prefrontal cortex; STG- Superior temporal gyrus regions of postmortem brain. The color bar represents the frequency of non-coding RNAs across different publications. See the Methods section for more detail.
Fig. 2.

(a) (i) Lymphocytes (2)
LCLs (2)
Hair follicle keratinocytes (2)
Fibroblasts, lymphocytes (1)
Fibroblasts, T-cells, lymphocytes (1)

Fibroblasts (35)

(a) (ii) mRNA and Retrovirus (1)
mRNA (4)
Episomal (11)
Episomal and Retroviral (1)
Episomal and Sendai virus (1)

Sendai virus (8)
Retroviral (3)
NA (1)
Lentiviral (11)
Lentiviral and Sendai virus (2)

(a) (iii) Astrocites (3)
Cortical Neurons (6)
NPCs (15)
Microglial cells (1)
Brain Organoids (4)
OPCs (1)
GABAergic Neurons (6)
Glutamatergic Neurons (9)

(b)

<table>
<thead>
<tr>
<th></th>
<th>Astrocytes</th>
<th>Cortical</th>
<th>Hippocampal</th>
<th>GABAergic (cortical and induced)</th>
<th>Glutamatergic (cortical and induced)</th>
<th>Microglial cells</th>
<th>Brain organoids</th>
<th>NPC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spike frequency</td>
<td><strong>1.3</strong></td>
<td><strong>1.3</strong></td>
<td><strong>1.3</strong></td>
<td><strong>1.3</strong></td>
<td><strong>1.3</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Excitatory synapse counts and activity</td>
<td><strong>1.3</strong></td>
<td><strong>1.3</strong></td>
<td><strong>1.3</strong></td>
<td><strong>1.3</strong></td>
<td><strong>1.3</strong></td>
<td><strong>1.3</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mitochondrial dysfunction</td>
<td><strong>1.3</strong></td>
<td><strong>1.3</strong></td>
<td><strong>1.3</strong></td>
<td><strong>1.3</strong></td>
<td><strong>1.3</strong></td>
<td><strong>1.3</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Synaptic structure/connexion</td>
<td><strong>1.3</strong></td>
<td><strong>1.3</strong></td>
<td><strong>1.3</strong></td>
<td><strong>1.3</strong></td>
<td><strong>1.3</strong></td>
<td><strong>1.3</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dendritic arborization</td>
<td><strong>1.3</strong></td>
<td><strong>1.3</strong></td>
<td><strong>1.3</strong></td>
<td><strong>1.3</strong></td>
<td><strong>1.3</strong></td>
<td><strong>1.3</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Neurite length</td>
<td><strong>1.3</strong></td>
<td><strong>1.3</strong></td>
<td><strong>1.3</strong></td>
<td><strong>1.3</strong></td>
<td><strong>1.3</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Proliferation/differentiation</td>
<td><strong>1.3</strong></td>
<td><strong>1.3</strong></td>
<td><strong>1.3</strong></td>
<td><strong>1.3</strong></td>
<td><strong>1.3</strong></td>
<td><strong>1.3</strong></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Figure 2.

(a) A pie chart showing the distribution of i. Different clinical tissues are used for deriving iPSCs in different SCZ-iPSC-based publications. ii. The type of reprogramming method used for deriving SCZ patient-specific iPSCs in different publications. iii. Neural cell types studied in different SCZ-iPSC-based publications

(b) A heatmap for the phenotypes observed during differentiation of SCZ-iPSCs into different cell types. The blue color denotes the downregulation (↓) and the green denotes an upregulation (↑) of disease phenotypes. No information or report is denoted by a black background. Differential gene expression observed in neural cell type has been denoted with both upregulation and downregulation (↑↓) symbol for DEGs and both blue and green colors has been used for DEGs. Both cortical glutamatergic neurons and induced glutamatergic neuron phenotypes have been unified in one cell type. A similar analysis has been done for GABAergic neurons. Proliferation-n stands for NPC proliferation while differentiation- N stands for neuronal subtype differentiation.