CellPress

Review

Molecular Risk Factors for Schizophrenia

Shira Modai[1] and Noam Shomron[1,*]

Schizophrenia (SZ) is a complex and strongly heritable mental disorder, which is also associated with developmental–environmental triggers. As opposed to most diagnosable diseases (yet similar to other mental disorders), SZ diagnosis is commonly based on psychiatric evaluations. Recently, large-scale genetic and epigenetic approaches have been applied to SZ research with the goal of potentially improving diagnosis. Increased computational analyses and applied statistical algorithms may shed some light on the complex genetic and epigenetic pathways contributing to SZ pathogenesis. This review discusses the latest advances in molecular risk factors and diagnostics for SZ. Approaches such as these may lead to a more accurate definition of SZ and assist in creating extended and reliable clinical diagnoses with the potential for personalized treatment.

Trends

Although genetic and epigenetic variations are involved in the etiology of schizophrenia, current diagnosis is based on psychiatric evaluations.

Recent experimental advances have provided a growing list of genes, polymorphisms, and genomic alterations with strong associations with schizophrenia.

Innovative and integrated large-scale screens on schizophrenia samples have strengthened the identification of potential genetic risk factors.

A more accurate molecular definition of schizophrenia may lead to successfully achieving personalized individual treatments for patients.

The Complexity of Schizophrenia

Approximately 1% of the world's population is affected by schizophrenia (SZ), a severe neuropsychiatric disorder, with strong genomic and environmental developmental associations [1]. Diagnosed patients have a high risk of mortality [2] and morbidity, including a 10% risk of suicide and substance abuse [3,4]. The *Diagnostic and Statistical Manual of Mental Disorders* (DSM) characterizes SZ as one or more mental abnormalities including delusions, hallucinations, disorganized thinking (speech), grossly disorganized or abnormal motor behavior (including catatonia), and negative symptoms (flat affect or poverty of speech) [5]. The onset of SZ occurs mostly between the ages of 15–35 years. Treatment is more effective during the initial stage of the disorder; however, as in other mental disorders, the patient is often evaluated only after presenting symptoms, which appear at later stages of pathogenesis (Box 1) [6].

The difficulty in modifying the traditional diagnosis protocol to one that is based on genetic tests is based on the fact that SZ is a complex disorder. Namely, there is no strong identification of one or more genetic mutations with a deterministic effect (also known as 'Mendelian') [7]. While some genetic contributions to SZ have been established through high heritability of the disease [8–11], it is likely that multiple genetic, and possibly **epigenetic** (see Glossary), variations are responsible for the majority of SZ cases. These variations individually may confer little to no risk; however, the combination of several of these variations may be sufficient to induce the onset of SZ. Furthermore, SZ is also characterized by high impact environmental risk factors [10,12], such as maternal stress, which sets its effect during perinatal and early childhood [13,14].

One methodology to achieve a more dimensional approach, when studying complex diseases, is taken by the Research Domain Criteria. This framework studies mental disorders by integrating multiple sets of information, such as genomics and clinical information to better understand human psychiatric conditions [15]. A caveat of this may be SZ diagnosis *per se*, under DSM definitions; the precise classification does not guarantee biological and etiological validity, but is sought for clinical utility. As a result, from a biological standpoint, SZ represents a collective

[1]Faculty of Medicine, Tel-Aviv University, Tel-Aviv, 69978, Israel

*Correspondence: nshomron@post.tau.ac.il (N. Shomron).

Box 1. Schizophrenia and Diagnosis

A patient with a familial history of SZ and with characteristic clinical symptoms will address a psychiatrist to seek a treatment. The psychiatrist will interview the patient and a diagnosis will be made based on DSM-5 and ICD-10 (the International Statistical Classification of Diseases and Related Health Problems – 10th revision). A few caveats include patients misdiagnosed as a result of subjective evaluation [90]; a current version of the DSM removed subclassifications of the disease [91] resulting in a wider definition of SZ; in the event that the clinical picture changes over time, this may lead to mistreatment or an unnecessary hospitalization [92,93]; and, the long-standing diagnostic questionnaire is updated periodically leading to shifting definitions of SZ. However, as a result of the recent genomic revolution some findings might be able to assist in the molecular definition of SZ features, subclassification, and in the ease of recategorizing SZ patients.

syndrome in which many biological diseases are included. Thus, we caution that from a clinical and conceptual perspective, the risk factors associated with SZ may eventually be valid for a multisyndrome disease.

Genetic and epigenetic studies may potentially define disease diagnosis, prognosis, treatment, or clinical stratification. Yet, in this review we focus on the latest developments regarding the potential risk factors and indicators of SZ. With the aid of these, it may be possible to achieve personalized treatment plans for each patient.

Genomics of Schizophrenia

Genetic changes include those that directly affect the DNA sequence. Epigenetic changes affect DNA structure or chemical modifications, and might also alter RNA levels. Genetic and epigenetic changes are collectively termed 'genomics', representing a comprehensive in-depth analysis. The first genetic abnormality, suspected to cause SZ, was identified in 1970 by cytogenetics; in a study of an 18-year-old male, the gene 'Disrupted in Schizophrenia 1' (*DISC1*) was identified [16]. The potential for screening genomic features is gaining ground through the years, as sequencing technologies advance. These methodologies encompass whole-exome and whole-genome sequencing, profiling of transcriptomes (RNAs) and DNA modifications, to name a few. Using these comprehensive screens, the association between the *DISC1* gene and SZ has been validated [17]. However, other genes are collectively being listed as putative identifiers of, or contributors to, SZ (Table 1, Key Table).

The Impact of Genetics on Schizophrenia

Studies aiming to link genetic variation to disease typically examine either **single nucleotide polymorphisms (SNPs)**, **copy number variations (CNVs)**, **single nucleotide variants (SNVs)**, or **insertions or deletions (Indels)** [18]. Most SZ studies are focused on screening only one type of variant. For instance, known associations of CNVs with SZ psychopathology are limited, although in some cases CNVs have been reported to alter the expression of a specific protein or group of proteins, leading to pathogenesis [7]. The 22q11.2 microdeletion in velocardiofacial syndrome may alone confer a risk for SZ (1% among SZ patients) [19]. However, it is likely that not one but a combination of multiple types of variations are responsible for the majority of SZ cases [7], emphasizing the importance of integrative studies.

When the analyzed data includes parental information, these variations can be classified as either inherited or *de novo*. The latter are of great importance as the fecundity in people with SZ is low (first child fertility incidence rate ratios among men is 0.1 [20]), and hence tend to be removed from the population in a natural selection process. An increased prevalence of rare *de novo* CNV mutations have been shown in SZ patients. Kirov *et al.*, for instance, detected 5.5% of patients with family history in a sample of 662 SZ proband–parents **trios**, compared with 2.2% among 2623 controls [21]. These were specifically located in neuronal related genes, which are known to contribute to the etiology of SZ, and might serve as potential SZ genetic markers [22]. Other studies have proven that NMDA receptor (NMDAR) dysfunction is a convergence point for the

Glossary

Biomarker: a portmanteau for 'biological marker', which refers to objective, accurate, and reproducible measurable indications of normal biological or pathogenic processes, as well as pharmacological responses to a therapeutic intervention. The biomarker is measured by its sensitivity and specificity.

ChIP techniques: structural variations involving chromatins, and the assemblies of condensed DNA on histones, may be detected using chromatin immunoprecipitation (ChIP) techniques. The integration of ChIP and a high-throughput method, either by sequencing (ChIP-seq) or microarray (ChIP-chip), can help identify open chromatin regions in DNA (associated with permissive transcription).

Epigenetics: changes in DNA structure, DNA modifications, or gene expression, not due to changes in the DNA sequence itself. This can be through chemical modifications such as DNA methylation, structural modifications such as histone or chromatin remodeling, or regulatory elements, such as DNA-binding proteins or miRNA expression. Epigenetic changes can either be heritable or non-heritable. It has been shown that these alterations can be affected environmentally.

Exome sequencing: the sequencing of the entire set of functional protein coding regions across the genome.

Genetic Variants:

Copy number variant (CNV): a large genetic variation, responsible for a change in the number of repeats of a large genomic DNA region (over 1000 base pairs), due to a deletion or a duplication.

***De novo* mutations:** mutations that were created in the affected individual rather than inherited from a parent.

Insertion or deletion (Indel): a variant that includes an addition or a deletion of nucleotides, which changes the total number of nucleotides and might cause a frameshift mutation. Indels have a wide range of size-based variability.

Single nucleotide polymorphism (SNP): a variation in a single nucleotide, which arises commonly within a population and differs between species or paired chromosomes.