Risk architecture of schizophrenia: the role of epigenetics

Dragan M. Svrakic\textsuperscript{a,b}, Charles F. Zorumski\textsuperscript{a}, Nenad M. Svrakic\textsuperscript{a}, Igor Zwir\textsuperscript{a}, and Claude R. Cloninger\textsuperscript{a}

\textbf{Purpose of review}
To systematize existing data and review new findings on the cause of schizophrenia and outline an improved mixed model of schizophrenia risk.

\textbf{Recent findings}
Multiple and variable genetic and environmental factors interact to influence the risk of schizophrenia. Both rare variants with large effect and common variants with small effect contribute to genetic risk of schizophrenia, with no indication for differential impact on its clinical features. Accumulating evidence supports a genetic architecture of schizophrenia with multiple scenarios, including additive polygenic, heterogeneity, and mixed polygenic-heterogeneity. The epigenetic mechanisms that mediate gene–environment (GxE) interactions provide a framework to incorporate environmental factors into models of schizophrenia risk. Environmental pathogens with small effect on risk have robust effects in the context of family history of schizophrenia. Hence, genetic risk for schizophrenia may be expressed in part as sensitivity to environmental factors.

\textbf{Summary}
We propose an improved mixed model of schizophrenia risk in which abnormal epigenetic states with large effects are superimposed on a polygenic liability to schizophrenia. This scenario can account for GxE interactions and shared family environment, which in many cases are not explained by a single structural variant of large effect superimposed on polygenes (the traditional mixed model).

\textbf{Keywords}
epigenetics, genetics, risk factors, schizophrenia

\textbf{INTRODUCTION}
According to family, twin, adoption and epidemiological studies of schizophrenia, multiple and variable genetic and environmental factors influence its development and expression [1–4]. Recent advances in genetic studies of schizophrenia and studies of epigenetic mechanisms that mediate gene–environment (GxE) interactions provide a basis for an integrative model of schizophrenia risk that accounts for genetic and environmental factors.

\textbf{GENETIC FACTORS IN SCHIZOPHRENIA RISK}
Lifetime risk of schizophrenia increases with the number of schizophrenia relatives: if one parent has schizophrenia, the risk for each child is 10–15%; if both parents have schizophrenia, the risk increases to 35–46% [1]. However, about 90% of individuals with schizophrenia have no schizophrenia parent, and about 60–80% have no first or second degree schizophrenia relatives. Hence, the majority of cases appear sporadic, even though the liability to schizophrenia is strongly (up to 80%) heritable [2]. Familial aggregation of schizophrenia is increased in risk environments [3], pointing to the importance of environmental factors in exposing the underlying polygenic liability. In other words, additional genomic, epistatic, and environmental (epigenetic) factors influence the risk of expression of illness in many cases,
Schizophrenia and related psychotic disorders

KEY POINTS

- Accumulating evidence supports a genetic architecture of schizophrenia with multiple scenarios, including additive polygenic, heterogeneity, and mixed polygenic-heterogeneity.
- The epigenetic mechanisms that mediate gene–environment (GxE) interactions provide a framework to incorporate environmental factors into models of schizophrenia risk.
- We outline an improved mixed model of schizophrenia risk in which abnormal epigenetic states with large effects are superimposed on a polygenic liability to schizophrenia; this scenario can account for GxE interactions and shared family environment, which in many cases are not explained by a single structural variant of large effect superimposed on polygenes (the traditional mixed model).
- As neuroplasticity from epigenetics modulates complex development of individuals, not groups or populations, research and treatment of people with schizophrenia must become more person-centered, that is, informed by their biology, psychology, environmental and life event history; still, common environmental factors affect populations, so prevention, risk prediction, and final common pathway models can be applied.

particularly in cases in which the liability is near the threshold for expression on average.

Common variants with small effects or rare variants with large effects?

Sequencing of the human genome has enabled genome-wide association studies (GWASs) of common genetic variation using single nucleotide polymorphisms (SNPs) as markers. So far, no common variant with odds ratio more than 1.2 has been identified. However, recent meta-analyses indicate that thousands of the same SNPs may cumulatively explain up to 30% of risk variance in multiple independent cohorts of schizophrenia [5]. Most recently, using a family-based sample to control for cryptic population stratification, enrichment across a large number of SNPs for the same sets of common alleles in distinct European populations has been shown [6**].

Larger structural lesions of the genome, such as copy number variants or CNVs (these are duplications and/or deletions of large DNA segments encompassing multiple genes), explain a higher portion of genetic variance in schizophrenia [7,8]. CNVs with replicated associations with schizophrenia (e.g., 1q21.1, 15q11.1, 15q13.3, 16p11.2, 16p13.1, 17p12, 22q11.2) are located in unstable genomic regions with high rates of mutations and account for only 2–4% of cases [7]. Most of the increased burden of CNVs in schizophrenia is conferred by gene variants within regions with moderate-to-low rates of mutations (less than one in 500 individuals) [7] explaining a higher number but not all cases of schizophrenia.

In summary, evidence shows that both rare variants with large effect and common variants with small effect contribute to genetic risk of schizophrenia with no indication for differential impact on its clinical features. Neither frequent SNPs with small average effect (polygenic model) nor rare CNVs and single gene lesions with large effect (heterogeneity model) can fully account for schizophrenia risk, which likely involves additional genetic (epistatic) and/or environmental (epigenetic) factors (e.g., to produce the steep rise in illness probability necessary to generate the observed recurrence risks to relatives) [9].

Is liability to schizophrenia a trade-off for advanced human evolution?

Unstable genomic regions with high mutation rates are evolutionarily recent (humans and chimps have the highest rates of CNVs) and correspond to the period of rapid brain evolution in hominids [10]. These flexible (unstable) regions of the genome are epigenetically most active [11] and, thus, represent targets of environmental influences that allow rapid adaptation to the surroundings. In humans, genes involved in neurodevelopment are overrepresented at sites with high rates of CNVs conferring schizophrenia risk [12]. Although controversial, brain plasticity, which has enabled humans to adapt to a spectrum of environments better than other species, may also be a liability for schizophrenia.

ENVIRONMENTAL FACTORS IN SCHizophRenia RISK

Recent reviews summarize environmental pathogens associated with increased risk of schizophrenia [3,13,14]. Sophisticated epidemiological designs demonstrate that many of these pathogens are causative (not just contributing or modifying), especially in the context of preexisting family history and/or schizotypal traits [3,15]. Environmental pathogens with documented association with schizophrenia risk include prenatal exposures (affecting early neurodevelopment) and postnatal exposures (affecting interactive experience-based cortical development).

Prenatal factors include maternal infections (e.g., influenza, toxoplasmosis, genitourinary infections) [14,16], maternal psychological problems
during pregnancy (e.g., major stress, depression) [17,18], hypoxia secondary to obstetric complications [19], and nutritional deficiencies and/or allergies, such as protein, choline, B vitamins, folate deficiencies [13], or high levels of immunoglobulin G directed at dietary gluten [20], among others. Prenatal pathogens rarely affect risk of schizophrenia in the absence of a family history of psychosis or schizotypal traits, but the risk is robustly greater in offspring with such histories [15,18].

Postnatal factors include cannabis use, psychosocial stress, migration, and urbanicity [3], among others. Similarly to prenatal exposures, postnatal factors are rarely sufficient to cause schizophrenia independently, but act primarily in the context of preexisting family history [3] or superimposed on pre-existing central nervous system (CNS) vulnerabilities relevant to schizophrenia (e.g., cannabis use superimposed on COMT or AKT1 polymorphisms) [21,22].

**EPGENETIC MECHANISMS MEDIATE GENE–ENVIRONMENT INTERACTIONS**

Specific models that integrate environmental factors into schizophrenia risk architecture have not been developed thus far. Recent evidence shows that most GxE interactions are mediated by epigenetic modulation of gene expression [23–25], but in some cases may involve activity-based rearrangements of existing postsynaptic protein networks and augmentation of preexisting vulnerabilities relevant to schizophrenia, as could be the case with cannabis use and COMT Val/Val polymorphisms [21]. Clearly, epigenetic mechanisms provide a framework to incorporate environmental factors into models of schizophrenia risk. Epigenomics focuses on regulations of gene activity that do not involve alterations in the nucleotide sequence, but are mediated mostly via methylation of gene promoters and/or covalent modifications of chromatin [23,25]. Multiple hormonal, enzymatic, and second-messenger cascades link the external environment with chromatin to modulate gene activity [24,25] in response to biochemical [25] and psychosocial exposures in animals [24–27] and humans [28].

Epigenetic programming of the genome is particularly active during gametogenesis and embryogenesis, when epigenetic marks for temporally and tissue-specific gene expression set in place to regulate key periods of neurodevelopment [23,25]. However, epigenetic modulations of DNA activity continue throughout lifespan [29], in response to changing or pathological environments [25], even in fully differentiated somatic cells and neurons [30]. This lifelong susceptibility to environmental influences provides a mechanism for rapid genome adaptations to environment. However, it also allows environmental pathogens to reach the cell nucleus and adversely affect the genome by creating abnormal epigenetic states that can dysregulate development and/or function and increase risk of many complex illnesses, including cancer and schizophrenia.

**EPGENETIC DYSREGULATION IN SCHIZOPHRENIA**

Epigenetic processes may be particularly important for understanding complex disorders like schizophrenia because they are consistently heritable despite having relatively weak and inconsistent association with individual genetic variants. Evidence for epigenetic dysregulation in schizophrenia includes exacerbation of psychosis during high methionine diets [31], hypermethylation of gamma-aminobutyric acid (GABA) genes resulting in GABA deficit in cortical neurons of schizophrenia patients [32], widespread DNA methylation differences in relevant neurobiological networks between monozygotic twins discordant for schizophrenia [33], and brain-wide methylation abnormalities in psychosis [34], among many others. Mill et al. [34] report aberrant DNA methylation in about 100 loci in schizophrenia, including genes regulating glutamatergic and GABAergic systems, stress responses, and neurodevelopment. Using a network-based approach to determine methylation patterns across different loci, they found a lower degree of modularity in schizophrenia, suggesting systemic epigenetic dysfunction rather than isolated missteps [34]. These and similar studies indicate that gene expression involving numerous neurodevelopmental genes, longitudinal patterns of gene activity, and modular networks of co-expressed genes (all regulated primarily by epigenetic mechanisms) are altered in schizophrenia.

Such massive involvement of neurodevelopmental genes suggests a nonspecific, global epigenetic effect and an early lesion, occurring at a time when both DNA methylation and neurodevelopment are highly active. Global epigenetic defects are unlikely due to transcription errors in copying DNA methylation marks in the course of rapid cell replication during embryogenesis (these are called primary epimutations) but rather to epigenetic dysregulation resulting from pathological GxE interactions. There is suggestive evidence, such as dermoglyphic irregularities in schizophrenia patients [35], that genes regulating early development of other ectodermal tissues may have been affected by the same pathogen, but complex neurodevelopment
is apparently more sensitive to the insult and the consequences are far greater. These early epigenetic marks are copied (propagated) through mitotic cell divisions throughout neurodevelopment as molecular precursors of evolving structural and functional CNS abnormalities.

**PRODROMAL CENTRAL NERVOUS SYSTEM AS AN EMERGING RISK INTERFACE: GENE X BRAIN X ENVIRONMENT INTERACTION**

Early aberrant neurodevelopment in schizophrenia is manifested postnatally as abnormalities in mental functioning, commonly called prodromal symptoms. We refer to the prodrome as a ‘risk interface’ and not an ‘intermediate phase’ of schizophrenia because there is no simple stepwise propagation of the preexisting pathology into manifest illness. In fact, only 16–40% of ultra high risk individuals, with positive family histories and early mental abnormalities, develop schizophrenia [36].

The prodromal brain is interposed between genes and environment, creating a complex interacting triad (genes–brain–environment) that must be incorporated into disease models of schizophrenia. Specifically, at every point in time, schizophrenia development will depend on the brain’s features at the previous time point. In an ongoing feedback–feedforward pattern, outcomes are modified by potentiating or protective factors, including self-correcting homeostasis. Over time, this early abnormal brain will either get worse and express schizophrenia, or remain the same, or possibly improve. Exceptions are observed in cases in which a strong genetic load drives a psychotic outcome, regardless of modifying influences. In the case of expressed schizophrenia, prodromal abnormalities amplify and disperse throughout the CNS, affecting all mental activity, including emotion, perception, motivation, and cognition. This is reflected as widespread structural and functional CNS abnormalities and aberrant connectivity of neural networks [37] involving multiple neurotransmitter systems.

**RISK ARCHITECTURE OF SCHIZOPHRENIA: TOWARD AN INTEGRATIVE MODEL**

There is accumulating evidence that schizophrenia reflects a complex, multifactorial, and probabilistic etiology involving variable interacting genetic and environmental factors with variable contributions in individual cases. In other words, research supports a composite model of schizophrenia risk architecture with multiple scenarios, specifically polygenic-threshold, heterogeneity, and mixed polygenic-heterogeneity models.

Recent GWASs have identified common genetic variants that segregate in schizophrenia populations supporting a polygenic model of schizophrenia risk and causal common variations [6**,38**]. Hence, in a number of cases, schizophrenia is manifest once the threshold of cumulative genetic burden is reached (discontinuous expression of a continuous liability). This polygenic-threshold model, however, does not account for the stable global incidence of schizophrenia despite negative selection pressure, although creative relatives of schizophrenia patients, who carry the polygenic liability but do not have schizophrenia [39*], could compensate reproductively for the low fertility of the patients.

In other cases, rare CNVs or single gene mutations with large effect have a major role in producing schizophrenia [8,40*,41]. Many different CNVs can cause schizophrenia but do so one or two at a time, consistently with the heterogeneity model [42*]. This model, however, does not account for the co-occurrence of different clinical subtypes of schizophrenia in the same family.

Most CNVs associated with schizophrenia are inherited from apparently healthy parents (i.e., cases appear sporadic), indicating incomplete penetrance, epistasis, or epigenetic effects. In other sporadic cases, de-novo CNVs are found [40*]. It is not clear whether these large genetic lesions act alone or superimposed on polygenic liability. Many sporadic cases can be explained by threshold effects on the expression of the liability, as supported by transmission patterns in children of monozygotic twins discordant for schizophrenia [43,44]. Nevertheless, at least some sporadic cases may be truly nonfamilial due to effects unique to the individual, including personal experiences or de-novo mutations [40*].

In yet other cases, polygenic liability to schizophrenia is potentiated by threshold effects of abnormal epigenetic states. As noted, environmental factors with small or no effect in the general population, like prenatal maternal infections, have robust effects in the context of preexisting family history of schizophrenia [16]. Likewise, familial aggregation of schizophrenia is increased in risk environments [3]. In other words, genetic risk for schizophrenia may be expressed in part as sensitivity to environmental factors [22**]. Here, abnormal epigenetic states with large effect are superimposed on preexisting polygenic liability in which each factor has a small average effect. This variant of the mixed model can account for GxE interactions and/or shared environmental effects, which in many cases are not explained by a single structural
variant of large effect superimposed on polygenes (the latter, or ‘traditional’, mixed model does not account for environmental factors in schizophrenia risk).

In summary, there are multiple factors that are large enough to affect the liability distribution, which is no longer a typical bell-shaped curve seen in polygenic disorders, but includes ‘bumps’ indicating admixture of multiple distributions contributing to liability (e.g., a large single gene mutation or a large CNV or abnormal epigenetic states).

Taken together, the above scenarios (also illustrated in Fig. 1) can account for epidemiological and clinical peculiarities associated with schizophrenia, such as its high global incidence and evolutionary resilience despite negative selection pressure, its sporadic appearance despite strong heritability, or co-occurrence of different clinical subtypes in the same family, among others.

**SHARED GENETIC LIABILITY FOR SCHIZOPHRENIA, AUTISM, AND MENTAL RETARDATION**

Genome-wide studies of structural polymorphisms are beginning to yield replicable results [7,45,46]. However, the five most consistent CNVs associated with schizophrenia are also consistently associated with mental retardation, autism, and epilepsy [7,45,46]. This shared genetic pathology is involved in key aspects of neurodevelopment, including synapse formation and maintenance, neuronal migration and survival, and regulation of basic excitatory and inhibitory mechanisms [45]. Hence, it may be thought of as a general liability for nonspecific brain dysfunction which is, as a core deficit, shared by autism, schizophrenia, mental retardation, epilepsy, and possibly other neurodevelopmental disorders. Phenotype-specific mechanisms that steer shared genetic liability toward autism, schizophrenia, or mental retardation must include other genomic factors (e.g., epistasis, protective genes, nonshared genetic polymorphisms) and/or nongenomic factors (e.g., nonshared and/or protective environments). Otherwise, the outcomes would be decided by neutral genetic differences or there would be no different phenotypes at all.

**Phenotype-specific factors in schizophrenia neurodevelopment**

Prenatal epigenetic influences appear to be global rather than gene-specific, affecting numerous genes regulating early neurodevelopment. However, different gestational periods may be associated with specific disturbances in fetal brain development and different adult psychopathology [47]. This may reflect a specific phase in neurodevelopment (e.g., early cell proliferation and differentiation vs. subsequent cell migration and synapse maturation) [47] or specific features of environmental pathogens (e.g., not all pro-inflammatory cytokines cross the placenta, and some, such as interleukin-6, do so early but not late in gestation) [48]. Hence, the same pathogen acting on the same genetic background at different phases of early neurodevelopment may lead to mental retardation, autism, or schizophrenia. Acting afterward, this may lead to schizotypy or even to normal outcomes. Postnatal pathogens, such as cannabis use or psychosocial stress, appear more outcome-specific for schizophrenia, as both usually occur after autism and mental retardation are already manifest.

Although genetic lesions affecting GABAergic and glutamatergic systems are shared among the...
Schizophrenia and related psychotic disorders

three phenotypes [45], the severity of this pathology may steer neurodevelopment toward schizophrenia. Indeed, multiple candidate genes affecting GABA (GAD1 and RELN) and glutamate (G72, DAAO, SRR, GRM3, Neuregulin1-ErbB4, DNTP1) [32,41] have been implicated in schizophrenia. Another example of severe glutamatergic pathology in schizophrenia is provided by Grant [49], who identified genomic abnormalities involving 46 N-methyl-D-aspartate receptor proteins in mental illnesses. Most abnormalities occur in schizophrenia (26 proteins), compared with mental retardation (18 proteins) and mood disorders (bipolar seven proteins, depression six proteins). In addition to classical studies demonstrating hypermethylation of GABA genes in schizophrenia [32], recent analyses of CNV intensities in genomic DNA in GABAergic interneurons in the hippocampus indicate transcriptional changes in GAD67 regulation that are circuitry-based and diagnosis-specific [50].

Moreover, analysis of dendritic spines forming glutamatergic synapses in CNVs associated with autism revealed an increase in spine or dendritic growth [51], whereas a similar analysis in schizophrenia suggests that a majority of schizophrenia-associated CNVs lead to a decrease in growth of dendrites or spines [52]. Therefore, although the molecular networks implicated in these distinct disorders may be related, particular mutations associated with each disease may produce different functional consequences [52].

Although the pathophysiology of schizophrenia likely involves the majority (if not all) of neurotransmitters and brain networks, system-level models postulate that concurrent hypogluta
matergic and hyperdopaminergic metastable states [53] may be specific for schizophrenia onset and clinical course.

CONCLUSION

The schizophrenia ‘spectrum’ disorder results from shared genomic liability that underlies a range of outcomes in biological relatives of schizophrenia patients [54]. This shared liability passes across generations with only some developing schizophrenia (rare, most severe cases), the majority developing less severe manifestations (e.g., lower IQ, schizotypy), and some escaping clinical symptoms altogether. Hence, our thinking must shift toward searching for the cause of extended schizophrenia phenotypes [3], not just schizophrenia proper. As neuroplasticity from epigenetics modulates complex development of individuals, not groups or populations, research and treatment of people with schizophrenia must become more person-centered, that is, informed by their biology, psychology, environmental and life event history. Still, common environmental factors affect populations, so prevention, risk prediction, and final common pathway models can be applied.

The ultimate purpose of defining gene–gene (GxG) and GxE interactions is to identify what they mean in terms of the pathways they involve. Reflective of complex and variable etiology and multidirectional GxG and GxE interactions, the same (different) genotype may lead to different (same) phenotypes and vice versa (this is commonly referred to as ‘multifinality’ and ‘equifinality’). Current methodologies have not solved the problem of interpretability (or semantics) of the GWAS results. For example, common variants with small individual effects might contribute more substantially to disease risk through nonadditive interactions among loci (epistasis). Because of the number of SNPs under consideration in a typical GWAS, it is impossible to identify interaction effects, unless additional assumptions are made which may invalidate the solution. Fortunately, some acceptable approaches, for example, specific network inference tools for specific applications, are beginning to emerge [52,55].

As gene expression is governed by allelic specificity, epigenetic status, noncoding RNAs, and interactions with co-involved gene products (proteins), optimal study designs should involve simultaneous analyses of genomic, epigenomic, and neuroproteomic data. Such comprehensive datasets can reveal sequential molecular processes involved in schizophrenia and ultimately lead to the understanding of different and variable biological pathways involved in schizophrenia. In leukemia research [56], integration of genome-wide epigenetic patterns with gene expression levels revealed hundreds of differentially expressed genes that distinguish dysregulated pathways that were missed by gene expression arrays alone. Whether this will be true of schizophrenia remains unknown, but less comprehensive approaches are unlikely to provide meaningful descriptions of this complex set of illnesses.

Acknowledgements

C.F.Z. was supported by MH077791 (from NIMH) and by the Bantly Foundation. I.Z. is supported, in part, by Grant TIN2009-13950, MYCIN, UGR, Spain.

Conflicts of interest

There are no conflicts of interest.
REFERENCES AND RECOMMENDED READING

Pages of particular interest, published within the annual period of review, have been highlighted as:
- of special interest
- of outstanding interest

Additional references related to this topic can also be found in the Current World Literature section in this issue (pp. 000–000).


This most recent GWAS provides compelling evidence that common genetic variants segregate in schizophrenia populations, supporting a polygenic model of schizophrenia risk and causal variation.

7. Sebat J, Levy DL, McCarthy SE. Rare structural variants in schizophrenia: one et al.

Disease-associated epigenetic effects may also interact with genetic variation to increase the risk of schizophrenia.

10. St Clair D. Copy number variation and schizophrenia. Schiz Bull 2009; 35:9–12.

Similarly to the Caspi et al. (2005) results, this study provides a example of augmentation of small genetic risk by environmental factors and highlights our mixed model of schizophrenia risk, wherein abnormal epigenetic states are superimposed on genetic background with small effect.


32. Guidotti A, Aita J, Davis JM, et al. Decreased reelin and glutamic acid decarboxylase 67 (GAD67) expression in schizophrenia and bipolar disorders. Arch Gen Psychiatry 2005; 57:1061–1069.

This study provides further evidence for a role in DNA methylation differences in mediating phenotypic differences between monozygotic twins and in the cause of both schizophrenia and bipolar disorder. Numerous loci demonstrated disease-associated DNA methylation differences between monozygotic twins discordant for schizophrenia affecting biological networks and pathways directly relevant to psychotic disorder and neurodevelopment.


This most recent GWAS confirmed the role of common genetic variation in schizophrenia in a substantial proportion of cases.

An excellent review of polygenic, heterogeneity, and mixed models of schizophrenia risk architecture.

This study demonstrates familial cosegregation of schizophrenia and creativity in disease-free siblings. This finding may help explain the stable global incidence of schizophrenia despite negative selection pressure, as creative relatives of schizophrenic patients, who carry the polygenic liability but do not have schizophrenia, could compensate reproductively for the low fertility of the patients.

This seminal study provides evidence that large recurrent CNVs are associated with increased risk of schizophrenia.
Schizophrenia and related psychotic disorders

50. Sheng G, Demers M, Subburaju S, Benes FM. Differences in the circuitry-based association of copy numbers and gene expression between the hippocampi of patients with schizophrenia and the hippocampi of patients with bipolar disorder. Arch Gen Psychiatry 2012; 69:550–561. Using both microarrays and quantitative real-time polymerase chain reaction, this work demonstrates that site-specific insertions and deletions of genomic DNA in GABA cells at a key locus of the hippocampal circuit are reflected in transcriptional changes in GAD67 regulation that are circuitry-based and diagnosis-specific.

53. Lisman JE, Pi HJ, Zhang Y. A thalamo-hippocampal-ventral tegmental area loop may produce the positive feedback that underlies the psychotic break in schizophrenia. Biol Psychiatry 2010; 68:17–24.