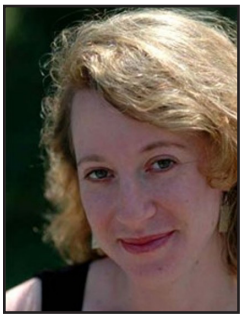
Unraveling the Complexity of Schizophrenia Genetics

Rebecca Birnbaum, M.D.[1,2] and Daniel R. Weinberger, M.D. [1-5]
[1]Lieber Institute for Brain Development, Johns Hopkins School of Medicine, Departments of Psychiatry[2], Neurology[3], Neuroscience[4] and The Institute of Genetic Medicine[5]

Schizophrenia is a psychiatric syndrome that is recognized and classified by its "phenomenology," its clinical presentation of positive, negative, and cognitive symptoms. Positive symptoms may include auditory hallucinations and paranoid delusions; negative symptoms may include diminished emotional expression, lack of social interaction, or avolition; cognitive symptoms may include executive dysfunction or deficits in working memory and attention. Schizophrenia occurs in about 1% of the world's population resulting in significant global morbidity and mortality. Notably during their lifetimes, compared to their healthy counterparts, patients with schizophrenia suffer from more physical illness and have shorter lifespans by about 15-20 years. For patients with schizophrenia, further compounding the psychological toll of disability is that they often must endure subtle and overt stigma in regular interpersonal interactions or entrenched within the very institutions from which they seek support.



Rebecca Birnbaum, M.D.

Daniel R. Weinberger, M.D.

Currently, schizophrenia medications are comprised of a small armamentarium of 'antipsychotics' with a single common feature: all available antipsychotic drugs inhibit a dopamine receptor pathway whose pathogenic relevance was serendipitously inferred in the 1960's and has since been the target of me-too drug development. At best, schizophrenia treatment ameliorates positive symptoms, but unfortunately has little to no effect on negative symptoms or cognitive symptoms, and overall is not curative. A single biomarker, blood test, or imaging study cannot be used to diagnose schizophrenia, but rather diagnosis at present remains entirely descriptive, just as it did in the early twentieth century, when the illness was first codified by classical German and Swiss psychopathologists and the name "schizophrenia" first coined by Eugene Bleuler. The quest for an understanding of the basic mechanisms of schizophrenia pathogenesis (origin and development of the disease) upon which a biologically-driven diagnostic and treatment system could be predicated, has challenged researchers for well over a century.

From this background, our understanding of the biological mechanisms of schizophrenia risk has steadily evolved over the past few decades, attributable largely to advances in human genetics and to genomic technologies. Twin studies comparing the concordance rates of illness in identical and fraternal twins led to the initial insight that schizophrenia is indeed highly heritable, meaning that differences in risk for illness were best explained by differences in individual genomes. This was followed by linkage analyses within family pedigrees that identified a handful of plausible schizophrenia susceptibility loci across the human genome, i.e. broad regions of the genome where susceptibility genes may be found. Studies focusing on individual genes contained within linkage regions or based on prevailing biological hypotheses, so called "candidate gene studies" produced a number of provocative findings, but their validity has been questioned because of inconsistencies across additional studies and uncertainty about the statistical evidence. A popular strategy for understanding how genes influence the brain and translate into the clinical syndrome involves identifying association of genes with aspects of brain structure or function that are found in healthy relatives of patients with schizophrenia, so-called intermediate phenotypes or "endophenotypes." This is based on the rationale that the effect of a risk gene is more penetrant (has a stronger effect) at the level of an intermediate phenotype, such as a cognitive test or a functional neuroimaging paradigm

associated with schizophrenia, rather than at the level of the clinical syndrome itself. A number of genes found to be associated with schizophrenia appear to influence the development and function of brain regions critically important for higher level cognition and behavioral controls, such as the prefrontal cortex and hippocampus.

More recently, advances in genomic technology and the development of a consensus strategy for finding genes associated with complex medical syndromes have enabled with high confidence the identification of genes related to risk of schizophrenia, catapulting the field forward with robust statistical evidence of specific risk-associated genetic loci (specific location of a gene or DNA sequence). These high powered genetic association studies termed "GWAS" (genome wide association study) survey millions of common genetic sequence variants called "SNPs" (single nucleotide polymorphisms) in large samples of cases and controls. Notably, a recent landmark schizophrenia genome-wide association study of 36,989 cases and 113,075 controls that was published by an international consortium comprised of investigators from over 50 medical centers identified 108 independent risk loci spread throughout the genome associated with schizophrenia. This report heralds a genetics breakthrough in understanding schizophrenia at the population genetics level and confirms the long-held view that schizophrenia is highly complex etiologically, with multiple and heterogeneous genetic risk factors. This biological insight of genetic complexity is not specific to the schizophrenia syndrome, but in fact is characteristic of many medical and neuropsychiatric diseases, including diabetes mellitus, macular degeneration, obesity, and even common to complex traits, such as height. Further, compounding the apparent genetic complexity is the interplay of genetic risk factors with environmental risk factors, only some of which have been identified in epidemiological studies to date, including prenatal stress, perinatal birth complication, marijuana use, or childhood residence in an urban compared to a rural environment.

Some recently identified schizophrenia risk loci contain genes that have already been investigated for many years as candidate genes based on biological hypotheses including: the DRD2 gene encoding the dopamine receptor D2; genes encoding glutamate receptor components (GRM3, GRIN2A, GRIA1); the SRR gene encoding serine racemase, an enzyme that synthesizes D-serine, a co-agonist at the glutamate NMDA receptor; and genes within the major histocompatibility complex (MHC) region encoding proteins with immunological function. However, there are many novel loci containing genes not previously investigated in schizophrenia biology studies that may yield insights into new pathogenic pathways and new drug target options, or conversely, may be false signals, artifacts from a large scale statistical analysis.

Now that the field of psychiatric genetics has emerged from a period of uncertain gene discovery to an overabundance of potential genetic risk factors, the challenges are daunting and the opportunities are vast. The 108 risk loci that were reported are just a starting point on a long and divergent path to validation and understanding mechanisms. Each of the genetic clues reported as "significant" and even thousands of loci reported as less significant will need to be thoughtfully characterized, by 'functional genomics' strategies, using expression analyses of genetic material derived from post-mortem brain tissue, or cell or animal modeling.

It is important to note that most of the reported common variants identified in the landmark schizophrenia genome-wide association study are located between genes or within non-coding regions of genes. Since they map to non-protein coding regions of the genome, they may be located in a regulatory region (e.g. so-called 5' or 3' UTR's or splice sites) or in a region that interacts with a regulatory region (such as an enhancer, insulator, or silencer) that would affect the transcriptional output of a gene. These regulatory regions do not change the amino acid sequence of the protein encoded by the gene but rather, influence gene assembly and expression. Therefore, one way to functionally characterize a risk locus is via an eQTL (expression quantitative trait locus) analysis, based on the hypothesis that the functional variant within any risk locus may exert its effect on a nearby gene by affecting the transcriptional output of the gene. Studying the associations between genetic variation and gene expression (eQTL analysis) may connect risk variants to their putative target genes or transcripts, and may be performed in human tissues, especially in postmortem brain.

It is also important to note that each of the 108 loci, many of which contain multiple genes, accounts for a tiny increment in individual risk across the populations studied, a very modest increase in the odds ratios of schizophrenia case status (generally less than 1.2 fold increases in risk). In actuality, the differences

in risk associated allele frequency between cases and controls are typically less than two percent. In other words, these loci cannot be used to stratify individuals based on their risk liability and are unlikely by themselves to have clinical impact. Statistically, these loci represent common denominators in genetic risk, identifying genetic variation that is relatively common across diverse populations of individuals with the diagnosis of schizophrenia. It is likely that within a given affected individual the risk architecture will include some of the genes in the 108 loci as well as genes that are not within these 108 loci, and that risk will also vary based on individual ancestries and individual environments.

Within a given genetic risk locus associated with schizophrenia the causal variant(s), or the causal gene(s), likely acts in a polygenic, broader biological context, additively and/or interactively, with other risk genes, and with environmental factors that influence risk pathways related to cell function and development. Therefore, multidimensional functional genomic analysis may be necessary to reveal the molecular basis of disease pathogenesis. Notably, the identification of *pathways* of risk may be especially clinically beneficial since they may lead to the identification of interacting genes outside of the top risk loci that may be "druggable", i.e. more optimal targets for therapeutic intervention. Recent pathways to date that have been identified by considering schizophrenia risk loci in aggregate include neuronal, immune, and epigenetic pathways.

One of the surprises from recent genetic association studies is that in a small percentage of cases with the diagnosis of schizophrenia, perhaps 2-3% carry a genetic mutation that has much greater impact on risk than the common variants. While these are relatively rare individuals, the strong biologic effect of the mutation may offer unique insight into biological events that underlie the development of the condition even in individuals without these rare mutations. This is work that is being pursued in a number of laboratories around the world.

Understanding the expression patterns of putative schizophrenia-associated genes in the brain can help to reveal the underlying neurobiology of schizophrenia risk. Assessing a risk gene's transcriptional activity (the step of turning DNA into RNA) could suggest the timing of its translation, and consequently the timing of cellular events critical to disease etiology and pathogenesis. Furthermore, the patterns of transcriptional activity of disease-associated genes and pathways might have implications for optimal timing of diagnostic predictions and potential therapeutic intervention. For example, if a given risk pathway is especially transcriptionally active during pre-natal development then diagnosis or a preventive intervention prenatally or peri-natally may be ultimately considered. Preliminary studies to date indeed indicate that genes associated with schizophrenia may be preferentially *fetally* expressed, suggesting that the genetics of schizophrenia is at least in part related to the genetics of brain development, as previously postulated in neurodevelopmental hypotheses of schizophrenia.

A novel application of the identification of common and rare genetic variants associated with schizophrenia may be the opportunity for pharmacogenomics, patient stratification based on schizophrenia risk genotype. Clinical trials and pharmacotherapy may be more efficient, if they are tailored to an individual specific risk genotype rather than a nebulous symptom or clinical presentation that is consequentially a mechanistic 'black box'. The notion of stratifying patients by genetic risk to improve therapeutic efficacy is a novel paradigm in psychiatry but already underway in other clinical specialties (e.g. Trastuzumab treatment for breast cancer patients who are positive for a protein called human epidermal growth factor receptor 2, HER2+, which promotes the growth of cancer cells, Ivacaftor for cystic fibrosis patients with a specific mutation that encodes a protein that is defective in chloride transport, CFTR G551D.

While there is still debate as to whether we have successfully traveled from famine to feast in clues to causal etiology related to schizophrenia, it is clear that in the past decade we have learned more about the causes of this syndrome than in all of past history. We now stand at a new starting point, a new precipice rooted in complex genetics, and on a path that may be biologically tractable, offering the promise of a refined understanding of schizophrenia and new treatment targets.

For further reading:

1. Schizophrenia Working Group of the Psychiatric Genomics Consortium. Biological Insights from 108 schizophrenia--□ associated genetic loci. Nature. 511, 421--□427 (2014).

2. Harrison P and Weinberger DR: Schizophrenia genes, gene expression, and neuropathology: On the matter of their convergence. Mol Psychiatry 10:40--68 (2005).

3. Weinberger D.R. and Levitt, P: The neurodevelopmental origins of Schizophrenia, In Schizophrenia, 3rd Edition, P Harrison and D.R. Weinberger (eds), Blackwell, Oxford (2011).

4. Birnbaum R., Jaffe, A.E., Hyde, T.M., Kleinman, J.E., Weinberger, D.R. Investigating The prenatal expression patterns of genes associated with neuropsychiatric disorders. Am J Psychiatry, 7, 758--□67 (2014).

5. Jaffe AE, Shin J, Collado-Torres L, Leek JT, Tao R, Li C, Gao Y, Jia Y, Maher BJ, Hyde TM, Kleinman JE, Weinberger DR. Developmental regulation of human cortex transcription and its clinical relevance at single base resolution. Nat Neurosci. 2015 Jan;18(1):154-61.