Can Genomics Help Usher Schizophrenia Into the Age of RDoC and DSM-6?

Ayman H. Fanous*1,2

1Mental Health Service Line, Washington VA Medical Center, Washington, DC; 2Department of Psychiatry, Georgetown University School of Medicine, Washington, DC

*To whom correspondence should be addressed; 50 Irving Street, NW Washington, DC 20422, US; tel: 202-745-8000 ext. 5-6553; fax: 202-518-4645; e-mail: ayman.fanous@va.gov

Introduction

In the last 7 years, we have witnessed the emergence of the first molecular insights into the etiopathogenic mechanisms of psychiatric illness. These have resulted from the first genome-wide association studies (GWASs) with sufficient power to detect small allelic effects. Two major factors have contributed to these advances. First, technological breakthroughs have made it possible to affordably interrogate the human genome in exquisite detail in large numbers. DNA microarrays have made it possible to rapidly genotype millions of single nucleotide polymorphisms (SNPs), providing information on common alleles, with the added benefit of also being used to detect structural copy number variants (CNVs), including deletions, insertions, and duplications. Next-generation sequencing methods have advanced such that the cost of sequencing whole genomes has diminished more than a million-fold. This has allowed the completion of mega-projects such as HapMap1 and the Thousand Genomes Project,2 which have in turn greatly benefitted studies of specific illnesses by allowing them to “fill in the gaps” in their own limited genomic data.

The second major factor has been the convening of large international consortia comprised of individual groups, which together have resulted in samples of tens of thousands of cases and controls.3 This has resulted from the realization that the small effect sizes of most common variants can only be detected using very large Ns, which could never be obtained in any single study. Schizophrenia has been one of the main beneficiaries of these advances. For the first time, numerous genetic markers have been reliably associated with illness across multiple populations. The Psychiatric Genomics Consortium (PGC) Schizophrenia Workgroup has by now amassed most of the existing GWAS samples in the field and continues to grow.4 The most recent “mega-analysis” has resulted in the identification of >100 significantly associated loci, of which 75% included protein coding genes.4 While many of the reported genes have been novel, with no previous suggestions of effect on schizophrenia risk, others have long been suspected. The latter include DRD2 and several genes involved in glutamatergic neurotransmission. At the level of rare variants (<1% minor allele frequency), a number of deleterious CNVs have been identified,5–8 and the accumulation of rare sequence variants has been in genes that have been implicated in GWASs and other genetic studies.9,10 While it is nearly certain that no discovered genetic risk factor will be deterministic, we are in a very different position now than when systems of psychiatric diagnosis and classification could only rely on expert consensus, which was used to generate operationalized criteria. In this article, I will describe recent and ongoing efforts to utilize the wealth of genomic information provided by GWA and other studies to clarify basic issues in the diagnosis and classification of schizophrenia.

Boundaries of Schizophrenia With Other Psychiatric Illnesses

Kraepelin’s most influential contribution was the distinction between Dementia Praecox and Manic-Depressive Illness on the basis of a chronic deteriorating vs relapsing-remitting course.11 This evolved into an often debated but relatively unchanged line of demarcation between schizophrenia and bipolar disorder (BPD), espoused by both clinicians and researchers. While earlier family studies largely failed to confirm familial coaggregation of the 2 disorders in general, there were some indications that psychotic BPD overlaps genetically with schizophrenia.12,13 This was supported by linkage studies reporting increased evidence of linkage when psychotic BPD was included in the definition of affection of schizophrenia.14 However, a very large registry-based family study in Sweden reported significantly increased risk of BPD in biological relatives of schizophrenia probands and vice versa.15 The first convincing molecular evidence of genetic
overlap was reported by the International Schizophrenia Consortium, which showed that polygenic risk scores (PRSs) based on the results of a GWAS of schizophrenia significantly predicted case-control status in independent BPD data sets but did not do so in nonpsychiatric disorders. This strongly suggested that at least some portion of common variants increasing risk for schizophrenia were shared with BPD.

As the PGC continued to grow in size and scope, further efforts by the Cross-Disorder Group examined genetic overlap at both the individual variant and polygenic levels across multiple disorders including schizophrenia, BPD, major depressive disorder (MDD), autism, and Attention Deficit Hyperactivity Disorder (ADHD). Using a model-fitting approach, a locus near NT5C2 was reported to increase risk for all 5 disorders. Schizophrenia was associated with 10 loci in that study, which was smaller than the most recent schizophrenia study. At 6 of them, a 2- or 3-disorder model fit best, most notably, CACN11C, which was associated with MDD, BPD, and schizophrenia. Further refinement of polygenic methods resulted in the more powerful Genome-Wide Complex Trait Analysis (GCTA) approach, which examines genome-wide genetic similarity among cases, which is then compared with genetic similarity among both cases and controls. This can then be used to calculate the risk to a disorder explained by all common variants taken together, called SNP-heritability. GCTA has been extended to similarly estimate SNP-coheritability or the proportion of disease risk due to common variants that is shared by 2 disorders. Among disorders represented in the PGC, the genetic correlations \( r_g \) of schizophrenia were .68, .43, and .16 to BPD, MDD, and autism, respectively. While schizophrenia and BPD clearly have clinical and biological overlap, they also comprise different clusters of symptoms at different times in their clinical course, which respond to different treatments. This suggests that a number of etiological mechanisms could be involved. Their diagnostic distinction is therefore therapeutically pragmatic. More recently, the PGC Schizophrenia and BPD Workgroups tested for variants distinguishing the two. This was done in a variation on traditional case-control analysis, by conducting GWAS of schizophrenia vs BPD as opposed to cases vs controls. While no individual variants reached genome-wide significance, several were moderately significant \( P < 10^{-5} \). More importantly, a polygenic score of schizophrenia vs BPD was seen to significantly predict schizophrenia vs BPD status in independent samples. Despite having the power to detect polygenic effects, more than likely, the study was underpowered to detect individual variant effects. However, it did suggest that a number of individual variants are disorder-specific. One caveat to this kind of study is that it assumes that the 2 disorders in question are themselves unitary, where they could in fact represent partially distinct genetic etiologies. Nevertheless, refinement and extension of analyses like this hold out hope of utilizing genetic information to assist in differential diagnosis, which might be especially useful in these 2 disorders, as they are sometimes difficult to distinguish cross-sectionally because of overlapping clinical features such as psychosis and agitation.

The distinction between schizophrenia and BPD at the level of common variants is reinforced by CNV findings, which have been very sparse in BPD. In contradistinction, CNVs at specific loci as well as genome-wide CNV burden have been associated with schizophrenia risk. On the other hand, schizophrenia does share reported CNV associations with both autism and intellectual disability. This suggests the intriguing possibility that negative symptom-like psychopathology, which schizophrenia shares with autism but not with BPD in general, is more likely than other clinical features to arise from CNVs. Because of their size, often encompassing many genes, CNVs might be more disruptive and lead to more severe brain structural and functional abnormalities and, hence, greater impairment.

**Clinical Heterogeneity of Schizophrenia**

The clinical heterogeneity of schizophrenia has been a subject of debate since the earliest modern descriptions. Kraepelin’s seminal distinction between *Dementia Praecox* and Manic-Depressive Illness was accompanied by his subsuming several heterogeneous conditions under the rubric of *Dementia Praecox*. These included catatonia, paranoia, hebephrenia, etc. However, it is sometimes forgotten that Kraepelin himself described up to 10 subtypes of *Dementia Praecox* in his influential textbook. Bleuler, working simultaneously with Kraepelin, posited that the psychotic disorders were a “Group of Schizophrenias.” Furthermore, his distinction between fundamental and accessory symptoms gave primacy to negative symptom-like clinical features as being more specific to schizophrenia. Much of the debate about the classification of schizophrenia since then has been a footnote to these 2 viewpoints. Diagnostic and Statistical Manual of Mental Disorders, Third Edition (DSM-3) carried over previous concepts about more distinct psychotic syndromes into the paranoid, catatonic, disorganized, and residual subtypes. Several decades of work have resulted in little empirical support for the validity, reliability, or longitudinal stability of these constructs, and there has been decreased usage of them in the research literature. This has led to their elimination in DSM-5. Meanwhile, clinicians and researchers have continued to document and attempt to explain the clinical heterogeneity of schizophrenia, with the development of numerous symptom rating scales and factor analytic and other studies attempting to validate symptom dimensions. While such studies have presented numerous models of psychopathology, the general distinctness of positive and negative symptoms has been generally
upheld, while disorganization has variably been conceptualized as independent or subsumed into the negative factor. More recently, the cognitive deficits prominently observed in the illness have been emphasized as contributing to psychosocial dysfunction and while associated with negative symptoms, might be a partially etiologically distinct feature of the illness.

Molecular genetic studies of clinical variation among schizophrenia cases have built on family studies, which have demonstrated sib-pair correlation for many clinical features as well as subtypes, suggesting that genes influence the pattern of clinical symptomatology seen in the illness. Linkage studies and the association studies that followed them have generally taken 3 approaches. In the first approach, affected individuals with a known genetic susceptibility factor are compared phenotypically to affected individuals without it. In essence, they go from genotype to phenotype, seeking to characterize the phenotypic effects of genetic risk factors, rather than to discover such risk factors. One linkage study reported that affected individuals from families linked 8p had higher levels of negative and lower levels of mood symptoms than those from other families. Association studies have focused on a number of candidate genes, selected either on the basis of a priori hypotheses, association mapping, and more recently, GWAS. A number of studies have compared patients carrying the high-risk variant to negative symptoms. Among genes identified by GWAS, the schizophrenia PRS correlated with the negative/disorganized dimension, but not the positive or affective dimensions, of the Lifetime Dimensions of Psychosis Scale. Although this correlation was significant, it explained a very small amount of the variance. However, we have seen the variance explained by PRS to be substantially larger in studies of larger samples and this study was relatively small by GWAS standards. In the larger but more heterogeneous PGC sample, this correlation was replicated and even more significant (AH Fanous et al, ISPG Abstracts, 2012). A recent attempt to replicate the initial MGS findings using 2 independent Irish samples showed significantly similar directions of effect among the top SNP associations using a sign test, for both positive and negative symptom dimensions (AC Edwards et al, submitted). Pathway analysis of these SNPs revealed differences between them. Positive, but not negative, symptoms were associated with pathways involved in addiction and dopamine synapse. On the other hand, negative, but not positive, symptoms were associated with immune pathways. In a similar analysis, the depressive dimension was uniquely associated with several pathways (AR Docherty et al, submitted). In the PGC, a number of phenotypic features are currently being examined. Manic symptoms in schizophrenia were shown to be predicted by the PRS for BPD, which had been based on independent BPD discovery samples. Phenotypes in ongoing analyses include the presence of lifetime suicide attempt, smoking behavior, and substance misuse. These studies necessarily rely on multiple stages of analysis as new sites join the consortium, with an incremental increase in sample size. Initial results suggest potential sex-specific effects on suicidal behavior (AH Fanous et al, ISPG Abstracts, 2013), as well as potentially novel regions influencing smoking quantity in schizophrenia that showed no evidence of involvement in the largest studies of smoking genetics in general population samples (TB Bigdeli et al, personal communication). However, these studies remain far less powered than recent case-control analyses.

The second, or case-only discovery approach, goes from phenotype to genotype. It aims to discover genetic risk factors for a specific trait (eg, symptom dimensions, comorbid conditions, course features, etc.), by examining only affected individuals. This approach is particularly suited for detecting modifier loci, defined as loci influencing the clinical form of illness, but not the risk to the illness itself. Modifier loci have been demonstrated in Mendelian disorders, such as cystic fibrosis, in which they have been shown to influence the rate of pulmonary function decline, a key index of severity, whereas CFTR is the only known disease susceptibility gene. In schizophrenia, this approach has yet to result in statistically significant linkage, although several suggestive linkages have been reported. In a GWAS of quantitative symptom dimensions in the Multicenter Genetic Studies of Schizophrenia (MGS) sample, 19 independent loci were moderately associated with positive, negative/disorganized, or affective symptoms. More interestingly, the schizophrenia PRS correlated with the negative/disorganized dimension, but not the positive or affective dimensions, of the Lifetime Dimensions of Psychosis Scale. These are defined as loci that affect susceptibility to a more-or-less phenotypically specific form of illness. Such genes have been demonstrated in somatic diseases. For example, FGFR2 is a susceptibility gene for estrogen receptor (ER)-positive, but not ER-negative breast cancer, which is associated with a significantly worse prognosis. In this approach, phenotypic subtypes can be determined a priori. However, in schizophrenia studies, they have often
been derived using empirical subtyping approaches, such as latent class analysis, based on lifetime ratings. This may be in part due to the limitations of subtyping based on operationalized criteria, including the heterogeneous and prevalent undifferentiated subtype. There have been a number of interesting findings using this approach. Deficit syndrome–like subtypes were significantly linked to 1q31 and suggestively linked to 20p, while a cognitive deficit subtype was linked to 6p. Ongoing work in the PGC has included empirical subtyping using cluster analysis of symptom dimensions. Subsequent GWAS was implemented using the resultant phenotypic definitions in a model-fitting approach, which determines whether association with one or more subtypes best fits the data, as previously published. Initial results suggest that individuals with more prominent negative and less prominent affective symptoms have significantly higher polygenic risk than other presentations (TB Bigdeli et al, personal communication). This is consistent with a previous study demonstrating higher polygenic scores in traditionally defined schizophrenia compared with spectrum conditions, as well as controls.

Research Domain Criteria: Occam’s Razor or Platonic Forms?

The Research Domain Criteria (RDoC) project is a major National Institute of Mental Health (NIMH) initiative aiming to explain mental illnesses on the basis of deficits in a small number of basic domains of function: Positive Valence, Negative Valence, Cognitive Systems, Social Systems, and Sleep/Arousal Systems. The putative biological substrates of these domains comprise increasingly complex “levels of analysis” (LOAs), which progressively superene on each other and serve as external validators. These include genes, molecules, cells, circuits, physiology, behavior, self-report, and paradigms. The underlying premises of the RDoC project include the theses posited above, namely (1) clinical and likely etiological heterogeneity within traditional definitions of illness; (2) clinical and likely etiological continuity across traditional diagnostic boundaries between disorders; and (3) etiological continuity of traits across boundaries of traditionally defined illness and health. To reconcile these observations in new systems of classification, it offers a parsimonious model whereby illness features, such as depressive symptoms, represent deficits in specific RDoC domains (eg, Negative Valence in the case of depressive symptoms). Such features can be observed in multiple disorders, including MDD, BPD, and schizophrenia. Furthermore, deficits in RDoC domains occur in a continuous gradation from severe symptoms (eg, melancholic features in depression) to those seen in milder forms of illness (eg, dysthymia) and further into personality/temperamental traits more or less normally distributed in the general population (eg, trait neuroticism). The model encompasses endophenotypes, biomarkers, and intermediate phenotypes, which can be observed at various LOAs.

The RDoC project holds out the promise of being “Occam’s Razor”—a parsimonious explanation of the almost limitless variation in psychopathology on the basis of 5 domains, each of which is undergirded by distinct biological processes. This could in turn provide the needed empirical evidence for more biologically driven diagnostic systems, biomarkers to assist in diagnosis, and the multi-level mechanistic causal relationships needed to design more effective therapeutics. In the case of schizophrenia, dimensional definitions of illness, such as the DSM-5’s 8 symptom domains, could be refined by understanding their structure and interrelationships at multiple LOAs. This could potentially shed light on the distinctness of features commonly grouped together such as hallucinations and delusions; anhedonia and restricted affect; and negative symptoms and conceptual disorganization. It could also clarify the etiological continuity of symptoms across disorders, such as delusions in schizophrenia and in psychotic form of BPD and MDD. As it currently stands, the RDoC matrix represents the commonly accepted facets of schizophrenia psychopathology but maps onto them in a complex, many-to-many relationship. For example, negative symptoms include asociality, a deficit in the Social Systems domain, as well as anhedonia, a deficit in the Positive Valence domain. In this article, Ford et al recently described how the cognitive, sensory, and emotive aspects of hallucinations fall under multiple constructs within 3 different RDoC domains.

While the potential benefits of validating this paradigm are clear, it is very early days in this “experiment toward classification.” RDoC is best conceptualized as a set of heuristic propositions to provide a framework for hypothesis generation and was never meant to be a classification system in itself. Ultimately, the parsimonious model may not hold. It may be the case that features seen in different disorders, such as anhedonia in depression and schizophrenia, have different etiologies. This is suggested by the fact that psychotic symptoms themselves can arise from etiologies in such disparate conditions as substance intoxication and numerous medical conditions. Furthermore, a key aspect of diagnostic definitions has been the temporal relationships of various signs and symptoms, which comprise episodes, which in turn comprise disorders. These temporal relationships may in fact represent critical etiological distinctions between disorders but are not addressed by RDoC. Nevertheless, RDoC presents an important opportunity to free our diagnostic and nosological thinking from the artificial constraints of consensus-based historical precedents and align it with biological insights acquired through contemporary scientific and technological breakthroughs. Many of these have been completely unexpected—evidence that our conceptual schemata with respect to etiology are simplistic at best and merit fundamental reexamination.
The kinds of genetic analyses described above will be critical building blocks of the evidentiary basis of RDoC constructs in schizophrenia. In fact, genes might have a privileged place in the RDoC matrix as the most basic LOA. As such, they would be expected to be more tractable analytically, as they are likely to have fewer causal inputs than more complex LOAs, including environmental ones. Disentangling the impact of more complex LOAs will also require accounting for the influence of genes. Future progress, of course, will be dependent on adequately powered studies. The largest available GWAS sample of dimensional traits in the PGC is only a fraction of that available for case-control analysis. The vast majority of GWAS have emphasized making reliable diagnoses but have not prioritized deeper clinical phenotyping. While this is understandable from the standpoint of the need for sensitive and specific determination of caseness, it has resulted in the use of numerous symptom inventories (PANSS, OPCRIT, etc.), which differ in their emphases, purposes, and granularity. This presents challenges in harmonizing such data across studies, and subsequently interpreting results to inform descriptive conventions, such as DSM-5’s dimensional classification. We anticipate substantial increases in power over the next few years with large samples becoming available for use in both discovery and replication of existing signals. Furthermore, we are at the cusp of the advent of large-scale whole-genome sequencing studies, which promise to provide information in unprecedented detail on coding and noncoding variation, both common and rare. This is in turn expected to be able to measure polygenic and single-variant contributions with far less error. Such studies have already begun to appear in the literature on somatic illnesses, such as an Icelandic study of type 2 diabetes in which rare variants were imputed into the genomes of more than 11 000 individuals.

Conclusions

The paradigm-shifting advances in genomics over the last decade have resulted in a number of biomarkers associated with schizophrenia, which have long been sought. While none of these are highly sensitive or specific, when taken together in the form of polygenic risk scores, they have provided insights into the etiological contours of schizophrenia. Recent studies have strongly suggested that BPD and schizophrenia share a number of genetic risk factors. However, genetic discontinuities also exist, telling against arguments for abolishing the separation of the 2 illnesses. They also suggest that the overall genetic risk of schizophrenia is slightly biased toward syndromes with more prominent negative symptoms. It is unclear whether this is a consequence of their being more stable and directly observable than other symptoms or a higher phenocopy rate in other psychotic presentations. Such studies are of necessity smaller than existing schizophrenia GWAS samples because of their greater phenotypic granularity, which requires greater time and expense. However, as the number of available samples is likely to increase in the next few years, we are likely to witness individual SNPs with genome-wide significant association to finer-grained phenotypes. Some of these may provide actionable targets for the development of novel therapeutics. One area not covered in this review is the genetic relationship between personality and other traits in the general population and schizophrenia risk. This is because there are as of yet no adequately powered GWAS of such traits, which is an unmet research need. In the decade to come, the RDoC initiative will undoubtedly increase the number of studies addressing issues of the diagnosis and classification of schizophrenia, whether directly or indirectly. It is hoped that this will provide a stronger empirical and biological basis for the development of DSM-6 as was available to benefit DSM-5.

References


