Current Research on the Genetic Contributors to Schizophrenia
Michael F. Pogue-Geile and Jessica L. Yokley
Current Directions in Psychological Science 2010 19: 214
DOI: 10.1177/0963721410378490

The online version of this article can be found at:
http://cdp.sagepub.com/content/19/4/214

Published by:
http://www.sagepublications.com

On behalf of:
Association for Psychological Science

Additional services and information for Current Directions in Psychological Science can be found at:

Email Alerts: http://cdp.sagepub.com/cgi/alerts
Subscriptions: http://cdp.sagepub.com/subscriptions
Reprints: http://www.sagepub.com/journalsReprints.nav
Permissions: http://www.sagepub.com/journalsPermissions.nav
Current Research on the Genetic Contributors to Schizophrenia

Michael F. Pogue-Geile¹,² and Jessica L. Yokley¹
¹ Department of Psychology, University of Pittsburgh, and ² Department of Psychiatry, University of Pittsburgh

Abstract
In this article, we review genetic research on schizophrenia to illustrate current strategies, findings, challenges, and future directions in the study of a relatively common, severe psychopathological phenotype. Family, twin, and adoption studies indicate that overall genetic effects on schizophrenia are both important and complex. Although efforts to identify specific causal genetic variants have utilized the full range of molecular and analytic techniques, results have been modest. Several putative common genetic variants of small effect appear to be implicated along with some extremely rare variants of potentially large effect. It seems clear that most of the genetic contributors to the liability to schizophrenia, as well as nongenetic ones, remain to be identified. New strategies give reason for optimism that our understanding of the causes of this tragic disorder will continue to increase.

Keywords
schizophrenia, genetics, GWAS, linkage

Schizophrenia is a psychopathological diagnosis whose current definition includes hallucinations, delusions, disorganized speech, emotional flattening, and bizarre behavior. This definition has a long and evolving history: Kraepelin proposed an influential definition of the disorder as far back as 1896, and revisions for the new fifth edition of the Diagnostic and Statistical Manual of Mental Disorders (DSM-V) are being debated today. Not only is the syndrome usually clinically severe and persistent, it is also relatively common and widespread, with a 1% lifetime risk in the general population worldwide.

Our aim here is to briefly review research on the genetic contributors to schizophrenia. Such research is not only important for schizophrenia but may also offer valuable insights into the complexities that lie in wait in the study of genetic causes of many other common psychopathologies, as well as of psychological individual differences in the normal range. Our focus on genetic contributors to schizophrenia does not imply that their effects are specific to that disorder—rather, numerous findings suggest some shared genetic effects with other psychopathologies.

Total Aggregated Genetic Effects on Schizophrenia

Logically, the first question to be asked concerning the causes of schizophrenia (or any phenotype—that is, measurable characteristic of interest) is the degree to which the total sum of all genetic effects contributes to causing the diagnosis. For example, if monozygotic (MZ) co-twins of schizophrenic patients who have been reared apart (and thus who share 100% of the patient’s genotype but not his or her environment) were not at increased risk for schizophrenia compared to the general population, then studies seeking to identify specific genetic variants for schizophrenia would be an exercise in futility because there would be no genetic effects to be found. The usual methods used to answer this question (i.e., family, twin, and adoption designs) have a long history in schizophrenia research (e.g., Pogue-Geile & Gottesman, 2006), and only final conclusions will be outlined here. Based on the familial risk data outlined in Table 1, aggregated genetic effects have been estimated to account for approximately 83% of the total variation in liability to schizophrenia (Cardno & Gottesman, 2000). Thus genetic effects appear to be the dominant overall cause of schizophrenia. It is also estimated from the pattern of risk across relatives that variation in multiple genes contributes to schizophrenia. In addition, because MZ co-twins of patients do not all have schizophrenia (i.e., MZ concordance rate of 48%) it seems clear that environmental effects that are largely...
not shared among relatives also play an important role (Gottesman & Bertelsen, 1989).

A general etiological hypothesis, the multifactorial threshold (MFT) model, provides an excellent overall fit to the risk data from twin and family studies. First proposed for schizophrenia by Gottesman and Shields in 1967, the MFT model includes independent, additive, equal-sized effects from many genetic variants along with environmental effects that are assumed to add together to form a continuous distribution of risk for schizophrenia, with a categorical threshold beyond which a clinical diagnosis of schizophrenia is produced. For example, if four genetic variants and two environmental experiences each independently and equally increased the risk for schizophrenia and the threshold for diagnosis was a total of three risk factors, then individuals who inherited and/or were exposed to any combination of three, four, five, or six of the six risk factors would be diagnosed with schizophrenia and those with any combination of zero, one, or two risk factors would not be diagnosed. This zero-to-six scale would reflect a continuous distribution of risk or liability to schizophrenia in the population. A model allowing for rare, large-effect genetic variants in some families (i.e., a mixed model) further improves the predictions of observed data on risk among relatives of patients. We would expect such a complex etiological situation for schizophrenia, with a categorical threshold beyond which a clinical diagnosis of schizophrenia is produced. For example, if four genetic variants and two environmental experiences each independently and equally increased the risk for schizophrenia and the threshold for diagnosis was a total of three risk factors, then individuals who inherited and/or were exposed to any combination of three, four, five, or six of the six risk factors would be diagnosed with schizophrenia and those with any combination of zero, one, or two risk factors would not be diagnosed. This zero-to-six scale would reflect a continuous distribution of risk or liability to schizophrenia in the population. A model allowing for rare, large-effect genetic variants in some families (i.e., a mixed model) further improves the predictions of observed data on risk among relatives of patients. We would expect such a complex etiological situation for schizophrenia, probably for most other common psychological phenotypes. And we should not be surprised that common psychological phenotypes whose definitions were initially developed based on clinical considerations and without reference to genetic criteria may turn out to be genetically complex.

Where Are Schizophrenia Liability Genes Located?

Composed of approximately 3.5 billion nucleotide base pairs, the human DNA sequence (i.e., genome) is a dauntingly large haystack in which to be looking for these numerous small needles. Therefore, initial attempts to screen for the general locations (loci) of genetic risk variants can be useful in narrowing areas to particular chromosomal “neighborhoods” that can then be searched more intensively. Initial studies attempting to discover the chromosomal locations of liability genes for schizophrenia employed some sort of linkage study. Utilizing polymorphic genetic markers (i.e., DNA sequences that vary among individuals) whose chromosomal locations are known, linkage studies seek to correlate the resemblance between relatives for genetic markers at a particular chromosomal location inherited identically by descent from a common ancestor with their phenotypic resemblance. That is, are relatives who resemble each other at a particular chromosomal location also similar phenotypically (e.g., both have schizophrenia)? If they are, then it is likely that near to that particular chromosomal location is genetic variation that contributes to causing schizophrenia. Genetic markers used in linkage studies do not actually cause schizophrenia but rather serve as signposts that may indicate the presence nearby of a schizophrenia risk gene. DNA sequences close to each other on a chromosome tend to be transmitted together from parents to offspring, because it is less likely that crossing-over (an exchange of paternal and maternal homologous chromosomal sections) occurred between their locations during meiosis, compared to sequences that are farther apart. Thus, relatives who share the same genetic markers at those chromosomal locations physically near to causal genetic variants should resemble each other phenotypically as well (e.g., both have schizophrenia) and vice versa.

Methodological developments now allow screening of the entire genome (using 300–400 genetic markers equally spaced across all chromosomes) in order to identify multiple locations harboring gene variants with modest effects. Meta-analyses of dozens of large, genome-wide studies have produced generally (although not unanimously) agreed-upon linkages at regions on chromosomes 1, 2, 6, 8, 13, and 22, as well as others (Lewis et al., 2003) that can become foci of more fine-grained efforts. However, even with quite large samples, linkage studies still have relatively little power to detect gene variants with small effects, and the chromosomal regions identified are quite large, often containing hundreds of genes.

Which Specific Genetic Variants Contribute to Schizophrenia Liability?

Identification of specific causal gene variants typically require allelic association studies, in which correlations between
particular DNA-sequence variants that differ among individuals (i.e., alleles) and the schizophrenia phenotype are examined. For example, at a particular polymorphic chromosomal location, is allele “A” or “B” more common among schizophrenia patients compared to controls? Valid phenotype–allele associations may arise either because the allele contributes causally to schizophrenia or because it is in linkage disequilibrium with an allele at a very nearby locus that does. Linkage disequilibrium denotes the situation in which alleles at different but very close loci tend to remain together on the same chromosome even across many generations, due to their lower likelihood of being reshuffled in the crossing over that occurs during meiosis. Because of linkage disequilibrium, allelic association studies have the advantage of being more statistically powerful for detecting small effects than are linkage studies, with the corresponding feature that they are sensitive to effects from loci only a very small distance on either side of the genetic marker. Association studies typically utilize single nucleotide polymorphisms (SNPs) whose chromosomal locations are known as genetic markers. SNPs are DNA sequences at a particular chromosomal location that differ across individuals by only a single nucleotide base pair at a single location. Millions of such SNPs spread throughout the human genome have been identified. Because the less-frequent allele of any particular SNP is by definition relatively common in the population (greater than 1% and ranging up to 50%), such techniques are best suited for detecting causal alleles that are also relatively common in the population.

Because they are only sensitive to effects from loci that are extremely close to the genetic markers, a very large number of SNPs are needed to screen any particular chromosomal region for associations with a phenotype. In order to reduce genotyping expense, until recently allelic association studies have typically employed some sort of candidate strategy to narrow the search to a manageable number of markers. Candidate strategies have been based either on hypotheses drawn from models of pathology or on chromosomal location information (positional candidates) provided by linkage studies. (As mentioned above, linkage studies are most useful for identifying a general chromosomal neighborhood that might harbor a risk locus—for example, regions of Chromosome 6). There have been hundreds of such association studies using SNPs located within candidate genes chosen because they might be relevant to schizophrenia’s hypothesized pathophysiology (e.g., genes coding for aspects of dopamine neurotransmission). However, to date, many such pathology candidate association results have been negative, and positive findings have been small and difficult to replicate, suggesting that our current models of schizophrenia pathology may not be very accurate or that the causes of the pathology lay elsewhere. In contrast, positional candidate strategies that use a large number of SNPs to screen genes in those chromosomal regions suggested by linkage studies have produced some promising associations with schizophrenia. Although by no means unanimous, variants in the following genes initially suggested by linkage studies have all had a number of positive replications (along with some negative):

1. DTNBP1/Dysbindin (Dystrobrevin-Binding Protein 1, on chromosome 6), and Neuregulin 1 (NRG1, on chromosome 8; Williams, Owen, & O’Donovan, 2009). It is important to note that even the most positive studies find only small associations between SNP alleles in these genes and schizophrenia, accounting for 1% to 2% of the liability to schizophrenia.

In contrast to candidate approaches, genome-wide association (GWA) studies have become practical only recently, as about 1 million SNP markers are needed to screen the entire genome for associations (Sullivan, 2009). However, new technologies using microchips that allow relatively inexpensive genotyping have now made such strategies realistic, although the statistical issues surrounding performing hundreds of thousands of tests are notable. These SNP-based GWA studies combine the statistical power and chromosomal precision of allelic association studies with the genome-wide screening of the earlier linkage studies and have successfully revealed several novel gene variants of small to moderate effects for other conditions. After much anticipation, in the past year results from several large GWA studies, including tens of thousands of schizophrenia cases and controls and hundreds of thousands of SNPs, have been reported (The International Schizophrenia Consortium, 2009; Shi et al., 2009; Stefansson et al., 2009). There were some points of agreement among the GWA studies with previous linkage results and with positional candidate association findings. For example, SNPs on Chromosome 6 in genes of the major histocompatibility complex (MHC), which controls immune functions, were consistent across some of the GWA studies and with prior linkage results. However, few of the favorite positional candidate genes discussed above were significant in the GWA studies, and replication across the GWA studies was far from complete, with only a few (and often not the same) SNPs in each study reaching genome-wide significance. Importantly, even the most significant SNPs only accounted for 1% to 2% of the liability for schizophrenia. Such results imply both that only small genetic effects contribute to schizophrenia risk and that most of the genetic liability is still unidentified, despite the fact that researchers have apparently looked everywhere! This latter phenomenon has been termed “missing” or “dark” heritability, analogous to the presumably common but undetectable “dark matter” of astronomy (Manolio et al., 2009), although recent GWA analyses suggest that perhaps, at least for schizophrenia and height, much of this “dark heritability” may be due to the accumulated impact of tens of thousands of very-small-effect polymorphisms along with possibly lower-than-expected correlations between the SNP markers and putative risk variants (the International Schizophrenia Consortium, 2009; Yang et al., 2010). In any case, this initial wave of GWA studies of schizophrenia has implicated some new gene variants of small effect and suggested that many of the genetic effects are quite small and are very numerous and perhaps that much causal genetic variation remains to be identified (see the Schizophrenia Research Forum, www.schizophreniaforum.org, for weekly updates).
Other Genetic Effects

As noted earlier, SNP-based association studies are most sensitive to relatively common risk alleles, which is appropriate if liability largely arises according to the “common disease-common variant” hypothesis, which, like the MFT model, states that schizophrenia results from combinations of common risk alleles, each with a small main effect. In the presence of evolutionarily negative selection effects (i.e., reduced reproductive rates among patients over time), main effects of common alleles should be small, whereas large effects should be rare because they are being selected out. A recent approach has focused on one class of such rare structural variants. Structural variants include copy number variants (CNVs)—that is, variation in the number of copies of certain DNA sequences—and micro-deletions—that is, small missing bits of DNA sequence at particular locations. Recent studies that have searched for such structural variants have produced some intriguing findings that suggest the total number of such variants is increased in individuals with schizophrenia compared to controls. Although each particular variant is still extremely rare among schizophrenia patients (i.e., much less than 1%), altogether they may occur in 1% to 3% of schizophrenia patients compared to almost 0% in controls (St. Clair, 2009). Although not accounting for a large percent of cases, such rare but large genetic effects appear to be part of the overall genetic architecture of schizophrenia.

Future Directions

Although research will certainly continue to follow up leads discussed above, there are also several other approaches that may improve on the modest results to date.

Improving the phenotype

Although the diagnosis of schizophrenia is useful clinically, it has certainly not mapped simply onto genetic effects. This suggests that “improvements” in the phenotype— that is, developing a definition or measurement of the phenotype that better reflects genetic causes—might clarify matters. Two general and potentially related approaches have been taken to this question. One strategy aims to resolve phenotypic variation among patients with schizophrenia and perhaps identify subgroups or dimensions that correlate better with genetic effects. Subtyping of schizophrenia to reduce heterogeneity has a long history, but so far, results from recent candidate allele studies have not been dramatic. However, using GWA studies to correlate phenotypic variation among schizophrenia patients with measured genetic variation may prove more useful.

A related phenotypic strategy aims to “extend” the schizophrenia phenotype by identifying characteristics that are more sensitive to genetic effects than is the overall diagnosis of schizophrenia itself. For example, although only a small number of individuals with a particular genetic risk allele may develop schizophrenia (e.g., 2% compared to 1% among those without it) perhaps almost all of those with the risk allele have attentional problems even if they are not schizophrenic. Measuring attentional problems in this situation would make identifying the schizophrenia risk allele much easier than focusing on the diagnosis alone. There is an important literature attempting to identify such “endophenotypes” (Gottesman & Gould, 2003) using risk for schizophrenia as the criterion for determining the potential usefulness of the endophenotype, with many promising suggestions, ranging from schizotypal (mild schizophrenia-like symptoms) personality traits (Pogue-Geile, 2003), neuropsychological deficits (Snitz, MacDonald, & Carter, 2006), and neural differences revealed by brain imaging (MacDonald, Thermenos, Barch, & Seidman, 2009), that are more common among nonschizophrenic relatives of patients than among controls. It is only recently, however, that such potential endophenotypes have been incorporated within multivariate, genome-wide linkage studies, with some early positive results suggesting novel potential risk loci with pleiotropic (joint) effects on both schizophrenia and cognitive function (e.g., Almasy et al., 2008). There have also been a number of studies investigating phenotypic correlates of putative schizophrenia risk alleles in the general population or among relatives of patients. Although having the potential to elaborate the pathological effects of putative risk alleles, such studies must rely on risk alleles being identified in the first place.

“Proximal” measures of pathophysiology

Although many strategies are being used to improve understanding of schizophrenia pathophysiology that could be used to “improve the phenotype,” as described above, there are some approaches that emphasize aspects of pathology very close to the genotype. Perhaps the most exciting are techniques to measure gene expression—that is, the extent to which genes are “turned on” and are producing their RNA transcripts. These techniques may be applied to hypothesized pathological tissue (e.g., brain tissue studied post-mortem) or even to blood cells in order to identify genes that are under- or over-expressed among patients compared to controls. Because gene expression can be environmentally modulated, such measures also reflect environmental effects. Although most efforts to date have relied on expression levels of candidate genes, genome-wide expression studies of virtually all human genes (over 20,000) can now be performed (Cookson, Liang, Abecasis, Moffatt, & Lathrop, 2009). This emerging technology opens numerous possibilities for identifying candidate genes for allelic association studies, although access to appropriate tissue for the expression study is a theoretical and practical challenge. A related strategy attempts to identify risk-associated variation in the “epigenome” (i.e., epigenetic effects). Environmental exposures may affect histones and DNA methylation patterns that in turn affect gene expression. Histones are chemical structures that surround DNA and whose configuration affects whether genes can be expressed or “turned on”; DNA methylation patterns similarly affect gene expression. Assays of
histone or methylation patterns in appropriate tissue may thus identify abnormalities in the epigenome of schizophrenia.

**Searching for dark heritability**

Although perhaps much of this dark heritability is due to the accumulated effects of tens of thousands of polymorphisms, each with very small effect, along with imprecision of SNP measurement, other approaches have also attempted to explain the small (1%-2% of liability) main and aggregate effects of most putative genetic variants suggested to date. One hypothesis is that genetic main effects are small because genes interact with other genes (called epistasis) and that it is the interactions, not the main or average effects, that are large. For example, perhaps only individuals with risk alleles at two (or more) loci develop schizophrenia but those with only one or the other do not. Although intuitively plausible, investigations to identify important interactions among a large number of alleles, such as are produced in a GWA study, suffer from enormous statistical complexities. In addition, statistical modeling of familial risk does not usually detect such non-additive epistatic effects. Nevertheless these approaches may hold promise if the problems of high dimensionality can be resolved either statistically or through the use of improved candidate hypotheses that are perhaps based on identification of interactions within gene networks.

A similar hypothesis concerns interactions between gene variants and environmental experiences. For example, it may be that a genetic effect is small on average but that among individuals with a particular environmental exposure, such as a viral infection, many more would develop schizophrenia. Again, this is a plausible scenario for which there are some early suggestive leads.

A final approach to illuminating the dark heritability of schizophrenia is *whole-genome sequencing*. Although they are massive, SNP-based GWA and CNV studies still detect only particular kinds of polymorphisms. Only by complete sequencing of individuals’ genomes will all the variation present be measurable. Although currently such whole-genome sequencing is too expensive, it is anticipated that within the relatively near future, costs will approach $1,000 per genome (Metzker, 2010). Of course, the statistical issues of comparing DNA sequences of approximately 3.5 billion nucleotides will be daunting.

**In Conclusion**

Genetic research on schizophrenia has grown exponentially in recent years and has exploited the rush of new molecular technologies and analytic techniques. In many ways it represents a prototype of modern genetic research on a psychological phenotype—a model for application of new techniques and of the resulting challenges. Although it is clear that molecular genetic and analytic techniques will continue to advance, it is equally clear that research on improving the phenotype and identifying environmental contributors will be needed to improve our understanding of the tragic problems associated with the disorder. Similarly, although theoretical genome-wide explorations may lead us to undiscovered causes, innovative hypotheses of pathophysiology have the potential to suggest both candidate phenotypes and genotypes that may lead us more directly to at least some of the needles in this large and important haystack.

**Recommended Reading**


**Declaration of Conflicting Interests**

The authors declared that they had no conflicts of interest with respect to their authorship or the publication of this article.

**Funding**

Preparation of this manuscript was supported in part by National Institutes of Health Grant: MH063480.

**References**


