Published in final edited form as:

Psychiatr Clin North Am. 2010 March ; 33(1): 35–66. doi:10.1016/j.psc.2009.12.003.

The Role of Genetics in the Etiology of Schizophrenia

PV Gejman, AR Sanders, and J Duan

Center for Psychiatric Genetics, NorthShore University HealthSystem Research Institute, Evanston, Illinois 60201, USA

Synopsis

Genome-wide experiments are rapidly changing our understanding of the molecular genetics of schizophrenia. These studies have discovered uncommon copy number variations (mainly deletions) associated with schizophrenia as well as common SNPs with alleles associated with schizophrenia. The aggregate data provide initial support for polygenic inheritance and for genetic overlap of schizophrenia with autism and with bipolar disorder. It is anticipated that as genetic discoveries accumulate, the application of a myriad of tools from systems biology will lead to a delineation of biological pathways involved in the pathophysiology of schizophrenia, and eventually to new therapies.

Keywords

evolution SNP; CNV; complex disorders; polygenic; GWAS

Introduction

The aim of this chapter is to introduce the reader to the genetics of schizophrenia: its background, the status of a variety of genetic findings, new developments (which are many since our last review 1), and current and future challenges. Schizophrenia is a devastating psychiatric disorder with a median lifetime prevalence of 4.0 per 1,000 and a morbid risk of 7.2 per 1,000 2. The age at onset is typically in adolescence or early adulthood 3, with onset after the fifth decade of life and in childhood both being rare 4·5. Although the prevalence for males and females is similar 2, the course of schizophrenia is often more severe and with earlier onset for males 3·6. The standardized mortality ratio (SMR; ratio of observed deaths to expected deaths) for all-cause mortality is 2.6 for patients with schizophrenia compared to the general population 2, with excess deaths mainly from suicide during the early phase of the disorder, and later from cardiovascular complications.

Schizophrenia commonly has a chronic course albeit with fluctuating patterns, and cognitive disability. Its hallmark is psychosis, mainly characterized by positive symptoms such as hallucinations and delusions that are frequently accompanied by negative (deficit) symptoms such as reduced emotions, speech, and interest, and by disorganization symptoms such as

Pablo V. Gejman, M.D. is the corresponding author, the person to receive proof and reprints.

Pablo V. Gejman, M.D., pgejman@gmail.com, Center for Psychiatric Genetics, NorthShore University HealthSystem Research Institute, 1001 University Place, Evanston, IL 60201 USA, Phone: 1-224-364-7550, Fax: 1-224-364-7570

Alan R. Sanders, M.D., alan.sanders.md@gmail.com, Affiliation, address, and fax the same, Phone: 1-224-364-7560

Jubao Duan, Ph.D., jduan69@gmail.com, Affiliation, address, and fax the same, Phone: 1-224-364-7564

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

disrupted syntax and behavior. Severe mood symptoms, up to and including manic and major depressive episodes, are present in many cases. There are no diagnostic laboratory tests for schizophrenia; instead, the diagnosis relies on clinical observation and self-report. It is then remarkable that ongoing epidemiological study over the last century using the clinical phenotype, but with variable ascertainment and assessment rules, has consistently shown the importance of genetic factors in schizophrenia.

The Phenotypic Conundrum

The definition of caseness is fundamental to research design decisions. Bipolar disorder, schizoaffective disorder, and schizophrenia share some phenotypic aspects in common, both in terms of symptoms and also therapeutics, with all responding to antipsychotic drugs. Emil Kraepelin defined dementia praecox as a group of psychotic conditions with a tendency toward poor prognosis⁷. He grouped under the term manic-depressive psychoses a set of conditions that included periodic and circular insanity, simple mania, and melancholia, which he thought did not result in deterioration. Kraepelin believed that dementia praecox and manic-depressive psychoses had specific and separate causes. However, reality proved to be more complex, and in 1933 Jacob Kasanin coined the term schizoaffective psychosis to refer to a disorder with mixed features of schizophrenia and affective disorder⁸. Compared to the general population, family studies show that the clinically intermediate diagnosis of schizoaffective disorder is more common in families ascertained from probands with schizophrenia as well as in families ascertained from probands with bipolar disorder^{9–14}. The diagnostic distinction between schizophrenia or bipolar disorder and schizoaffective disorder is not reliable¹⁵. The specific time criterion for affective symptoms relative to the schizophrenic symptoms is not well defined and varies in different modern classifications^{16,17}.

Complex Genetics

Our knowledge of the molecular mechanisms of schizophrenia pathophysiology remains very incomplete. False starts and research dead ends have taught the field the need for cautiousness; the biological complexity of schizophrenia is much higher than was anticipated. This complexity also applies to simple Mendelian disorders, which although easily analyzed by studying pedigrees, can present unexpectedly intricate biology. Yet, the architecture of schizophrenia is incommensurably more difficult than simple genetic disorders. The idea that one or a few common major gene effects explain schizophrenia was empirically tested in genome-wide linkage scans but results mostly fell short of genome-wide significance¹⁸. That schizophrenia is very complex should not be surprising. First, the brain is more complicated than any other organ; the number of neuronal interconnections and permutations thereof in humans is enormous ($\sim 2 \times 10^{10}$ neocortical neurons and $\sim 10^{14}$ synapses)^{19,20}, and our knowledge of the physiological basis of higher brain functions is very incomplete. Second, the absence of well-defined, focal, and specific microscopic neuropathology has contributed to making schizophrenia particularly impervious to molecular progress, but this is starting to change as we discuss below.

Schizophrenia belongs to a group of pathologies known as complex genetic disorders. Our understanding of complex genetic disorders is still evolving as new experiments uncover novel mechanisms of disease. It is commonly thought that many genes are involved in each disorder with each gene conferring only a small effect on the phenotype. The individual risk variants are thus without diagnostic predictive value, and any estimations of risk are probably going to change in the future as large epidemiological samples become available for analysis²¹. Epistatic interactions between these genes and among their products, and interactions with environmental risk factors are considered highly plausible. However, the study of genetic interactions utilizing genome-wide data remains largely unexplored because of need to correct

for an enormous number of statistical comparisons. Our knowledge is shifting from oligogenic models to a polygenic model of schizophrenia, but its genetic architecture still remains largely unknown. The current evidence strongly suggests that the mutation frequency spectrum comprises a mix of many common and rare mutations. The idea that complex disorders do not result from abnormal function of individual genes but from dysfunction of entire molecular networks, the concept of system disorder, is making strong inroads in the literature (see, for example, 22). Whether this applies to schizophrenia is still an empirical question that remains to be addressed.

It has traditionally been assumed that changes in DNA sequence are solely responsible for the transmission of schizophrenia. However, twin studies show that it is also conceivable that an epigenetic mechanism may contribute to the transmission of schizophrenia. The possibility of a role for epigenetics, i.e., changes in phenotype not explained by DNA sequence, was raised first as an explanation of the incomplete concordance for schizophrenia in monozygotic twins (see, for example 23), but still remains little tested due to methodological difficulties 24.

Evidence For Environmental Factors

The longstanding and influential belief that the incidence of schizophrenia is unaffected by place and time has been recently disproven, opening a remarkably productive period for the study of schizophrenia epidemiology. New epidemiological results show specific circumstances where risk for schizophrenia is increased, including various obstetric complications 25·26, urban birth or residence, famines, migrant status, and seasonal effects (via prenatal infections, e.g., influenza) 2. Other epidemiological evidence strongly suggests that advanced paternal age 27·28 (see more in “Darwinian Paradox”), along with cerebral hypoxia and other severe pregnancy and perinatal complications 29·30 are also environmental risk factors. Overall, the landscape of environmental risks is fertile, and growing rapidly, pointing to a myriad of risk factors acting early during development. Yet, the individual effects of environmental risks, even those that are biologically catastrophic such as famines 31·32, are relatively small. Although the specific pathophysiological connections between environmental risk factors and schizophrenia remain largely tentative, epidemiological findings can potentially provide strong guidance to molecular genetic experiments, e.g., screening specific genes involved in prenatal nutrition and performing serologic assays from epidemiological samples 33. While the prevalence of some other complex disorders are rising, such as obesity 34 and diabetes 35, there is no evidence for such a rapid change for schizophrenia (a registry analysis from Denmark has even detected a possible trend towards decreased incidence 36). Finally, it is likely that additional environmental factors associated with increased risk for schizophrenia still remain to be discovered, and that an understanding of gene-environment interactions will be necessary to unravel the biology of schizophrenia.

The Evidence for Genetic Factors

The modern twin and adoption studies were instrumental in rejecting psychological hypotheses of schizophrenia causation ³⁷ and became the main foundation for the search of molecular genetic risk factors.

Familial clustering is a characteristic of schizophrenia

Ernst Rüdin, who was a disciple of Kraepelin, but later infamously became the main scientific leader of Nazi eugenics 38, conducted the first systematic family study for a psychiatric disorder 39. He realized that the data would not fit a model of simple monogenic Mendelian transmission, but missed the evidence for additional complexity. Many family studies of schizophrenia were conducted since then, with the available evidence showing that the child of a parent with schizophrenia has an elevated empirical risk about tenfold over the general

population risk (for a review, see 40). The risk of a disease in a type of relative compared to that in the general population is often called λ (if the risk is conferred by an allelic variant, it is further specified as an allele specific λ). The relative risk to siblings resulting from having a proband with the illness is called λ_s ⁴¹. Common disorders have a smaller λ_s than rare disorders, even with similar overall genetic effects. For example, the respective λ_s for the autosomal dominant Huntington disease (assuming population prevalence 0.0001), the autosomal recessive cystic fibrosis (assuming population prevalence 0.0004), and autism are 5,000, 625, and 60–100^{42,43}, though the λ_s for major adult psychiatric disorders of the adult typically are under 10 (λ_s is ~10 for schizophrenia). The risk for schizophrenia to a relative of an affected proband decays much more rapidly than the proportion of genes shared between them, which is also inconsistent with a simple Mendelian model⁴⁰. Still most cases of schizophrenia in the general population are sporadic^{44,45}, which may seem surprising at first glance. However, assuming polygenic inheritance (which explains the molecular findings of schizophrenia better than other models, see below), for a disease with a prevalence of 1% and 90% heritability, more sporadic than familial cases are expected⁴⁴.

Twin studies

Differences between monozygotic (identical) twins are attributed to the environment, and differences between dizygotic (fraternal) twins to both hereditary and environmental factors in twin studies. The concordance rate, the probability that a second twin will develop a disorder if the proband (first examined) twin has the disorder, is commonly used. Heritability is the proportion of variance explained by genetic factors. The concordance rates of schizophrenia for monozygotic twins have been found to be about 40 to 50%, and heritability estimates are around 80%^{46,47}. The reader should note that heritability *per se* is not an estimation of the cause of the disease, but rather of the cause of the variation of the disease in a particular population⁴⁸. Studies from Denmark and Finland finding concordances consistent with older studies have employed population registries^{49,50}, which present two major advantages⁴⁶: systematic ascertainment and an estimation of the population risk for the studied trait. Contemporary studies based on hospital registries from Germany, UK, and Japan also yielded similar results^{51,52}. The risk of schizophrenia and schizophrenia-related disorders is similar for the offspring of both the unaffected and the affected monozygotic twins^{53,54}, which suggests that the unaffected twins do carry a heritable genetic risk for schizophrenia without expressing the disease (supporting either or both, epigenetics and non-shared environments). It has recently been proposed that DNA methylation differences might be the cause of monozygotic twin discordance⁵⁵, and also might provide a mechanism for a variety of environmental risk factors for schizophrenia^{33,56}.

Adoption Studies

Such studies allow dissection of genetic from environmental contributions to a disorder in ways that twin studies cannot (see review 57 which also explores methodological strengths and weaknesses of these approaches). The high-risk adoptees approach evaluates adopted away offspring of parents with schizophrenia to see if risk for schizophrenia (or often also schizophrenia spectrum disorders) is elevated. These studies have found an elevated risk for psychosis in such offspring, whether the parents had schizophrenia onset before or after adoption, and whether the rearing environment was foster parents or institutional^{58–64}. Consistent with the risk traveling with the biological rather than the adoptive relationship, it was shown that the risk was similar for offspring of schizophrenia mothers, whether they were raised by the biological (schizophrenic) parent or an adoptive (non-schizophrenic) parent⁵⁹, and that offspring of mothers without schizophrenia did not have an increased risk when raised by psychotic adoptive parents⁶⁰. Furthermore, adoption studies can yield some insight into gene-environment interactions, for example by comparing communication deviance in

adoptive parents of high-risk adoptees 65. The adoptees' family approach starts with schizophrenic adoptees and matched control adoptees, and evaluates their adoptive and biological families for illness. These studies have shown elevated rates of schizophrenia and schizophrenia spectrum disorders in biological families of schizophrenic adoptees compared to biological families of control adoptees, coupled with low and equivalent such rates in adoptive families of both types of adoptees 66^{–70}.

Darwinian Paradox

Schizophrenia has long been known to be associated with decreased fertility 39·71, which is explained by the behavioral and social characteristics of schizophrenia. Fertility is substantially compromised in both genders 72·73, though more markedly in males. Decreased fertility is anticipated to increase because of the delayed marriage patterns in Western societies, while age of onset for schizophrenia has not changed. It is expected that natural selection would decrease the population frequencies of disorder genes that diminish fertility. However, the prevalence of schizophrenia remains high — much higher than for Mendelian disorders. How schizophrenia circumvents the effect of natural selection (sometimes called a “Darwinian paradox”) remains an enigma and multiple hypotheses have been proposed – see review 74. Fananas et al 75 proposed that the relatives of schizophrenics might have a compensatory increase in fertility, but preliminary data did not replicate in larger samples 72·73·76. Lack of evidence for increased fertility in relatives of schizophrenics weakens alternative explanations, such as heterozygote advantage (either homozygote shows reduced fitness compared to the heterozygote such as with sickle cell anemia 77) and antagonist pleiotropy (an allele might reduce fitness for one trait while increasing fitness of a related trait) 74.

Another proposed explanation is that the clinical phenotype might have poor correlation with the underlying genetic susceptibility (i.e., genotype), and it has been suggested that endophenotypic variables (sometimes called intermediate phenotypes) such as structural and functional neuroimaging characteristics constitute a better index of the underlying gene effects than the clinical phenotype 78. There are two problems with this argument: First, a large body of genetic epidemiology is based on the clinical phenotype. Second, none of the proposed endophenotypes has been proven yet to be more heritable than the aggregate clinical phenotype 79.

It is also conceivable that the alleles conferring susceptibility to schizophrenia might be maintained in the population against negative selection by a high mutation rate 80. Advanced paternal age would then be a risk factor as spermatogonia replicate many more times over life than oocytes and the age of fathers is greater than expected for some autosomal dominant diseases due to new mutations 81. A study of an epidemiological sample of 87,907 individuals born in Jerusalem between 1964 and 1976 found that the relative risk of schizophrenia increased continuously with the age of the fathers to a maximum of 2.96 in offspring of fathers aged 50 and 54 years 27. This finding has been replicated in larger samples from different populations, especially for older fathers (see review 28), and found to be a stronger effect in sporadic (family history negative) cases 82, as would be predicted for *de novo* mutations. As reviewed 74, paternal age effects challenge neutral and balancing selection (such as heterozygote advantage and antagonist pleiotropy) explanations of schizophrenia's Darwinian paradox, but are expected under a mutation-selection model 83. In addition, polygenic mutation-selection balance (where deleterious mutations have yet to go extinct: many older mutations of milder individual effect that are removed more slowly from the population, and rarer recent mutations of larger effect that have not had time to diminish over generations) is consistent other important aspects of schizophrenia (e.g., its prevalence, reproductive fitness costs, and its expression via the body's most complex organ with an enormous “mutational target size”) 74, and also with

recent findings detailed below from genome wide association studies (GWAS), especially support for the importance of polygenic inheritance.

First Modern Association Studies

Before the availability of GWAS, most gene association studies consisted of tests of candidate gene involvement. Close to 800 genes have been tested for association (see www.schizophreniaforum.org/res/sczgene; 84). This makes schizophrenia one of the most studied disorders through a candidate gene approach. Unfortunately, none of them as of today can be considered fully established. As samples frequently lacked sufficient statistical power, the problem of non-replication has been far from trivial. In a comprehensive study of some of the most cited candidate genes (e.g., *DISC1*, *DTNBP1*, *NRG1*, *DRD2*, *HTR2A* (5-hydroxytryptamine [serotonin] receptor 2A), and *COMT*), each of 14 genes were tested by genotyping a sample of 1,870 cases and 2,002 screened controls of European ancestry (EA)⁸⁵. A total of 789 single nucleotide polymorphisms (SNPs), including tags for common variation in each gene (tag SNPs are SNPs that are correlated with many other nearby SNPs, for which they are proxies), SNPs previously reported as associated, and SNPs located in functional domains of genes was genotyped, but no association was found (Figure 1), which clearly contradicts ORs predicted from the analysis of smaller samples (the effect size can be conceptualized as the strength of the association between a marker and the disorder, and it can be expressed as the odds ratio, OR, which is the odds for an event, here, possessing a risk allele, in the risk group, i.e., cases, divided by the odds in the non-risk group, i.e., controls). The dilemma for the field is interpreting the reasons for the abundance of positive and negative associations with candidate genes. It is likely that the use of small sample sizes and inadequate or loose statistical thresholds are behind many of the unreplicated observations. Other potential causes of false positives are multiple analyses and selective reporting⁸⁶. It is possible that genetic heterogeneity in some specific cases would preclude a replication, but it would seem unlikely that this would be a robust general argument (for a detailed discussion of heterogeneity, see 87). Multiplicative epistasis (where the individual gene effects might not be detectable but the product of the effects might become detectable) is another largely unexplored possibility that could in principle explain non-replication, and of course environmental variation is another source of heterogeneity. Furthermore, very provocative work by Richter et al.⁸⁸ suggests that increased standardization (such as in experiments designed to decrease heterogeneity, which allows and frequently requires smaller sample sizes) can actually decrease reproducibility in animal behavioral experiments, challenging long held ideas. Recent schizophrenia GWAS results (where each candidate gene is typically more comprehensively tested than in most candidate gene experiments), overall, have not supported most associations to classical candidate genes (Table 1, also see supplementary data file 3 from 89), a pattern consistent with the general results of GWAS in complex disorders (www.genome.gov/gwastudies; 90). Although some candidate genes have been replicated (e.g., *APOE4* in Alzheimer's disease), most discovered associations from GWAS were either in genes that were not previously suspected to be involved in the disease, or in regions of the genome with no obvious genes. Still, the evidence supporting some candidate genes is difficult to ignore (for example, *ERRB4*, the receptor for *NRG1* has shown high significance in an African American (AA) GWAS sample⁸⁹). Additional research and the analysis of cumulative data, with particular attention to both quality control and statistical rigor, will be required for definitive conclusions. The interpretation of data can be treacherous. For example, the reader should be aware that gene pathway analyses do not specifically confirm individual gene associations. While an association with *ERRB4* may suggest potential involvement of *NRG1* as both belong to the same biological pathway, it does not confirm, by itself, participation of *NRG1* in schizophrenia, which would still require a genome-wide association signal between *NRG1* and schizophrenia.

GWAS

Genome-wide studies, in combination with system biology approaches, yield comprehensive information and have been demonstrated to be more useful to deal with complex phenotypes. In direct opposition to candidate gene studies, GWAS interrogate, one at a time, markers of common variation across the human genome, investigating all genes and the majority of the non-genic regions, whether they were previously implicated by pathophysiological hypotheses or not. The large number of tests in a GWAS makes the method highly susceptible to false positive hits; therefore, the estimation of an appropriate genome-wide significance threshold is fundamental. The genome-wide significance threshold, for a value of 5% significance assuming tests for all common SNPs, has been estimated to be around $p < 5 \times 10^{-8}$ ^{91–93}. Due to their more comprehensive coverage of the human genome, GWAS have been more successful than any previous approach to find new susceptibility loci for complex disorders. According to www.genome.gov/gwastudies (as of 09/20/2009), 732 genes were reported associated to one or more complex disease phenotypes at genome-wide significant levels ($p < 5 \times 10^{-8}$)^{90,94}, and many of these associations have already been replicated. GWAS are based on linkage disequilibrium (LD), a non-random statistical association of alleles at two or more loci, which is characteristically associated with short physical distance between genetic markers.

To a large extent, GWAS was made possible by the Human Genome Project (www.ornl.gov/sci/techresources/Human_Genome/home.shtml). Major improvements in SNP genotyping and DNA sequencing were spinoffs from the human genome project, and microarrays made possible rapid and accurate genome-wide genotyping resulting in a map of common genetic variation in a reference set of individuals of European, Asian, and African descent (HapMap project). The majority of the markers used for GWAS are tag SNPs; thus the most significant associated SNP in a GWAS may reflect an indirect association (i.e., be in LD with a causative variant). The Affymetrix 6.0 and Illumina 1M SNP arrays include ~1M common SNPs and probes for analysis of copy number variants (CNVs), with their SNPs assaying ~80% of the common variation in the genome for EA samples⁹⁵. However, the estimated number of common (minor allele frequency, MAF > 1%) SNPs is ~10M, but our genotyping capabilities are not sufficiently developed yet to genotype every SNP in a very large clinical sample (deep re-sequencing technology and new arrays may soon overcome this difficulty). In the meantime, imputation, the computational prediction of genotypes from non-genotyped SNPs, is used to extend GWAS map coverage^{96,97}. By design, the main assumption under GWAS is the common disease/common variant hypothesis (CDCV)^{98,99}.

Recent complex disorders GWAS show two main characteristics: First, common loci with small effects are typically reported (ORs = 1.1–1.5), an empirical confirmation that a large body of epidemiological studies predicting multiple small common genetic effects for complex disorders were correct (including for schizophrenia^{100,101}), since loci with larger effects are rapidly eliminated from the population through selection. Second, most studies have tended to detect new susceptibility loci, and only very large samples obtained from combining studies are powered to show robust replication. This is because the power to detect one out of many possible risk loci is much larger than the power to detect specific disorder alleles²¹. Furthermore, if only small effects are found, many genes would be predicted to underlie the pathophysiology of most complex genetic disorders. On the other hand, it is important to also emphasize some main GWAS limitations. The reader should be aware that the statistical power of GWAS to detect an association with rare alleles (i.e., SNPs or CNVs with MAF < 1%) is very limited, that for the detection of rare variants re-sequencing is more useful than GWAS, and that the study of gene-gene interactions (epistasis), although widely expected to be a significant

source of heritability, is strictly limited by the statistical power of currently existing samples contrasted to the large number of such tests.

GWAS have already yielded genome-wide significant results for schizophrenia, which we now discuss in more detail, though the reader should note that the small individual ORs do not permit prediction of caseness from specific individual susceptibility loci. Seven GWAS for schizophrenia have been published (Table 1). The sizes of the investigated discovery samples have ranged from 322 to >16,161, but even the largest studies did not yield a genome-wide significant result before combined testing of independent samples. This was not unexpected. The collective experience of GWAS for complex disorders shows that a typical susceptibility locus has an OR of 1.1–1.3, which often necessitates extremely large samples for detection. A sample with a total N of 5,334 such as the Molecular Genetics of Schizophrenia (MGS) EA sample (most investigated samples have been smaller), has adequate statistical power only to detect very common risk alleles (30–60% frequency, log additive effects) with genotypic relative risks ~1.3⁸⁹. To reach sufficient statistical power, the combined analysis of independent datasets is useful. Although the diagnostic spectrum of the final combined sample is naturally wider than for the component datasets, combining datasets has been remarkably successful for a variety of complex disorders including schizophrenia^{89·102·103} (see below, meta-analysis). Different samples often were typed with different platforms, but imputation largely overcomes the limitation that many SNPs from different platforms do not overlap. These results suggest that schizophrenia, despite the very high reported heritability, is among the most complex of human genetic disorders. An additional analysis of the International Schizophrenia Consortium (ISC) and MGS samples¹⁰² supported a polygenic model for schizophrenia susceptibility, involving a set of hundreds of genes, each with unquantified but very small individual effects¹⁰⁰ (see below, polygenic section). Finally, rapidly mounting evidence shows that cases have more rare (<1%) and large (>100kb) CNVs than controls.

Meta-analysis of GWAS Data and the Major Histocompatibility Complex (MHC) Locus

The initial attempts to map schizophrenia to the MHC started in the 1970s¹⁰⁴, only a few years after the discovery of the human HLA system¹⁰⁵. Many attempts since then had been made, and some yielded suggestive evidence (see for example, ¹⁰⁶), but definitive evidence of MHC involvement was only recently obtained from a combined analysis of GWAS data. Three GWAS studies published jointly in 2009 (ISC, MGS, and Schizophrenia Genetics Consortium, SGENE), reaching a total EA sample of 8,008 cases and 19,077 controls, performed a meta-analysis of schizophrenia GWAS for the first time^{89·102·103}. The meta-analysis combined the *p*-values for all imputed and genotyped SNPs from the most significant regions of each study. This analysis generated a genome-wide significant association at the MHC region on chromosome 6.

The MHC signal extends over much of the MHC region, from ~26 Mb to ~33 Mb (Figure 2). The strongest evidence (rs13194053, $p=9.54 \times 10^{-9}$) for association from the meta-analysis was observed near a cluster of histone genes and several immune-related genes, including *butyrophilin subfamily 3 member A2* and *A1* (*BTN3A2* and *BTN3A1*) and *protease serine 16* (*PRSS16*), but each individual dataset tends to have a different location for its best findings. The MHC region has a very high gene density and long-range LD blocks¹⁰⁷ – the human genome is structured in many “blocklike” islands of LD generated by a great variation of recombination rates. Blocks from regions of low recombination are long and are interspersed with interblock regions of higher recombination^{108,109}. The location of the causative variation remains indeterminate but it could be in one or more genes or a nongenic region within the MHC. In the MGS sample (and undoubtedly similarly in the other two samples), ~50% of the top 1,000 highest ranking GWAS SNPs were intergenic, located outside the 10 kb region on

either side of a gene, although many of these may represent a genic region signal due to LD. Even at the MHC locus, rs13194053, the SNP with the most significant association from the meta-analysis, is ~29 kb away from its closest gene, *HIST1H2AH* (*histone cluster 1, H2ah*)^{89,102,103}. A functional role for many intergenic regions would not be surprising. Many of these regions contain highly conserved sequences believed to have a regulatory function¹¹⁰. The associated variants, or variants in LD with them, in intergenic regions may then alter expression of upstream or downstream genes. Moreover, most of the human genome is transcribed, with some transcripts serving as regulatory RNAs, but the function of most transcripts still undefined¹¹¹.

The genes in the MHC region have many different biological functions, but genes with an immune function predominate. Histones regulate DNA transcription by chromatin modification through histone methylation or acetylation^{112–114}, and have a role as antimicrobial agents – histones disrupt the bacterial cell membrane and interfere with microbial gene expression¹¹⁵. In human placenta, histones (H2A and H2B) neutralize bacterial endotoxins as part of an infection barrier¹¹⁶. This raises the possibility that genetic variation in histones might underlie a differential placental susceptibility to infections, and that one or more haplotypes spanning histones might increase the susceptibility to schizophrenia through this mechanism. A Danish registry study reported an increased risk of autoimmune disorders (thyrotoxicosis, intestinal malabsorption, acquired hemolytic anemia, chronic active hepatitis, interstitial cystitis, alopecia areata, myositis, polymyalgia rheumatica, and Sjögren's syndrome) for schizophrenics, and a history of any autoimmune disorder (of 29 evaluated) was found associated with a 45% increase in risk for schizophrenia¹¹⁷. The MHC region has been implicated in many genetic disorders with immune-related abnormalities¹¹⁸, including type 1 diabetes (T1D), multiple sclerosis (MS), Crohn's disease (CD), and rheumatoid arthritis (RA), among many others per www.genome.gov/gwastudies⁹⁰. It is noteworthy that rs3800307 (found on the DRB1*03-DQA1*0501-DQB1*0201 haplotype), a SNP in complete LD ($r^2=1$) with rs13194053, which reached genome-wide significant association with schizophrenia in the combined GWAS meta-analysis^{89,102,103}, is associated with T1D¹¹⁹. In addition, rs3131296 at *NOTCH4* is in strong LD ($r^2>0.73$) with the classical HLA allele DRB1*03 and other SNPs that are associated with several autoimmune disorders (T1D, celiac disease, systemic lupus erythematosus, etc.), albeit with opposite alleles¹⁰³. Finally, the MGS GWAS showed some evidence, with $p=3.5\times10^{-5}$ in the EA data set and $p=1.9\times10^{-6}$ in the EA plus AA data set, for association with schizophrenia at the chromosome 1p22.1 *FAM69A-EVI-RPL5* gene cluster⁸⁹ which has been implicated in MS¹²⁰.

Other genes in the same region are involved in chromatin structure (*high mobility group nucleosomal binding domain 4*, *HMGN4*), transcriptional regulation (*activator of basal transcription 1*, *ABT1*; *zinc finger protein 322A*, *ZNF322A*; *zinc finger protein 184*, *ZNF184*), G-protein-coupled receptor signaling (*FKSG83*), and the nuclear pore complex (*nuclear pore membrane protein 121-like 2*, *POM121L2*). The SGENE-plus (their GWAS set) and follow-up samples (i.e., an extended SGENE dataset which added a follow-up EA sample of 4,999 cases and 15,555 controls) analysis reported an independent association (i.e., in weak LD with rs13194053 at the histone gene cluster) at *NOTCH4* (*Notch homolog 4* [*Drosophila*], rs3131296, $p=2\times10^{-8}$), located at 32.28 Mb on chromosome 6, and the combined meta-analysis of SGENE-plus and follow-up samples, along with MGS and ISC samples, gave a $p=2.3\times10^{-10}$ there¹⁰³. See Table 2 for a list of genes mentioned in the text and their functions.

Non-MHC loci that have also achieved genome-wide significance

The combined analysis of SGENE-plus GWAS samples and replication samples uncovered associations with *neurogranin* (*NRGN*) and with *transcription factor 4* (*TCF4*) that subsequently reached genome-wide significance in the combined analysis of SGENE-plus and

replication samples along with ISC and MGS samples (Table 1) ¹⁰³. *NRGN* encodes a postsynaptic protein kinase substrate that binds calmodulin, mediating N-methyl-D-aspartate (NMDA) receptor signaling that is important for learning and memory, and relevant to the proposed glutamate pathophysiology of schizophrenia ^{121,122}. *TCF4* is a neuronal transcriptional factor essential for brain development, specifically neurogenesis ¹²³. Mutations in *TCF4* cause Pitt–Hopkins syndrome, a neurodevelopmental disorder characterized by severe motor and mental retardation, including absent language development, microcephaly, epilepsy, and facial dysmorphisms ^{124–126}. It is also of interest that homozygous and compound-heterozygote deletions and mutations in *CNTNAP2* and *NRXN1* can symptomatically resemble Pitt–Hopkins syndrome along with autistic behavior ¹²⁷, and both *NRXN1* (via the 2p16.3 CNV, Table 3) and *CNTNAP2* (a rarer CNV ¹²⁸) have previously been implicated in schizophrenia (and also autism spectrum disorders and epilepsy, as reviewed in ¹²⁷). Another new schizophrenia susceptibility gene from schizophrenia GWAS is *zinc finger protein 804A* (*ZNF804A*), which was identified by a two-stage study, with a GWAS discovery phase using 479 cases and 2,937 controls, followed with 6,829 cases and 9,897 controls for loci with a discovery $p < 10^{-5}$ ¹²⁹. A combined $p = 1.61 \times 10^{-7}$ was obtained for SNP rs1344706 in the initial report, and the association evidence was supported in later large GWAS of schizophrenia ^{89, 102,103}. Subsequently, rs1344706 in *ZNF804A* was reported to be associated with altered neuronal connectivity in the dorsolateral prefrontal cortex in a functional magnetic resonance imaging study of healthy controls ¹³⁰.

Polygenic Contributions to Schizophrenia

Many genetic variants, each with a very small effect, combined together, make substantial contributions to disorder risk under a polygenic model, first hypothesized for schizophrenia four decades ago ¹⁰⁰. Simulations show that even a disorder with 1,000 risk loci with low mean relative risks ($RR=1.04$), when evaluated in a large scale (10K cases and 10K controls) GWAS would still allow prediction of individual disorder risk with an accuracy >0.75 by using 75 loci explaining ~50% of the risk variance ¹³¹. The first empirical test of the polygenic hypothesis of schizophrenia by the ISC used their GWAS (discovery data set) to define a large set of very small effect common variants as “score” alleles with increasingly liberal association significance thresholds ¹⁰². With the set of score alleles, the ISC generated an aggregate risk score for each individual in independent target GWAS data sets of schizophrenia, using the MGS EA and AA data sets as well as a UK sample ^{89,129}. Aggregate risk scores in cases were found to be significantly higher than in controls in each of the GWAS data sets of schizophrenia, and also in GWAS data sets of bipolar disorder ^{132,133}, but not in control GWAS data sets of non-psychiatric disorders: T1D; type 2 diabetes, T2D; hypertension; CD; RA; and coronary artery disease ^{102,133}. Collectively, ISC concluded that thousands of common polygenic variants with very small individual effects explain about one-third of the total variation in genetic liability to schizophrenia ¹⁰². In an independent bioinformatics-based study ¹³⁴, schizophrenia candidate genes selected from literature mining were found to be enriched in the list of genes with small p -values from independent schizophrenia GWAS data sets. The polygenic model under the assumption of less common ($MAF < 5\%$) causal alleles did not fit well with the observed schizophrenia GWAS data ¹⁰², however, both simulated and empirical data indicate that the spectrum of risk alleles for common disorders includes both common and rare variants ^{98,135}. Furthermore, despite the substantial variation in liability to schizophrenia possibly explained by polygenic variants (~30%), coupled with contributions from common and small-effect variants individually detectable in GWAS (e.g., MHC variants, *TCF4*, etc., above) and rare and large-effect CNVs (see below), the problem of how to explain the substantial missing heritability remains fundamental. Missing heritability here refers to heritability that is unexplained after well-powered GWAS have been conducted. Although it has been argued that the heritability of some behavioral traits and disorders may have been overestimated ¹³⁶, this seems unlikely for schizophrenia given the large body of high quality

evidence that is available, and other reasons seem more plausible (see an excellent review on the topic 137)

Rare CNVs and Schizophrenia

CNVs are stretches of genomic deletions and duplications ranging from 1 kb to several Mb, and thus are likely to have larger phenotypic effects than SNPs. Only rare (<1%) and large (>100kb) CNVs have thus far been implicated in schizophrenia^{138–144}, as reflected by overall CNV burden and individual CNV loci. Supporting evidence for association of specific rare and large CNVs with schizophrenia is emerging at 1q21.1, 2p16.3 (*NRXN1*), 15q11.2, 15q13.2, 16p11.2, and 22q11.21^{138–144} (Table 3). The 3 Mb deletion at 22q11.21 (22qDS) has been known to cause velocardiofacial syndrome (VCFS), and increases the risk for schizophrenia 145–147. An epidemiological study found that more than 30% of 22qDS carriers develop psychosis, about 80% of this manifesting as schizophrenia 147, which represents the largest known individual risk factor for the development of schizophrenia, besides having an identical twin with schizophrenia. The 22q11.21 CNV was the only CNV reaching genome-wide significance in some schizophrenia GWAS 142 and the only one not found in controls (case %/control% was 0.30/0.00). Each of these other CNVs was found over-represented in schizophrenia cases in at least one study (Table 3), with the supporting evidence remaining consistent for CNVs at the 1q21.1 deletion (0.24/0.02), 2p16.3 (*NRXN1*) deletion (0.10/0.02), 15q13.2 deletion (0.20/0.02), and 16p11.2 duplication (0.35/0.03), and plausible for the 15q11.2 deletion (0.65/0.22). All the CNVs (except for 22q11.21) initially found only in schizophrenia cases were also found in healthy controls in later studies, suggesting that the penetrance of these rare CNVs may be relatively low (Table 3). The deletions at *NRXN1*, which only involve this single gene, suggest that exon disruptions rather than deletions of other parts of *NRXN1* are associated with schizophrenia 148. With the rest of the rare and large CNVs implicated in schizophrenia spanning multiple genes, specific gene effects, including possibly genes presenting pleiotropy (see next section), will be difficult to disentangle.

Pleiotropy and Overlap with Bipolar Disorder and Autism

Pleiotropy refers to the common phenomenon of variation in a gene simultaneously affecting different phenotypes. While examples abound in model organisms (e.g., flies 149), evidence for pleiotropy in humans is also available, such as genes for body weight and height 150, and also for disorders such as T2D 151 and prostate cancer 152. The molecular genetic overlaps between schizophrenia and bipolar disorder, and between schizophrenia and autism are consistent with pleiotropy; but shared genetic loci may actually determine an aspect (somewhat in isolation from the overall phenotype) shared by two disorders such as psychosis in schizophrenia and in bipolar disorder 153.

Although schizophrenia and bipolar disorder are classified as different psychiatric disorders in most contemporary classifications such as the Diagnostic and Statistical Manual of Mental Disorders Fourth Edition, DSM-IV 154, a distinction historically referred to as the Kraepelinian dichotomy⁷, they share some similarities, such as peak onset in early adulthood, prevalence around 1% generally similar worldwide, psychotic symptoms in more than half of bipolar disorder type 1 subjects and the same treatment of psychosis (in both cases with antipsychotic medications), substance use comorbidity, increased suicide risk (and common severe mood disorder in schizophrenia), and sometimes a difficult differential diagnosis with schizoaffective disorder. Overlap of susceptibility genes has been postulated^{155–157} for more circumscribed aspects such as psychosis proneness 153, 158 especially for “positive symptoms” such as hallucinations and delusions, or might reflect a wider range of susceptibility to higher brain dysfunction such as pleiotropy found with CNVs (an overlapping CNV for schizophrenia and bipolar disorder has been reported for the 16p11.2 duplication in a meta-analysis with an

association $p=4.8 \times 10^{-7}$ for schizophrenia and $p=0.017$ for bipolar disorder 159). Family studies have shown that schizoaffective disorder is more common in families ascertained from probands with schizophrenia or with bipolar disorder type 1 than in the general population 9–14. A recent meta-analysis of family studies found familial coaggregation of schizophrenia and bipolar disorder as well 160. The largest family study reported to date, comprised of almost 36,000 schizophrenia and over 40,000 bipolar disorder Swedish probands, concluded that familial coaggregation between schizophrenia and bipolar disorder was ~63% due to additive genetic effects common to both disorders 161. Recent GWAS SNP data strongly support a genetic overlap between schizophrenia and bipolar disorder, which were shown to share polygenic common variants with very small effect sizes (see polygenic section above) 102. A gene-wide analysis found a significant excess of genes showing associations with both disorders 162. *ZNF804A* and *CACNA1C* (*calcium channel, voltage-dependent, L type, alpha 1C subunit*) are two such genes that are shared by both disorders: *ZNF804A* was initially identified as a schizophrenia susceptibility gene, and *CACNA1C* was initially identified as a bipolar disorder susceptibility gene, in respective GWAS 129,163.

Interestingly, the strongest SNP association of $p=4.6 \times 10^{-7}$ in the MGS GWAS EA data set was found with *centaurin gamma 2* (*CENTG2*, also known as *AGAP1*), a gene that has been implicated in autism 164. It is also noteworthy that our exploratory analysis of MGS GWAS data combining both EA and AA ancestries (3,967 cases and 3,626 controls) showed a $p=1.9 \times 10^{-6}$ association with *fragile X mental retardation, autosomal homolog 1* (*FXR1*) 89, and the association reached genome-wide significance when the ISC and SGENE-plus GWAS datasets were included in the analysis along with MGS EA+AA (data not shown). *FXR1* is a paralog of *fragile X mental retardation 1* (*FMR1*), dysfunction of which causes the FMR syndrome that includes autism as a common feature 165. Schizophrenia and autism also share some clinical features such as social interaction and communication impairments, and some negative/deficit symptoms 166. This is more noticeable for childhood-onset schizophrenia where developmental delays may be more marked than in adult-onset schizophrenia with subtler such findings 167. However, autism remains an exclusion criterion for schizophrenia in the DSM-IV unless prominent hallucinations and delusions are present for at least a month 154. A study of 129 adults with autism spectrum disorders, ASD, found that 7% had psychotic bipolar disorder and 8% had schizophrenia or other psychotic disorders 168. The current diagnostic hierarchy, which largely treats ASD and schizophrenia as mutually exclusive, could mask some ASD/schizophrenia comorbidity since an autism case might be less likely to be diagnosed with schizophrenia in adulthood 167, even in the presence of overt and chronic psychosis. A two-fold increase of schizophrenia in the parents of autism cases 169, a risk ratio of 3.44 for autism when prenatal parental history of a schizophrenia-like psychosis was present in a nationwide Danish study 170, and an increased risk of schizophrenia in autism patients 168,171 suggest overlapping risk factors between schizophrenia and autism. Genetic overlap of schizophrenia and autism (and other conditions) has recently received indirect support since a number of schizophrenia cases carrying rare and large CNVs have comorbidities such as learning disabilities, mental retardation, autism and autism spectrum disorders, and seizures or epilepsy, and/or such disorders' own CNV scans have implicated the same CNV loci (Table 3). Besides the aforementioned conditions, others are over-represented in schizophrenia cases in general, such as seizures 172,173.

Challenges for the Field of Schizophrenia Genetics

After over a quarter century of molecular genetics work in schizophrenia, advances in biotechnology and statistics applied to the study of large and well characterized clinical samples have made possible the discovery of individual susceptibility loci with subsequent replication. A comprehensive discussion of what comes next after a successful GWAS is outside the scope of this manuscript. We have selected for discussion a handful of issues that have been

instrumental to generate progress until now, and are a foundation for further progress. First, the social environment where science is conducted has deeply changed during the last years. Of fundamental importance is the accentuated stance of new openness in the field of schizophrenia genetics in consonance with the instituted NIH policy of wide GWAS data sharing. De-identified phenotypic and genotypic data from GWAS studies funded by NIH are to be submitted to a centralized NIH GWAS data repository, the database of Genotypes and Phenotypes (dbGaP, www.ncbi.nlm.nih.gov/gap) hosted by the National Center for Biotechnology Information (NCBI), and studies supported the Wellcome Trust Case-Control Consortium are also deposited at a database (WTCCC, www.wtccc.org.uk). Data, in both cases, are accessible by application to access committees. The new NIH policy (grants.nih.gov/grants/gwas) has already created extraordinary opportunities to access data from independent research groups before publication¹⁷⁴. For example, the MGS Genetic Association Information Network (GAIN) genotypic/phenotypic sample (www.genome.gov/19518582) has been accessed 140 times for a large diversity of genetics research projects as of 11/17/09. The Psychiatric GWAS Consortium (PGC) 175·176 continues this new tradition of openness. The PGC is comprised of five groups: schizophrenia, bipolar disorder, major depressive disorder, attention deficit hyperactivity disorder (ADHD), and autism. A primary goal of the PGC is to perform disease specific and cross-disorder analyses from combined GWAS datasets composed of all qualifying samples for each of the disorders.

The method for following up GWAS results needs to be thorough; replication and fine mapping of associated regions are necessary for further progress (see informative review 177). The preferred approach is to combine GWAS data from independent studies, but when some of the samples do not have GWAS data, focused genotyping is still useful, although less informative. The analysis of combined data is important because most clinical samples do not carry the power to detect effect sizes typically uncovered in well-powered GWAS 178, and the estimated ORs tend to be inflated because only top-ranking associations are reported 179; a less biased estimation of ORs requires the systematic combination of GWAS and focused replication studies. Data can be meta-analyzed with a variety of methods (see comprehensive review 177). For example, three consortia 89·102·103 meta-analyzed a set of their most significant p -values ($p < 0.001$) from their independent GWAS uncovering a genome-wide significant locus at the MHC. SNPs other than genome-wide significant ($p < 5 \times 10^{-8}$) ones merit inclusion in confirmation experiments: while some genome-wide significant SNPs from a single study might not be confirmed in replication studies, other SNPs very highly ranking in the primary study, though not achieving genome-wide significance (e.g., SNPs with $p < 1 \times 10^{-5}$), might surpass that threshold in a replication experiment. Association signals in an extended LD block that spans many genes (e.g., the MHC locus implicated in schizophrenia) make it hard to disentangle which gene/s is/are likely to be causal. Populations of non-European ancestry might have some non-overlapping susceptibility loci and it is fundamental to investigate these differences, as they can also be informative about different environmental risk factors. An important characteristic of African populations (e.g., AA) is reduced LD, which would facilitate the narrowing of the associated genomic intervals; existing limitations of CNV and SNP map coverage, and imputation of AA datasets are currently being addressed (for examples, see 180·181).

It is important to bear in mind that given the GWAS SNP design (where SNPs are selected because they are common and are informative of many other SNPs, not because of their functional properties), in most cases the associated SNPs are probably not the causal SNPs. As previously mentioned, we have noted that in the MGS GWAS and in the combined sample (MGS, ISC, and SGENE-plus 89·102·103), the vast majority of the strongest common SNP associations were not located in coding sequences where such a signal would be easier to interpret, but are in intergenic regions (over half of these SNPs >10 kb from a gene, almost all with no clear association to any known gene, i.e., via LD) or of unclear function, e.g., intronic,

but not near a splice site, or known or putative regulatory site. Although the causal SNPs should be in LD with the GWAS associated SNPs, the causative genes may be close to the statistically associated locus, but may also be farther removed, even on a different chromosome. For example if the causal variant was a *trans*-acting factor that regulates transcription, the regulated gene/s might be located on a different chromosome. The integration of genome-wide transcription data (expression quantitative trait loci, eQTL, currently detected by microarray expression data) and GWAS data (DNA variation data) can help close this gap by linking the GWAS statistical results and biology and is expected to lead to discoveries of mechanisms of disease susceptibility otherwise obscured to either method in isolation. The approach has been proven successful in asthma¹⁸². Interestingly, within the MHC region implicated in schizophrenia, there are more than 10 *cis*-eQTL (*cis* meaning nearby on the same chromosome) in the eQTL database, which uses expression data from lymphoblastoid cell lines (LCLs) of asthma patients (Figure 2)¹⁸³, and the SNP showing the most significant association with schizophrenia, rs13194053 with $p=9.54 \times 10^{-9}$, is in strong LD ($r^2=0.43$) with a SNP showing the strongest association with *BTN3A2* expression (Figure 2).

The selection of tissue for gene expression study is critical, and brain is not always necessarily the best choice of tissue. Epidemiological evidence strongly suggests that some of the primary genetic mechanisms leading to schizophrenia might reside in other tissues than brain – for example, an autoimmune mechanism that would compromise the brain – in such a case, the symptoms of schizophrenia would still reflect brain dysfunction, but would be removed from the primary noxae (investigations of these leads remain to be thoroughly explored). A more explicit example would be if a genetic abnormality affecting an immunological response to a virus contributed to schizophrenia risk, studying the brain transcription characteristics of a neurotransmitter system would only reflect secondary (or even terminal) neural changes to the primary immune response (which might be more apparent in immune tissue such as lymphocytes).

Establishing causal mechanisms may require, in addition to statistical testing of association, the functional characterization of implicated genes and variants in simple cell models (and in model organisms) targeting phenotypes with a high probability of association with the studied disorder – among other potential advantages, in the absence of buffering effects present in multicellular systems, *in vitro* effects are expected to be amplified (which may make detectable an effect that is very small in the whole/intact organism). Dendrou et al.¹⁸⁴ studied cell-specific protein phenotypes for *IL2RA*, a locus associated with T1D. By taking advantage of a large collection of normal donors from whom fresh, primary cells could be analyzed (the experimental subjects can be recalled for repeated measurements; this resource is known as the Cambridge BioResource) it was very elegantly demonstrated that elevated CD25 expression is associated with *IL2RA* haplotypes that protect from T1D¹⁸⁴.

It is still premature to conclude whether the genetic architecture of schizophrenia is like mental retardation where thousands of individual genetic disorders have been cataloged, or whether some widely speculated upon, but still little investigated mechanism such as epigenetics (which influences phenotypes through the regulation of gene expression) or gene-environment interactions will explain the bulk of the missing heritability for the disorder. Basic genomics research has produced major breakthroughs during recent years such as discovery of microRNAs, long-range promoters, epigenetic factors, and variable copy number variations, and many more will probably be made as our knowledge of the genome is rapidly increasing. It should not be surprising if still unknown genetic mechanisms will, at the end, explain a substantial proportion of the heritability of schizophrenia. Nonetheless, the task of defining the spectrum of molecular genetic mechanisms in schizophrenia is now at the forefront of our field. Some immediate research efforts will, in large measure, focus on whole genome re-sequencing and genome-wide gene transcription and epigenetic analyses. Rapid progress in biotechnology

¹⁸⁵ is making the study of rare variants in many genes or large genomic regions in larger samples increasingly feasible – proof of principle is provided by the 1,000 genomes project (www.1000genomes.org), which is designed to build a deep catalog of human genetic variation. The design of experiments aimed at fine mapping of regions of association and the precision of imputation will both benefit from this project.

It is anticipated that as genetic discoveries accumulate, the application of a myriad of tools from systems biology (e.g., genomics, transcriptomics, proteomics, etc.) will lead to a delineation of biological pathways involved in the pathophysiology of schizophrenia, and eventually to new therapies (developments in treatment still lag compared to discoveries of new genetic associations for complex disorders, see ¹⁸⁶, but this situation is expected to change as biological research makes inroads into still purely statistical associations). There is, however, a strong temptation to accept the simplest observations (i.e., those with immediate biological connotation) as the most meaningful and the only ones that merit follow-up. For example, Mitchell and Porteous stated: “Occam’s razor and statistical probability both argue that the co-inheritance of one or just a few risk genes by any individual case is the more likely explanation for the majority of incidence” ¹⁸⁷. They continued: “Haven’t we learnt more about disease mechanism and potential routes to the treatment of Alzheimer’s disease from the rare variant examples of amyloid beta (A4) precursor protein (APP), presenilin-1 (PS1) and presenilin-2 (PS2) than from the archetypal common variant example of apolipoprotein E, isoform 4 (ApoE4)?” These arguments appear necessarily true at first sight, however, as previously discussed ¹⁸⁸, an explanation may superficially appear more complicated than need to be, but only if considered apart from its evolutionary context ¹⁸⁹. Research in model organisms (e.g., *Drosophila*) shows that most phenotypes are the result of complicated genetic architectures: multiple genes, often showing pleiotropy (thus likely associated with multiple traits) and epistasis, and even single mutation effects differing with genetic background and environment ^{190,191}, and this landscape will probably be true for complex human behavioral traits as well. Explanations relying on single genes are unlikely to capture the fundamental complexity of most human complex traits, and all the associated genetic variation needs to be pursued to understand the pathophysiology of a complex disorder. A task of utmost importance is the integration of the spectrum of mutations found in schizophrenia into a system that takes into account constantly changing environments and evolution.

References

1. Gejman, PV.; Owen, MJ.; Sanders, AR. Psychiatric Genetics. In: Tasman, A.; Kay, J.; Lieberman, J., editors. Psychiatry. Hoboken: Wiley; 2003.
2. McGrath J, Saha S, Chant D, et al. Schizophrenia: a concise overview of incidence, prevalence, and mortality. *Epidemiol Rev* 2008;30:67. [PubMed: 18480098]
3. Messias EL, Chen CY, Eaton WW. Epidemiology of schizophrenia: review of findings and myths. *Psychiatr Clin North Am* 2007;30:323. [PubMed: 17720026]
4. Girard C, Simard M. Clinical characterization of late- and very late-onset first psychotic episode in psychiatric inpatients. *Am J Geriatr Psychiatry* 2008;16:478. [PubMed: 18515692]
5. Remschmidt H, Theisen FM. Schizophrenia and related disorders in children and adolescents. *J Neural Transm Suppl* 2005;121
6. Leung A, Chue P. Sex differences in schizophrenia, a review of the literature. *Acta Psychiatr Scand Suppl* 2000;401:3. [PubMed: 10887978]
7. Kraepelin, E. Manic-Depressive Insanity and Paranoia. Edinburgh: E. & S. Livingstone; 1899. (English translation 1921)
8. Kasanin J. The acute schizoaffective psychoses. *Am J Psychiatry* 1933;90:97.
9. Maier W, Lichtermann D, Minges J, et al. Continuity and discontinuity of affective disorders and schizophrenia. Results of a controlled family study. *Arch Gen Psychiatry* 1993;50:871. [PubMed: 8215813]

10. Maier W, Lichtermann D, Franke P, et al. The dichotomy of schizophrenia and affective disorders in extended pedigrees. *Schizophr Res* 2002;57:259. [PubMed: 12223257]
11. Gershon ES, DeLisi LE, Hamovit J, et al. A controlled family study of chronic psychoses. Schizophrenia and schizoaffective disorder. *Arch Gen Psychiatry* 1988;45:328. [PubMed: 3355320]
12. Gershon ES, Hamovit J, Guroff JJ, et al. A family study of schizoaffective, bipolar I, bipolar II, unipolar, and normal control probands. *Arch Gen Psychiatry* 1982;39:1157. [PubMed: 7125846]
13. Kendler KS, McGuire M, Gruenberg AM, et al. The Roscommon Family Study. I. Methods, diagnosis of probands, and risk of schizophrenia in relatives. *Arch Gen Psychiatry* 1993;50:527. [PubMed: 8317947]
14. Valles V, Van Os J, Guillamat R, et al. Increased morbid risk for schizophrenia in families of in-patients with bipolar illness. *Schizophr Res* 2000;42:83. [PubMed: 10742646]
15. Heckers S. Is schizoaffective disorder a useful diagnosis? *Curr Psychiatry Rep* 2009;11:332. [PubMed: 19635243]
16. Spitzer RL, Endicott J, Robins E. Research diagnostic criteria: rationale and reliability. *Arch Gen Psychiatry* 1978;35:773. [PubMed: 655775]
17. Nurnberger JI Jr, Blehar MC, Kaufmann CA, et al. Diagnostic interview for genetic studies. Rationale, unique features, and training. NIMH Genetics Initiative. *Arch Gen Psychiatry* 1994;51:849. [PubMed: 7944874]
18. Ng MY, Levinson DF, Faraone SV, et al. Meta-analysis of 32 genome-wide linkage studies of schizophrenia. *Mol Psychiatry* 2009;14:774. [PubMed: 19349958]
19. Sporns O, Tononi G, Kotter R. The human connectome: A structural description of the human brain. *PLoS Comput Biol* 2005;1:e42. [PubMed: 16201007]
20. Drachman DA. Do we have brain to spare? *Neurology* 2005;64:2004. [PubMed: 15985565]
21. Kraft P, Hunter DJ. Genetic risk prediction--are we there yet? *N Engl J Med* 2009;360:1701. [PubMed: 19369656]
22. Schadt EE. Molecular networks as sensors and drivers of common human diseases. *Nature* 2009;461:218. [PubMed: 19741703]
23. Petronis A. The origin of schizophrenia: genetic thesis, epigenetic antithesis, and resolving synthesis. *Biol Psychiatry* 2004;55:965. [PubMed: 15121478]
24. Roth TL, Lubin FD, Sodhi M, et al. Epigenetic mechanisms in schizophrenia. *Biochim Biophys Acta* 2009;1790:869. [PubMed: 19559755]
25. Byrne M, Agerbo E, Bennedsen B, et al. Obstetric conditions and risk of first admission with schizophrenia: a Danish national register based study. *Schizophr Res* 2007;97:51. [PubMed: 17764905]
26. Mittal VA, Ellman LM, Cannon TD. Gene-environment interaction and covariation in schizophrenia: the role of obstetric complications. *Schizophr Bull* 2008;34:1083. [PubMed: 18635675]
27. Malaspina D, Harlap S, Fennig S, et al. Advancing paternal age and the risk of schizophrenia. *Arch Gen Psychiatry* 2001;58:361. [PubMed: 11296097]
28. Torrey EF, Buka S, Cannon TD, et al. Paternal age as a risk factor for schizophrenia: how important is it? *Schizophr Res* 2009;114:1. [PubMed: 19683417]
29. Clarke MC, Harley M, Cannon M. The role of obstetric events in schizophrenia. *Schizophr Bull* 2006;32:3. [PubMed: 16306181]
30. Cannon M, Jones PB, Murray RM. Obstetric complications and schizophrenia: historical and meta-analytic review. *Am J Psychiatry* 2002;159:1080. [PubMed: 12091183]
31. Susser ES, Lin SP. Schizophrenia after prenatal exposure to the Dutch Hunger Winter of 1944–1945. *Arch Gen Psychiatry* 1992;49:983. [PubMed: 1449385]
32. St Clair D, Xu M, Wang P, et al. Rates of adult schizophrenia following prenatal exposure to the Chinese famine of 1959–1961. *Jama* 2005;294:557. [PubMed: 16077049]
33. Brown AS, Susser ES. Prenatal nutritional deficiency and risk of adult schizophrenia. *Schizophr Bull* 2008;34:1054. [PubMed: 18682377]
34. James WP. The epidemiology of obesity: the size of the problem. *J Intern Med* 2008;263:336. [PubMed: 18312311]

35. Wild S, Roglic G, Green A, et al. Global prevalence of diabetes: estimates for the year 2000 and projections for 2030. *Diabetes Care* 2004;27:1047. [PubMed: 15111519]
36. Munk-Jorgensen P, Mortensen PB. Is schizophrenia really on the decrease? *Eur Arch Psychiatry Clin Neurosci* 1993;242:244. [PubMed: 8461351]
37. Neill J. Whatever became of the schizophrenogenic mother? *Am J Psychother* 1990;44:499. [PubMed: 2285075]
38. Gejman PV. Ernst Rudin and Nazi euthanasia: another stain on his career. *Am J Med Genet* 1997;74:455. [PubMed: 9259387]
39. Rüdin, E. *Zur Vererbung und Neuenstehung der Dementia Praecox*. Berlin: Springer; 1916.
40. Gottesman, II.; Shields, J. *Schizophrenia: The Epigenetic Puzzle*. Cambridge, England: Cambridge University Press; 1982.
41. Risch N. Linkage strategies for genetically complex traits. I. Multilocus models. *Am J Hum Genet* 1990;46:222. [PubMed: 2301392]
42. Birren, B.; Green, ED.; Klapholz, S., et al. *Genome Analysis: A Laboratory Manual: Mapping Genome*. Vol. 4. Cold Spring Harbor Laboratory Press; 1998.
43. Smalley SL. Genetic influences in childhood-onset psychiatric disorders: autism and attention-deficit/hyperactivity disorder. *Am J Hum Genet* 1997;60:1276. [PubMed: 9199546]
44. Yang J, Visscher PM, Wray NR. Sporadic cases are the norm for complex disease. *Eur J Hum Genet*. 2009
45. Kendler KS. Sporadic vs familial classification given etiologic heterogeneity: I. Sensitivity, specificity, and positive and negative predictive value. *Genet Epidemiol* 1987;4:313. [PubMed: 3692132]
46. Cardno AG, Gottesman I. Twin studies of schizophrenia: From bow-and-arrow concordances to Star Wars Mx and functional genomics. *Am J Med Genet* 2000;97:12. [PubMed: 10813800]
47. Sullivan PF, Kendler KS, Neale MC. Schizophrenia as a complex trait: evidence from a meta-analysis of twin studies. *Arch Gen Psychiatry* 2003;60:1187. [PubMed: 14662550]
48. Rees J. Complex disease and the new clinical sciences. *Science* 2002;296:698. [PubMed: 11976444]
49. Klaning U, Mortensen PB, Kyvik KO. Increased occurrence of schizophrenia and other psychiatric illnesses among twins. *Br J Psychiatry* 1996;168:688. [PubMed: 8773810]
50. Cannon TD, Kaprio J, Lonnqvist J, et al. The genetic epidemiology of schizophrenia in a Finnish twin cohort. A population-based modeling study. *Arch Gen Psychiatry* 1998;55:67. [PubMed: 9435762]
51. Franzek E, Beckmann H. Different genetic background of schizophrenia spectrum psychoses: a twin study. *Am J Psychiatry* 1998;155:76. [PubMed: 9433342]
52. Cardno AG, Marshall EJ, Coid B, et al. Heritability estimates for psychotic disorders: the Maudsley twin psychosis series. *Arch Gen Psychiatry* 1999;56:162. [PubMed: 10025441]
53. Gottesman II, Bertelsen A. Confirming unexpressed genotypes for schizophrenia. Risks in the offspring of Fischer's Danish identical and fraternal discordant twins. *Arch Gen Psychiatry* 1989;46:867. [PubMed: 2802925]
54. Kringlen E, Cramer G. Offspring of monozygotic twins discordant for schizophrenia. *Arch Gen Psychiatry* 1989;46:873. [PubMed: 2802926]
55. Mill J, Tang T, Kaminsky Z, et al. Epigenomic profiling reveals DNA-methylation changes associated with major psychosis. *Am J Hum Genet* 2008;82:696. [PubMed: 18319075]
56. Ellman LM, Susser ES. The promise of epidemiologic studies: neuroimmune mechanisms in the etiologies of brain disorders. *Neuron* 2009;64:25. [PubMed: 19840544]
57. Ingraham LJ, Kety SS. Adoption studies of schizophrenia. *Am J Med Genet* 2000;97:18. [PubMed: 10813801]
58. Heston LL. Psychiatric disorders in foster home reared children of schizophrenic mothers. *Br J Psychiatry* 1966;112:819. [PubMed: 5966555]
59. Higgins J. Effects of child rearing by schizophrenic mothers: a follow-up. *J Psychiatr Res* 1976;13:1. [PubMed: 966188]
60. Wender PH, Rosenthal D, Kety SS, et al. Crossfostering. A research strategy for clarifying the role of genetic and experiential factors in the etiology of schizophrenia. *Arch Gen Psychiatry* 1974;30:121. [PubMed: 4808731]

61. Rosenthal D, Wender PH, Kety SS, et al. Schizophrenics' offspring raised in adoptive homes. *J Psychiatr Res* 1968;6:377.
62. Tienari P. Interaction between genetic vulnerability and family environment: the Finnish adoptive family study of schizophrenia. *Acta Psychiatr Scand* 1991;84:460. [PubMed: 1776499]
63. Tienari P, Sorri A, Lahti I, et al. The Finnish adoptive family study of schizophrenia. *Yale J Biol Med* 1985;58:227. [PubMed: 4049906]
64. Tienari P, Wynne LC, Moring J, et al. The Finnish adoptive family study of schizophrenia. Implications for family research. *Br J Psychiatry* 1994;20(Suppl)
65. Wahlberg KE, Wynne LC, Oja H, et al. Gene-environment interaction in vulnerability to schizophrenia: findings from the Finnish Adoptive Family Study of Schizophrenia. *Am J Psychiatry* 1997;154:355. [PubMed: 9054783]
66. Kendler KS, Gruenberg AM, Kinney DK. Independent diagnoses of adoptees and relatives as defined by DSM-III in the provincial and national samples of the Danish Adoption Study of Schizophrenia. *Arch Gen Psychiatry* 1994;51:456. [PubMed: 8192548]
67. Kendler KS, Gruenberg AM, Strauss JS. An independent analysis of the Copenhagen sample of the Danish adoption study of schizophrenia. II. The relationship between schizotypal personality disorder and schizophrenia. *Arch Gen Psychiatry* 1981;38:982. [PubMed: 7283669]
68. Kety SS, Wender PH, Jacobsen B, et al. Mental illness in the biological and adoptive relatives of schizophrenic adoptees. Replication of the Copenhagen Study in the rest of Denmark. *Arch Gen Psychiatry* 1994;51:442. [PubMed: 8192547]
69. Kety, SS.; Rosenthal, D.; Wender, PH., et al. The types and prevalence of mental illness in the biological and adoptive families of adopted schizophrenics. In: Rosenthal, D.; Kety, SS., editors. *The Transmission of Schizophrenia*. Oxford, England: Pergamon Press; 1968. p. 345
70. Kety, SS.; Rosenthal, D.; Wender, PH., et al. Mental illness in the biological and adoptive families of adopted individuals who have become schizophrenic: a preliminary report based upon psychiatric interviews. In: Fieve, R.; Rosenthal, D.; Brill, H., editors. *Genetic Research in Psychiatry*. Baltimore: The Johns Hopkins University Press; 1975. p. 147
71. Kallmann, FJ. *The genetics of schizophrenia: A study of heredity and reproduction in the families of 1,087 schizophrenics*. New York, NY: J. J. Augustin Publisher; 1938.
72. Svensson AC, Lichtenstein P, Sandin S, et al. Fertility of first-degree relatives of patients with schizophrenia: a three generation perspective. *Schizophr Res* 2007;91:238. [PubMed: 17275261]
73. Haukka J, Suvisaari J, Lonnqvist J. Fertility of patients with schizophrenia, their siblings, and the general population: a cohort study from 1950 to 1959 in Finland. *Am J Psychiatry* 2003;160:460. [PubMed: 12611825]
74. Keller MC, Miller G. Resolving the paradox of common, harmful, heritable mental disorders: which evolutionary genetic models work best? *Behav Brain Sci* 2006;29:385. [PubMed: 17094843]
75. Fananas L, Bertranpetit J. Reproductive rates in families of schizophrenic patients in a case-control study. *Acta Psychiatr Scand* 1995;91:202. [PubMed: 7625196]
76. MacCabe JH, Koupil I, Leon DA. Lifetime reproductive output over two generations in patients with psychosis and their unaffected siblings: the Uppsala 1915–1929 Birth Cohort Multigenerational Study. *Psychol Med* 2009;39:1667. [PubMed: 19265569]
77. Allison AC. Notes on sickle-cell polymorphism. *Ann Hum Genet* 1954;19:39. [PubMed: 13208024]
78. Gottesman II, Gould TD. The endophenotype concept in psychiatry: etymology and strategic intentions. *Am J Psychiatry* 2003;160:636. [PubMed: 12668349]
79. Greenwood TA, Braff DL, Light GA, et al. Initial heritability analyses of endophenotypic measures for schizophrenia: the consortium on the genetics of schizophrenia. *Arch Gen Psychiatry* 2007;64:1242. [PubMed: 17984393]
80. Book JA. Schizophrenia as a gene mutation. *Acta Genet Stat Med* 1953;4:133. [PubMed: 13137881]
81. Friedman JM. Genetic disease in the offspring of older fathers. *Obstet Gynecol* 1981;57:745. [PubMed: 7231827]
82. Sipos A, Rasmussen F, Harrison G, et al. Paternal age and schizophrenia: a population based cohort study. *BMJ* 2004;329:1070. [PubMed: 15501901]

83. Crow JF. The origins, patterns and implications of human spontaneous mutation. *Nat Rev Genet* 2000;1:40. [PubMed: 11262873]
84. Allen NC, Bagade S, McQueen MB, et al. Systematic meta-analyses and field synopsis of genetic association studies in schizophrenia: the SzGene database. *Nat Genet* 2008;40:827. [PubMed: 18583979]
85. Sanders AR, Duan J, Levinson DF, et al. No significant association of 14 candidate genes with schizophrenia in a large European ancestry sample: implications for psychiatric genetics. *Am J Psychiatry* 2008;165:497. [PubMed: 18198266]
86. Ioannidis JP. Why most discovered true associations are inflated. *Epidemiology* 2008;19:640. [PubMed: 18633328]
87. McCarthy MI, Abecasis GR, Cardon LR, et al. Genome-wide association studies for complex traits: consensus, uncertainty and challenges. *Nat Rev Genet* 2008;9:356. [PubMed: 18398418]
88. Richter SH, Garner JP, Wurbel H. Environmental standardization: cure or cause of poor reproducibility in animal experiments? *Nat Methods* 2009;6:257. [PubMed: 19333241]
89. Shi J, Levinson DF, Duan J, et al. Common variants on chromosome 6p22.1 are associated with schizophrenia. *Nature* 2009;460:753. [PubMed: 19571809]
90. Hindorf LA, Sethupathy P, Junkins HA, et al. Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. *Proc Natl Acad Sci U S A* 2009;106:9362. [PubMed: 19474294]
91. Dudbridge F, Gusnanto A. Estimation of significance thresholds for genomewide association scans. *Genet Epidemiol* 2008;32:227. [PubMed: 18300295]
92. Hoggart CJ, Clark TG, De Iorio M, et al. Genome-wide significance for dense SNP and resequencing data. *Genet Epidemiol* 2008;32:179. [PubMed: 18200594]
93. Pe'er I, Yelensky R, Altshuler D, et al. Estimation of the multiple testing burden for genomewide association studies of nearly all common variants. *Genet Epidemiol* 2008;32:381. [PubMed: 18348202]
94. Manolio TA, Collins FS, Cox NJ, et al. Finding the missing heritability of complex diseases. *Nature* 2009;461:747. [PubMed: 19812666]
95. Li M, Li C, Guan W. Evaluation of coverage variation of SNP chips for genome-wide association studies. *Eur J Hum Genet* 2008;16:635. [PubMed: 18253166]
96. Halperin E, Stephan DA. SNP imputation in association studies. *Nat Biotechnol* 2009;27:349. [PubMed: 19352374]
97. Nothnagel M, Ellinghaus D, Schreiber S, et al. A comprehensive evaluation of SNP genotype imputation. *Hum Genet* 2009;125:163. [PubMed: 19089453]
98. Pritchard JK. Are rare variants responsible for susceptibility to complex diseases? *Am J Hum Genet* 2001;69:124. [PubMed: 11404818]
99. Reich DE, Lander ES. On the allelic spectrum of human disease. *Trends Genet* 2001;17:502. [PubMed: 11525833]
100. Gottesman II, Shields J. A polygenic theory of schizophrenia. *Proc Natl Acad Sci U S A* 1967;58:199. [PubMed: 5231600]
101. Risch N. Genetic linkage and complex diseases, with special reference to psychiatric disorders. *Genet Epidemiol* 1990;7:3. [PubMed: 2184091]
102. Purcell SM, Wray NR, Stone JL, et al. Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature* 2009;460:748. [PubMed: 19571811]
103. Stefansson H, Ophoff RA, Steinberg S, et al. Common variants conferring risk of schizophrenia. *Nature* 2009;460:744. [PubMed: 19571808]
104. Worden FG, Childs B, Matthyse S, et al. Frontiers of psychiatric genetics. *Neurosci Res Program Bull* 1976;14:8. [PubMed: 1083493]
105. Bach ML, Bach FH, Joo P. Leukemia-associated antigens in the mixed leukocyte culture test. *Science* 1969;166:1520. [PubMed: 17655047]
106. Wei J, Hemmings GP. The NOTCH4 locus is associated with susceptibility to schizophrenia. *Nat Genet* 2000;25:376. [PubMed: 10932176]

107. Traherne JA. Human MHC architecture and evolution: implications for disease association studies. *Int J Immunogenet* 2008;35:179. [PubMed: 18397301]
108. Daly MJ, Rioux JD, Schaffner SF, et al. High-resolution haplotype structure in the human genome. *Nat Genet* 2001;29:229. [PubMed: 11586305]
109. Gabriel SB, Schaffner SF, Nguyen H, et al. The structure of haplotype blocks in the human genome. *Science* 2002;296:2225. [PubMed: 12029063]
110. Kleinjan DA, van Heyningen V. Long-range control of gene expression: emerging mechanisms and disruption in disease. *Am J Hum Genet* 2005;76:8. [PubMed: 15549674]
111. Birney E, Stamatoyannopoulos JA, Dutta A, et al. Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. *Nature* 2007;447:799. [PubMed: 17571346]
112. Adegbola A, Gao H, Sommer S, et al. A novel mutation in JARID1C/SMCX in a patient with autism spectrum disorder (ASD). *Am J Med Genet A* 2008;146A:505. [PubMed: 18203167]
113. Costa E, Dong E, Grayson DR, et al. Reviewing the role of DNA (cytosine-5) methyltransferase overexpression in the cortical GABAergic dysfunction associated with psychosis vulnerability. *Epigenetics* 2007;2:29. [PubMed: 17965595]
114. Shi Y. Histone lysine demethylases: emerging roles in development, physiology and disease. *Nat Rev Genet* 2007;8:829. [PubMed: 17909537]
115. Kawasaki H, Iwamuro S. Potential roles of histones in host defense as antimicrobial agents. *Infect Disord Drug Targets* 2008;8:195. [PubMed: 18782037]
116. Kim HS, Cho JH, Park HW, et al. Endotoxin-neutralizing antimicrobial proteins of the human placenta. *J Immunol* 2002;168:2356. [PubMed: 11859126]
117. Eaton WW, Byrne M, Ewald H, et al. Association of schizophrenia and autoimmune diseases: linkage of Danish national registers. *Am J Psychiatry* 2006;163:521. [PubMed: 16513876]
118. Shiina T, Inoko H, Kulski JK. An update of the HLA genomic region, locus information and disease associations: 2004. *Tissue Antigens* 2004;64:631. [PubMed: 15546336]
119. Viken MK, Blomhoff A, Olsson M, et al. Reproducible association with type 1 diabetes in the extended class I region of the major histocompatibility complex. *Genes Immun* 2009;10:323. [PubMed: 19295542]
120. Oksenberg JR, Baranzini SE, Sawcer S, et al. The genetics of multiple sclerosis: SNPs to pathways to pathogenesis. *Nat Rev Genet* 2008;9:516. [PubMed: 18542080]
121. Wang H, Feng R, Phillip Wang L, et al. CaMKII activation state underlies synaptic labile phase of LTP and short-term memory formation. *Curr Biol* 2008;18:1546. [PubMed: 18929487]
122. Harrison PJ, Weinberger DR. Schizophrenia genes, gene expression, and neuropathology: on the matter of their convergence. *Mol Psychiatry* 2005;10:40. [PubMed: 15263907]
123. Gulacsi AA, Anderson SA. Beta-catenin-mediated Wnt signaling regulates neurogenesis in the ventral telencephalon. *Nat Neurosci* 2008;11:1383. [PubMed: 18997789]
124. Flora A, Garcia JJ, Thaller C, et al. The E-protein Tcf4 interacts with Math1 to regulate differentiation of a specific subset of neuronal progenitors. *Proc Natl Acad Sci U S A* 2007;104:15382. [PubMed: 17878293]
125. Kalscheuer VM, Feenstra I, Van Ravenswaaij-Arts CM, et al. Disruption of the TCF4 gene in a girl with mental retardation but without the classical Pitt-Hopkins syndrome. *Am J Med Genet A* 2008;146A:2053. [PubMed: 18627065]
126. Brockschmidt A, Todt U, Ryu S, et al. Severe mental retardation with breathing abnormalities (Pitt-Hopkins syndrome) is caused by haploinsufficiency of the neuronal bHLH transcription factor TCF4. *Hum Mol Genet* 2007;16:1488. [PubMed: 17478476]
127. Zweier C, de Jong EK, Zweier M, et al. CNTNAP2 and NRXN1 Are Mutated in Autosomal-Recessive Pitt-Hopkins-like Mental Retardation and Determine the Level of a Common Synaptic Protein in Drosophila. *Am J Hum Genet* 2009;85:655. [PubMed: 19896112]
128. Friedman JI, Vrijenhoek T, Markx S, et al. CNTNAP2 gene dosage variation is associated with schizophrenia and epilepsy. *Mol Psychiatry* 2008;13:261. [PubMed: 17646849]
129. O'Donovan MC, Craddock N, Norton N, et al. Identification of loci associated with schizophrenia by genome-wide association and follow-up. *Nat Genet* 2008;40:1053. [PubMed: 18677311]

130. Esslinger C, Walter H, Kirsch P, et al. Neural mechanisms of a genome-wide supported psychosis variant. *Science* 2009;324:605. [PubMed: 19407193]
131. Wray NR, Goddard ME, Visscher PM. Prediction of individual genetic risk to disease from genome-wide association studies. *Genome Res* 2007;17:1520. [PubMed: 17785532]
132. Sklar P, Smoller JW, Fan J, et al. Whole-genome association study of bipolar disorder. *Mol Psychiatry* 2008;13:558. [PubMed: 18317468]
133. WTCCC. Wellcome Trust Case Control Consortium: Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature* 2007;447:661. [PubMed: 17554300]
134. Sun J, Jia P, Fanous AH, et al. A multi-dimensional evidence-based candidate gene prioritization approach for complex diseases-schizophrenia as a case. *Bioinformatics* 2009;25:2595. [PubMed: 19602527]
135. Kathiresan S, Willer CJ, Peloso GM, et al. Common variants at 30 loci contribute to polygenic dyslipidemia. *Nat Genet* 2009;41:56. [PubMed: 19060906]
136. Kamin LJ, Goldberger AS. Twin studies in behavioral research: a skeptical view. *Theor Popul Biol* 2002;61:83. [PubMed: 11895384]
137. Herold C, Steffens M, Brockschmidt FF, et al. INTERSNP: Genome-wide Interaction Analysis Guided by A Priori Information. *Bioinformatics*. 2009
138. Kirov G, Gumus D, Chen W, et al. Comparative genome hybridization suggests a role for NRXN1 and APBA2 in schizophrenia. *Hum Mol Genet* 2008;17:458. [PubMed: 17989066]
139. Walsh T, McClellan JM, McCarthy SE, et al. Rare structural variants disrupt multiple genes in neurodevelopmental pathways in schizophrenia. *Science* 2008;320:539. [PubMed: 18369103]
140. Xu B, Roos JL, Levy S, et al. Strong association of de novo copy number mutations with sporadic schizophrenia. *Nat Genet* 2008;40:880. [PubMed: 18511947]
141. Stefansson H, Rujescu D, Cichon S, et al. Large recurrent microdeletions associated with schizophrenia. *Nature* 2008;455:232. [PubMed: 18668039]
142. ISC. Rare chromosomal deletions and duplications increase risk of schizophrenia. *Nature* 2008;455:237. [PubMed: 18668038]
143. Kirov G, Grozeva D, Norton N, et al. Support for the involvement of large copy number variants in the pathogenesis of schizophrenia. *Hum Mol Genet* 2009;18:1497. [PubMed: 19181681]
144. Need AC, Ge D, Weale ME, et al. A genome-wide investigation of SNPs and CNVs in schizophrenia. *PLoS Genet* 2009;5:e1000373. [PubMed: 19197363]
145. Shprintzen RJ, Goldberg R, Golding-Kushner KJ, et al. Late-onset psychosis in the velo-cardio-facial syndrome. *Am J Med Genet* 1992;42:141. [PubMed: 1308357]
146. Pulver AE, Nestadt G, Goldberg R, et al. Psychotic illness in patients diagnosed with velo-cardio-facial syndrome and their relatives. *J Nerv Ment Dis* 1994;182:476. [PubMed: 8040660]
147. Murphy KC, Jones LA, Owen MJ. High rates of schizophrenia in adults with velo-cardio-facial syndrome. *Arch Gen Psychiatry* 1999;56:940. [PubMed: 10530637]
148. Rujescu D, Ingason A, Cichon S, et al. Disruption of the neurexin 1 gene is associated with schizophrenia. *Hum Mol Genet* 2009;18:988. [PubMed: 18945720]
149. Mackay TF. The genetic architecture of complex behaviors: lessons from *Drosophila*. *Genetica* 2009;136:295. [PubMed: 18758968]
150. Weedon MN, Lango H, Lindgren CM, et al. Genome-wide association analysis identifies 20 loci that influence adult height. *Nat Genet* 2008;40:575. [PubMed: 18391952]
151. Zeggini E, Scott LJ, Saxena R, et al. Meta-analysis of genome-wide association data and large-scale replication identifies additional susceptibility loci for type 2 diabetes. *Nat Genet* 2008;40:638. [PubMed: 18372903]
152. Thomas G, Jacobs KB, Yeager M, et al. Multiple loci identified in a genome-wide association study of prostate cancer. *Nat Genet* 2008;40:310. [PubMed: 18264096]
153. Schulze TG, Hedeker D, Zandi P, et al. What is familial about familial bipolar disorder? Resemblance among relatives across a broad spectrum of phenotypic characteristics. *Arch Gen Psychiatry* 2006;63:1368. [PubMed: 17146011]

154. APA. Diagnostic and Statistical Manual of Mental Disorders. 4. Washington, DC: American Psychiatric Association; 1994. (DSM-IV)
155. Berrettini WH. Are schizophrenic and bipolar disorders related? A review of family and molecular studies. *Biol Psychiatry* 2000;48:531. [PubMed: 11018225]
156. Craddock N, Owen MJ. Rethinking psychosis: the disadvantages of a dichotomous classification now outweigh the advantages. *World Psychiatry* 2007;6:84. [PubMed: 18235858]
157. van Os J, Kapur S. Schizophrenia. *Lancet* 2009;374:635. [PubMed: 19700006]
158. Goes FS, Zandi PP, Miao K, et al. Mood-incongruent psychotic features in bipolar disorder: familial aggregation and suggestive linkage to 2p11-q14 and 13q21-33. *Am J Psychiatry* 2007;164:236. [PubMed: 17267786]
159. McCarthy SE, Makarov V, Kirov G, et al. Microduplications of 16p11.2 are associated with schizophrenia. *Nat Genet* 2009;41:1223. [PubMed: 19855392]
160. Van Snellenberg JX, de Candia T. Meta-analytic evidence for familial coaggregation of schizophrenia and bipolar disorder. *Arch Gen Psychiatry* 2009;66:748. [PubMed: 19581566]
161. Lichtenstein P, Yip BH, Bjork C, et al. Common genetic determinants of schizophrenia and bipolar disorder in Swedish families: a population-based study. *Lancet* 2009;373:234. [PubMed: 19150704]
162. Moskvina V, Craddock N, Holmans P, et al. Gene-wide analyses of genome-wide association data sets: evidence for multiple common risk alleles for schizophrenia and bipolar disorder and for overlap in genetic risk. *Mol Psychiatry* 2009;14:252. [PubMed: 19065143]
163. Green EK, Grozeva D, Jones I, et al. The bipolar disorder risk allele at CACNA1C also confers risk of recurrent major depression and of schizophrenia. *Mol Psychiatry*. 2009
164. Wassink TH, Piven J, Vieland VJ, et al. Evaluation of the chromosome 2q37.3 gene CENTG2 as an autism susceptibility gene. *Am J Med Genet B Neuropsychiatr Genet* 2005;136B:36. [PubMed: 15892143]
165. Bassell GJ, Warren ST. Fragile X syndrome: loss of local mRNA regulation alters synaptic development and function. *Neuron* 2008;60:201. [PubMed: 18957214]
166. Konstantareas MM, Hewitt T. Autistic disorder and schizophrenia: diagnostic overlaps. *J Autism Dev Disord* 2001;31:19. [PubMed: 11439750]
167. Rapoport J, Chavez A, Greenstein D, et al. Autism spectrum disorders and childhood-onset schizophrenia: clinical and biological contributions to a relation revisited. *J Am Acad Child Adolesc Psychiatry* 2009;48:10. [PubMed: 19218893]
168. Stahlberg O, Soderstrom H, Rastam M, et al. Bipolar disorder, schizophrenia, and other psychotic disorders in adults with childhood onset AD/HD and/or autism spectrum disorders. *J Neural Transm* 2004;111:891. [PubMed: 15206005]
169. Daniels JL, Forssen U, Hultman CM, et al. Parental psychiatric disorders associated with autism spectrum disorders in the offspring. *Pediatrics* 2008;121:e1357. [PubMed: 18450879]
170. Larsson HJ, Eaton WW, Madsen KM, et al. Risk factors for autism: perinatal factors, parental psychiatric history, and socioeconomic status. *Am J Epidemiol* 2005;161:916. [PubMed: 15870155]
171. Mouridsen SE, Rich B, Isager T, et al. Psychiatric disorders in individuals diagnosed with infantile autism as children: a case control study. *J Psychiatr Pract* 2008;14:5. [PubMed: 18212598]
172. Cascella NG, Schretlen DJ, Sawa A. Schizophrenia and epilepsy: is there a shared susceptibility? *Neurosci Res* 2009;63:227. [PubMed: 19367784]
173. Hyde TM, Weinberger DR. Seizures and schizophrenia. *Schizophr Bull* 1997;23:611. [PubMed: 9365998]
174. Manolio TA, Rodriguez LL, Brooks L, et al. New models of collaboration in genome-wide association studies: the Genetic Association Information Network. *Nat Genet* 2007;39:1045. [PubMed: 17728769]
175. Cichon S, Craddock N, Daly M, et al. Genomewide association studies: history, rationale, and prospects for psychiatric disorders. *Am J Psychiatry* 2009;166:540. [PubMed: 19339359]
176. Psychiatric_GWAS_Consortium_Steering_Committee. A framework for interpreting genome-wide association studies of psychiatric disorders. *Mol Psychiatry* 2009;14:10. [PubMed: 19002139]
177. Ioannidis JP, Thomas G, Daly MJ. Validating, augmenting and refining genome-wide association signals. *Nat Rev Genet* 2009;10:318. [PubMed: 19373277]

178. Spencer CC, Su Z, Donnelly P, et al. Designing genome-wide association studies: sample size, power, imputation, and the choice of genotyping chip. *PLoS Genet* 2009;5:e1000477. [PubMed: 19492015]
179. Bowden J, Dudbridge F. Unbiased estimation of odds ratios: combining genomewide association scans with replication studies. *Genet Epidemiol* 2009;33:406. [PubMed: 19140132]
180. Hao K, Chudin E, McElwee J, et al. Accuracy of genome-wide imputation of untyped markers and impacts on statistical power for association studies. *BMC Genet* 2009;10:27. [PubMed: 19531258]
181. McElroy JP, Nelson MR, Caillier SJ, et al. Copy number variation in African Americans. *BMC Genet* 2009;10:15. [PubMed: 19317893]
182. Moffatt MF, Kabesch M, Liang L, et al. Genetic variants regulating ORMDL3 expression contribute to the risk of childhood asthma. *Nature* 2007;448:470. [PubMed: 17611496]
183. Dixon AL, Liang L, Moffatt MF, et al. A genome-wide association study of global gene expression. *Nat Genet* 2007;39:1202. [PubMed: 17873877]
184. Dendrou CA, Plagnol V, Fung E, et al. Cell-specific protein phenotypes for the autoimmune locus IL2RA using a genotype-selectable human bioresource. *Nat Genet* 2009;41:1011. [PubMed: 19701192]
185. Mardis ER. New strategies and emerging technologies for massively parallel sequencing: applications in medical research. *Genome Med* 2009;1:40. [PubMed: 19435481]
186. O'Rahilly S. Human genetics illuminates the paths to metabolic disease. *Nature* 2009;462:307. [PubMed: 19924209]
187. Mitchell KJ, Porteous DJ. GWAS for psychiatric disease: is the framework built on a solid foundation? *Mol Psychiatry* 2009;14:740. [PubMed: 19626023]
188. Sullivan PF, Gejman PV. Response to Mitchell & Porteus. *Mol Psychiatry*. in press.
189. Crick, F. *What Mad Pursuit: A Personal View of Scientific Discovery*. New York: Basic Books; 1988.
190. Mackay TF, Anholt RR. Of flies and man: *Drosophila* as a model for human complex traits. *Annu Rev Genomics Hum Genet* 2006;7:339. [PubMed: 16756480]
191. Ayroles JF, Carbone MA, Stone EA, et al. Systems genetics of complex traits in *Drosophila melanogaster*. *Nature Genetics* 2009;41:299. [PubMed: 19234471]
192. Lencz T, Morgan TV, Athanasiou M, et al. Converging evidence for a pseudoautosomal cytokine receptor gene locus in schizophrenia. *Mol Psychiatry* 2007;12:572. [PubMed: 17522711]
193. Sullivan PF, Lin D, Tzeng JY, et al. Genomewide association for schizophrenia in the CATIE study: results of stage 1. *Mol Psychiatry* 2008;13:570. [PubMed: 18347602]
194. Sebat J, Lakshmi B, Malhotra D, et al. Strong association of de novo copy number mutations with autism. *Science* 2007;316:445. [PubMed: 17363630]
195. Weiss LA, Shen Y, Korn JM, et al. Association between microdeletion and microduplication at 16p11.2 and autism. *New England Journal of Medicine* 2008;358:667. [PubMed: 18184952]
196. Christian SL, Brune CW, Sudi J, et al. Novel submicroscopic chromosomal abnormalities detected in autism spectrum disorder. *Biol Psychiatry* 2008;63:1111. [PubMed: 18374305]
197. Kumar RA, KaraMohamed S, Sudi J, et al. Recurrent 16p11.2 microdeletions in autism. *Hum Mol Genet* 2008;17:628. [PubMed: 18156158]
198. Marshall CR, Noor A, Vincent JB, et al. Structural variation of chromosomes in autism spectrum disorder. *Am J Hum Genet* 2008;82:477. [PubMed: 18252227]
199. Brunetti-Pierri N, Berg JS, Scaglia F, et al. Recurrent reciprocal 1q21.1 deletions and duplications associated with microcephaly or macrocephaly and developmental and behavioral abnormalities. *Nat Genet* 2008;40:1466. [PubMed: 19029900]
200. de Kovel CG, Trucks H, Helbig I, et al. Recurrent microdeletions at 15q11.2 and 16p13.11 predispose to idiopathic generalized epilepsies. *Brain*. 2009
201. Dibbens LM, Mullen S, Helbig I, et al. Familial and sporadic 15q13.3 microdeletions in idiopathic generalized epilepsy: precedent for disorders with complex inheritance. *Hum Mol Genet* 2009;18:3626. [PubMed: 19592580]
202. Helbig I, Mefford HC, Sharp AJ, et al. 15q13.3 microdeletions increase risk of idiopathic generalized epilepsy. *Nat Genet* 2009;41:160. [PubMed: 19136953]

203. Mefford HC, Sharp AJ, Baker C, et al. Recurrent rearrangements of chromosome 1q21.1 and variable pediatric phenotypes. *N Engl J Med* 2008;359:1685. [PubMed: 18784092]
204. Sharp AJ, Mefford HC, Li K, et al. A recurrent 15q13.3 microdeletion syndrome associated with mental retardation and seizures. *Nat Genet* 2008;40:322. [PubMed: 18278044]

Figure 1.

Quantile-quantile plot of observed vs. expected p -values for tag SNPs for fourteen schizophrenia candidate genes. Open circles represent the relationship between the expected (X-axis) and observed (Y-axis) p -values for pointwise nominal Armitage trend tests for the 433 SNPs that represent tags (at $r^2 > 0.8$) for common SNPs in each of 14 tested genes (*RGS4*, *DISC1*, *DTNBP1*, *STX7*, *TAAR6*, *PPP3CC*, *NRG1*, *DRD2*, *HTR2A*, *DAOA*, *AKT1*, *CHRNA7*, *COMT*, and *ARVCF*). The solid line represents the null expectation. The observed distribution is within the 95% confidence interval of the null expectation, consistent with a lack of evidence in the tested sample for association with schizophrenia in the tested candidate genes. The lowest p -values are slightly below the line (less significant than expected), but still within the confidence interval. Reprinted with permission⁸⁵.

Figure 2.

The combined p -values of three datasets (MGS, ISC, and SGENE) at the MHC region (~26~33Mb). The positions of the three histone gene clusters and *NOTCH4* are indicated in the RefSeq track (top of the graph). The relative location of the maximum peak in each data set is indicated by a line below the p -value peaks. *Cis*-eQTL (with LOD > 5)¹⁸³ are shown below the graph with association $-\log p$ -values.

Table 1

Top genes or genomic regions identified in recent schizophrenia GWAS.

First author and year	Sample (case/control)	Gene or region	Lowest <i>p</i> -values	OR	Reference
Lenz 2007	178/144 (EA)	<i>CSF2RA</i> , <i>SHOX</i>	3.7×10^{-7}	3.23	192
Sullivan 2008	738/733 (EA)	<i>AGBL1</i>	1.71×10^{-6}	6.01	193
O'Donovan 2008	Discovery: 479/2,937 (EA) Follow up: 6,829/9,897 (EA)	<i>ZNF804A</i>	1.61×10^{-7}	1.12	129
Need 2009	Discovery: 871/863 (EA) Follow up: 1,460/12,995 (EA)	<i>ADAMTSL3</i>	1.35×10^{-7}	0.68	144
Purcell 2009 (ISC)	3,322/3,587 (EA)	MHC region ^a <i>MYO18B</i>	9.5×10^{-9} 3.4×10^{-7}	0.82	102
Stefansson 2009 (SGENE)	Discovery: 2,663/13,498 (EA) Follow up: 4,999/15,555 (EA)	MHC region ^b <i>NRGN</i> ^b <i>TCF4</i> ^b	1.4×10^{-12} 2.4×10^{-9} 4.1×10^{-9}	1.16 ^c 1.15 1.23	103
Shi 2009 (MGS)	2,681/2,653 (EA) 1,286/973 (AA)	MHC region ^a <i>CENTG2</i> (in EA only) <i>ERBB4</i> (in AA only)	9.5×10^{-9} 4.59×10^{-7} 2.14×10^{-6}	0.88 1.23 0.73	89

^aCombined analysis of ISC, SGENE-plus (GWAS set), and MGS.

^bCombined analysis of ISC and MGS, along with SGENE-plus and follow up samples.

^cOR (odds ratio) is for common allele of the associated SNP, which is different from that in ISC and MGS.

Table 2

Genes mentioned in review, and their functions.

Gene symbol	Gene description	Chromosome	Function
<i>ABT1</i>	<i>activator of basal transcription 1</i>	6	transcriptional regulation
<i>AKT1</i>	<i>v-akt murine thymoma viral oncogene homolog 1</i>	14	critical mediator of growth factor-induced neuronal survival
<i>ARVCF</i>	<i>armadillo repeat gene deletes in VCFS</i>	22	catenin family member (adherens junction complex formation)
<i>BTN3A1</i>	<i>butyrophilin subfamily 3 member A1</i>	6	immunoglobulin superfamily
<i>BTN3A2</i>	<i>butyrophilin subfamily 3 member A2</i>	6	immunoglobulin superfamily
<i>CACNA1C</i>	<i>calcium channel, voltage-dependent, L type, alpha 1C subunit</i>	12	regulates muscle contraction, hormone or neurotransmitter release, cell cycle
<i>CENTG2</i>	<i>centaurin gamma 2</i>	2	protein trafficking in the endosomal-lysosomal system
<i>CHRNA7</i>	<i>cholinergic receptor, nicotinic, alpha 7</i>	15	ligand-gated ion channel mediating synaptic transmission a member of the neurexin family which functions in the vertebrate
<i>CNTNAP2</i>	<i>contactin associated protein-like 2</i>	7	nervous system as cell adhesion molecules and receptors
<i>COMT</i>	<i>catechol-O-methyltransferase</i>	22	catecholamine neurotransmitter metabolism; VCFS region (22q deletion syndrome)
<i>DAOA</i>	<i>D-amino acid oxidase activator</i>	13	activator of the enzyme DAAO, which degrades D-serine
<i>DISC1</i>	<i>disrupted in schizophrenia 1</i>	1	neurite outgrowth and cortical development; disrupted in t(1;1)(q42.1;q14.3)
<i>DRD2</i>	<i>dopamine receptor D2</i>	11	G-protein coupled receptor for dopamine; inhibits adenylyl cyclase
<i>DRD4</i>	<i>dopamine receptor D4</i>	11	G-protein coupled receptor for dopamine; inhibits adenylyl cyclase
<i>DTNBP1</i>	<i>dystrobrevin binding protein 1</i>	6	component of the biogenesis of lysosome-related organelles complex 1
<i>ERBB4</i>	<i>v-erb-a erythroblastic leukemia viral oncogene</i>	2	Receptors for neuregulins; cell differentiation
<i>EV1</i>	<i>ecotropic viral integration site 5</i>	1	cell cycle progression
<i>FAM69A</i>	<i>family with sequence similarity 69, member A</i>	1	unknown; implicated in MS
<i>FKSG83</i>	<i>hypothetical protein</i>	6	G-protein-coupled receptor signaling
<i>FMR1</i>	<i>fragile X mental retardation 1</i>	23	may be involved in nucleus → cytoplasm mRNA trafficking
<i>FXR1</i>	<i>fragile X mental retardation, autosomal homolog 1</i>	3	RNA-binding protein; embryonic and postnatal development of muscle tissue
<i>HIST1H2AH</i>	<i>histone cluster 1, H2ah</i>	6	chromatin modification, transcription regulation, host defense
<i>HTR2A</i>	<i>5-hydroxytryptamine (serotonin) receptor 2A</i>	13	G-protein coupled receptor for serotonin; activates phosphoinositide hydrolysis
<i>HMGNA4</i>	<i>high mobility group nucleosomal binding domain 4</i>	6	involved in chromatin structure
<i>MYO18B</i>	<i>myosin XVIIIIB</i>	22	regulate muscle-specific genes; intracellular trafficking
<i>NOTCH4</i>	<i>Notch homolog 4 (Drosophila)</i>	6	receptor for Jagged1, Jagged2 and Delta1; cell-fate determination
<i>NRG1</i>	<i>neuregulin 1</i>	8	signaling protein that mediates cell-cell interactions; roles in growth & development

Gene symbol	Gene description	Chromosome	Function
<i>NRGN</i>	<i>neurogranin</i>	11	postsynaptic protein kinase substrate; learning and memory; glutamate signaling
<i>POM121L2</i>	<i>nuclear pore membrane protein 121-like 2</i>	6	nuclear pore complex
<i>PPP3CC</i>	<i>protein phosphatase 3 (formerly 2B), catalytic subunit, gamma isoform</i>	8	Ca ⁺⁺ -dependent modifier of phosphorylation status
<i>PRSS16</i>	<i>protease serine 16</i>	6	alternative antigen presenting
<i>RGS4</i>	<i>regulator of G-protein signaling 4</i>	1	GTPase activating protein for G alpha subunits of G proteins
<i>RPL5</i>	<i>ribosomal protein L5</i>	1	rRNA maturation and formation of the 60S ribosomal subunits
<i>STX7</i>	<i>syntaxin 7</i>	6	endosomal soluble N-ethylmaleimide-sensitive factor attachment protein receptors (SNARE)
<i>TAAAR6</i>	<i>trace amine associated receptor 6</i>	6	Trace amines are endogenous amine compounds, chemically similar to classic biogenic amines like dopamine
<i>TCF4</i>	<i>transcription factor 4</i>	18	neuronal transcriptional factor; neurogenesis
<i>ZNF184</i>	<i>zinc finger protein 18</i>	6	transcriptional regulation
<i>ZNF322A</i>	<i>zinc finger protein 322A</i>	6	transcriptional regulation
<i>ZNF804A</i>	<i>zinc finger protein 804A</i>	2	transcription factor; neuronal connectivity in the dorsolateral prefrontal cortex

Notes: Genes are listed alphabetically by symbol. All chromosome 6 genes are in the MHC except *DTNBP1*.

Summary of recent genome-wide CNV studies of schizophrenia.

Table 3

CNVs in case:control (NA = data not available)									
Study	Sample	Major findings	1q21.1	2p16.3 (NRXN1)	15q11.2	15q13.2	16p11.2	22q11.21	Reference
Kirov 2008	93 trios	Two CNVs likely to be pathogenic	NA	1 del	NA	1 dup	NA	NA	138
Walsh 2008	150 cases/268 controls; 92 childhood onset cases	Rare CNVs in 15% cases vs. 5% controls	1:0 del	1:0 del	NA	NA	2:0 dup	NA	139
Xu 2008	359 trios as screening sample; 152 cases/159 controls	In sporadic cases, frequency of rare de novo CNVs was 10% vs. 1.3% in controls	1:0 del	NA	NA	NA	NA	3:0 del	140
Stefansson 2008	1,433 cases/33,250 controls; 3 CNVs (1q21.1, 15q11.2 and 15q13.3) were followed up in 3,285 cases/7,951 controls	Three rare CNVs (1q21.1, 15q11.2, and 15q13.3) showed nominal association	11:8 del	0:2 del	26:79 del	7:8 del	2:11 del	8:0 del	141
Stone 2008	3,391 cases/3,181 controls	Rare (<1%) and large CNVs (>100kb) are enriched in cases (1.15-fold); 3 regions (1q21.1, 15q13.2, and 22q11.21) showed significant association	10:1 del	5:6 del	26:11 del	9:0 del	5:1 dup	13:0 del	142
Kirov 2009	471 cases/2,792 controls	Large CNVs (>1Mb) were 2.26 times over-represented in cases	0:2 del	1:3 del	4:14 del	0:0 del	NA	2:0 del	143
Need 2009	1,013 cases/1,084 controls	Large CNVs (>2Mb) are enriched in cases	1:0 del	3:1 del	NA	NA	NA	4:0 del	144
Summary frequency of most implicated CNV in schizophrenia (case vs. control)		CNV	del	del	del	del	dup	del	
		case	0.24%	0.10%	0.65%	0.20%	0.35%	0.30%	
		control	0.02%	0.02%	0.22%	0.02%	0.03%	0.00%	
Other disorders reported in schizophrenia CNV carriers of the most implicated CNV in the studies above			MR						
Other disorders reported associated with these CNVs in other studies			ASD, LD						
Other disorders reported associated with these CNVs in other studies			ASD, MR, epilepsy, microcephaly, cardiac abnormalities						
Other disorders reported associated with these CNVs in other studies			ASD, MR, ASD, MR, epilepsy						
Other disorders reported associated with these CNVs in other studies			ASD, MR, bipolar disorder						
Other disorders reported associated with these CNVs in other studies			ASD, MR, VCFS						
Other disorders reported associated with these CNVs in other studies			LD						
Other disorders reported associated with these CNVs in other studies			141–143						

Note: Study lists first author and year. The summary frequency for 16p11.2 also includes an unlisted study 159 that only examined 16p11.2 CNV (i.e., not genome-wide) and found 21 dup in 4,551 schizophrenia cases and 2 dup in 6,391 controls (the sample included MGS-GAIN samples used in our GWAS 89). MR = mental retardation; ASD = autism or autism spectrum disorders; LD = learning disability (dyslexia and others); epilepsy (includes seizures). The 22q11.21 CNV deletion, of course, also includes DiGeorge or velocardiofacial syndrome (VCFS).