Genomewide Association Studies: History, Rationale, and Prospects for Psychiatric Disorders

Objective: The authors conducted a review of the history and empirical basis of genomewide association studies (GWAS), the rationale for GWAS of psychiatric disorders, results to date, limitations, and plans for GWAS meta-analyses.

Method: A literature review was carried out, power and other issues discussed, and planned studies assessed.

Results: Most of the genomic DNA sequence differences between any two people are common (frequency >5%) single nucleotide polymorphisms (SNPs). Because of localized patterns of correlation (linkage disequilibrium), 500,000 to 1,000,000 of these SNPs can test the hypothesis that one or more common variants explain part of the genetic risk for a disease. GWAS technologies can also detect some of the copy number variants (deletions and duplications) in the genome. Systematic study of rare variants will require large-scale resequencing analyses. GWAS methods have detected a remarkable number of robust genetic associations for dozens of common diseases and traits, leading to new pathophysiological hypotheses, although only small proportions of genetic variance have been explained thus far and therapeutic applications will require substantial further effort. Study design issues, power, and limitations are discussed. For psychiatric disorders, there are initial significant findings for common SNPs and for rare copy number variants, and many other studies are in progress.

Conclusions: GWAS of large samples have detected associations of common SNPs and of rare copy number variants with psychiatric disorders. More findings are likely, since larger GWAS samples detect larger numbers of common susceptibility variants, with smaller effects. The Psychiatric GWAS Consortium is conducting GWAS meta-analyses for schizophrenia, bipolar disorder, major depressive disorder, autism, and attention deficit hyperactivity disorder. Based on results for other diseases, larger samples will be required. The contribution of GWAS will depend on the true genetic architecture of each disorder.

Since 2005 (1), genomewide association studies (GWAS [jē’wēd wōs]) have produced strongly significant evidence that specific common DNA sequence differences among people influence their genetic susceptibility to more than 40 different common diseases (2). Many of these findings implicate previously unsuspected candidate genes and new pathophysiological hypotheses. The method is feasible because millions of human DNA sequence variations have been catalogued and new technologies have been developed that can assay more than one million variants rapidly and accurately. The first GWAS reports have appeared for psychiatric disorders, and nearly 50 GWAS are completed or will be reported soon for attention deficit hyperactivity disorder (ADHD), autism, bipolar disorder, major depressive disorder, and schizophrenia. Additional studies are in progress. We formed an international consortium of psychiatric GWAS investigators to carry out rapid meta-analyses of these five disorders to maximize power. In the present overview, we describe GWAS methods, their rationale, and current results for nonpsychiatric and psychiatric disorders as well as limitations and uncertainties.

Candidate Genes, Linkage, and Linkage Disequilibrium

Genetic Epidemiology

Before any molecular genetic study is undertaken, the methods of genetic epidemiology are used to identify a phenotype (observable disease or trait) that is at least partially heritable. An introduction to these methods is available online (http://www.dorak.info/epi/genetepi.html). Briefly, twin, family, and population-based studies are used to estimate heritability, define the most heritable phenotype, and explore interactions between genetic and environmental factors. The current diagnostic definitions of major psychiatric disorders are based in part on twin and family data. Epidemiological data are also critical for defining appropriate comparison groups for molecular studies. The data for psychiatric disorders suggest that
most of the heritable risk is due to interactions of combinations of genetic risk variants, each with a relatively small effect on risk.

**Candidate Genes**

When the pathophysiology of a disease is known (e.g., an enzyme deficiency), it may be straightforward to define candidate genes and to determine which DNA sequence variants predict who becomes ill. For psychiatric disorders, pathophysilogies are unknown. Most candidate gene hypotheses are based on the effects of psychiatric medications on monoamine neurotransmission, focusing particularly on several functional polymorphisms in dopaminergic or serotonergic pathways (i.e., sequence variants that alter relevant receptor proteins or enzymes) (3, 4). None has been shown to be associated with a psychiatric disorder with a level of significance that would lead to general acceptance of a finding.

**Positional Methods**

The alternative strategy is to localize disease-related sequence variation based entirely on its location or position in the genome. Before GWAS, available methods included the genomewide linkage study and linkage disequilibrium mapping (of which GWAS is a large-scale example). (See Table 1 for definitions and Table 2 for a timeline of critical developments.)

Genomewide linkage studies became feasible in the 1980s, with genomewide "maps" (7) of hundreds of DNA sequence variations (markers). Linkage analysis (reviewed in [15]), of families with multiple ill members, exploits within-family correlations between illness and the alternative sequences (alleles) of the markers that are closest to the disease-related gene(s). Linkage studies led to the discovery of mutations (mostly rare dominant or recessive) for more than 1,600 diseases (see the Online Mendelian Inheritance in Man [http://www.ncbi.nlm.nih.gov/Omim/mimstats.html]). They have been less successful for complex (multifactorial/multigenic) disorders. In psychiatric linkage studies (catalogued online [https://slep.unc.edu]), small samples of pedigrees were initially examined in the hope of discovering simpler genetic mechanisms that would provide clues to pathophysiology. Then, larger studies (involving hundreds of families) searched for genes with smaller effects. There are diverse opinions regarding the past success and future prospects of these studies. Statistically significant linkages have been reported but have been difficult to replicate, presumably because linkage is much less powerful when risk variants have small effects and there is heterogeneity in the underlying genetic factors in different families. Meta-analyses have supported linkage for some disorders (16–18).

GWAS relies instead on population-wide correlation between two sequence variants. Most variants are single nucleotide polymorphisms (SNPs) (almost always just two alternative nucleic acids at a genomic position). SNP variants that are reasonably common are mutations that occurred thousands of generations ago and then spread due to chance or natural selection. When a second SNP mutation occurred very close to an earlier one (separated by up to tens of thousands of base pairs), then both variant alleles are almost always transmitted to the same offspring in subsequent generations. Link-
age disequilibrium is this nonrandom association of two alleles. Approximately 20 years ago, it was proposed that linkage disequilibrium could be exploited to “map” or identify disease genes, such as in linkage candidate regions (or in recently isolated populations in which linkage disequilibrium spans long distances) (19). If one SNP increases the risk of a common disease, then there will be a statistical association in the population between disease and that SNP (direct association) and several nearby SNPs (indirect association due to linkage disequilibrium).

Linkage disequilibrium mapping studies have identified plausible positional candidate genes in regions of linkage or of cytogenetic abnormalities associated with psychiatric disorders, and these genes have suggested new mechanistic hypotheses (20). For example, as of April 2008, there were 1,291 published studies of 690 schizophrenia candidate genes (see http://www.schizophreniaforum.org/res/sczgene/default.asp). A recent meta-analysis of these studies (3) identified four “strong” psychiatric candidate gene associations based on epidemiological criteria for meta-analysis but not at what is currently understood to be a genomewide level of statistical significance.

**Common SNPs, HapMap, and GWAS**

Risch and Merikangas (21) noted that small genetic effects could be detected with greater power by association analyses and proposed that genomewide linkage disequilibrium mapping (i.e., GWAS) could be applied if technologies were developed to study SNP frequencies in all genes, contrasting ill case subjects versus comparison subjects or case subjects and their parents (associated alleles are transmitted to ill offspring more often than expected by chance). Lander (9) proposed the common-disease common-variant hypothesis. Comparing any two people, most sequence differences are ancient, “common” SNPs (by convention, varying on at least 5% of chromosomes in a population), which Lander argued must confer at least some (not all) of the genetic risk for common diseases. He proposed cataloguing them and studying their association with disease in large samples. SNPs become common because they are neutral or favorable with respect to survival (e.g., evolutionary pressures can rapidly increase frequencies of adaptive SNPs in gene regulating regions). However, some have mildly harmful effects, perhaps depending on environmental conditions (e.g., preserving fat during an ice age but leading to obesity in the fast food era). The common-disease common-variant GWAS strategy assumed that many different common SNPs have small effects on each disease and that some could be found by testing enough SNPs in enough people.

How many SNPs should be tested? Studies of small regions revealed linkage disequilibrium blocks in which

### TABLE 2. Timeline of Positional Genetic Methods From Linkage To Genomewide Association Studies

<table>
<thead>
<tr>
<th>Year</th>
<th>Development</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>1980</td>
<td>Proposal to create a genomewide map of DNA markers for human linkage analysis (5).</td>
<td>Following the discovery of restriction fragment length polymorphism markers, it was proposed that once restriction fragment length polymorphisms throughout the genome were available, it would be possible to search any genomic region, or the entire human genome, for evidence of genetic linkage.</td>
</tr>
<tr>
<td>1983</td>
<td>Linkage mapping and identification of the Huntington’s disease gene (6).</td>
<td>The first of the many Mendelian disorders for which genetic linkage was detected followed by identification of specific disease mutations in the linkage region.</td>
</tr>
<tr>
<td>1987</td>
<td>First human linkage map (7).</td>
<td>The first genomewide map of approximately 400 restriction fragment length polymorphisms were supplanted by short tandem repeat markers and then SNPs.</td>
</tr>
<tr>
<td>1993</td>
<td>First genomewide linkage study of a psychiatric disorder (8).</td>
<td>Psychiatric genomewide linkage studies (catalogued online [<a href="https://slep.unc.edu">https://slep.unc.edu</a>]) produced some convergent linkage evidence, but no definitive evidence for susceptibility genes.</td>
</tr>
<tr>
<td>1996</td>
<td>Common-disease common-variant hypothesis (9).</td>
<td>The HapMap Project grew out of the need to develop a dense set of genetic markers to test this hypothesis.</td>
</tr>
<tr>
<td>2001</td>
<td>Draft of the complete human genome sequence (10).</td>
<td>The genome sequence set the stage for all future progress. It stimulated critical advances in genomic sequencing technology and set a new standard of immediate public release of government-supported genomic research data.</td>
</tr>
<tr>
<td>2002–2007</td>
<td>International HapMap Project (11, 12) (<a href="http://www.hapmap.org">www.hapmap.org</a>).</td>
<td>The project discovered and genotyped (in 270 individuals from three populations) 1.3 million SNPs in Phase I plus 2.1 million in Phase II—approximately 25% to 35% of common SNPs in these populations, providing good genomewide coverage. It spurred advances in SNP assays, making genomewide association studies possible. “HapMap III” provided genotypes in an expanded data set for the Illumina 1M and Affymetrix 6.0 (900K) SNP sets.</td>
</tr>
<tr>
<td>2002</td>
<td>First published genomewide association study (13).</td>
<td>This study of myocardial infarction used few SNPs (65,761) and cases (94) by current standards.</td>
</tr>
<tr>
<td>2005–2007</td>
<td>Availability of high-throughput array-based SNP assays.</td>
<td>Affymetrix and Illumina arrays became available, initially with approximately 100,000 SNPs, and currently with up to approximately 1 million SNPs per array plus additional probes for analysis of copy number. These have made it possible to carry out genomewide association studies for many diseases and samples.</td>
</tr>
<tr>
<td>2005</td>
<td>First year with multiple genomewide association study publications.</td>
<td>The first small studies using denser SNP sets produced strong associations for macular degeneration (1) and Crohn’s disease (14), demonstrating the feasibility and power of genomewide association studies.</td>
</tr>
<tr>
<td>2007</td>
<td>Initiation of the 1,000 Genomes Project (<a href="http://www.1000genomes.org">www.1000genomes.org</a>).</td>
<td>This project aims to extend the HapMap to all SNPs with 1% frequency in diverse populations, functional SNPs of lower frequencies, and sequence-level data on structural variants, utilizing multiple high-throughput sequencing technologies.</td>
</tr>
</tbody>
</table>
common SNPs are highly correlated (usually <10,000–30,000 base pairs in African populations or 30,000–50,000 base pairs in the newer European and Asian populations) (22). This motivated the HapMap Project (www.hapmap.org [12]), which has validated approximately 4 million SNPs, including 2.8 million of the estimated 10 million common SNPs in major world populations, while creating competition among biotechnology companies to develop high-throughput genotyping technologies. Sequencing and genotyping studies showed that sets of 500,000 (European populations) to 1,000,000 (African populations) SNPs could “tag” (serve as proxies for) approximately 80% of common SNPs (23). Over the last 3 years, the Affymetrix and Illumina companies have developed “chips” (arrays of assays on glass slides) that assay large SNP sets with high accuracy (0%–2% missing data; <0.5% errors), at low cost (approximately $500 [U.S.] per subject; an approximate 2,000-fold reduction in cost per genotype in 10 years), and rapidly (>1,000 DNA specimens per week in some labs). Thus, the GWAS era has arrived.

Rare SNPs

Common SNPs are unlikely to explain all of the genetic risk for common disorders. An evolutionary model of complex diseases (24) predicts roles for common SNPs and for multiple rare variants (such as SNPs) in some genes (multiple rare variant hypothesis). A rare variant is usually defined by a frequency <1%, although many are so rare that they are found in only one individual in a sample (25). Most variants carried by any one person are common, but if one sequences a chromosomal region in many people, one finds more rare SNP sites. The most deleterious variants die out or remain rare due to natural selection (i.e., they reduce survival). They are found in functional regions (i.e., among the SNPs in exons [protein coding regions]) that alter amino acid sequence (nonsynonymous SNPs) or in promoters (sequences that regulate gene expression) (26, 27). However, there are other poorly understood functional regions. Many noncoding regions are highly conserved across species, suggesting that they have a function. Gene expression can be altered by common, synonymous exonic SNPs (no coding change) and by SNPs in introns (noncoding gene segments) (28). Indeed, most genomic DNA is apparently transcribed into RNA and thus could have unknown regulatory functions (29). Most rare SNP associations will be missed by current GWAS methods, but it is expected that the 1,000 Genomes Project (www.1000genomes.org) will discover most SNPs with 1%–5% frequencies, which would permit an extension of systematic GWAS methods to these less common SNPs. Linkage could detect a locus with rare pathogenic variants in many families.

Rare SNP associations are more likely to be detected by resequencing of relevant regions in hundreds or thousands of individuals. (By convention, resequencing, which is sometimes now referred to as “medical sequencing,” determines an individual’s DNA sequence versus sequencing of an organism’s genome.) Botstein and Risch (30) encouraged the systematic study of nonsynonymous SNPs in common diseases. Multiple rare pathogenic variants have been discovered by resequencing genes influencing lipid metabolism (31) and hypertension (32) and also genes for which GWAS have already detected common-SNP associations (33–35). It is anticipated that advances in resequencing technologies will make it feasible to search systematically for rare variant effects in parts of the genome (e.g., linkage regions, all exons, all promoters) and eventually genomewide.

Copy Number Variants

GWAS technologies can also detect more of the copy number variants in the genome than was possible with older cytogenetic methods, by analysis of the relative intensities of the fluorescent labels used in the assays. Copy number variants are deletions and duplications of DNA segments of diverse sizes and population frequencies. For example, large deletions on chromosome 22q11 cause velocardiofacial/DiGeorge syndrome, and 20% of patients with this syndrome also develop schizophrenia (36). Copy number variants tend to arise in regions with repetitive DNA sequences. Some copy number variants are common and transmitted from generation to generation, while others recurrently arise de novo. Similar to rare SNPs, rare copy number variants are more likely to be harmful. (Other structural variants, such as inversions and translocations, remain difficult to detect.) Large genomewide copy number variant scans show that copy number variants are more common than previously recognized (37).

Structural variation has not been as comprehensively studied as SNPs because copy number variant detection is less accurate, biological confirmation remains costly, and smaller copy number variants (<100,000 base pairs) are less reliably detected. However, technologies are rapidly improving. Significant copy number variant findings are now being reported for psychiatric disorders.

GWAS Study Design

Study design issues are summarized in Table 3. A GWAS sample, selected based on a well-defined heritable phenotype, might include case subjects (ill) and comparison subjects, subjects with a range of values for a continuous phenotypic variable, or probands and both of their parents (trios) or other constellations of relatives. Samples are often limited to a single ancestry (European, Asian, etc.) because some SNPs have markedly different frequencies across populations (and some are not observed in every population) so that some associations can best be detected in homogeneous samples. Each subject is genotyped using a GWAS SNP array. Extensive “quality control” (data cleaning) is required to detect problems that can result in false negative or false positive findings, such as SNPs and DNA specimens that give poor quality results or...
TABLE 3. Genomewide Association Study Design Issues and Requirements

<table>
<thead>
<tr>
<th>Issue</th>
<th>Requirement</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phenotype</td>
<td>Well-defined, adequately heritable disorder (e.g., schizophrenia) or trait (e.g., high-density cholesterol level or neuroticism score).</td>
<td>Power depends on the frequency and effect size for individual variants, not overall heritability. Case/comparison subjects have more power per subject but are prone to mismatch biases (e.g., ancestry). For more common disorders, comparison subjects with the disorder may be excluded to avoid false negative results (40). Samples up to tens of thousands of subjects have proven useful, but some common risk variants cannot feasibly be detected.</td>
</tr>
<tr>
<td>Sample type</td>
<td>Ill case subjects and comparison subjects or subjects with a range of trait scores (e.g., highest and lowest) or case subjects and their parents or other relatives.</td>
<td></td>
</tr>
<tr>
<td>Comparison subjects</td>
<td>Match for ancestry, other relevant attributes (e.g., age for an Alzheimer’s disease study) or environmental exposures (e.g., “ever smoked” for a study of nicotine dependence).</td>
<td>For more common disorders, comparison subjects with the disorder may be excluded to avoid false negative results (40). Samples up to tens of thousands of subjects have proven useful, but some common risk variants cannot feasibly be detected.</td>
</tr>
<tr>
<td>Sample size</td>
<td>Depends on the actual frequency and genetic effect of risk variants in the sample.</td>
<td></td>
</tr>
<tr>
<td>SNPs</td>
<td>300,000 to 1,000,000 common SNPs, depending on ancestry of the sample.</td>
<td></td>
</tr>
<tr>
<td>Multiple testing</td>
<td>P value correction for multiple, partially correlated genotyped SNPs, plus imputed data for all HapMap SNPs to permit cross-study comparison and meta-analysis (40, 41).</td>
<td>Match case/comparison subjects for ancestry; apply statistical correction for population differences (38). Requires powerful computers or computer clusters and software (76). Without adequate quality control, spurious highly “significant” findings are common.</td>
</tr>
<tr>
<td>Population substructure</td>
<td>World populations differ in frequencies of many SNPs. Case subject-comparison subject ancestry differences can create false positive and negative results.</td>
<td></td>
</tr>
<tr>
<td>Data management</td>
<td>Billions of data points to manage.</td>
<td></td>
</tr>
<tr>
<td>Quality control</td>
<td>Extensive quality control analyses are required to exclude poorly performing SNPs and DNA specimens, identify duplicate or closely related specimens, and more subtle assay and sample problems.</td>
<td>Copy number variant detection is less specific, sensitive, or accurate than SNP genotype detection. Biological confirmation is needed.</td>
</tr>
<tr>
<td>Detection of copy number variants</td>
<td>Computational methods to detect copy number change from intensities of fluorescent labels in assays; additional nonpolymorphic assays can be added to improve copy number variant detection.</td>
<td>Biological confirmation is needed.</td>
</tr>
</tbody>
</table>

unexpected relatedness among subjects. Case-control differences in ancestry (“population substructure”) can also confound association test results, but this can be corrected statistically based on correlations among SNP genotypes that reflect ancestry (38). Most studies then test each SNP for association of genotypes to the phenotype and impute the genotypes of other HapMap SNPs based on the correlations among SNPs in HapMap data (39–41).

Selection of comparison groups is critical beyond the problem of ancestral matching. It is ideal to recruit case subjects and comparison subjects systematically from the same population. This is not always feasible for very large samples of a clinically severe disorder, but comparison subjects must be sufficiently comparable with case subjects to avoid systematic biases. Depending on the phenotype, it might be important to match for variables such as age (e.g., for an Alzheimer’s disease study) or sex. Information about known gene-environment interactions should be considered (e.g., in studies of substance dependence, comparison subjects are usually selected who have used the substance but did not become dependent). When the phenotype is relatively uncommon (e.g., 5% prevalence), little power is lost by studying comparison subjects without clinical screening, but for more common disorders, power is increased if ill individuals are excluded from the comparison group (40). It is reassuring that in the United Kingdom Wellcome Trust Case Control Consortium GWAS of seven common diseases, robust results were obtained when association was tested using comparison groups recruited from blood donors or from a population-based birth cohort.

Statistical Power of GWAS

A key factor in the recent success of GWAS has been the assembling of large samples with adequate statistical power to detect small effects of common SNPs on disease risks. Figure 1 illustrates the reasons for this. Large genotypic relative risks (e.g., 5- to 10-fold increase in risk to carriers) would have produced large linkage signals. Early GWAS analyses of several hundred case subjects were powered to search for risk alleles with genotypic relative risks above 2. Only a few such effects were detected (1). The more typical GWAS has included 1,000 to 2,000 case subjects plus a similar number of comparison subjects, with power to detect risk alleles that are reasonably common and have genotypic relative risks of 1.5 to 2. The small number of robust findings suggested the need to detect smaller genotypic relative risks (2).

This led to much larger GWAS analyses in collaborative samples, which have proven to be remarkably successful for many diseases. Most of the new, highly significant findings have been for alleles with genotypic relative risks of 1.1 to 1.4 (mostly between 1.12 and 1.20). In this range (Figure 1), good or excellent power requires samples of 8,000 to 20,000 case subjects (plus comparison subjects), depending on genotypic relative risk and allele frequency (i.e., larger than any sample collected by a single research group to date).

GWAS Findings for Nonpsychiatric Disorders and Lessons for Psychiatry

Over the past 3 years, many highly significant GWAS findings have been reported for nonpsychiatric disorders. Table 4 summarizes a systematic listing of GWAS findings.
provided by the National Institute for Human Genome Research (http://www.genome.gov/GWAstudies/ [accessed Nov. 2008, refs. 42–44]), restricted to findings with p values \( \leq 5 \times 10^{-8} \). There are 200 distinct findings listed for 59 disorders or traits. Some may be false positives due to chance (every p value is an estimate of the probability of a false positive result) or to technical problems such as genotyping or analytical errors. However, many of these findings have already been replicated in independent samples, and most robust p values replicate. These results far exceed all previous robust associations for complex disorders. This confirms that common SNPs explain part of the genetic risk for these disorders, as predicted by the common-disease common-variant hypothesis. There are almost certainly also many common SNPs with smaller effects on risk as well as rare and very rare SNPs and copy number variants with diverse effect sizes.

**Sample size.** Most initial GWAS samples included 500 to 3,000 case subjects (plus comparison subjects) or as many as 10,657 subjects for a continuous trait. One or more replication samples were usually then studied via collaboration, totaling 2,000 to 8,000 subjects (case and comparison subjects or case subjects and family members). For studies with at least 1,000 case subjects, most findings involved common alleles (20%-80%), with odds ratios (estimates of genotypic relative risk) between 1.1 and 1.4 (i.e., the range within which there was some power).

Findings for type 2 diabetes illustrate the importance of sample size. In late 2007, there were 11 strong candidate genes. Of these, six were discovered by GWAS, four were identified based on mechanistic hypotheses, and one, TCF7L2, was identified by linkage disequilibrium mapping of a linkage region, although TCF7L2 SNPs did not explain the linkage (47). The TCF7L2 gene has an overall odds ratio of 1.37. It was detected by most (not all) studies. Other type 2 diabetes loci have allelic odds ratios between 1.1 and 1.2, requiring from 10,000 to well over 20,000 total subjects for 80% power. Each locus was missed by most single studies. For example, in the Wellcome Trust Case Control Consortium study (2,000 case subjects; 3,000 comparison subjects), these 11 SNPs were ranked from 2 to 26,017 in their strength of association (47). Zeggini et al. (48) combined more than 60,000 subjects to study type 2 diabetes findings that did not quite reach genomewide significance previously. Six SNPs (implicating eight different genes) have since achieved p values \( \leq 5 \times 10^{-8} \), with odds ratios from 1.09 to 1.15.

**Novel etiological hypotheses.** Most findings have implicated novel genes or regions and suggested new mechanisms. For example, SNPs in FTO ("fat mass and obesity associated" gene) are strongly associated with common obesity (49, 50). This was surprising because FTO knockout mice are not obese. Mechanisms are under study, including a role in adipocyte lipolysis (51). As noted by Todd (52), implicating a gene in disease requires both compel-
epigenetic effects. The results suggest that the largest pos-
ter interactions that have not been adequately modeled, and
to rare SNPs, copy number variants, other unsuspected
sampling problems). The remaining variance may be due
to other common SNPs (those that have smaller effects
than can be detected with current sample sizes or are not
characterized by the arrays or are missed because of technical or
false discovery rate (45) or Bayes Factor (41, 46) criteria. Shown for each category is the number of distinct findings (defined as one or more
SNPs in a single chromosomal band for a specific disease or trait) with a p value \( \leq 5 \times 10^{-8} \), counting only once those findings reported more
than once. Of the 200 findings, 95 had p values \( < 10^{-12} \), and 58 had p values \( < 10^{-15} \). Some SNPs or regions have produced findings for different
orders or traits (see article). In many cases, there are additional studies or meta-analyses (not included in this tabulation of genome-
wide association study reports) that contain additional findings or updated significance levels. “Physical traits” includes nondisease traits such as hair and eye color and height. “Plasma values” includes studies of potentially disease-related values (other than lipids) such as C-reactive
protein, glucose, and IgE. “Other” includes studies related to eye, skin, or pulmonary diseases; obesity-related traits; aging; and other traits.

Data are summarized online (http://www.genome.gov/GWAstudies/ [accessed November 2008]).

<table>
<thead>
<tr>
<th>Type of Disease or Trait</th>
<th>Unique Findings With ( p \leq 5 \times 10^{-8} )</th>
<th>Number of Disorders or Traits</th>
</tr>
</thead>
<tbody>
<tr>
<td>Autoimmune</td>
<td>12</td>
<td>3</td>
</tr>
<tr>
<td>Bone density</td>
<td>10</td>
<td>1</td>
</tr>
<tr>
<td>Cancer</td>
<td>37</td>
<td>8</td>
</tr>
<tr>
<td>Cardiovascular</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>Diabetes type 1</td>
<td>10</td>
<td>1</td>
</tr>
<tr>
<td>Diabetes type 2</td>
<td>10</td>
<td>1</td>
</tr>
<tr>
<td>Gastrointestinal</td>
<td>25</td>
<td>5</td>
</tr>
<tr>
<td>Lipid levels</td>
<td>13</td>
<td>3</td>
</tr>
<tr>
<td>Neurological</td>
<td>9</td>
<td>6</td>
</tr>
<tr>
<td>Physical traits</td>
<td>28</td>
<td>7</td>
</tr>
<tr>
<td>Plasma values</td>
<td>22</td>
<td>10</td>
</tr>
<tr>
<td>Other</td>
<td>19</td>
<td>10</td>
</tr>
<tr>
<td>Total</td>
<td>200</td>
<td>59</td>
</tr>
</tbody>
</table>

\( a \) There is no definitive p value threshold that predicts true positive genomewide association study findings. Interpretation rests on consistency
of replication and/or meta-analysis of cumulative data. A p value threshold of \( 5 \times 10^{-8} \) has been used throughout this review, based on three
estimates that assumed that all common SNPs have been tested (42–44), but other thresholds can be defended. Other approaches include

Insights into phenotypes. FTO also exemplifies the importance of phenotypic variables. Type 2 diabetes is common in obese individuals. FTO SNPs are associated with type 2 diabetes, but this is due to the association between type 2 diabetes and body mass index (50). The association of FTO with type 2 diabetes is not found when type 2 diabetes case and comparison subjects are matched for body
mass index (53). Surprising relationships among phenotypes have also been discovered. For example, SNPs on chromosome 8q24.21 are associated with prostate, breast, and colorectal cancers, which were not previously thought to be genetically related (54). The region contains no known genes, and thus it would have been ignored without a GWAS strategy. It is now being intensively studied.

Thus, GWAS have been remarkably successful for many common diseases. Large multicenter samples have usually been required, and larger samples have detected more associations. Only a small part of the genetic risk for any one disease has been explained, but these discoveries have suggested new disease mechanisms and targets for therapy and prevention, although direct therapeutic applications will require substantial additional effort to characterize the biological mechanisms and develop new treatments. Some of the unexplained variance is likely due to other common SNPs (those that have smaller effects than can be detected with current sample sizes or are not tagged by the arrays or are missed because of technical or sampling problems). The remaining variance may be due to rare SNPs, copy number variants, other unsuspected genomic mechanisms, gene-gene or gene-environment interactions that have not been adequately modeled, and epigenetic effects. The results suggest that the largest pos-
sible samples should be studied by GWAS for each of the major psychiatric disorders to test the hypothesis that common SNPs or detectable copy number variants are involved in etiology. Positive findings could lead to important etiological discoveries.

GWAS of Psychiatric Disorders

GWAS findings are now emerging for psychiatric disorders (Table 5). The early findings include replicated copy number variant associations for schizophrenia and autism, a genomewide significant association for bipolar disorder that emerged when several data sets were combined, and a significant association in a combined schizophrenia-bipolar data set.

For schizophrenia, four genomewide studies of copy number variants (55–58) have produced two types of replicated findings. First, two large studies (55, 56) found two rare deletions that are significantly associated with schizophrenia on chromosomes 1q21.1 (0.2% of case subjects) and 15q13.3 (0.3% of case subjects). The case:comparison subjects ratio (approximately 10) suggests major effects on risk, but it is unknown which deleted genes or sequences are responsible or whether they account for all of the subjects’ genetic risks. These deletions are also found (but probably less frequently) in individuals with mental retardation and/or autism and are typically de novo (not inherited from parents) (55). The well-known chromosome 22q11 deletions were also significantly associated with schizophrenia (0.2%–0.4% of case subjects across studies versus 0% of comparison subjects). Second, the three studies (56–58) that tested such a hypothesis showed that schizophrenia case subjects have a small but significant increase in their total genomewide count of rare long copy number variants, suggesting that there are other patho-
<table>
<thead>
<tr>
<th>Item</th>
<th>Disorder</th>
<th>Initial Sample (case subjects/comparison subjects)</th>
<th>Other Data</th>
<th>Genomewide Significant Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Studies of association to SNP genotypes (individual genotyping)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wellcome Trust Case Control Consortium (41)</td>
<td>Bipolar disorder</td>
<td>1,868/2,938 (United Kingdom)</td>
<td>Replication sample (for best results): 409 U.S. trios; 365 case subjects/351 comparison subjects (Scottish)</td>
<td>p=9.1×10&lt;sup&gt;−9&lt;/sup&gt;</td>
</tr>
<tr>
<td>Sklar et al. (59)</td>
<td>Bipolar disorder</td>
<td>1,461/2,008 (United States, United Kingdom; Systematic Treatment Enhancement Program for Bipolar Disorder (STEP-BD)-University College of London)</td>
<td>Wellcome Trust Case Control Consortium (41) plus Sklar (59) (see numbers above) plus ED-DUB-STEP2 (1,098 case subjects/1,267 comparison subjects)</td>
<td>p=9.96×10&lt;sup&gt;−9&lt;/sup&gt;; ZNF804A gene (odds ratio=1.12; frequency of the associated allele in comparison subjects=0.59)</td>
</tr>
<tr>
<td>Ferreira et al. (60)</td>
<td>Bipolar disorder</td>
<td>4,387/6,209</td>
<td>With bipolar disorder included:</td>
<td>p=0.0008; increased novel copy number variants in case subjects (15%) versus comparison subjects (5%) (p=0.03 in child-onset schizophrenia)</td>
</tr>
<tr>
<td>Lencz et al. (61)</td>
<td>Schizophrenia</td>
<td>178/144 (United States)</td>
<td></td>
<td>p=0.0008; increased noninherited copy number variants in sporadic case subjects (9.9%) versus comparison subjects (1.26%)</td>
</tr>
<tr>
<td>Sullivan et al. (62)</td>
<td>Schizophrenia</td>
<td>738/733 (United States)</td>
<td></td>
<td>p=3×10&lt;sup&gt;−5&lt;/sup&gt;; increased copy number variants (&lt;1% frequency; &gt;100Kb) in case subjects (1.14% per subject) versus comparison subjects (0.99). Genomewide significant evidence for association of copy number variants on chromosome 1q21.1, 22q11.2, 15q13.3.</td>
</tr>
<tr>
<td>O’Donovan et al. (63)</td>
<td>Schizophrenia</td>
<td>479/2,937 (United Kingdom) plus 1,865 Wellcome Trust Case Control Consortium bipolar disorder case subjects</td>
<td></td>
<td>p=0.044 (uncorrected); increased 16p11.2 deletions in case subjects (0.6%) versus comparison subjects (0%)</td>
</tr>
<tr>
<td><strong>Studies of association to copy number variants</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Walsh et al. (57)</td>
<td>Schizophrenia</td>
<td>150/268</td>
<td>Replication sample (for best results): 83 childhood-onset schizophrenia patients plus parents</td>
<td>p=0.0008; increased copy number variants in case subjects versus comparison subjects (9.9%) versus comparison subjects (1.26%)</td>
</tr>
<tr>
<td>Xu et al. (58)</td>
<td>Schizophrenia</td>
<td>152/159</td>
<td>Sporadic case subjects</td>
<td>p=3×10&lt;sup&gt;−5&lt;/sup&gt;; increased copy number variants (&lt;1% frequency; &gt;100Kb) in case subjects (1.14% per subject) versus comparison subjects (0.99). Genomewide significant evidence for association of copy number variants on chromosome 1q21.1, 22q11.2, 15q13.3.</td>
</tr>
<tr>
<td>Stone et al. (56)</td>
<td>Schizophrenia</td>
<td>3,381/3,191</td>
<td></td>
<td>Genomewide significant evidence for association of copy number variants on chromosome 1q21.1, 22q11.2, 15q13.3.</td>
</tr>
<tr>
<td>Sebat et al. (64)</td>
<td>Autism</td>
<td>118 (sporadic)/196</td>
<td>Some comparison subjects from autism families; some Autism Genetics Resource Exchange families</td>
<td>Genomewide significant evidence for association of increased de novo copy number variants in case subjects (10%) versus comparison subjects (1%); (note: &gt;1 comparison subject per family)</td>
</tr>
<tr>
<td>Kumar 2008 (65)</td>
<td>Autism</td>
<td>180/372</td>
<td>Replication sample (for best results): 532 case subjects/465 comparison subjects</td>
<td>p=0.044 (uncorrected); increased 16p11.2 deletions in case subjects (0.6%) versus comparison subjects (0%)</td>
</tr>
<tr>
<td>Marshall et al. (66)</td>
<td>Autism</td>
<td>427/500</td>
<td>Replication sample (for best results): 1,152 additional comparison subjects</td>
<td>Genomewide significant evidence for association of increased de novo copy number variants in case subjects (7%) versus comparison subjects (1%). Increased 16p11.2 deletions in case subjects (approximately 1%) versus comparison subjects (0%); (p=0.002)</td>
</tr>
<tr>
<td>Weiss et al. (67)</td>
<td>Autism</td>
<td>751 multiplex Autism Genetics Resource Exchange families (1,441 case subjects) plus 2,814 comparison subjects</td>
<td>Replication samples (for best results): 512 case subjects/434 comparison subjects; 299 case subjects/18,834 comparison subjects</td>
<td>Increased 16p11.2 copy number variants in case subjects (1.1%) versus comparison subjects (0.05%); significant in all three samples</td>
</tr>
<tr>
<td>Christian et al. (68)</td>
<td>Autism</td>
<td>397/372</td>
<td>Case subjects from Autism Genetics Resource Exchange families</td>
<td>11.6% of case subjects had a copy number variant unique to case subjects</td>
</tr>
</tbody>
</table>

<sup>a</sup>The copy number variant studies of Sebat et al., Weiss et al., and Christian et al. all used some families from the Autism Genetics Resource Exchange repository and thus are not entirely independent. SNP studies all used SNP arrays with 500,000 SNPs (Affymetrix 500K or 5.0) or 900,000 SNPs (Affymetrix 6.0). Some studies used more than one type of array. Copy number variant studies used array-based comparative genomic hybridization and/or genomewide association study SNP arrays (Affymetrix or Illumina), with additional confirmation of some or all results using additional methods (e.g., quantitative polymerase chain reaction, karyotyping, and other methods). Studies using pooled genotyping are not included but are cited in the article text.
genic copy number variants that are too rare to detect singly. Other GWAS studies of psychiatric disorders are summarized in Table 5.

At the time of this writing, there have been three published small schizophrenia GWAS analyses using individual genotyping (61–63), with samples that included 178 to 738 case subjects. Two additional studies (69, 70) used pooled genotyping. To date, no genomewide significant finding has emerged for schizophrenia alone. However, when the 12 “best” SNPs from a GWAS of 479 case subjects and 2,937 Wellcome Trust Case Control Consortium comparison subjects were genotyped in an additional 7,308 schizophrenia case subjects and 12,834 comparison subjects, and the 1,868 Wellcome Trust Case Control Consortium bipolar disorder case subjects were added to the analysis, a genomewide significant p value was seen for a SNP in a gene of unknown function (zinc finger protein 804A [ZNF804A]) (63). This will require replication for these disorders, both separately and combined. It illustrates the potential importance of cross-diagnosis analyses, although such analyses will increase the problem of multiple testing and thus require very large samples for confirmation.

For autism, three studies (65–67) have reported association with a rare (1% of case subjects), large, high-penetrance deletion on chromosome 16p11.2. There is also support for the hypothesis that there is an excess of rare, mostly de novo, copy number variants in approximately 10% of case subjects, although the role of these rare copy number variants in autism remains to be proven (64, 65, 68). Autism GWAS analyses of common SNPs have yet to be reported.

For bipolar disorder, three individual studies (41, 59, 60), of 1,000 to 2,000 case subjects each, failed to detect significant association, but the three data sets combined produced a p value of 9.1×10^−8 for ankyrin-G (ANK3), the product of which links membrane proteins, such as voltage-dependent sodium channels, to the axonal cytoskeleton. The larger analysis did not replicate a significant association (for DGKH [diacylglycerol kinase, eta]) that had been reported in a smaller study using pooled genotyping (71).

Among the reports that will be available in the near future are the four psychiatric GWAS reports, supported by the Genetic Association Information Network (fnih.org), on schizophrenia, bipolar disorder, major depression, and ADHD. Details and preliminary results are available online (http://www.ncbi.nlm.nih.gov/sites/entrez?db=gap). We are not permitted to summarize these analyses pending the initial publications by the investigators. The Genetic Association Information Network is an example of a new emphasis on rapid public sharing of genetic data to accelerate the process of discovery.

**The Psychiatric GWAS Consortium**

The first set of psychiatric GWAS analyses have demonstrated that this methodology can work for psychiatric disorders. The pattern observed in the bipolar disorder studies is particularly encouraging because it is consistent with what has happened for nonpsychiatric diseases: combining several smaller samples produced a significant result as well as several other findings with modestly significant p values in each individual study, which could prove to be significant as more data become available (60).

These results support our expectation that multiple definitive association findings will be detected for many psychiatric disorders, often requiring large samples. We therefore organized the Psychiatric GWAS Consortium, which includes almost all known GWAS studies to date for schizophrenia, bipolar disorder, major depressive disorder, ADHD, and autism (Table 6), contributed by 121 investigators from 61 institutions at the time of this writing (as listed in the Acknowledgments). The Psychiatric GWAS Consortium has the following three specific objectives:

1. **Within-disorder meta-analyses of all available GWAS data.** These diagnoses are based on definitions that produced maximum heritability estimates in genetic epidemiological studies. Thus, disorder-specific analyses represent our strongest hypotheses.

2. **Cross-disorder analyses, including analyses of combinations of disorders and of phenotypes observed in two or more disorders (such as depression or psychosis) based on the recommendations of an expert committee. Because data are insufficient to determine what common cross-disorder etiological factors might exist, alternative phenotypes should be explored. GWAS analyses have produced surprising cross-disorder associations, such as those found for cancers (54) and inflammatory bowel diseases (86), which could also exist for psychiatric disorders given the many common symptoms.**

3. **Analyses of comorbidities such as alcohol, nicotine, and illicit drug use disorders, which disorders can be studied across multiple case groups.**

Additional exploratory analyses will be carried out by analysts from participating research groups, generating new hypotheses that can be tested as more samples become available. All GWAS data used by the Psychiatric GWAS Consortium will become available to the scientific community through data repositories (unless prohibited by the original consents or institutional review board decisions).

A central analytical team in consultation with participating analysts will conduct uniform quality control analyses and imputation of nongenotyped HapMap SNPs (to permit combining of data). The disorder-specific workgroups will design their own primary meta-analyses, with additional workgroups to define other phenotypic and cross-disorder analyses. Analyses will account for ethnic substructure within samples and appropriate pairing of case and comparison groups.

Depending on the genetic architecture of each disorder, one or more primary analyses could have sufficient power to detect genomewide significant evidence for associa-
For example, the largest analyses, with approximately 10,000 case subjects and 10,000 comparison subjects, would have 80% power to detect a SNP with a genotypic relative risk of 1.152 and p value <5x10^-8—assuming direct association with an allele with a frequency of 0.25 and log-additive inheritance—or 57% power for indirect association with an r^2 value of 0.8. Power would be reduced for smaller samples, less common alleles, and recessive effects. If there are many risk alleles in the genome with a sufficient effect size, there would be substantial power to detect at least one. Meta-analyses will be completed in 2009. Updated results will be made available on the Psychiatric GWAS Consortium website (http://pgc.unc.edu).

**Discussion**

There is a compelling rationale for applying GWAS methods to very large samples for major psychiatric disorders. Given that the pathophysiology of these disorders is unknown, genomewide studies provide an unbiased way to search the genome for causative factors. Many successful GWAS analyses have combined data from diverse clinical samples and SNP arrays to obtain replicable findings that point to new hypotheses about disease mechanisms and treatment targets. The first significant psychiatric GWAS findings have been reported (Table 5), using large collaborative samples. It is hoped that meta-analyses can produce multiple robust findings for psychiatric disorders.

GWAS SNP arrays “cover” ≥80% of common HapMap SNPs, and regional resequencing data suggest that most unknown common SNPs are also being tested indirectly. Within these limitations, GWAS methods test the common-disease common-variant hypothesis. Copy number variants are also detected but less systematically or accurately. The Psychiatric GWAS Consortium meta-analyses will have reasonable power to detect common SNP associations for each disorder within the limitations illustrated in Figure 1. However, it is possible that very few significant associations might be detected for some disorders or none. How far should we go with GWAS?

Past experience suggests that for some disorders as many as 20,000 to 30,000 case subjects and a similar number of comparison subjects (or case subjects plus their parents) could be required to obtain highly robust findings. More data sets will be genotyped in the near future, and the National Institute of Mental Health plans to collect additional large schizophrenia and bipolar disorder samples (http://grants.nih.gov/grants/guide/rfa-files/ RFA-MH-08-131.html). This raises important questions regarding resource allocation. For example, the next phase of genetic studies will involve a combination of increasingly large GWAS analyses (for common SNP and copy number variant associations) and resequencing studies (for rare variants). It is not known how these and other research investments should be optimally balanced.

To the extent that resources are available, we encourage a long-term view, avoiding the well-known pattern of initial exuberance followed by disillusionment. The logic of GWAS has been clear for more than 10 years (23). Results have been remarkably consistent with expectations in the sense that common SNP associations have been discovered for many common disorders, particularly those that have been studied with larger sample sizes. It is true that initial GWAS results have explained only a small part of the etiological variance for each disease, and it seems certain that studies of copy number variants and rare SNPs will also be critical in elucidating disease mechanisms. However, it is likely that common SNPs explain a larger portion of the variance than that which can be determined with existing sample sizes, with many common SNPs, each with small effects, contributing collectively to a major portion of genetic risk (24). As the number of associations increases, the biological pathways underlying risk for each disease become clearer. GWAS methods should be applied systematically to major psychiatric disorders in large samples. The following caveats are important to consider:

<table>
<thead>
<tr>
<th>Disorder</th>
<th>Samples</th>
<th>Case Subjects</th>
<th>Comparison Subjects</th>
<th>Trios or Families</th>
<th>Prevalence Rate (%)</th>
<th>Heritability (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADHD</td>
<td>6</td>
<td>1,418</td>
<td>0</td>
<td>2,443</td>
<td>4–12</td>
<td>70–80</td>
</tr>
<tr>
<td>Autism</td>
<td>6</td>
<td>652</td>
<td>6,000</td>
<td>4,661</td>
<td>0.3–0.6</td>
<td>90–100</td>
</tr>
<tr>
<td>Bipolar disorder</td>
<td>10</td>
<td>7,075</td>
<td>10,559</td>
<td>0</td>
<td>0.3–1.5</td>
<td>73–93</td>
</tr>
<tr>
<td>Major depressive disorder</td>
<td>9</td>
<td>12,926</td>
<td>9,618</td>
<td>0</td>
<td>5–18</td>
<td>31–42^b</td>
</tr>
<tr>
<td>Schizophrenia</td>
<td>11</td>
<td>9,588</td>
<td>13,500</td>
<td>650</td>
<td>0.2–1.1</td>
<td>73–90</td>
</tr>
<tr>
<td>Total</td>
<td>42</td>
<td>31,659</td>
<td>26,945</td>
<td>7,772</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

^a Data shown are expected combined sample sizes for meta-analysis of genomewide association study data by the Psychiatric GWAS Consortium by the end of 2009. Data are reported for subjects of European-ancestry only; a small number of African American samples are also available for schizophrenia and bipolar disorder. The case subjects are all independent (although independence is tested using genotypes). For each disorder, comparison subjects used in more than one study are counted once; also, comparison subjects used for more than one disorder are counted once in the Total, which is therefore less than the sum of the rows for the disorders. The column for "Trios or Families" includes a sample of multiply-affected families for schizophrenia and trio or sib-pair families (with parents) for ADHD and autism. References for prevalence and heritability are as follows: ADHD (72–74), autism (75, 76), bipolar disorder (77, 78), major depressive disorder (79–82), and schizophrenia (83, 84).

^b For major depression, higher estimates have been obtained in clinical samples (85) or using repeated interviews (81).
1. Some disorders might not be amenable to GWAS (e.g., if all risk alleles have very low genotypic relative risks or if genetic risks are conferred by multiple rare SNPs or by copy number variants too small to be detected reliably). Discoveries for these disorders might only be possible with larger-scale resequencing studies.

2. Current diagnostic categories might be inadequate. Endophenotypic variables (neuroimaging, electrophysiological, neuropsychological, biochemical, and other markers) might better index the underlying gene effects (87), although none has yet proven more heritable than diagnostic categories. These measures are not usually available in large data sets.

3. Genetic heterogeneity reduces power. A low frequency risk allele is an example of heterogeneity (i.e., most case subjects do not share that risk factor). Power (Figure 1) is best for frequencies above approximately 20% and poor for frequencies greatly below 10% unless genotypic relative risk is high. Heterogeneity might be increased in large multicenter samples. For example, despite the generally high interrater reliability for these disorders, research groups can have diagnostic “biases,” some of which could correlate with specific risk alleles. However, power increases with sample size despite some degree of misclassification, and classification is imperfect for many medical disorders for which there are GWAS findings.

4. More needs to be learned about the selection of comparison subjects for psychiatric GWAS studies. It remains possible that some findings will be confounded by systematic biases in comparison groups, such as under-representation of developmental disabilities. The field will need much larger comparison groups ascertained by diverse methods and from multiple ethnic populations.

5. For some disorders there might be no detectable main effects of SNPs, only higher order gene-gene or gene-environment interactions. However, main effects are often detectable even if interactions are erroneously excluded. Explicit tests of interactions (88) or data mining might prove informative.

6. GWAS assays do not interrogate all common variants. For each array type, some assays perform poorly and some common SNPs are not or cannot be tagged.

7. Improved methods will be needed to provide more systematic information about copy number variants and their relationship to disease. Associated copy number variant regions will require resequencing studies of large numbers of subjects without copy number variants to determine whether these regions also contain rare, highly penetrant associated variants.

8. There are probably unknown genetic mechanisms. We have only recently recognized the importance of copy number variants, micro RNAs, long-range promoters, and epigenetic factors (genomic effects other than sequence changes such as DNA methylation patterns) (89). The discovery that most of the genome is transcribed suggests that many types of functional sequence are undiscovered (12).

Bearing these risks and caveats in mind, we conclude that GWAS methods have discovered a remarkable set of robust common SNP association findings for a broad range of diseases (90), now including an initial set of SNP and copy number variant associations for psychiatric disorders. It is reasonable to predict that studies of sufficiently large samples can produce definitive discoveries of genetic risk factors for psychiatric disorders and that these discoveries will contribute to the definitive identification of pathophysiological mechanisms for the first time.

This article was written by the Psychiatric GWAS Consortium Coordinating Committee, whose members (presented in alphabetical order) take responsibility for its content: Sven Cichon, Ph.D. (University of Bonn, Germany); Nick Craddock, M.D., Ph.D. (Cardiff University); Mark Daly, Ph.D. (Harvard Medical School, Broad Institute); Stephen V. Faraone, Ph.D. (State University of New York Upstate Medical University); Pablo V. Gejman, M.D. (Northshore University HealthSystem and Feinberg School of Medicine of Northwestern University); John Kelsoe, M.D. (University of California, San Diego); Thomas Lehner, Ph.D., M.P.H. (NIH); Douglas F. Levinson, M.D. (Stanford University); Audra Moran, M.A. (NARSAD, Ex Officio); Pamela Sklar, M.D., Ph.D. (Massachusetts General Hospital, Broad Institute); and Patrick F. Sullivan, M.D. (University of North Carolina at Chapel Hill). Received Sept. 11, 2008; revision received Nov. 17, 2008; accepted Nov. 20, 2008 (doi: 10.1176/appi.ajp.2008.08091354). From the Department of Psychiatry and Behavioral Sciences, Stanford University, Palo Alto, Calif. Address correspondence and reprint requests to Dr. Levinson, Department of Psychiatry and Behavioral Sciences, Stanford University, 710 Welch Rd., Suite A-3325, Palo Alto, CA 94304; dlevei@stanford.edu (e-mail).

Dr. Faraone receives research support from or has served on the advisory boards of Shire, Eli Lilly, Pfizer, McNeil, and NIH. Dr. Kelsoe is a founder of and holds equity in Psysomics, Inc. Dr. Sullivan has received unrestricted research support from Eli Lilly for genetic research in schizophrenia. Drs. Cichon, Craddock, Daly, Gejman, Lehner, Levinson, Sklar, and Sullivan and Ms. Moran report no competing interests.

Supported by NIH grant MH-085520. Statistical analyses were conducted using the Genetic Cluster Computer, which is supported by the Netherlands Scientific Organization (NWO 480–05–003, PI Danielle Posthuma), along with a supplement from the Dutch Brain Foundation.

ADHD Working Group: Stephen Faraone, Chair (SUNY-UMU); Richard Anney (Trinity College Dublin); Jan Buïtelaar (Radboud University); Josephine Elia (Children’s Hospital of Philadelphia); Barbara Franke (Radboud University); Michael Gill (Trinity College Dublin); Hakon Hakonarson (CHOP); Lindsay Kent (St. Andrews University); James McGough (UCLA); Eric Mick (Massachusetts General Hospital/ Harvard University); Laura Nisenbaum (Eli Lilly); Susan Smalley (UCLA); Anita Thapar (Cardiff University); Richard Todd, deceased (Washington University/St. Louis, MO); and Alexandre Todorov (Washington University/St. Louis, MO).

Autism Working Group: Bernie Devlin, Chair (University of Pittsburgh); Mark Daly, Co-Chair (Massachusetts General Hospital/ Harvard University); Richard Anney (Trinity College Dublin); Dan Arking (Johns Hopkins University); Joseph D. Buxbaum (Mt. Sinai School of Medicine, New York); Aravinda Chakravarti (Johns Hopkins University); Edwin Cook (University of Illinois); Michael Gill (Trinity College Dublin); Leena Peltonen (University of Helsinki); Joseph Piven (University of North Carolina-Chapel Hill); Guy Rouleau (University of Montreal); Susan Santangelo (Massachusetts General Hospital/Harvard University); Gerard Schellenberg (University of Washington); Steve Scherer (University of Toronto); James Sutcliffe (Vanderbilt University); Peter Sztat-
mari (McMaster University); and Veronica Vieland (Columbus Children’s Research Institute).

Bipolar Disorder Working Group: John Kelsoe, Co-Chair (UCSD); Pamela Sklar, Co-Chair, (Massachusetts General Hospital/Harvard University); Ole A. Andreassen (University of Oslo, Norway); Douglas Blackwood (University of Edinburgh, Scotland); Michael Boehnke (University of Michigan); Hugh Gumley (University College London); Jun Li (University of Michigan); Eric Mick (Massachusetts General Hospital/Harvard University); Valentina Moskville (Cardiff University); Pierandrea Muglia (GlaxoSmithKline); Walter Muir (University of Edinburgh, Scotland); Markus Noethen (University of Bonn, Germany); John Nurnberger (Indiana University); Shaun Purcell (Massachusetts General Hospital/Harvard University); Marcella Rietschel (CIMH, Mannheim); Douglas Ruderfer (Massachusetts General Hospital/Harvard University); Nicholas Schork (UCSD); Thomas Schulze (CIMH, Mannheim); Laura Scott (University of Michigan); Michael Steffens (University of Bonn, Germany); Ruchi Upmanyu (GlaxoSmithKline); and Thomas Wienerk (University of Bonn, Germany).

Cross-Disorder Working Group: Jordan Smoller, Co-Chair (Massachusetts General Hospital/Harvard University); Nicholas Craddock, Co-Chair (Cardiff University); Kenneth Kendler, Co-Chair (Virginia Commonwealth University); John Nurnberger (Indiana University); Roy Perils (Massachusetts General Hospital/Harvard University); Shaun Purcell (Massachusetts General Hospital/Harvard University); Marcella Rietschel (CIMH, Mannheim); Susan Santangelo (Massachusetts General Hospital/Harvard University); and Anita Thapar (Cardiff University).

Minor Depressive Disorder Working Group: Patrick Sullivan, Chair (University of North Carolina-Chapel Hill); Douglas Blackwood (University of Edinburgh, Scotland); Dorret Boomsma (Vrije University, Amsterdam); Rene Breuer (CIMH, Mannheim, Germany); Sven Cichon (University of Bonn, Germany); William Coryell (University of Iowa); Eco de Geus (Vrije University, Amsterdam); Steve Hamilton (UCSF); Witte Hoogendijk (Vrije University, Amsterdam); Stafam Kloiber (MPI-P Munich); William B. Lawson (Howard University); Douglas Levinson (Stanford University); Cathryn Lewis (IOP, London); Susanne Lucae (MPI-P Munich); Nick Martin (QIMR); Patrick McGrath (Columbia University); Peter Mcguinn (IOP, London); Pierandrea Muglia (GlaxoSmithKline); Walter Muir (University of Edinburgh, Scotland); Markus Noethen (University of Bonn, Germany); James Offord (Pfizer); Brenda Penn (Vrije University, Amsterdam); James B. Potash (Johns Hopkins University); Marcella Rietschel (CIMH, Mannheim, Germany); William A. Scheffner (Rush University); Thomas Schulze (CIMH, Mannheim); Susan Slagter (Mayo Clinic); Federica Tozzi (GlaxoSmithKline); Myrna M. Weissman (Columbia University); and Veronica Vieland (Columbus Children’s Research Institute).

The authors thank the Psychiatric GWAS Consortium colleagues for their contributions. The authors also thank NARSAD for infrastructure support.

References


10. Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Feuk L, Dunham I, Sherry S, Birren B, Weissman D, Petitt P, Garrett S, Golub J, Edwards J, Yao J, Keinan A, Caccamo J, Rafferty J, Grimsby J, Green T,餮 X, Horn P, Alkan C, Brudno M, Li Y, Ainsley NJ, Member MC, Richmond D, Heidenblut M, Goodier J, Sargan D, Grady J, Trifonov V, COTHAM, Mannheim, Germany; Sibyle Schwab (University of Western Australia); Pamela Sklar (Massachusetts General Hospital/Harvard University); David St. Clair (University of Aberdeen); Patrick Sullivan (University of North Carolina-Chapel Hill); Jaana Susvisaari (University of Helsinki); Edwin van den Oord (Virginia Commonwealth University); Naomi Wray (QIMR); and Dieter Wildenmann (University of Western Australia).

Statistical Analysis and Computational Working Group: Mark Daly, Chair (Massachusetts General Hospital/Harvard University); Philip Awadalla (University of Montreal); Bernie Devlin (University of Pittsburgh); Frank Dudbridge (MRC-BSU); Arnoldo Frigessi (University of Oslo, Norway); Elizabeth Holliday (QCMHR/University of Queensland); Peter Holmans (Cardiff University); Todd Lencz (UJ); Douglas Levinson (Stanford University); Cathryn Lewis (IOP, London); Danyu Lin (University of North Carolina-Chapel Hill); Valentina Moskvin (Cardiff University); Bryan Movry (QCMHR/University of Queensland); Ben Neale (Massachusetts General Hospital/Harvard University); Eve Pickering (Pfizer Pharmaceuticals Group); Danielle Posthuma (Vrije University Amsterdam); Shaun Purcell (Massachusetts General Hospital/Harvard University); John Rice (Washington University/St. Louis, MO); Stephan Ripke (MPI-P Munich); Nicholas Schork (UCSD); Jonathan Sebat (CSHL); Michael Steffens (University of Bonn, Germany); Jennifer Stone (Massachusetts General Hospital/Harvard University); Jung-Ying Tzeng (NCSU); Edwin van den Oord (Virginia Commonwealth University); and Veronica Vieland (Columbus Children’s Research Institute).

The authors thank the Psychiatric GWAS Consortium colleagues for their contributions. The authors also thank NARSAD for infrastructure support.


90. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de Bakker PI, Daly MJ, Sham PC: PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet 2007; 81:559–575
91. Purcell S, Cherny SS, Sham PC: Genetic Power Calculator: design of linkage and association genetic mapping studies of complex traits. Bioinformatics 2003; 19:149–150