Molecular Psychiatry (2009) 14, 10–17

www.nature.com/mp

PERSPECTIVE

A framework for interpreting genome-wide association studies of psychiatric disorders

The Psychiatric GWAS Consortium Steering Committee

Genome-wide association studies (GWAS) have yielded a plethora of new findings in the past 3 years. By early 2009, GWAS on 47 samples of subjects with attention-deficit hyperactivity disorder, autism, bipolar disorder, major depressive disorder and schizophrenia will be completed. Taken together, these GWAS constitute the largest biological experiment ever conducted in psychiatry (59 000 independent cases and controls, 7700 family trios and >40 billion genotypes). We know that GWAS can work, and the question now is whether it will work for psychiatric disorders. In this review, we describe these studies, the Psychiatric GWAS Consortium for meta-analyses of these data, and provide a logical framework for interpretation of some of the conceivable outcomes.
Molecular Psychiatry (2009) **14,** 10–17; doi:10.1038/mp.2008.126; published online 11 November 2008

Keywords: genome-wide association; attention-deficit hyperactivity disorder; autism; bipolar disorder; major depressive disorder; schizophrenia

Genome-wide association studies (GWAS) have yielded confirmed associations for complex diseases in almost all medical specialties. Since 2005, there have been genetic breakthroughs in cardiology, endocrinology, gastroenterology, hepatology, infectious disease, oncology, ophthalmology, neurology, pulmonology and rheumatology. The success of GWAS for nonpsychiatric complex diseases has been exceptional – as of 1 August 2008, there were 197 associations where the initial GWAS finding replicated in one or more additional samples with a *P*-value $<5\times10^{-8}$ (a defensible choice for genome-wide significance).[1] Recent papers describing significant and replicated associations for autism,[2] bipolar disorder[3] and schizophrenia[4–6] suggest that psychiatry has joined this list. By the end of 2008, GWAS on 42 samples of subjects of European ancestry with attention-deficit hyperactivity disorder (ADHD), autism, bipolar disorder, major depressive disorder and schizophrenia will have been completed. Taken together, these GWAS constitute the largest biological experiment ever conducted in psychiatry: >59 000 independent cases and controls, >7700 family trios, on the order of 500 000 single nucleotide polymorphism (SNP) genotypes per subject, and a total of >40 billion genotypes.

The purpose of this article is to consider the 'big picture' and to provide a logical framework for the possible outcomes of these studies. This is not a review of GWAS *per se* as many excellent reviews of this technically and statistically intricate methodological approach are available.[7–12] This is also not a review of the advantages and disadvantages of different study designs and sampling strategies for the dissection of complex psychiatric traits. We would like to consider how the dozens of GWAS papers that will soon be in the literature can be synthesized: what can integrated mega-analyses (meta-analysis is based on summary data (for example, odds ratios) from all available studies whereas 'mega-analysis' uses individual-level genotype and phenotype data) of all available GWAS data tell us about the etiology of these psychiatric disorders? This is an exceptional opportunity as positive or negative results will enable us to learn hard facts about these critically important psychiatric disorders. We suggest that it is not a matter of 'success versus failure' or 'optimism versus pessimism' but rather an opportunity for systematic and logical approaches to empirical data whereby both positive and appropriately qualified negative findings are informative.

The Psychiatric GWAS Consortium

The studies that comprise the Psychiatric GWAS Consortium (PGC; http://pgc.unc.edu) are shown in Table 1. GWAS data for ADHD, autism, bipolar disorder, major depressive disorder and schizophrenia from 42 samples of European subjects should be available for mega-analyses by early 2009 (>59 000 independent cases and controls and >7700 family trios). To our knowledge, the PGC will have access to the largest set of GWAS data available.

Correspondence: Dr P Sullivan, Department of Genetics, CB#7264, 4109D Neurosciences Research Building, University of North Carolina, Chapel Hill, NC, 27599-7264, USA.
E-mail: pfsulliv@med.unc.edu
Received 27 October 2008; accepted 27 October 2008; published online 11 November 2008

Table 1 Summary of GWAS samples expected by the end of 2008

Disorder	*Samples*	*Cases*[a]	*Controls*[b]	*Trios*
ADHD	6	1418	0	2443
Autism	6	652	6000	4661
Bipolar disorder	10	7075	10 559	0
Major depressive disorder	9	12 926	9618	0
Schizophrenia	11	9588	13 500	650
Total[c]	**42**	**31 659**	**26 945**	**7772**

Abbreviation: ADHD, attention-deficit hyperactivity disorder.
[a]All cases are independent.
[b]Controls within each disorder are independent.
[c]Column totals excludes duplicates and are not necessarily a simple sum of the rows.

A major change in human genetics in the past 5 years has been in the growth of controlled-access data repositories, and individual phenotype and genotype data are now available for many of the studies in Table 1. When the PGC mega-analyses are completed, most data will be available to researchers via the NIMH Human Genetics Initiative (http://nimhgenetics.org). Although the ready availability of GWAS data is a benefit to the field by allowing rapid application of a wide range of analytic strategies to GWAS data, there are potential disadvantages. GWAS mega-analysis is complex and requires considerable care and expertise to be done validly. For psychiatric phenotypes, there is the additional challenge of working with disease entities based largely on clinical description, with unknown biological validity and having both substantial clinical variation within diagnostic categories as well as overlaps across categories.[13] Given the urgent need to know if there are replicable genotype–phenotype associations, a new type of collaboration was required.

The purpose of the PGC is to conduct rigorous and comprehensive within- and cross-disorder GWAS mega-analyses. The PGC began in early 2007 with the principal investigators of the four GAIN GWAS,[14] and within six months had grown to 110 participating scientists from 54 institutions in 11 countries. The PGC has a coordinating committee, five disease-working groups, a cross-disorder group, a statistical analysis and computational group, and a cluster computer for statistical analysis. It is remarkable that almost all investigators approached agreed to participate and that no one has left the PGC. Most effort is donated but we have obtained funding from the NIMH, the Netherlands Scientific Organization, Hersenstichting Nederland and NARSAD.

The PGC has two major specific aims. (1) Within-disorder mega-analyses: conduct separate mega-analyses of all available GWAS data for ADHD, autism, bipolar disorder, major depressive disorder, and schizophrenia to attempt to identify genetic variation convincingly associated with any one of these five disorders. (2) Cross-disorder mega-analyses: the clinically-derived DSM-IV and ICD-10 definitions may not directly reflect the fundamental genetic architecture.[15] There are two subaims. (2a) Conduct mega-analysis to identify genetic variation convincingly associated with conventional definitions of two or more disorders. This nosological aim could assist in delineating the boundaries of this set of disorders. (2b) An expert working group will convert epidemiological and genetic epidemiological evidence into explicit hypotheses about overlap among these disorders, and then conduct mega-analyses based on these definitions (for example, to examine the lifetime presence of idiopathic psychotic features without regard to diagnostic context).

We anticipate that interim mega-analyses for Aims 1 and 2a will be completed by late 2008 and final mega-analyses by mid-2009. Updated results will be posted on the PGC web site (http://pgc.unc.edu). The standard of evidence applied will be strict and conform to that recommended in human genetics.[9] Statistical power should be superior to any earlier study in psychiatric genetics. As a rough illustration, the minimum detectable genotypic relative risk for Aim 1 is 1.161 (assuming 10 000 cases for one disorder, 10 000 controls, $\alpha = 1 \times 10^{-8}$,[1] minor allele frequency = 0.25, and log-additive model). For the nosological Aim 2a, the minimum detectable genotypic relative risk is 1.100 (same assumptions but with 25 000 cases and 25 000 controls). Effect sizes from the literature for other complex traits are usually larger than these minima.[16]

A framework for interpreting GWAS of psychiatric disorders

Approach

We emphasize that the framework presented here is not the only possible formulation. Indeed, the recent history of human genetics has been full of unanticipated empirical results and it is quite possible that the future will eventuate differently from the framework presented below. At the same time, we believe it important that we think about these studies in an integrated and systematic manner.

The goal of the PGC is to identify convincing genetic variation-disease associations. A convincing association would be extremely unlikely to result

from chance, show consistent effect sizes across all or almost all samples and be impervious to vigorous attempts to disprove the finding (for example, by investigating sources of bias, confirmatory genotyping, and so on). Careful attention will be paid to the impact of potential sources of heterogeneity[17] with the goal of assessing its impact without minimizing its presence.

Biological plausibility is not an initial requirement for a convincing statistical association, as there are many examples in human genetics of previously unsuspected candidate genes nonetheless showing highly compelling associations. For example, multiple SNPs in intron 1 of the *FTO* gene were associated with body mass index in 13 cohorts with 38 759 participants[18] and yet '*FTO*' does not appear in an exhaustive 116 page compilation of genetic studies of obesity.[19] Some strong associations are in gene deserts: multiple studies have found convincing association between prostate cancer and a region on 8q24 that is ~250 kb from the nearest annotated gene.[20] Both of these examples are being intensively investigated and we suspect that a compelling mechanistic 'story' will emerge in the near future. The presence of a compelling association without an obvious biological mechanism establishes a priority research area for molecular biology and neuroscience of a psychiatric disorder.

The PGC will use mega-analysis as the main analytic tool as individual-level data will be available from almost all samples. To wield this tool appropriately, a number of preconditions must be met. First, genotype data from different GWAS platforms must be made comparable as the direct overlap between platforms is often modest. This requires meticulous quality control for the inclusion of both SNPs and subjects and attention to the factors that can cause bias (for example, population stratification, cryptic relatedness or genotyping batch effects). Genotype harmonization can be accomplished using imputation ([21,22], for example) so that the same set of ~2 million[23,24] directly or imputed SNP genotypes are available for all subjects. Second, phenotypes need to be harmonized across studies. This is one of the most crucial components of the PGC and we are fortunate to have world experts directing the work. Third, the mega-analyses will assess potential heterogeneity of associations across samples.

An interpretive framework

A decision-tree schematic of the potential outcomes of the PGC mega-analyses is shown in Figure 1. Note that many of the possibilities in Figure 1 are not mutually exclusive and different disorders may take different paths through this framework. It is possible that there eventually will be dozens or hundreds of sequence variants strictly associated with these disorders with frequencies ranging from very rare to common.

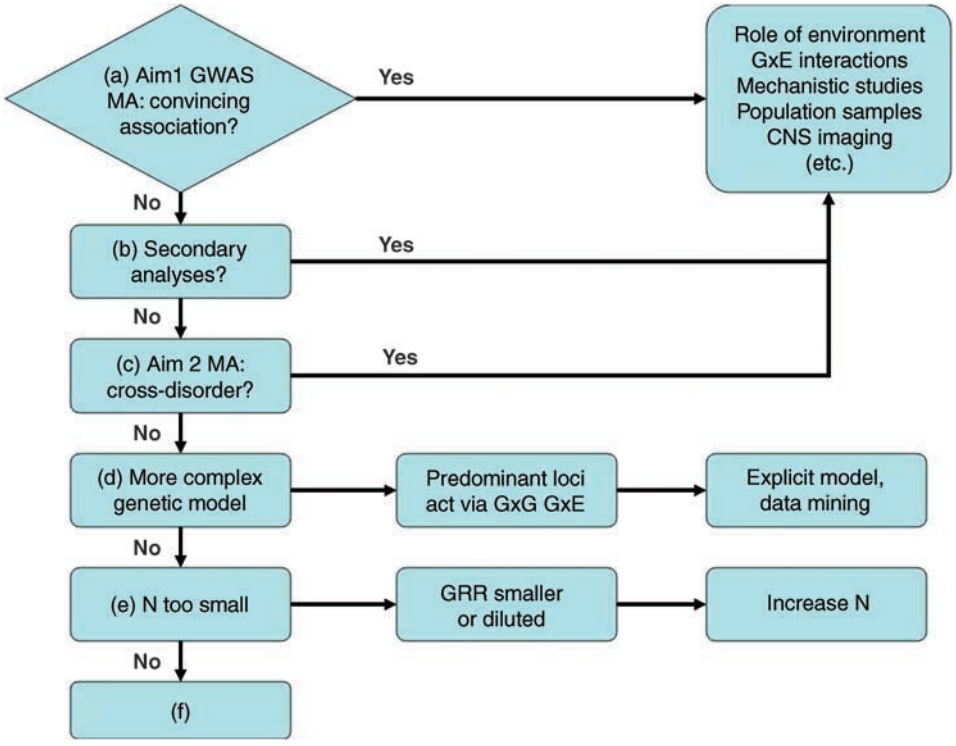

Figure 1 Flowchart of a conceivable set of outcomes for the planned Psychiatric GWAS Consortium (PGC) mega-analyses for five critically important psychiatric disorders. See text for explanations. Abbreviations: GWAS, genome-wide association study; MA, mega-analysis; GxE, gene–environment interaction; CNS, central nervous system; GxG, gene–gene interaction or epistasis;, *N*, sample size; GRR, genetic relative risk.

Possibility 'a'. GWAS mega-analysis for the within-disorder Aim 1 (that is, conventional definitions of ADHD, autism, bipolar disorder, major depressive disorder and schizophrenia) identifies one or more compelling associations with an SNP, haplotype or copy number variant. This is the most fortuitous outcome and a 'Holy Grail' of psychiatric genetics. This would constitute a historical advance if any such regions were identified. It is difficult to overemphasize how important such a landmark would be to patients, families, clinicians and researchers. In this instance, a wealth of studies become possible including identifying environmental main effects and gene–environment interaction, mechanistic studies of how genetic variation might lead to a clinical outcome, brain imaging studies of cases with and without the genetic variant, investigation of the variant in unaffected population samples and in different case groups, and so on. Autism,[2] bipolar disorder[3] and schizophrenia[4–6] now each have likely examples of possibility 'a'.

Bottom-line. What is the utility of a possibility 'a' finding? The immediate implications are described above. On an intermediate scale, such an association gives biologists and neuroscientists an excellent starting point for attempting to determine how genetic variation in an associated region (along with nearby genes or related pathways) changes vulnerability to these psychiatric disorders. For clinical researchers, such associations could provide a useful way to stratify clinical samples to investigate relationships with clinically relevant phenomena (for example, brain structure/function or treatment outcome). For epidemiologists, the study of how environmental factors could interact with such an association to alter disease risk is of definite importance. The long-term goals of this research are to develop a comprehensive understanding of how genetic and environmental factors act and interact over development to produce these disorders. Such knowledge could then be used in a public health context rationally to prevent disease, to identify cases earlier to initiate treatment, and to minimize morbidity and mortality and to derive more efficacious treatments.

Possibility 'b'. If the Aim 1 mega-analyses do not identify any compelling associations, the analytic plans for each disorder-working group specify several ways to identify putatively more homogeneous subgroups (for example, recurrent, early-onset major depressive disorder or schizophrenia with prominent negative symptoms). It is conceivable that these secondary analyses could identify compelling associations (after appropriate correction for multiple testing).

Possibility 'c'. Aim 2 cross-disorder mega-analyses. From the perspective of the genetic architecture of these five disorders, it is possible that our clinically derived nomenclature (DSM-IV or ICD-10) has imposed divisions between disorders where, from a genetic perspective, similarities outweigh differences. In this instance, Aim 2 mega-analyses using conventional or expert-refined definitions of illness or subphenotypes might detect compelling associations that transcend traditional diagnostic boundaries.

Possibility 'd'. The statistical model for the relationship between genetic variation and a psychiatric phenotype used in the mega-analyses is perhaps the simplest conceivable model: genetic variation acts in an additive manner to increase the probability of a disorder. This additive model has worked exceptionally well in human genetics as it is relatively robust to misspecification of the true model. In other words, if the true model is dominant, if the genotyped SNP is in linkage disequilibrium with a causal variant, or if the true model involves a gene–gene or gene–environment interaction, an additive model can detect the signal even if the true model is not explicitly specified. However, it is possible to imagine scenarios where interactions have no detectable additive main effect (for example, a genetic effect increases risk in the presence of an environmental risk factor to which half the population is exposed and decreases risk by the same amount in its absence). In this instance, explicitly modeling the interaction or the use of nontraditional data mining methods could help.

Possibility 'e'. It is possible that the minimum detectible genetic relative risks afforded by the currently available sample sizes are too small. This could be because effect sizes are smaller for psychiatric disorders or due to etiological heterogeneity or diagnostic imprecision leading to dilution of the genetic relative risk. In this instance, increasing the available sample sizes with GWAS genotyping would be important. Collection of considerably larger numbers of cases for the major psychiatric disorders is ongoing in the United States and Europe and could resolve this issue on the 1–3 year horizon.

The 'f'' Possibility. The box labeled 'f'' in Figure 1 represents the most bedeviling and yet the most intriguing possibility in the framework. The likelihood of option 'f'' is the outcome about which GWAS pundits debate with variable degrees of optimism and pessimism. In our view, we cannot know whether we have reached option 'f'' until we have carefully worked through a flowchart like that in Figure 1 which could take 2–4 years. For the sake of argument, assume that option 'f' becomes tenable for one or more of ADHD, autism, bipolar disorder, major depressive disorder and schizophrenia. What does this tell us about the fundamental nature of a disorder?

First, we learn what these disorders are not. GWAS is predicated upon a number of assumptions: if one or more does not hold, option 'f' could occur. It is disappointing to learn what a disorder is not (as opposed to the positive identification of a compelling association), but such knowledge is crucial so that the next generation of research studies can be appropriately focused. Indeed, this sort of iterative process describes the history of psychiatric genetics to date which has been one of such exclusions. For example, 43 years ago it was assumed that schizophrenia was caused by a single dominant gene with 25% penetrance[25]—results from 31 genome-wide linkage studies[26] have shown that there is no common effect of this magnitude and subsequent study designs have taken this datum into account. Knowledge, for example, that there is no single locus conferring a genetic relative risk of more than 1.075 under a set of assumptions (genetic model, allele frequency and heterogeneity) for a disorder will assist in figuring out what to do next. As another example, learning that there are no loci (under a set of assumptions) that confer risk for bipolar disorder-schizophrenia, major depressive-bipolar disorder or ADHD-bipolar disorder will provide important evidence in regard to nosological debates about these disorders. Thus, even if we do not gain the positive knowledge of a compelling genetic association, option 'f' enables us to learn.

Second, assumptions about the fundamental genetic architecture of these disorders may be incorrect. The most crucial aspects of the genetic architecture are the number of genomic regions involved and, for each locus, the genetic model by which it impacts phenotypic risk and its relative risk. These factors combine to form an 'allelic spectrum' (Figure 2). The allelic spectrum is bounded by 'common disease/common variant' (CDCV) and 'multiple rare variant' (MRV) models. The CDCV model postulates the existence of genetic variants that are relatively common (perhaps 5–50%) that confer modest risk (for example, relative risks 1.1–1.5). The MRV model holds that complex traits result from many different mutations each of which is individually rare (a few percent at most and perhaps orders of magnitude less common) but with very strong effect (for example, relative risks >10). These models are not dichotomous and the allelic spectrum is more appropriate. For example, there are compelling data implicating eight MODY (maturity onset diabetes of the young) genes under a MRV model and ~15 CDCV variants for type 2 diabetes. Moreover, genes containing CDCV could also have as yet undiscovered MRV. Relatively subtle CDCV alterations establish the importance of a gene in the pathophysiology of a disease and it is mechanistically plausible that rare MRV alterations with strong effects could occur. Framework option 'f' could result if the genetic causes of a psychiatric disorder are exclusively from the left side of the allelic spectrum. Note that studies of copy number variation (CNVs) in large samples affords partial protection under this scenario (for example, 22q11 deletion syndrome is a MRV for schizophrenia with prevalence of 0.3% and genetic relative risk >20).[5] If these disorders are caused by a large number of different mutations, individual genomic sequencing is likely to be required. This is not now economically feasible but is anticipated on the 5-year horizon.

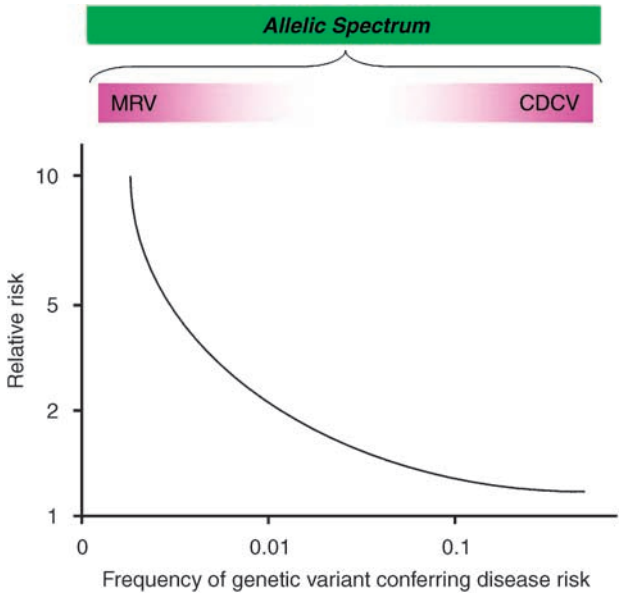

Figure 2 Depicts illustrative combinations of disease mutation frequency and genetic relative risk for a complex trait. The multiple rare variant (MRV) model is a loose descriptor for effects at the upper left of the graph and the common disease/common variant (CDCV) model for those at the lower right. The term 'allelic spectrum' encompasses both.

Third, GWAS assume a simple additive genetic model. However, this is only one slice through the possible parameter space and the true model for any these psychiatric disorders might not conform to an additive model. It is possible that some true models lie in the less detectable parts of the parameter space – for example, a very large number of contributing loci all with subtle effects, more complex single locus effects (for example, dominance or heterosis), interactions (intralocus, gene–gene or gene–environment interactions) and more intricate models involving parent-of-origin or epigenetic effects. Given the striking advances in human genetics in the past decade, it is prudent to note the possibility of a novel genetic mechanism. The implication here is that dissecting the genetic basis of framework option 'f' disorders will be quite difficult and possibly intractable based on human studies.

Fourth, GWAS currently assess only a subset of genetic variants. Available GWAS platforms directly measure SNPs and CNVs and indirectly assess untyped variants in 'useful' linkage disequilibrium with them. Although the current generation of GWAS platforms contains ~1 million SNPs and provides 'adequate' coverage of >90% of the genome for those of European or East Asian ancestry, important genetic

variation could lie in a poorly covered region. Moreover, SNPs and CNVs are only two types of genetic variation and there are many other types of genetic variants: important genetic variations that are not SNPs or CNVs might not be assessed by current GWAS platforms. This is a technical limitation, which future generations of GWAS platforms could remedy.

Fifth, the fundamental pathophysiology may not map onto the definition of a disorder. Whether this is or is not the case has been the subject of considerable debate for over a century and was a driving force for the repeated revisions of the DSM criteria since 1980. The limitations of DSM-IV for etiological studies must be kept in mind: 'our highest priority has been to provide a helpful guide to clinical practice' (page xv) and 'these diagnostic criteria … reflect a consensus of current formulations of evolving knowledge' (page xxvii).[27] The implication is that different ascertainment strategies or study designs may be needed (for example, on the basis of a component phenotype rather than at the disorder level).

Sixth, GWAS assume that the degree of etiological heterogeneity for each disorder is not prohibitive. Two extreme positions can be defined: (a) a single etiological type of disease and (b) as many types of disease as there are affected individuals. For each of the five disorders being considered in the PGC, the former is unlikely and the latter implies that searches for any type of etiological factor (genetic or otherwise) will be very difficult. Searches for genetic risk factors might succeed only if the degree of heterogeneity is not too awful and if one or a few etiologies predominate (for example, there might be 10 etiological 'types' of schizophrenia but one accounts for half of all cases). Indeed, it is possible that a relatively modest degree of heterogeneity can make detection of a subtle association improbable. If it were possible to index heterogeneity directly (for example, using an endophenotype),[28] this problem could be addressed; however, there are currently only experimental ways to index heterogeneity. Indeed, a goal of genetic studies is to identify genetic markers that can improve diagnostic classification by indexing heterogeneity. Careful attention to phenotypic measures (rather than reliance on diagnoses) could prove to be important for identifying and replicating susceptibility genes.

Finally, it is unlikely but formally possible that the genetic epidemiological studies whose results provide the rationale for GWAS are substantially incorrect. Family, adoption and twin studies all make a variety of assumptions. It is possible that there are undetected biases or novel etiological factors that led to overestimations of the importance of genetic effects.

Conclusions

GWAS has the potential to yield considerable insights but it is no panacea and may well perform differently for psychiatric disorders. Even if these psychiatric GWAS efforts are successful, the outcomes will be complex. GWAS may help us learn that clinical syndromes are actually many different things—for example, proportions of individuals with schizophrenia might evidence associations with rare CNVs of major effect,[5,6] with more common genetic variation in dozens (perhaps hundreds) of genomic regions, between genetic variation strongly modified by environmental risk factors, and some proportion may be genetically indistinguishable from the general population. Moreover, as fuel to long-standing 'lumper versus splitter' debates in psychiatric nosology, empirical data might show that some clinical disorders or identifiable subsets of subjects might overlap considerably.

The critical advantage of GWAS is the search of a 'closed' hypothesis space. If the large amount of GWAS data being generated are analyzed within a strict and coherent framework, it should be possible to establish hard facts about the fundamental genetic architecture of a set of important psychiatric disorders—which might include positive evidence of what these disorders are or exclusionary evidence of what they are not. Whatever the results, these historically large efforts should yield hard facts about ADHD, autism, bipolar disorder, major depressive disorder and schizophrenia that may help guide the next era of psychiatric research.

Acknowledgments

This article was written early in the history of the PGC by its coordinating committee who takes responsibility for its content. We wish to acknowledge our enduring debt to our PGC colleagues for their many contributions to this work. The members of the PGC Coordinating Committee are: Sven Cichon (University of Bonn), Nick Craddock (Cardiff University), Mark Daly (Harvard Medical School, Broad Institute), Stephen V Faraone (SUNY Upstate Medical University), Pablo V Gejman (Evanston Northwestern Healthcare and Feinberg School of Medicine, Northwestern University), John Kelsoe (University of California, San Diego), Thomas Lehner (NIMH), Douglas F Levinson (Stanford University), Audra Moran (NARSAD, *ex officio*), Pamela Sklar (Massachusetts General Hospital, Broad Institute) and Patrick F Sullivan (University of North Carolina at Chapel Hill). The PGC is supported by NIMH MH085520 (http://www.nimh.gov). Statistical analyses are conducted on 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 (http://www.hersenstichting.nl, PI Roel Ophoff). We are indebted to NARSAD for infrastructure support.

ADHD Working Group: Stephen Faraone (chair, SUNY UMU), Richard Anney (Trinity College Dublin), Jan Buitelaar (Radboud U), Josephine Elia (CHOP), Barbara Franke (Radboud U), Michael Gill

(Trinity College Dublin), Hakon Hakonarson (CHOP), Lindsey Kent (St Andrews U), James McGough (UCLA), Eric Mick (MGH/Harvard), Laura Nisenbaum (Eli Lilly), Susan Smalley (UCLA), Anita Thapar (Cardiff U), Richard Todd (Wash U/St Louis, deceased) and Alexandre Todorov (Wash U/St Louis). Autism Working Group: Bernie Devlin (chair, U Pittsburgh), Mark Daly (co-chair, MGH/Harvard), Richard Anney (Trinity College Dublin), Dan Arking (Johns Hopkins U), Joseph D Buxbaum (Mt SInai School Medicine), Aravinda Chakravarti (Johns Hopkins U), Edwin Cook (U Illinois), Michael Gill (Trinity College Dublin), Leena Peltonen (U Helsinki), Joe Piven (UNC-CH), Guy Rouleau (U Montreal), Gerard Schellenberg (U Washington), Steve Scherer (U Toronto), James Sutcliffe (Vanderbilt U), Peter Szatmari (McMaster U) and Veronica Vieland (Columbus Children's Research Institute).

Bipolar Disorder Working Group: John Kelsoe (co-chair, UCSD), Pamela Sklar (co-chair, MGH/Harvard), Ole A Andreassen (U Oslo), Douglas Blackwood (U Edinburgh), Michael Boehnke (U Mich), Margit Burmeister (U Mich), Sven Cichon (U Bonn), Aiden Corvin (Trinity College Dublin), Nicholas Craddock (Cardiff U), Manuel Ferreira (MGH/Harvard), Matthew Flickinger (U Mich), Tiffany Greenwood (UCSD), Weihua Guan (U Mich), Hugh Gurling (UCL), Jun Li (U Mich), Eric Mick (MGH/Harvard), Valentina Moskvina (Cardiff U), Pierandrea Muglia (GSK), Walter Muir (U Edinburgh), Markus Noethen (U Bonn), John Nurnberger (Indiana U), Shaun Purcell (MGH/Harvard), Marcella Rietschel (CIMH, Mannheim), Douglas Ruderfer (MGH/Harvard), Nicholas Schork (UCSD), Thomas Schulze (CIMH, Mannheim), Laura Scott (U Mich), Michael Steffens (U Bonn), Ruchi Upmanyu (GSK) and Thomas Wienker (U Bonn).

Major Depressive Disorder Working Group: Patrick Sullivan (chair, UNC-CH), Douglas Blackwood (U Edinburgh), Dorret Boomsma (Vrije U Amsterdam), Sven Cichon (U Bonn), Bryan Dechairo (Pfizer), Eco de Geus (Vrije U Amsterdam), Steve Hamilton (UCSF), Witte Hoogendijk (Vrije U Amsterdam), Doug Levinson (Stanford U), Cathryn Lewis (IOP London), Susanne Lucae (MPI-P Munich), Stuart MacGregor (QIMR), Nick Martin (QIMR), Patrick McGrath (Columbia U), Peter McGuffin (IOP London), Pierandrea Muglia (GSK), Walter Muir (U Edinburgh), Markus Noethen (U Bonn), James Offord (Pfizer), Brenda Penninx (Vrije U Amsterdam), Marcella Rietschel (CIMH, Mannheim), Thomas Schulze (CIMH, Mannheim), Susan Slager (Mayo Clinic), Federica Tozzi (GSK), Peter Visscher (QIMR), AHM Willemsen (Vrije U Amsterdam) and Naomi Wray (QIMR).

Schizophrenia Working Group: Pablo Gejman (chair, ENH/NU), Ole A Andreassen (U Oslo), Douglas Blackwood (U Edinburgh), Sven Cichon (U Bonn), Aiden Corvin (Trinity College Dublin), Mark Daly (MGH/Harvard), Ayman Fanous (Washington VAMC/Georgetown/VCU), Michael Gill (Trinity College Dublin), Hugh Gurling (UCL), Peter Holmans (Cardiff U), Christina Hultman (Karolinska Institutet), Kenneth Kendler (VCU), Sari Kivikko (National Public Health Institute), Todd Lencz (LIJ), Doug Levinson (Stanford U), Anil Malhotra (LIJ), Bryan Mowry (QCMHR/UQueensland), Markus Noethen (U Bonn), Mike O'Donovan (Cardiff U), Roel Ophoff (UCLA), Michael Owen (Cardiff U), Leena Peltonen (U Helsinki), Marcella Rietschel (CIMH, Mannheim), Alan Sanders (ENH/NU), Thomas Schulze (CIMH, Mannheim), Pamela Sklar (MGH/Harvard), David St Clair (U Aberdeen), Patrick Sullivan (UNC-CH), Jaana Suvisaari (U Helsinki) and Edwin van den Oord (VCU).

Cross-Disorder Working Group: Jordan Smoller (co-chair, MGH/Harvard), Nicholas Craddock (co-chair, Cardiff U), Kenneth Kendler (co-chair, VCU), John Nurnberger (Indiana U), Roy Perlis (MGH/Harvard), Shaun Purcell (MGH/Harvard), Marcella Rietschel (CIMH, Mannheim), Susan Santangelo (MGH/Harvard) and Anita Thapar (Cardiff U).

Statistical Analysis and Computational Working Group: Mark Daly (chair, MGH/Harvard), Phillip Awadalla (U Montreal), Bernie Devlin (U Pittsburgh), Frank Dudbridge (MRC-BSU), Arnoldo Frigessi (U Oslo), Elizabeth Holliday (QCMHR/U Queensland), Peter Holmans (Cardiff U), Todd Lencz (LIJ), Doug Levinson (Stanford U), Cathryn Lewis (IOP London), Danyu Lin (UNC-CH), Valentina Moskvina (Cardiff U), Bryan Mowry (QCMHR/UQueensland), Ben Neale (MGH/Harvard), Eve Pickering (Pfizer), Danielle Posthuma (Vrije U Amsterdam), Shaun Purcell (MGH/Harvard), John Rice (Wash U/St Louis), Stephan Ripke (MPI-P Munich), Nicholas Schork (UCSD), Jonathan Sebat (CSHL), Michael Steffens (U Bonn), Jennifer Stone (MGH/Harvard), Jung-Ying Tzeng (NCSU), Veronica Vieland (Columbus Children's Research Institute) and Edwin van den Oord (VCU).

References

1 Pe'er I, Yelensky R, Altshuler D, Daly MJ. Estimation of the multiple testing burden for genomewide association studies of nearly all common variants. *Genet Epidemiol* 2008; **32**: 381–385.
2 Weiss LA, Shen Y, Korn JM, Arking DE, Miller DT, Fossdal R *et al.* Association between microdeletion and microduplication at 16p11.2 and autism. *N Engl J Med* 2008; **358**: 667–675.
3 Ferreira M, O'Donovan M, Meng Y, Jones I, Ruderfer D, Jones L *et al.* Collaborative genome-wide association analysis of 10,596 individuals supports a role for Ankyrin-G (ANK3) and the alpha-1C subunit of the L-type voltage-gated calcium channel (CACNA1C) in bipolar disorder. *Nat Genet* 2008; e-pub ahead of print 2008 August 17.
4 O'Donovan M, Craddock N, Norton N, Williams H, Peirce T, Moskvina V *et al.* Identification of novel schizophrenia loci by genome-wide association and follow-up. *Nat Genet* 2008; e-pub ahead of print 2008 July 30.
5 International Schizophrenia Consortium. Greater burden of rare copy number variants in schizophrenia. *Nature* 2008; **455**: 237–241.
6 Stefansson H, Rujescu D, Cichon S, Pietilainen OP, Ingason A, Steinberg S *et al.* Large recurrent microdeletions associated with schizophrenia. *Nature* 2008; **455**: 232–236.
7 Craddock N, O'Donovan MC, Owen MJ. Genome-wide association studies in psychiatry: lessons from early studies of

non-psychiatric and psychiatric phenotypes. *Mol Psychiatry* 2008; **13**: 649–653.
8 McCarthy MI, Abecasis GR, Cardon LR, Goldstein DB, Little J, Ioannidis JP *et al.* Genome-wide association studies for complex traits: consensus, uncertainty and challenges. *Nat Rev Genet* 2008; **9**: 356–369.
9 Chanock SJ, Manolio T, Boehnke M, Boerwinkle E, Hunter DJ, Thomas G *et al.* Replicating genotype-phenotype associations. *Nature* 2007; **447**: 655–660.
10 Manolio TA, Brooks LD, Collins FS. A HapMap harvest of insights into the genetics of common disease. *J Clin Invest* 2008; **118**: 1590–1605.
11 Pearson TA, Manolio TA. How to interpret a genome-wide association study. *JAMA* 2008; **299**: 1335–1344.
12 Altshuler D, Daly M. Guilt beyond a reasonable doubt. *Nat Genet* 2007; **39**: 813–815.
13 Craddock N, O'Donovan MC, Owen MJ. Phenotypic and genetic complexity of psychosis. *Br J Psychiatry* 2007; **190**: 200–203.
14 Manolio TA, Rodriguez LL, Brooks L, Abecasis G, Ballinger D, Daly M *et al.* New models of collaboration in genome-wide association studies: the Genetic Association Information Network. *Nat Genet* 2007; **39**: 1045–1051.
15 Craddock N, Owen MJ. Rethinking psychosis: the disadvantages of a dichotomous classification now outweigh the advantages. *World Psychiatry* 2007; **6**: 20–27.
16 Hindorff LA, Junkins HA, Manolio TA. A catalog of published genome-wide association studies. Available at: www.genome.gov/26525384. Accessed 8/1/2008.
17 Ioannidis JP, Patsopoulos NA, Evangelou E. Heterogeneity in meta-analyses of genome-wide association investigations. *PLoS ONE* 2007; **2**: e841.
18 Frayling TM, Timpson NJ, Weedon MN, Zeggini E, Freathy RM, Lindgren CM *et al.* A common variant in the FTO gene is associated with body mass index and predisposes to childhood and adult obesity. *Science* 2007; **316**: 889–894.
19 Rankinen T, Zuberi A, Chagnon YC, Weisnagel SJ, Argyropoulos G, Walts B *et al.* The human obesity gene map: the 2005 update. *Obesity (Silver Spring)* 2006; **14**: 529–644.
20 Witte JS. Multiple prostate cancer risk variants on 8q24. *Nat Genet* 2007; **39**: 579–580.
21 Marchini J, Howie B, Myers S, McVean G, Donnelly P. A new multipoint method for genome-wide association studies by imputation of genotypes. *Nat Genet* 2007; **39**: 906–913.
22 Lin DY, Hu Y, Huang BE. Simple and efficient analysis of disease association with missing genotype data. *Am J Hum Genet* 2008; **82**: 444–452.
23 Altshuler D, Brooks LD, Chakravarti A, Collins FS, Daly MJ, Donnelly P. A haplotype map of the human genome. *Nature* 2005; **437**: 1299–1320.
24 Frazer KA, Ballinger DG, Cox DR, Hinds DA, Stuve LL, Gibbs RA *et al.* A second generation human haplotype map of over 3.1 million SNPs. *Nature* 2007; **449**: 851–861.
25 Huxley J, Mayr E, Osmond H, Hoffer A. Schizophrenia as a Genetic Morphism. *Nature* 1964; **204**: 220–221.
26 Konneker T, Barnes T, Furberg H, Losh M, Bulik CM, Sullivan PF. A searchable database of genetic evidence for psychiatric disorders. *Am J Med Genet B Neuropsychiatr Genet* 2008; **147**: 671–675.
27 American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*. 4th edn. American Psychiatric Association: Washington, DC, 1994.
28 Gottesman II, Gould TD. The endophenotype concept in psychiatry: etymology and strategic intentions. *Am J Psychiatry* 2003; **160**: 636–645.