

Editorial

Genome-wide association studies
in psychiatry: what have we learned?

Ann L. Collins and Patrick F. Sullivan

**Summary**

Genome-wide association studies (GWAS) have been the focus of considerable effort in psychiatry. These efforts have markedly increased knowledge of the genetic basis of psychiatric disorders, and yielded empirical data on genetic architecture critical to addressing long-standing debates in the field. There is now a clear path to increased knowledge of the 'parts lists' for these disorders.

Declaration of interest

P.F.S. was on the scientific advisory board for Expression Analysis, Durham, North Carolina, USA, and received unrestricted support for genetic research in schizophrenia from Eli Lilly.

Ann L. Collins (pictured) is a research associate in the Department of Genetics at the University of North Carolina at Chapel Hill. Patrick F. Sullivan is a distinguished professor in the Departments of Genetics and Psychiatry at the University of North Carolina at Chapel Hill.

Genome-wide association studies (GWAS) usually contrast the frequencies of genetic variants between cases and controls for a large set of genetic markers distributed across the genome (see Corvin *et al* for an introduction to these methods¹). These studies are notable for the large numbers of genetic markers (usually 500 000 to 1 million) and large sample sizes (often >10 000 participants). Since 2005, 1050 GWAS of human diseases and biometrical traits have been published.² These studies of 575 phenotypes implicated 2881 genetic variants at stringent levels of significance, often with compelling independent replication. Since 2007, 115 GWAS that focused on psychiatric disorders have appeared, most on Alzheimer's disease, bipolar disorder and schizophrenia, along with efforts in attention-deficit hyperactivity disorder (ADHD), autism, major depressive disorder, and licit and illicit drug use and dependence. Given these efforts, it is timely to ask what we have gained from this body of work.

Culture shifts

The efforts of GWAS have directly or indirectly led to important changes in the conduct of genetic studies. First, large-scale collaboration has become the norm in psychiatric genetics. Because the sample sizes needed for GWAS discovery and replication are beyond the reach of single groups, multiple consortia have emerged to foster scientific discovery. As an example, the Psychiatric Genomics Consortium (PGC)³ has approximately 300 investigators and more than 75 000 participants with GWAS data under analysis, and may be the largest consortium in the history of psychiatry.

Second, prompt sharing of results and full genomic data is now standard. This welcome change maximises progress towards understanding the genetic basis of critically important psychiatric disorders, and fosters reproducibility by allowing the independent evaluation of claims of association. For example, anyone can view or obtain the results from the PGC studies, and qualified investigators can obtain individual-level data in order to conduct additional analyses (see Appendix).

Third, as critically, uncompromising statistical rigor is now required. Genomic studies must explicitly account for the 10^5 – 10^6 statistical comparisons. Thresholds for declaring significance are severe but appropriate, and most journals require

replication in independent samples. Genome-wide association studies can observe dozens of associations with $P \sim 10^{-6}$ – 10^{-7} resulting from the play of chance. Most investigators now understand that 'intriguing biology' is irrelevant to establishing a robust genetic association. For example, we observed multiple associations of major depressive disorder with *PCLO* (piccolo) with $P \sim 10^{-7}$.⁴ The biology of *PCLO* is fascinating (among its other functions, *PCLO* 'touches' serotonin); however, this 'intriguing' association did not withstand the test of replication.^{4,5} The community standard in human genetics now requires significance well beyond chance plus replication in independent samples. The biology of a gene does not play a role in establishing the association.

Discoveries of genetic loci

The central goal of psychiatric genetics is to discover loci that are robustly and repeatedly associated with a disorder and thereby gain insight into aetiology (Table 1). Approximately 11 copy number variants (CNVs) have been discovered (included in this review as most were discovered using GWAS technology). These CNVs are rare (<0.1% in controls), potent (odds ratios (ORs) of 4–20), and often non-specific risk factors for psychiatric disorders.⁶ In addition, GWAS have strongly implicated common variation in approximately 30 different genomic loci for psychiatric disorders along with loci for nicotine (*CHRNA3*, *BNDF*, *CYP2A6*) and alcohol consumption (*ADH1B*, *ALDH2*, *AUTS2*). These genetic variants are relatively common (allele frequencies >5%) and subtly increase disease risk (ORs of 1.10–1.25).

Associations popular in the literature before 2007 (such as *COMT*, *DRD3*, *DRD2*, *HTR2A*, *NRG1*, *BDNF*, *DTNBP1* and *SLC6A4*) have generally not fared well in GWAS.⁷ The reasons why these candidate genes did not replicate are unclear. Although they cannot be definitely excluded from playing an aetiological role, their involvement now seems far less likely. A few associations have stood the test of time, and are limited to loci with unusually large effect sizes (Alzheimer's disease–*APOE*, schizophrenia–22q11.21, and alcohol dependence–alcohol metabolic genes).

More complex analyses

The basic analytic model used in most GWAS is very simple and considers single genetic markers in isolation. This simple model is not optimal given empirical data that psychiatric disorders are

Table 1 Empirical findings for psychiatric disorders ^a			
Finding, location or SNP	Band, nearest gene or finding	Class	Disorder
CNV			
chr1:145.0-148.0	1q21.1	Rare, potent	SCZ
chr2:50.1-51.2	2p16.3	Rare, potent	ASD, SCZ
chr3:195.7-197.3	3q29	Rare, potent	SCZ
chr7:72.7-74.1	7q11.23	Rare, potent	ASD
chr7:158.8-158.9	7q36.3	Rare, potent	SCZ
chr15:23.6-28.4	15q11.2	Rare, potent	ASD
chr15:30.9-33.5	15q13.3	Rare, potent	ADHD, ASD, SCZ
chr16:15.4-16.3	16p13.11	Rare, potent	ADHD
chr16:29.5-30.2	16p11.2	Rare, potent	ASD, SCZ
chr17:34.8-36.2	17q12	Rare, potent	ASD, SCZ
chr22:18.7-21.8	22q11.21	Rare, potent	ASD, SCZ
SNP			
rs3818361	<i>CR1</i>	Common, subtle	AD
rs744373	<i>BIN1</i>	Common, subtle	AD
rs9349407	<i>CD2AP</i>	Common, subtle	AD
rs11767557	<i>EPHA1</i>	Common, subtle	AD
rs11136000	<i>CLU</i>	Common, subtle	AD
rs610932	<i>MS4A cluster</i>	Common, subtle	AD
rs3851179	<i>PICALM</i>	Common, subtle	AD
rs3764650	<i>ABCA7</i>	Common, subtle	AD
rs2075650	<i>APOE, TOMM40</i>	Common, notably strong	AD
rs3865444	<i>CD33</i>	Common, subtle	AD
rs12576775	<i>ODZ4</i>	Common, subtle	BIP
rs4765913	<i>CACNA1C</i>	Common, subtle	BIP
rs1064395	<i>NCAN</i>	Common, subtle	BIP
rs1625579	<i>MIR137</i>	Common, subtle	SCZ
rs2312147	<i>VRK2</i>	Common, subtle	SCZ
rs1344706	<i>ZNF804A</i>	Common, subtle	SCZ
rs17662626	<i>PCGEM1</i>	Common, subtle	SCZ
rs13211507	<i>MHC</i>	Common, subtle	SCZ
rs7004635	<i>MMP16</i>	Common, subtle	SCZ
rs10503253	<i>CSMD1</i>	Common, subtle	SCZ
rs16887244	<i>LSM1</i>	Common, subtle	SCZ
rs7914558	<i>CNNM2</i>	Common, subtle	SCZ
rs11191580	<i>NT5C2</i>	Common, subtle	SCZ
rs11819869	<i>AMBRA1</i>	Common, subtle	SCZ
rs12807809	<i>NRGN</i>	Common, subtle	SCZ
rs12966547	<i>CCDC68</i>	Common, subtle	SCZ
rs9960767	<i>TCF4</i>	Common, subtle	SCZ
rs1344706	<i>ZNF804A</i>	Common, subtle	SCZ+BIP
rs2239547	<i>ITIH3-ITIH4</i>	Common, subtle	SCZ+BIP
rs10994359	<i>ANK3</i>	Common, subtle	SCZ+BIP
rs4765905	<i>CACNA1C</i>	Common, subtle	SCZ+BIP
Heritability			
	0.40	Common variation	BIP
	0.30	Common variation	SCZ
Pathway analysis			
	miR-137 network	Common variation	SCZ
	Cholesterol, innate immune	Common variation	AD
	Calcium signalling	Common variation	BIP
	Post-synaptic signalling	Rare CNVs	BIP, SCZ
Burden			
	Increased, multiple studies	Rare CNVs	ASD, SCZ
	Increased	Rare CNVs	ADHD, BIP, MDD
	Increased	Common variation	SCZ, BIP, MDD
<p>CNV, copy number variation; SNP, single nucleotide polymorphism; AD, Alzheimer's disease; ADHD, attention-deficit hyperactivity disorder; ASD, autism spectrum disorder; BIP, bipolar disorder; MDD, major depressive disorder; SCZ, schizophrenia.</p> <p>a. See Tables 2 and 3 in Sullivan <i>et al.</i>⁶ for full citations. The CNV and SNP findings meet genome-wide significance in large samples. Most are likely secure but some may not stand the test of time. The heritability, pathway and burden results have replicated in multiple samples and/or represent consistent results from different analytical methods. Genomic locations are NCBI Build 37/UCSC hg19.</p>			

polygenic, and analyses of sets of markers could provide further insight by better reflecting the fundamental genetic architecture. Of the ways in which multiple markers can be analysed in combination, three have been the focus of particular effort (Table 1). First, large sets of genetic markers can be used to estimate heritability. Unlike twin or family studies whose assumptions continue to be criticised, these approaches yield assessments of

heritability based directly on the genome. For schizophrenia and bipolar disorder, these results confirm that substantial proportions of the variance in liability (about 25%, or approximately a third to a half of the heritability) are accounted for by the current generation of genotyping arrays.⁸ Thus, the critical assumption that has driven a generation of genetic studies now seems particularly secure.

Second, pathway analyses evaluate whether associations in predefined sets of genes have smaller *P*-values than expected.⁶ These analyses have provided new ideas about the biology of these disorders as hypotheses for future work, and include: micro-RNA miR-137 and multiple genes containing binding sites for miR-137 in schizophrenia; calcium signalling in bipolar disorder; cholesterol metabolism and the innate immune response in Alzheimer's disease; and post-synaptic signalling in schizophrenia and bipolar disorder (via CNVs).

Third, it is possible to compare the 'burden' of a type of genetic variant between cases and controls.⁶ Individuals with autism or schizophrenia have consistently been reported to have a greater burden of rare CNVs than controls, and similar findings may hold for ADHD, bipolar disorder and major depressive disorder. For common genetic variation, people with schizophrenia, bipolar disorder and major depressive disorder have greater burden of risk alleles than controls. For schizophrenia, this finding is highly replicable and highly significant ($P < 10^{-25}$).⁶

Genetic architecture

Genome-wide association studies have provided real data about the genetic basis of psychiatric disorders. Genetic architecture refers to the number of loci conferring risk for a disorder and their frequencies, effect sizes, modes of action and interactions with other genetic loci and environmental factors. In the many, mostly philosophical debates on this topic in the past century, two extreme views have been articulated: psychiatric disorders are caused by rare mutations of strong effect with most cases having a different causal variant *v.* a causal model consisting of the cumulative effects of many common variants of relatively subtle effects.

Where the results are sufficient to afford the ability to judge, the answer is that both common and rare variants have roles. This general conclusion applies to Alzheimer's disease (rare mutations with very strong effects (*APP*, *PSEN1*, *PSEN2*) and ten common variants of far more subtle effects) and bipolar disorder and schizophrenia (although no Mendelian-like mutations have been identified, rare CNVs play a role along with multiple common variants). For autism spectrum disorders, the rare variant catalogue is more complete (rare Mendelian syndromes with autistic features like Rett syndrome, karyotype abnormalities in approximately 5% of individuals, and CNVs in 5–10% of individuals). A recent series of papers in *Nature* used exome sequencing to identify just three candidate genes (*SCN2A*, *KATNAL2*, *CHD8*) containing rare *de novo* variants. Indeed, the hypothesis that autism results only from many different Mendelian-like mutations could be rejected. The role of common variation in autism is currently unknown as the available GWAS samples are small by current standards.

Common variation has been implicated for alcohol and nicotine consumption. For ADHD, several rare CNVs have been reported. For anorexia nervosa, obsessive-compulsive disorder, post-traumatic stress disorder and Tourette syndrome, the published data are sparse, and the roles of common and rare variation are unknown. For all of these disorders, GWAS sample sizes are not large by current standards, and considerably smaller than the sample sizes that were required to identify robust and replicable findings for other biomedical diseases.

For major depressive disorder, a large GWAS mega-analysis (approximately 19 000 participants) failed to identify findings of genome-wide significance. In context, nearly all GWAS with sample sizes above 11 000 participants identified at least one genome-wide significant finding.

What have we learned?

For most researchers, the rare *v.* common variant debate is settled. Unsurprisingly, where there are reasonable amounts of data, the answers contain elements of both models. Although some have wished for these disorders to conform to a classical medical genetic model whereby the aetiology of complex psychiatric disorders would resolve into a series of highly penetrant mutations, such hopes have now been demonstrated to be inconsistent with results for Alzheimer's disease, autism and schizophrenia. The 'many Mendelians' model now seems to be very unlikely.

Psychiatric disorders are polygenic. The rare variant and the common variant results indicate that many different loci are involved in Alzheimer's disease, bipolar disorder, schizophrenia, autism spectrum disorders and drug consumption. A parsimonious hypothesis is that these variants encode or regulate multicomponent biological pathways.⁹ Several intriguing hypotheses have emerged from GWAS data that suggest novel mechanisms underlying these disorders.

There are more discoveries to be made. From the experiences of other biomedical disorders for which GWAS have been conspicuously successful (for example type 2 diabetes mellitus or inflammatory bowel disease), we can confidently project that larger studies will yield more robust and replicable findings. To our knowledge, this is the first time in the history of psychiatry where there is a clear path to increasing our fundamental understanding of these disorders.

Ann L. Collins, PhD, Department of Genetics, University of North Carolina at Chapel Hill, North Carolina; **Patrick F. Sullivan**, MD, FRANZCP, Department of Genetics and Department of Psychiatry, University of North Carolina at Chapel Hill, North Carolina, USA

Correspondence: Patrick F. Sullivan, Department of Genetics, CB#7264, 5097 Genomic Medicine Building, University of North Carolina, Chapel Hill, NC 27599-7264, USA. Email: pfsullivan@med.unc.edu

First received 6 Aug 2012, final revision 17 Aug 2012, accepted 21 Aug 2012

Funding

A multitude of sources have funded the research summarised in this article, including governmental funds and philanthropy but considerable effort has been donated by individual researchers. The PGC has been funded by the NIMH via MH085520 and MH094421.

Acknowledgements

We are indebted to the tens of thousands of individuals who chose to participate in the work summarised here, and to the hundreds of scientific colleagues who have donated thousands of person-hours to these efforts.

Appendix

Web resources

Genome-wide results for PGC analyses for ADHD, bipolar disorder, major depressive disorder and schizophrenia can be freely downloaded (<https://pgc.unc.edu>) and visualised in a genomic regional context (www.broadinstitute.org/mpg/ricopili). Full results and individual data are available by application to the NIMH Repository (www.nimhgenetics.org), the database of Genotypes and Phenotypes (www.ncbi.nlm.nih.gov/dbgap), and/or the Wellcome Trust (www.wtccc.org.uk). The top findings for published GWAS are also available (www.genome.gov/26525384).

References

- 1 Corvin A, Craddock N, Sullivan PF. Genome-wide association studies: a primer. *Psychol Med* 2010; **40**: 1063–77.

- 2 National Human Genome Research Institute. *A Catalog of Published Genome-Wide Association Studies*. National Human Genome Research Institute, no date (<http://www.genome.gov/26525384>).
- 3 Psychiatric GWAS Consortium. A framework for interpreting genomewide association studies of psychiatric disorders. *Mol Psychiatry* 2009; **14**: 10–7.
- 4 Sullivan P, de Geus E, Willemsen G, James MR, Smit JH, Zandbelt T, et al. Genomewide association for major depressive disorder: a possible role for the presynaptic protein piccolo. *Mol Psychiatry* 2009; **14**: 359–75.
- 5 Major Depressive Disorder Working Group of the PGC. A mega-analysis of genome-wide association studies for major depressive disorder. *Mol Psychiatry* 2012; April 3 (Epub ahead of print).
- 6 Sullivan PF, Daly MJ, O'Donovan M. Genetic architectures of psychiatric disorders: the emerging picture and its implications. *Nat Rev Genet* 2012; **13**: 537–51.
- 7 Collins AL, Kim Y, Sklar P, O'Donovan MC, Sullivan PF. Hypothesis-driven candidate genes for schizophrenia compared to genome-wide association results. *Psychol Med* 2012; **42**: 607–16.
- 8 Visscher PM, Brown MA, McCarthy MI, Yang J. Five years of GWAS discovery. *Am J Hum Genet* 2012; **90**: 7–24.
- 9 Sullivan PF. Schizophrenia as a pathway disease. *Nat Med* 2012; **18**: 210–1.

reflection

Karl Jaspers

Andrew Sims

Karl Jaspers cast dappled sunshine onto the vagaries of human experience, illuminating and allowing us to make sense of other people, especially our patients, as they are rather than through the distorting prism of our own preconceptions. This has given greater understanding of our patients, improving clinical practice. Jaspers' distinction between form and content in the subjective experience of those psychiatrically distressed has greatly helped us to formulate their complaints. Other leaders of psychiatric thought have been more glamorous, with more literary *éclat* and greater hold on the imagination of the general public, but none has helped so much in developing that unique skill of the psychiatrist, empathy.

Jaspers' *General Psychopathology* has an unusual place among psychiatric texts. In Germany, *Allgemeine Psychopathologie* has a pre-eminent position, with eight editions, and it is well known by all psychiatrists. In the USA, Jaspers is considered more as a philosopher, and his work has not influenced clinicians to a great extent, except indirectly through the successive editions of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM). I would contend that, although *General Psychopathology* is not widely read now by British psychiatrists, it is the bedrock of much of our clinical practice and has profoundly influenced diagnosis: it is fundamental to both the major nosological systems, ICD and DSM. For this, we are indebted to the English translation of the seventh edition by Hoenig and Hamilton in 1963.

When I started my psychiatric training in Manchester in 1967, John Hoenig was senior lecturer in psychiatry (he subsequently became professor at Memorial University, Newfoundland) and Marian Hamilton was lecturer in psychiatric social work, both in the small Department of Psychiatry based at Gaskell House, near Manchester Royal Infirmary. There were two conceptual strands dominating Manchester psychiatry at that time: epidemiology, presided over powerfully by Professor Neil Kessel, newly arrived from Edinburgh, and descriptive psychopathology, championed by Hoenig and the major interest of Professor E. W. Anderson, recently retired from the Department. Both approaches were clamouring for the allegiance of trainees, and I believe I was very fortunate to have had those two ideologies, and the dynamic tension engendered, as fundamental to my subsequent psychiatry, although the conflict at the time of acquiring them was not comfortable!

Since then, for better or worse, I have thought and taught that descriptive psychopathology (largely based on Jaspers' monumental work) and psychiatric epidemiology are the foundations of clinical psychiatry. This synthesis of dissimilar partners has influenced my research, for example, on the outcome of non-psychotic disorders, my clinical practice with diverse types of psychiatric patients, the postgraduate training of future psychiatrists, and also my attitudes and opinions in the public arena when I was President of the Royal College of Psychiatrists.

I am not much given to hero worship but Karl Jaspers is certainly up there with the greatest in my personal constellation of psychiatric luminaries. Through the English translation he has had as much significance for British-influenced psychiatry around the world as any native English speaker or British psychiatrist.

A series of 'Reflections on Karl Jaspers' commemorates the centenary of the first publication of his *Allgemeine Psychopathologie* in 1913.

The British Journal of Psychiatry (2013)
202, 4. doi: 10.1192/bjp.bp.111.098541