LETTERS

Common polygenic variation contributes to risk of schizophrenia and bipolar disorder

The International Schizophrenia Consortium*

Schizophrenia is a severe mental disorder with a lifetime risk of about 1%, characterized by hallucinations, delusions and cognitive deficits, with heritability estimated at up to 80%[1,2]. We performed a genome-wide association study of 3,322 European individuals with schizophrenia and 3,587 controls. Here we show, using two analytic approaches, the extent to which common genetic variation underlies the risk of schizophrenia. First, we implicate the major histocompatibility complex. Second, we provide molecular genetic evidence for a substantial polygenic component to the risk of schizophrenia involving thousands of common alleles of very small effect. We show that this component also contributes to the risk of bipolar disorder, but not to several non-psychiatric diseases.

We genotyped the International Schizophrenia Consortium (ISC) case-control sample for up to ~1 million single nucleotide polymorphisms (SNPs), augmented by imputed common HapMap SNPs. In the genome-wide association study (GWAS; genomic control $\lambda_{GC} = 1.09$; Supplementary Table 1 and Supplementary Figs 1–3), the most associated genotyped SNP ($P = 3.4 \times 10^{-7}$) was located in the first intron of myosin XVIIIB (*MYO18B*) on chromosome 22. The second strongest association comprised more than 450 SNPs on chromosome 6p spanning the major histocompatibility complex (MHC; Fig. 1). There is some evidence for between-site heterogeneity in both allele frequencies and odds ratios (Table 1). We observed associations consistent with previous reports in the 22q11.2 deletion region and *ZNF804A* (ref. 3) (Supplementary Table 2, Supplementary Fig. 2 and section 5 and 6 in Supplementary Information).

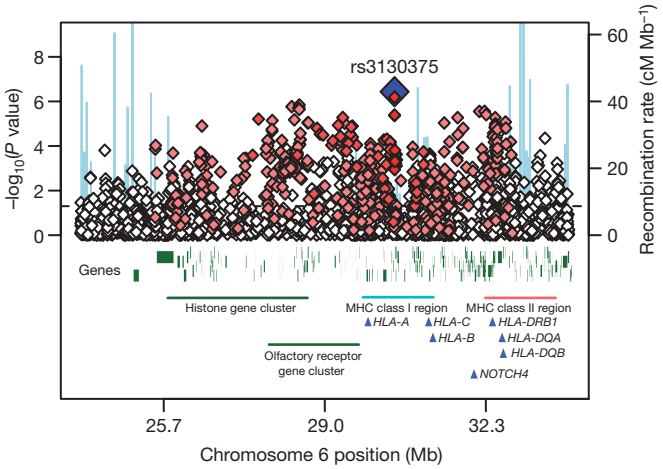

Figure 1 | Association results across the MHC region. Results are shown as $-\log_{10}$(*P* value) for genotyped SNPs. The most associated SNP is shown as a blue diamond. The colour of the remaining markers reflects r^2 with rs3130375, light pink, $r^2 > 0.1$, red, $r^2 > 0.8$. The recombination rate from the CEU HapMap (second *y* axis) is plotted in light blue.

The best imputed SNP, which reached genome-wide significance (rs3130297, $P = 4.79 \times 10^{-8}$, T allele odds ratio = 0.747, minor allele frequency (MAF) = 0.114, 32.3 megabases (Mb)), was also in the MHC, 7 kilobases (kb) from *NOTCH4*, a gene with previously reported associations with schizophrenia[4]. We imputed classical human leukocyte antigen (HLA) alleles; six were significant at $P < 10^{-3}$, found on the ancestral European haplotype[5] (Table 1, Supplementary Table 3 and section 3 in Supplementary Information). However, it was not possible to ascribe the association to a specific HLA allele, haplotype or region (Supplementary Table 3 and Supplementary Fig. 4).

We exchanged GWAS summary results with the Molecular Genetics of Schizophrenia (MGS) and SGENE consortia for genotyped SNPs with $P < 10^{-3}$. There were 8,008 cases and 19,077 controls of European descent in the combined sample (see refs 6, 7 and section 7 in Supplementary Information). Our top genotyped MHC SNP (rs3130375) had $P = 0.086$ and $P = 0.14$ in MGS and SGENE, respectively. Considering the combined results for genotyped and imputed SNPs across the MHC region more broadly, rs13194053 had a genome-wide significant combined $P = 9.5 \times 10^{-9}$ (ISC, MGS and SGENE: $P = 3 \times 10^{-4}$, 1×10^{-2} and 1×10^{-4}, respectively; C allele

Table 1 | MHC association for the most significant genotyped SNP rs3130375

a MHC association for rs3130375 by sample

Sample	Ancestry	Frequency (rs3130375, A allele) Cases	Controls	*P* value
University of Aberdeen	Scottish	0.132	0.168	0.0060
University of Edinburgh	Scottish	0.137	0.135	0.8930
University College London*	British	0.132	0.143	0.4836
Trinity College Dublin	Irish	0.110	0.170	0.0012
Cardiff University	Bulgarian	0.077	0.084	0.5602
Portuguese Island Collection	Portuguese	0.048	0.061	0.3510
Karolinska Institutet (5.0)	Swedish	0.043	0.119	0.0004
Karolinska Institutet (6.0)	Swedish	0.089	0.142	0.0040

b MHC association for classical HLA alleles with $P < 1 \times 10^{-3}$

HLA allele	Frequency†	Odds ratio	*P* value
*HLA-A*0101*	0.103	0.785	4×10^{-5}
*HLA-C*0701*	0.113	0.778	5×10^{-5}
*HLA-B*0801*	0.068	0.757	3×10^{-5}
*HLA-DRB*0301*	0.121	0.768	3×10^{-6}
*HLA-DQB*0201*	0.210	0.857	4×10^{-4}
*HLA-DQA*0501*	0.205	0.798	6×10^{-7}

Total sample Cochran–Mantel–Haenszel $P = 4 \times 10^{-7}$; Breslow–Day heterogeneity test $P = 0.012$ (d.f. = 6).
* SNP failed genotyping quality control in UCL. Allele frequency for UCL based on imputed genotypes.
† Frequency is estimated population frequency.

*Lists of authors and their affiliations appear at the end of the paper.

odds ratio = 0.82, 0.88 and 0.78) and was in linkage disequilibrium with rs3130375 ($r^2 = 0.35$ in HapMap). Across the region, 11 other SNPs had $P < 10^{-7}$ at 27.1–27.3 Mb and 32.7 Mb (Supplementary Table 5).

Our second approach was to evaluate whether common variants have an important role en masse, directly testing the classic theory of polygenic inheritance[8], previously hypothesized to apply to schizophrenia[9]. Although our GWAS analysis did not identify a large number of strongly associated loci, there could still be potentially thousands of very small individual effects that collectively account for a substantial proportion of variation in risk. We summarized variation across nominally associated loci into quantitative scores, and related the scores to disease state in independent samples[10]. Although variants of small effect (for example, genotypic relative risk (GRR) = 1.05) are unlikely to achieve even nominally significant P values, increasing proportions will be detected at increasingly liberal significance thresholds (P_T), for example, $P_T < 0.1$ or $P_T < 0.5$. Using such thresholds, we defined large sets of 'score alleles' in a discovery sample, to generate aggregate risk scores for individuals in independent target samples. We use the term score, instead of risk, as we cannot differentiate the minority of true risk alleles from unassociated variants.

We performed the score analyses on a reduced set of SNPs to facilitate analysis and interpretation. After filtering on MAF, genotyping rate and linkage disequilibrium (independent of association with schizophrenia), we obtained a subset of 74,062 autosomal SNPs in approximate linkage equilibrium (Supplementary Tables 6 and 7). In each discovery sample, we selected sets of score alleles at different association test P_T thresholds. For each individual in the target sample, we calculated the number of score alleles they possessed, each weighted by the log odds ratio from the discovery sample. To assess whether the aggregate scores reflect schizophrenia risk, we tested for a higher mean score in target cases compared to controls (sections 9–11 in Supplementary Information and Supplementary Table 7).

We selected males (2,176 cases, 1,642 controls) and females (1,146 cases, 1,945 controls) to form arbitrary discovery and target samples (Supplementary Table 8). Score alleles designated in the discovery sample were significantly enriched among target cases, and the effect was larger for increasingly liberal P_T thresholds. The score on the basis of all SNPs with male discovery $P_T < 0.5$ ($n = 37{,}655$ SNPs) was highly correlated with schizophrenia in target females ($P = 9 \times 10^{-19}$), explaining ~3% of the variance (Nagelkerke's pseudo R^2 from logistic regression), with higher scores in cases. The results were not driven by only a few highly associated regions (section 12 in Supplementary Information).

We eliminated several possible confounders, with emphasis on subtle population stratification (Supplementary Tables 9–15). Defining score alleles in British Isles samples and testing in target samples from Sweden, Portugal and Bulgaria, and vice versa, we observed a similar pattern of results. It is unlikely that the same substructure is overrepresented in the corresponding phenotype class when discovery and target samples are from distinct populations. The effect is also stronger for SNPs within annotated genes (Supplementary Table 16).

We used independent GWAS samples to replicate the polygenic component, to examine whether this component is shared with bipolar disorder[11], and to demonstrate specificity by considering non-psychiatric diseases. We used the entire ISC for the discovery sample, considering the five most informative P_T thresholds from the intra-ISC analyses. The independent target samples were the MGS European-American (MGS-EA), the MGS African-American (MGS-AA) and the UK sample described previously by O'Donovan *et al*[3]. The ISC-derived score was highly associated with disease in both European schizophrenia samples (Fig. 2, Supplementary Fig. 6 and Supplementary Table 17). The MGS-EA had a significantly higher mean $P_T < 0.5$ score in cases compared to controls ($P = 2 \times 10^{-28}$, $R^2 = 3.2\%$), as did the smaller O'Donovan sample ($P = 5 \times 10^{-11}$,

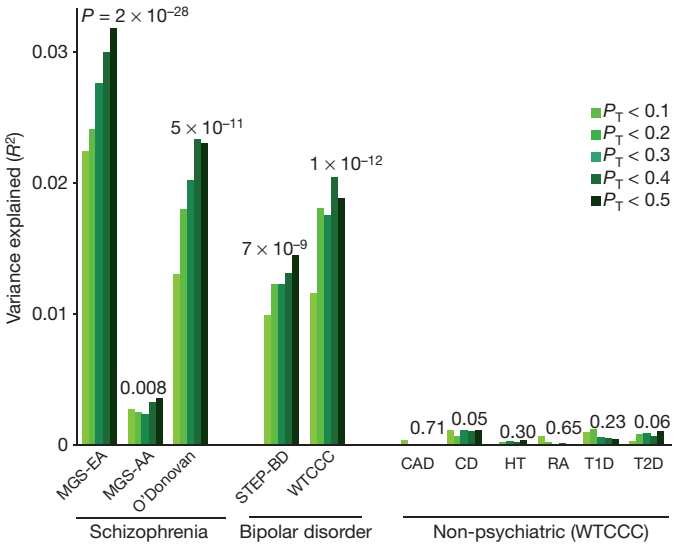

Figure 2 | Replication of the ISC-derived polygenic component in independent schizophrenia and bipolar disorder samples. Variance explained in the target samples on the basis of scores derived in the entire ISC for five significance thresholds ($P_T < 0.1$, 0.2, 0.3, 0.4 and 0.5, plotted left to right in each study). The y axis indicates Nagelkerke's pseudo R^2; the number above each set of bars is the P value for the $P_T < 0.5$ target sample analysis. CAD, coronary artery disease; CD, Crohn's disease; HT, hypertension; RA, rheumatoid arthritis; T1D, type I diabetes; T2D, type II diabetes. Numbers for cases/controls: MGS-EA 2,687/2,656; MGS-AA 1,287/973; O'Donovan 479/2,938; STEP-BD 955/1,498; WTCCC 1,829/2,935; CAD 1,926/2,935; CD 1,748/2,935; HT 1,952/2,935; RA 1,860/2,935; T1D 1,963/2,935; and T2D 1,924/2,935.

$R^2 = 2.3\%$). Aggregate differences in allele frequencies and patterns of linkage disequilibrium between Europeans and African-Americans are expected to lead to an attenuated effect. Still, MGS-AA cases carried more of the European-derived score alleles than the MGS-AA controls ($P = 0.008$; $R^2 = 0.4\%$).

The ISC-derived score alleles were also associated with bipolar disorder in two independent samples. Both samples, STEP-BD[12] and WTCCC[13], had higher mean $P_T < 0.5$ scores in cases than in controls ($P = 7 \times 10^{-9}$, $R^2 = 1.9\%$, and $P = 1 \times 10^{-12}$, $R^2 = 1.4\%$, respectively) indicating a substantial, shared genetic component.

To test disease specificity, we selected all six non-psychiatric WTCCC samples (coronary artery disease, Crohn's disease, hypertension, rheumatoid arthritis, type I and type II diabetes). Controls are shared among the WTCCC case samples, including bipolar disorder. In contrast to schizophrenia and bipolar disorder, there was no association ($P > 0.05$) between the ISC-derived schizophrenia scores and these non-psychiatric diseases, for any P_T threshold.

We next investigated the genetic models consistent with our data. The total additive genetic variance (V_A) reflects the number of causal alleles, as well as their frequency and effect size distributions. However, the variance explained by the markers that tag these causal alleles (V_M) will be attenuated, reflecting the average extent of linkage disequilibrium between marker and causal allele. In our target samples, the variance explained by the observed score alleles (V_S) will be further attenuated by sampling variation and P_T threshold, such that $V_S \leq V_M \leq V_A$.

We used simulation to estimate possible values for V_M and V_A, by identifying models that produced profiles of V_S across P_T threshold that were similar to those observed in the ISC data, as indexed by the target sample R^2. Under a variety of genetic models, we simulated discovery and target data sets of comparable sample size to the ISC. On the basis of the empirical allele frequency distribution, we simulated marker SNPs, varying the proportion that were in linkage disequilibrium with causal variants, for which we varied allele frequency (uniform, U-shaped) and effect size distributions (fixed

GRR values, exponential GRR values, or fixed variance explained) as well as the extent of linkage disequilibrium (section 16 in Supplementary Information).

From a broad range of models, a subset produced results consistent with the ISC data (Fig. 3 and Supplementary Fig. 7). Among these, all led to similar estimates of V_M (mean 34%, range 32% to 36%). In models in which the causal alleles were imperfectly tagged ($r^2 < 1$), estimates of V_A can be considerably larger. Therefore, our estimate that common polygenic variation accounts for one-third of the total variation in schizophrenia risk is a lower bound for the true value, which could be much higher. Figure 3b shows seven examples from the range of consistent models, detailed in Supplementary Table 18.

The simulated models consistent with our observed results all indicated a substantial number of common variants, whereas models that invoked only a few common variants of large effect or only rare variants were not able to account for our findings. For example, if $V_M \approx 34\%$ arose from only 100 common causal alleles, with GRR values at the tagging marker between ~1.2–1.5, most would be detected at $P_T < 0.01$, and so the variance explained would decline, not increase, as more SNPs were added (Fig. 3c and Supplementary Table 19). It is possible that an observed GRR of ~1.05 could represent a large effect of a weakly tagged rare variant, for example, a tenfold effect of a 1/10,000 variant in complete linkage disequilibrium ($D' = 1$, but low r^2) with a genotyped SNP. However, as this would only hold for low frequency markers (MAF < ~0.1), we stratified our analysis by score allele frequency (Fig. 4a). For simulated models in which all causal variants were of low frequency (<0.05), a stratified analysis revealed the expected, skewed distribution (Fig. 4c and section 17 in Supplementary Information), which was more pronounced for rarer causal alleles, for example, 1/1,000 (data not shown). In contrast, models in which causal alleles followed a uniform frequency distribution provided a closer fit to our data (Fig. 4b; although note some enrichment in the second quintile, of ~13–35% score alleles). Moreover, rare variants are likely to be population specific and if recurrent, in linkage disequilibrium with different common alleles within and between populations. As such, they could not account for the observation of disease variation that is largely shared across our different populations.

Decreased reproductive fitness in schizophrenia[14] suggests that risk alleles of large to moderate effect will be under negative selection and therefore very rare[15,16]. This is not inconsistent with our results, because

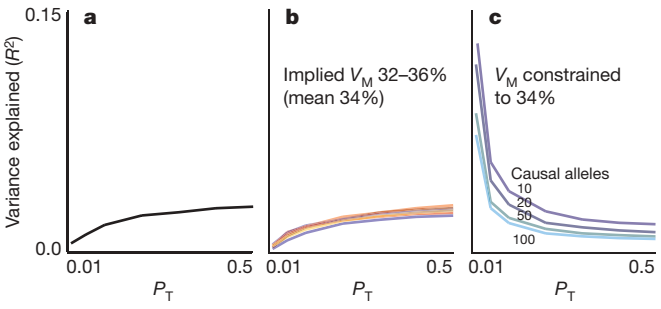

Figure 3 | Observed and simulated profiles of target sample variance explained. **a**, The observed variance explained is shown (R^2, black line). **b**, A subset of models that produced results consistent with the observed data is shown. All yielded similar estimates of the total variance explained by the SNPs that tag the causal variants, V_M, with a mean value of 34%. The seven models (shown as percentage SNPs, mean GRR/variance explained (V) per causal allele, linkage disequilibrium, and frequency model) were: M_1: 6.25%, GRR = 1.05, $r^2 = 1$, empirical; M_2: 25%, GRR = 1.025, $r^2 = 1$, empirical; M_3: 12%, GRR = 1.05, $r^2 < 1$, uniform; M_4: 32%, GRR = 1.04, $r^2 < 1$, U-shaped; M_5: 11%, $V = 0.00006$, $r^2 = 1$, empirical; M_6: 25%, GRR(exponential) = 1.025, $r^2 < 1$, uniform; M_7: 100%, GRR(exponential) = 1.012, $r^2 < 1$, uniform. **c**, Four inconsistent models with fewer variants of larger effect are shown.

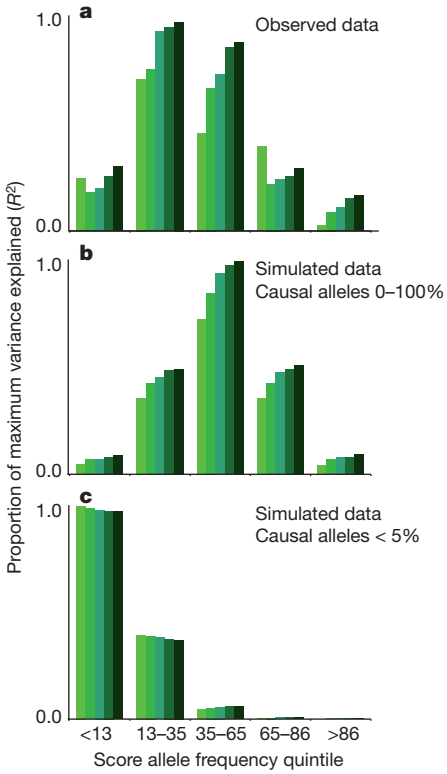

Figure 4 | Analysis stratified by score allele frequency. **a**, The observed data for the ISC/MGS-EA comparison is shown. The *y* axis is the target sample pseudo R^2, scaled within each figure as a proportion of the maximum value observed for five significance thresholds ($P_T < 0.1$, 0.2, 0.3, 0.4 and 0.5, plotted left to right in each quintile). **b**, **c**, Shown are results for simulated data: the common variant model, with a uniform frequency distribution for causal risk-increasing alleles (**b**) and a multiple rare variant model, in which the collective frequency of rare variants at a locus that all reside on the same haplotypic background with respect to the genotyped SNP was bounded at a maximum of 5% (**c**).

the common variants indexed by our polygenic score will not be subjected to strong selection, by virtue of their very small individual effect sizes. Our results do not exclude important contributions of rare variants for schizophrenia[15], because rare variants are expected as part of the allele frequency/effect size spectrum of a polygenic model. We and others recently reported higher genome-wide rates of rare copy number variants in schizophrenia[17–19]. However, our results indicate that medical sequencing and studies of structural variation to identify rare, highly penetrant variants will not alone fully characterize the genetic risk factors.

In conclusion, our molecular genetic data strongly support a polygenic basis to schizophrenia that (1) involves common SNPs, (2) explains at least one-third of the total variation in liability, (3) is substantially shared with bipolar disorder, and (4) is largely not shared with several non-psychiatric diseases. We also identified variants in the MHC region that received support in two independent studies, although the population specificity and extensive linkage disequilibrium will make follow-up challenging.

A highly polygenic model suggests that genetically influenced individual differences across domains of brain development and function may form a diathesis for major psychiatric illness, perhaps as multiple growth and metabolic pathways influence human height[20]. Our results may also reflect heterogeneity, such that some patients have aetiologically distinct diseases. The shared genetic liability between schizophrenia and bipolar disorder, previously suggested by clinical and genetic epidemiology[11,21], opens up the possibility of genetically based refinements in diagnosis. However, the scores derived here have little value for individual risk prediction, meaning that application

to clinical genetic testing for schizophrenia would be unwarranted. In the future, measures of polygenic burden, along with known risk loci and non-genetic factors such as season of birth, life stress, obstetrical complications, viral infections and epigenetics, could open new avenues for studying gene–gene and gene–environment interactions.

Increasing the discovery sample size should substantially refine the polygenic scores derived here. The variance explained by the observed score increases from ~3% to over 20% in extended simulations of 20,000 case/control pairs, as will soon be available by international meta-analytic efforts such as the Psychiatric GWAS Consortium[22–24] (section 18 in Supplementary Information and Supplementary Fig. 8). Furthermore, analyses that focus on gene pathways, clinical features and non-additivity may increase the variance captured by the score and identify genes or biological systems that are either shared by, or unique to, schizophrenia and bipolar disorder.

We identified fewer unambiguously associated variants than studies of some non-psychiatric diseases of comparable size[25]. Nonetheless, for other diseases replicated variants typically account for only a modest fraction of risk. The nature of this 'missing heritability' is a general problem now faced by complex disease geneticists[26]. For schizophrenia, our data point to a genetic architecture that includes many common variants of small effect. The extent to which similar models characterize genetic variation within and across other complex diseases remains to be investigated.

METHODS SUMMARY

Cases satisfied criteria for schizophrenia. Clinical characteristics and copy number variation have been described previously[17]. DNA was extracted from whole blood, with approval from institutional review boards. Genotypes were called using the Birdseed/Birdsuite algorithm[27] and analyses were performed with PLINK version 1.05 (ref. 28). Association analyses used a Cochran–Mantel–Haenszel test and logistic regression with covariates for sample site and ancestry. In the simulations, we generated data sets with pairs of unobserved variants and marker SNPs in varying degrees of within-pair linkage disequilibrium, on the basis of the effective number of independent SNPs in the ISC, and assuming Hardy–Weinberg equilibrium and linkage equilibrium between different pairs of SNPs. We considered a large grid of possible values for allele frequency and effect size distributions, also varying the proportion of non-null variants and the linkage disequilibrium between causal allele and observed marker. We retained models that produced similar profiles of target sample R^2 compared to the original ISC analysis, for the same range of P_T thresholds, and calculated the indicated total genetic variance under these models, assuming additivity within and across loci. See Supplementary Information for details.

Received 11 February; accepted 8 June 2009.
Published online 1 July; corrected 6 August 2009 (see full-text HTML version for details).

1. Cardno, A. G. & Gottesman, I. I. Twin studies of schizophrenia: from bow-and-arrow concordances to star wars Mx and functional genomics. *Am. J. Med. Genet.* **97,** 12–17 (2000).
2. Sullivan, P. F., Kendler, K. S. & Neale, M. C. Schizophrenia as a complex trait: evidence from a meta-analysis of twin studies. *Arch. Gen. Psychiatry* **60,** 1187–1192 (2003).
3. O'Donovan, M. C. *et al.* Identification of loci associated with schizophrenia by genome-wide association and follow-up. *Nature Genet.* **40,** 1053–1055 (2008).
4. Wei, J. & Hemmings, G. P. The *NOTCH4* locus is associated with susceptibility to schizophrenia. *Nature Genet.* **25,** 376–377 (2000).
5. Horton, R. *et al.* Variation analysis and gene annotation of eight MHC haplotypes: the MHC Haplotype Project. *Immunogenetics* **60,** 1–18 (2008).
6. Stefansson, H. *et al.* Common variants conferring risk of schizophrenia. *Nature* doi:10.1038/nature08186 (this issue).
7. Douglas, J. S. *et al.* Common variants on chromosome 6p22.1 are associated with schizophrenia. *Nature* doi:10.1038/nature08192 (this issue).
8. Fisher, R. A. The correlation between relatives on the supposition of Mendelian inheritance. *Philos. Trans. R. Soc. Edinb.* **52,** 399–433 (1918).
9. Gottesman, I. I. & Shields, J. A polygenic theory of schizophrenia. *Proc. Natl Acad. Sci. USA* **58,** 199–205 (1967).
10. Wray, N. R., Goddard, M. E. & Visscher, P. M. Prediction of individual genetic risk to disease from genome-wide association studies. *Genome Res.* **17,** 1520–1528 (2007).
11. Craddock, N., O'Donovan, M. C. & Owen, M. J. Genes for schizophrenia and bipolar disorder? Implications for psychiatric nosology. *Schizophr. Bull.* **32,** 9–16 (2006).
12. Sklar, P. *et al.* Whole-genome association study of bipolar disorder. *Mol. Psychiatry* **13,** 558–569 (2008).
13. The Wellcome Trust Case Control Consortium. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature* **447,** 661–678 (2007).
14. Svensson, A. C., Lichtenstein, P., Sandin, S. & Hultman, C. M. Fertility of first-degree relatives of patients with schizophrenia: a three generation perspective. *Schizophr. Res.* **91,** 238–245 (2007).
15. McClellan, J. M., Susser, E. & King, M. C. Schizophrenia: a common disease caused by multiple rare alleles. *Br. J. Psychiatry* **190,** 194–199 (2007).
16. Craddock, N., O'Donovan, M. C. & Owen, M. J. Phenotypic and genetic complexity of psychosis. Invited commentary on. Schizophrenia: a common disease caused by multiple rare alleles. *Br. J. Psychiatry* **190,** 200–203 (2007).
17. International Schizophrenia Consortium. Rare chromosomal deletions and duplications increase risk of schizophrenia. *Nature* **455,** 237–241 (2008).
18. Walsh, T. *et al.* Rare structural variants disrupt multiple genes in neurodevelopmental pathways in schizophrenia. *Science* **320,** 539–543 (2008).
19. Xu, B. *et al.* Strong association of *de novo* copy number mutations with sporadic schizophrenia. *Nature Genet.* **40,** 880–885 (2008).
20. Weedon, M. N. *et al.* Genome-wide association analysis identifies 20 loci that influence adult height. *Nature Genet.* **40,** 575–583 (2008).
21. Lichtenstein, P. *et al.* Common genetic determinants of schizophrenia and bipolar disorder in Swedish families: a population-based study. *Lancet* **373,** 234–239 (2009).
22. Psychiatric GWAS Consortium Steering Committee. A framework for interpreting genome-wide association studies of psychiatric disorders. *Mol. Psychiatry* **14,** 10–17 (2009).
23. Psychiatric GWAS Consortium Coordinating Committee. Genomewide association studies: history, rationale and prospects for psychiatric disorders. *Am. J. Psychiatry* **166,** 540–556 (2009).
24. Cross Disorder Phenotype Group of the Psychiatric GWAS Consortium. Dissecting the phenotype in genome-wide association studies of psychiatric illness. *Br. J. Psychiatry* doi:10.1192/bjp.bp.108.063156 (in the press).
25. Manolio, T. A., Brooks, L. D. & Collins, F. S. A. HapMap harvest of insights into the genetics of common disease. *J. Clin. Invest.* **118,** 1590–1605 (2008).
26. Maher, B. The case of the missing heritability. *Nature* **456,** 18–21 (2008).
27. Korn, J. M. *et al.* Integrated genotype calling and association analysis of SNPs, common copy number polymorphisms and rare CNVs. *Nature Genet.* **40,** 1253–1260 (2008).
28. Purcell, S. *et al.* PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **81,** 559–575 (2007).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank the patients and families who contributed to these studies. We also thank E. Lander, N. Patterson and members of the Medical and Population Genetics group at the Broad Institute of Harvard and Massachusetts Institute of Technology for valuable discussion, and members of the Broad Biological Samples and Genetic Analysis Platforms for sample management and genotyping. We particularly thank D. Levinson and P. Gejman for allowing access to the MGS samples, and J. Shi for analytic support with the MGS samples. The group at the Stanley Center for Psychiatric Research at the Broad Institute was supported by the Stanley Medical Research Institute (E.M.S.), the Sylvan C. Herman Foundation (E.M.S.), and MH071681 (P.S.). The Cardiff University group was supported by a Medical Research Council (UK) Programme grant and the National Institutes of Mental Health (USA) (CONTE: 2 P50 MH066392-05A1). The group at the Karolinska Institutet was supported by the Swedish Council for Working Life and Social Research (FO 184/2000; 2001-2368). The Massachusetts General Hospital group was supported by the Stanley Medical Research Institute (P.S.), MH071681 and MH077139 (P.S.) and a Narsad Young Investigator Award (S.M.P.). The group at the Queensland Institute of Medical Research was supported by the Australian National Health and Medical Research Council (grants 389892, 442915, 496688 and 496674) and thanks S. Gordon for data preparation. The Trinity College Dublin group was supported by Science Foundation Ireland, the Health Research Board (Ireland), the Stanley Medical Research Institute and the Wellcome Trust; Irish controls were supplied by J. McPartlin from the Trinity College Biobank. The work at the University of Aberdeen was partly funded by GlaxoSmithKline and Generation Scotland, Genetics Health Initiative. University College London clinical and control samples were collected with support from the Neuroscience Research Charitable Trust, the Camden and Islington Mental Health and Social Care Trust, East London and City Mental Heath Trust, the West Berkshire NHS Trust, the West London Mental Health Trust, Oxfordshire and Buckinghamshire Mental Health Partnership NHS Trust, South Essex Partnership NHS Foundation Trust, Gloucestershire Partnership NHS Foundation Trust, Mersey Care NHS Trust, Hampshire Partnership NHS Trust and the North East London Mental Health Trust. The collection of the University of Edinburgh cohort was supported by the Wellcome Trust Clinical Research Facility (Edinburgh) and grants from The Wellcome Trust, London and the Chief Scientist Office of the Scottish Government. The group at the University of North Carolina, Chapel Hill, was supported by MH074027, MH077139 and MH080403, the Sylvan C. Herman Foundation (P.F.S.) and the Stanley Medical Research Institute (P.F.S.). The group at the University of Southern California thanks the patients and their families for their collaboration, and acknowledges the support of the National Institutes of Mental Health and the Department of Veterans Affairs.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to P.S. (sklar@chgr.mgh.harvard.edu) and S.M.P. (shaun@pngu.mgh.harvard.edu).

The International Schizophrenia Consortium

Manuscript preparation Shaun M. Purcell[1,2,3,4], Naomi R. Wray[5], Jennifer L. Stone[1,2,3,4], Peter M. Visscher[5], Michael C. O'Donovan[6], Patrick F. Sullivan[7], Pamela Sklar[1,2,3,4]; **Data analysis** Shaun M. Purcell[1,2,3,4] (Leader), Jennifer L. Stone[1,2,3,4], GWAS analysis subgroup: Patrick F. Sullivan[7], Douglas M. Ruderfer[1,2,3,4], Andrew McQuillin[8], Derek W. Morris[9], Colm T. O'Dushlaine[9], Aiden Corvin[9], Peter A. Holmans[6], Michael C. O'Donovan[6], Pamela Sklar[1,2,3,4], Polygene analyses subgroup: Naomi R. Wray[5], Stuart Macgregor[5], Pamela Sklar[1,2,3,4], Patrick F. Sullivan[7], Michael C. O'Donovan[6], Peter M. Visscher[5]; **Management committee** Hugh Gurling[8], Douglas H. R. Blackwood[10], Aiden Corvin[9], Nick J. Craddock[6], Michael Gill[9], Christina M. Hultman[11,12], George K. Kirov[6], Paul Lichtenstein[11], Andrew McQuillin[8], Walter J. Muir[10], Michael C. O'Donovan[6], Michael J. Owen[6], Carlos N. Pato[13], Shaun M. Purcell[1,2,3,4], Edward M. Scolnick[2,3], David St Clair[14], Jennifer L. Stone[1,2,3,4], Patrick F. Sullivan[7], Pamela Sklar[1,2,3,4] (Leader); **Cardiff University** Michael C. O'Donovan[6], George K. Kirov[6], Nick J. Craddock[6], Peter A. Holmans[6], Nigel M. Williams[6], Lyudmila Georgieva[6], Ivan Nikolov[6], N. Norton[6], H. Williams[6], Draga Toncheva[16], Vihra Milanova[17], Michael J. Owen[6]; **Karolinska Institutet/University of North Carolina at Chapel Hill** Christina M. Hultman[11,12], Paul Lichtenstein[11], Emma F. Thelander[11], Patrick Sullivan[7]; **Trinity College Dublin** Derek W. Morris[9], Colm T. O'Dushlaine[9], Elaine Kenny[9], Emma M. Quinn[9], Michael Gill[9], Aiden Corvin[9]; **University College London** Andrew McQuillin[8], Khalid Choudhury[8], Susmita Datta[8], Jonathan Pimm[8], Srinivasa Thirumalai[18], Vinay Puri[8], Robert Krasucki[8], Jacob Lawrence[8], Digby Quested[19], Nicholas Bass[8], Hugh Gurling[8]; **University of Aberdeen** Caroline Crombie[15], Gillian Fraser[15], Soh Leh Kuan[14], Nicholas Walker[20], David St Clair[14]; **University of Edinburgh** Douglas H. R. Blackwood[10], Walter J. Muir[10], Kevin A. McGhee[10], Ben Pickard[10], Pat Malloy[10], Alan W. Maclean[10], Margaret Van Beck[10]; **Queensland Institute of Medical Research** Naomi R. Wray[5], Stuart Macgregor[5], Peter M. Visscher[5]; **University of Southern California** Michele T. Pato[13], Helena Medeiros[13], Frank Middleton[21], Celia Carvalho[13], Christopher Morley[21], Ayman Fanous[13,22,23,24], David Conti[13], James A. Knowles[13], Carlos Paz Ferreira[25], Antonio Macedo[26], M. Helena Azevedo[26], Carlos N. Pato[13]; **Massachusetts General Hospital** Jennifer L. Stone[1,2,3,4], Douglas M. Ruderfer[1,2,3,4], Andrew N. Kirby[2,3,4], Manuel A. R. Ferreira[1,2,3,4], Mark J. Daly[2,3,4], Shaun M. Purcell[1,2,3,4], Pamela Sklar[1,2,3,4]; **Stanley Center for Psychiatric Research and Broad Institute of MIT and Harvard** Shaun M. Purcell[1,2,3,4], Jennifer L. Stone[1,2,3,4], Kimberly Chambert[3,4], Douglas M. Ruderfer[1,2,3,4], Finny Kuruvilla[4], Stacey B. Gabriel[4], Kristin Ardlie[4], Jennifer L. Moran[4], Mark J. Daly[2,3,4], Edward M. Scolnick[3,4], Pamela Sklar[1,2,3,4]

[1]Psychiatric and Neurodevelopmental Genetics Unit, [2]Center for Human Genetic Research, Massachusetts General Hospital, 185 Cambridge Street, Boston, Massachusetts 02114, USA. [3]Stanley Center for Psychiatric Research, The Broad Institute of Harvard and MIT, Cambridge, Massachusetts 02142, USA. [4]The Broad Institute of Harvard and MIT, Cambridge, Massachusetts 02142, USA. [5]Queensland Institute of Medical Research, 300 Herston Road, Brisbane, Queensland 4006, Australia. [6]MRC Centre for Neuropsychiatric Genetics and Genomics, Department of Psychological Medicine, School of Medicine, Cardiff University, Cardiff C14 4XN, UK. [7]Departments of Genetics, Psychiatry, and Epidemiology, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599, USA. [8]Molecular Psychiatry Laboratory, Research Department of Mental Health Sciences, University College London Medical School, Windeyer Institute of Medical Sciences, 46 Cleveland Street, London W1T 4JF, UK. [9]Neuropsychiatric Genetics Research Group, Department of Psychiatry and Institute of Molecular Medicine, Trinity College Dublin, Dublin 2, Ireland. [10]Division of Psychiatry, University of Edinburgh, Royal Edinburgh Hospital, Edinburgh EH10 5HF, UK. [11]Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, SE-171 77 Stockholm, Sweden. [12]Department of Neuroscience, Psychiatry, Ulleråker, Uppsala University, SE-750 17 Uppsala, Sweden. [13]Center for Genomic Psychiatry, University of Southern California, Los Angeles, California 90033, USA. [14]Institute of Medical Sciences, [15]Department of Mental Health, University of Aberdeen, Aberdeen AB25 2ZD, UK.[16]Department of Medical Genetics, University Hospital Maichin Dom, Sofia 1431, Bulgaria. [17]Department of Psychiatry, First Psychiatric Clinic, Alexander University Hospital, Sofia 1431, Bulgaria. [18]West Berkshire NHS Trust, 25 Erleigh Road, Reading RG3 5LR, UK. [19]Department of Psychiatry, University of Oxford, Warneford Hospital, Headington, Oxford OX3 7JX, UK. [20]Ravenscraig Hospital, Inverkip Road, Greenock PA16 9HA, UK. [21]State University of New York – Upstate Medical University, Syracuse, New York 13210, USA. [22]Washington VA Medical Center, Washington DC 20422, USA. [23]Department of Psychiatry, Georgetown University School of Medicine, Washington DC 20057, USA. [24]Department of Psychiatry, Virginia Commonwealth University, Richmond, Virginia 23298, USA. [25]Department of Psychiatry, Sao Miguel, 9500-310 Azores, Portugal. [26]Department of Psychiatry University of Coimbra, 3004-504 Coimbra, Portugal.