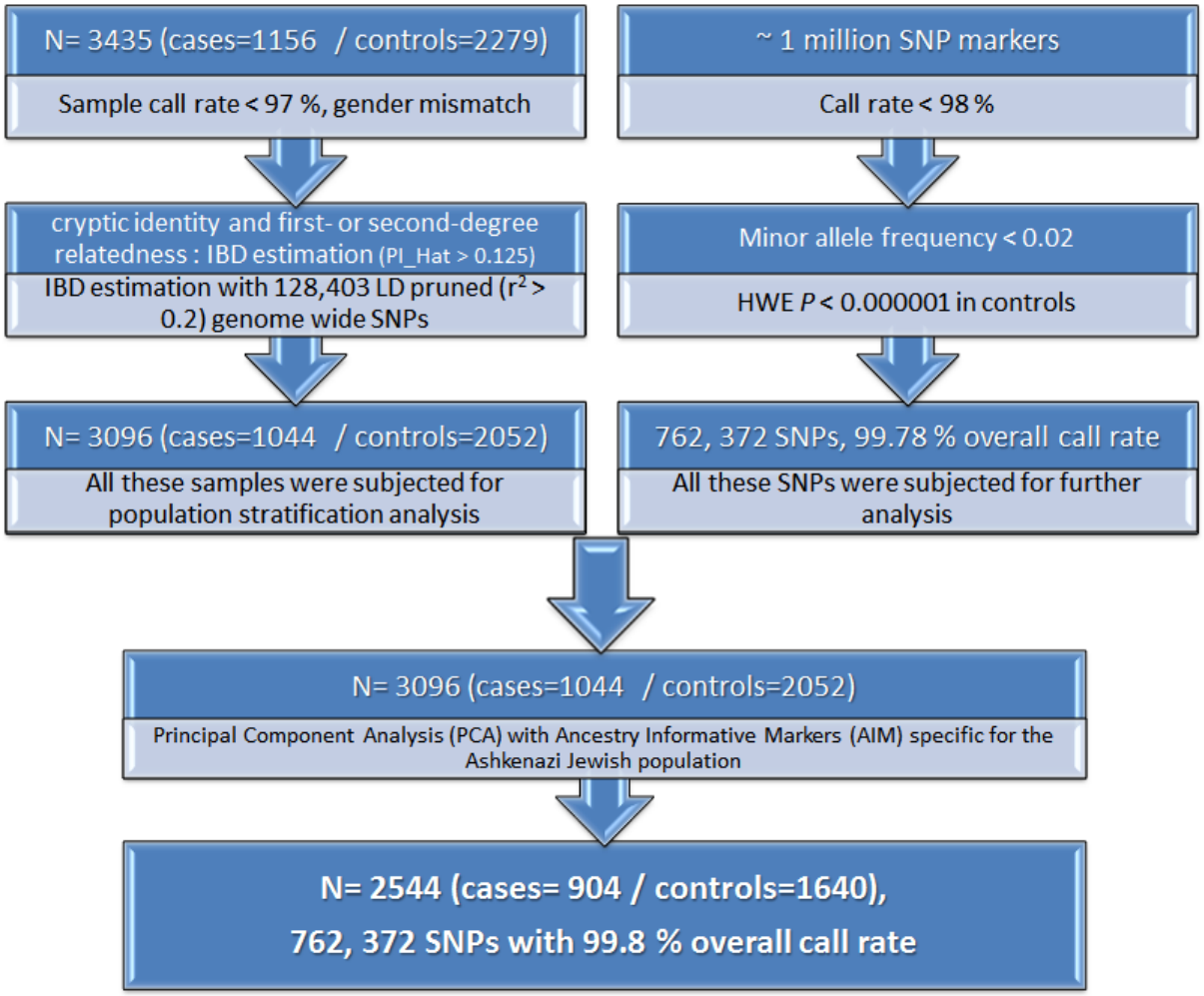
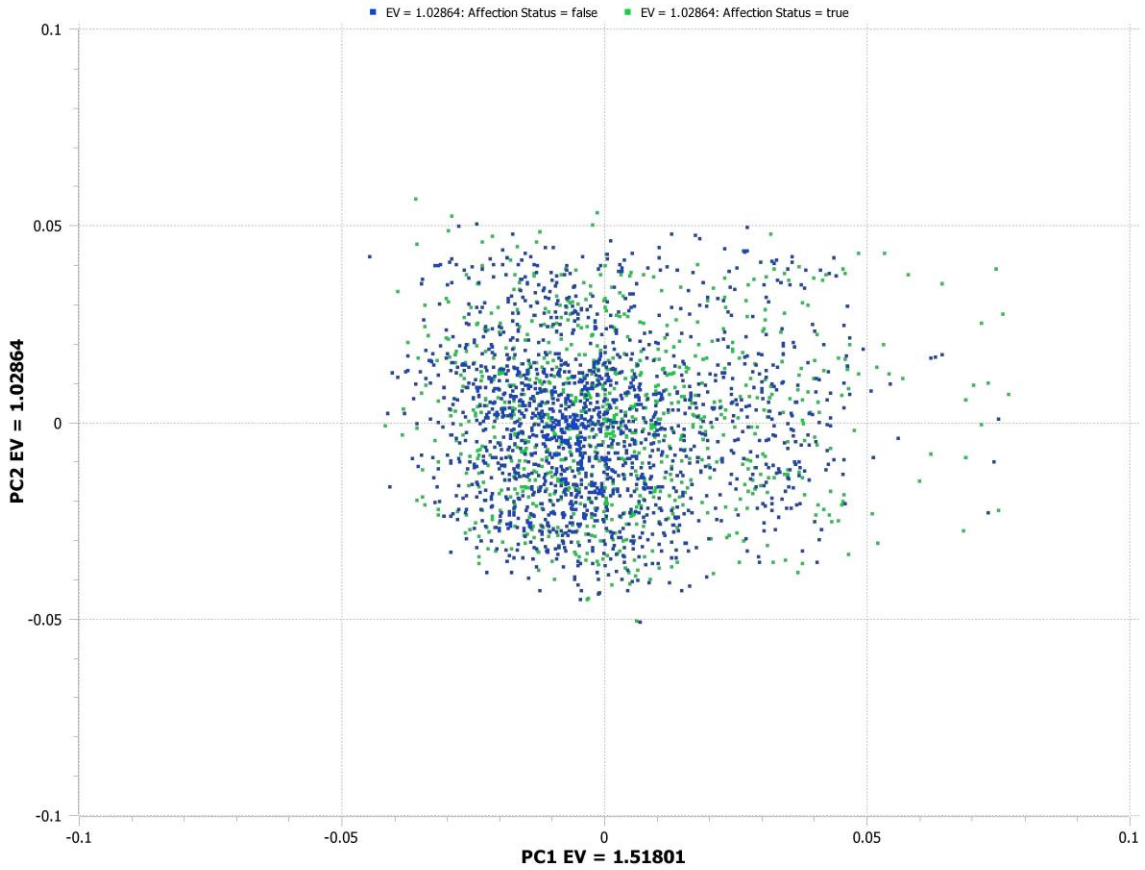
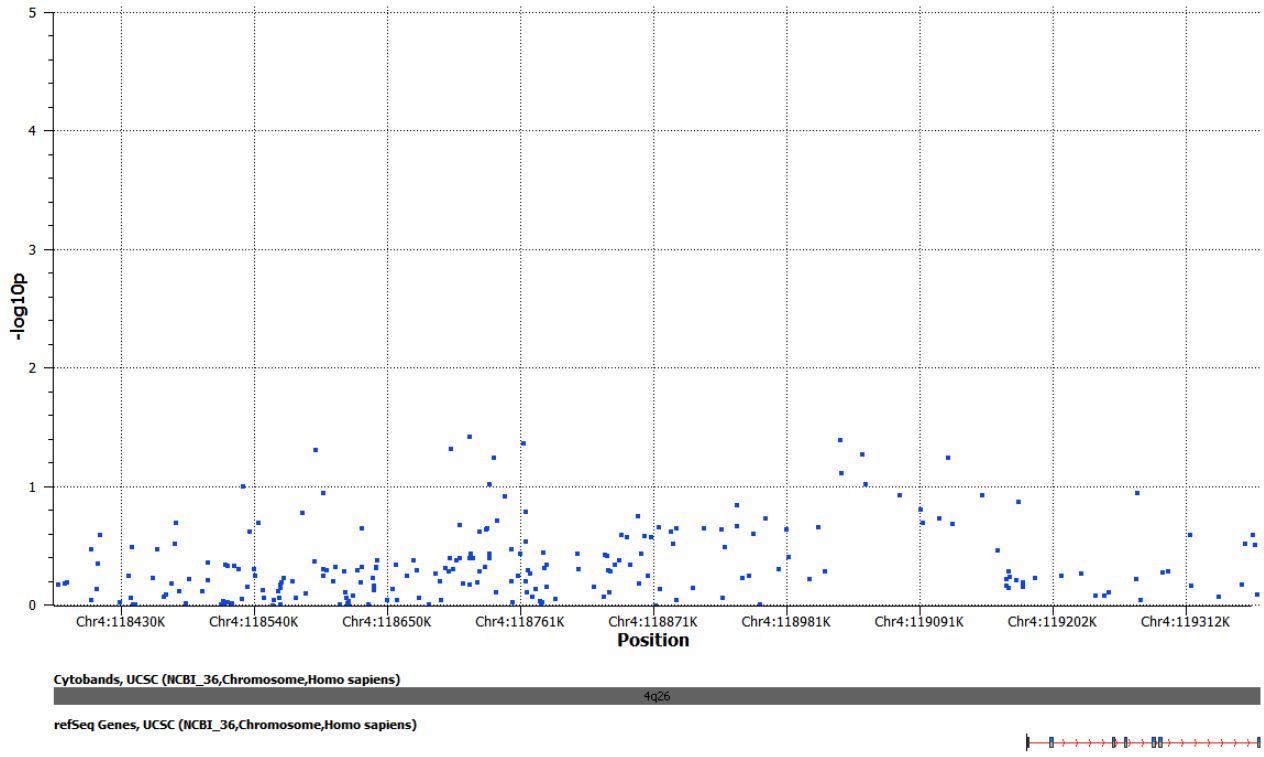
Supplementary Figures



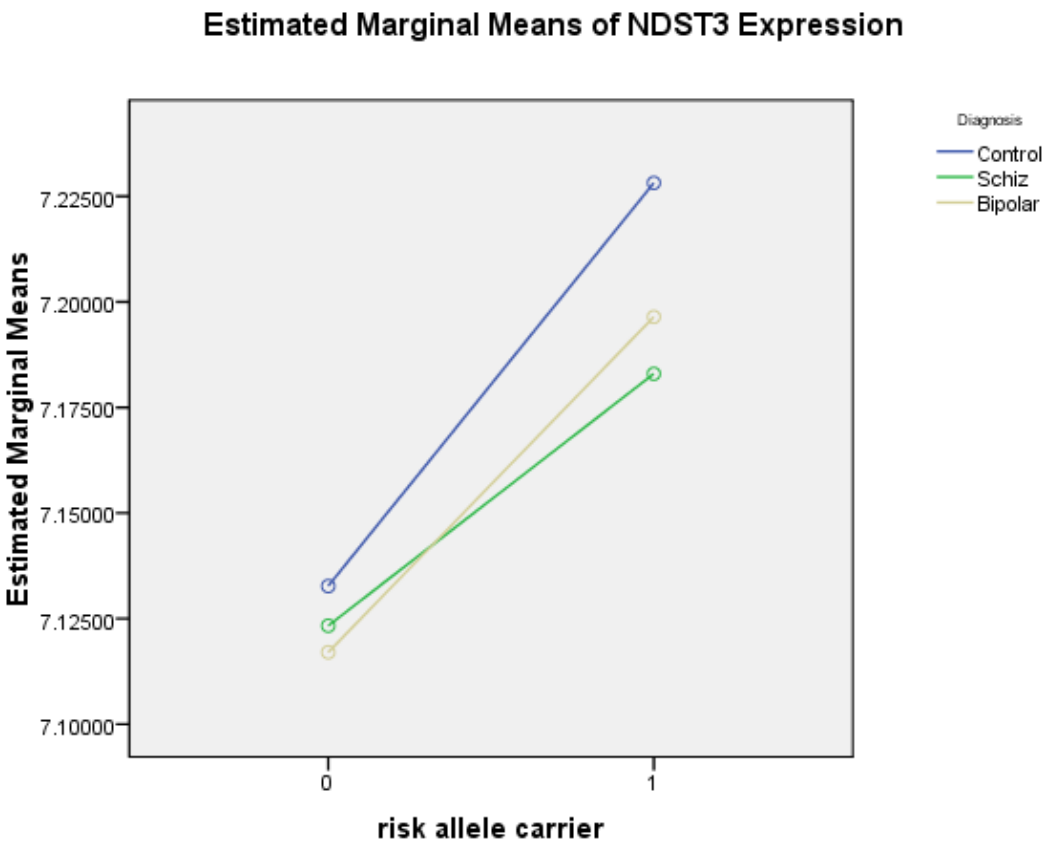
Supplementary Figure S1: Overview of the quality control workflow. These criteria were applied to the Ashkenazi Jewish discovery samples**.**



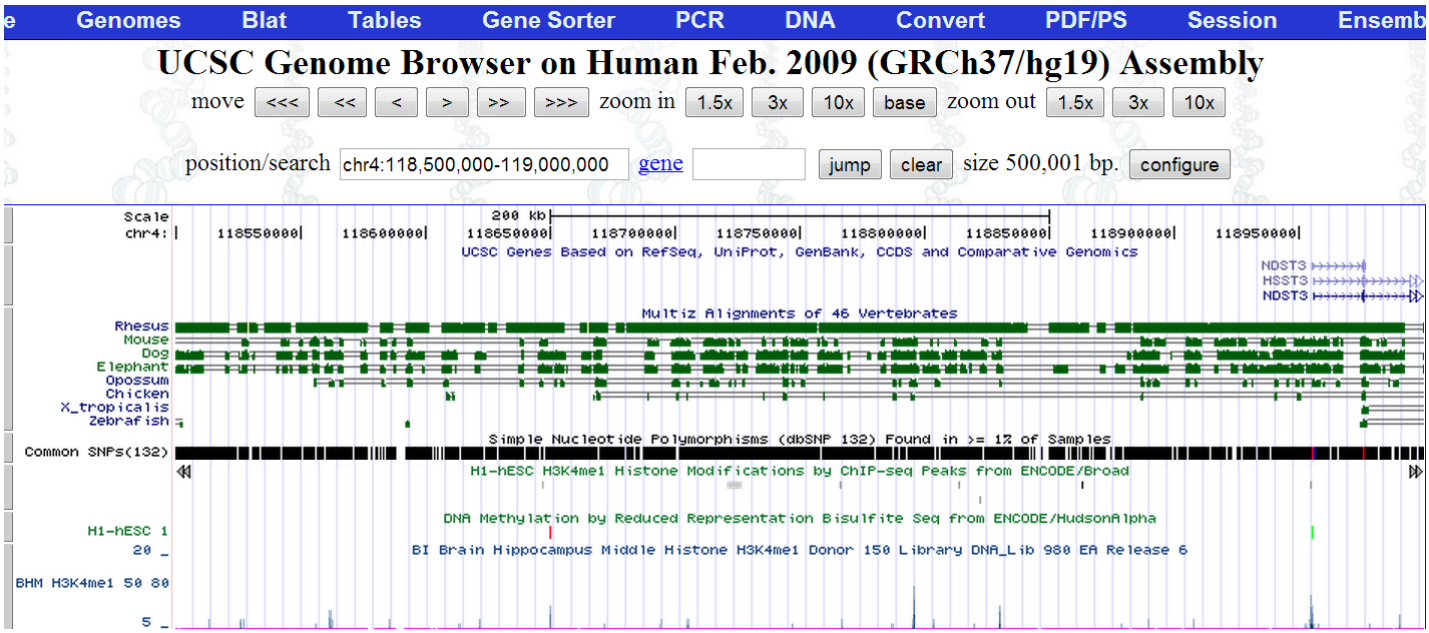
Supplementary Figure S2: Principal component analysis plot of 2544 cases and controls. The plot represents principal components 1 and 2 for 762,372 high-quality, genome-wide SNPs. Green dots represent cases and blue dots represent controls.



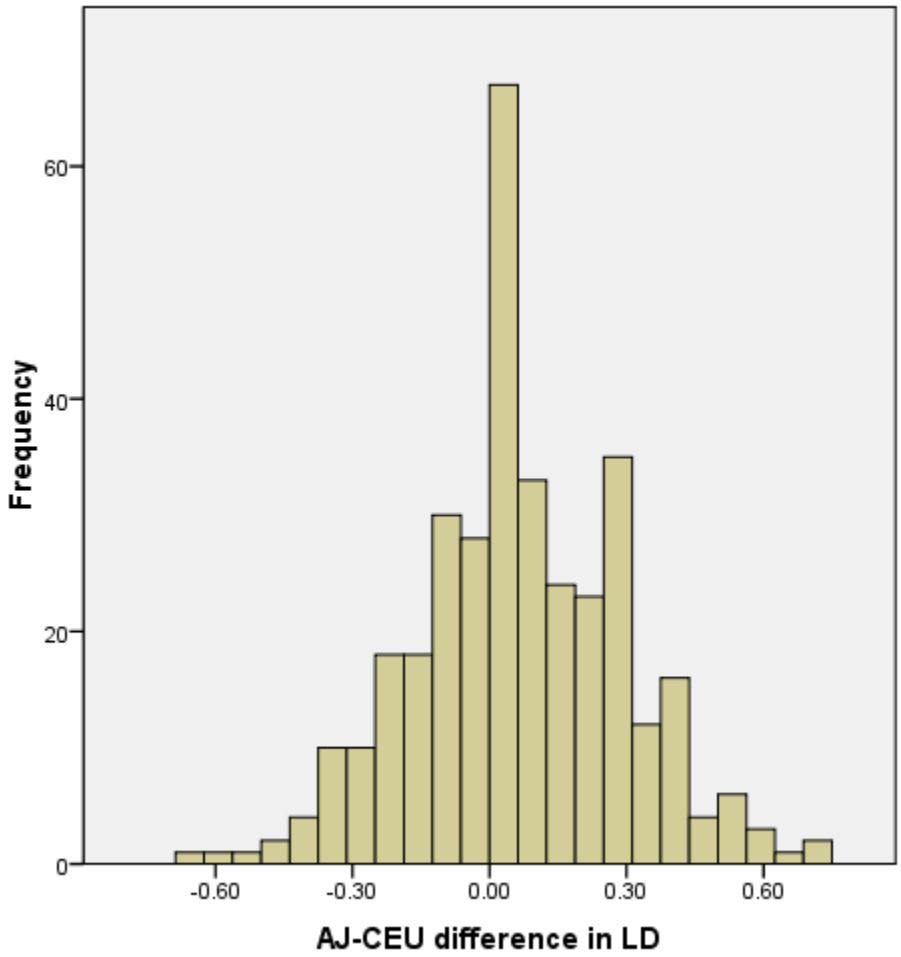
Supplementary Figure S3: Conditional analysis results after controlling for effects of rs11098403, within a 1MB window of this SNP. A case-control logistic regression (additive model) was used to generate the p-values. X-axis is labeled in hg18 coordinates.



Supplementary Figure S4: Expression of *NDST3* in postmortem cerebellar tissue as a function of genotype at rs11098403 and diagnostic group. For these analyses, heterozygotes were combined with the risk allele homozygotes; cell sizes were too small to permit modeling of additive allelic effects in each diagnostic group. Carriers of the schizophrenia risk allele (homozygous GG and heterozygous AG) demonstrate higher expression compared to non-carriers (AA homozygotes) in all three groups. There was no significant main effect of diagnostic group, and no genotype * diagnosis interaction effect. Y-axis represents estimated marginal mean expression values derived from ANCOVA analysis, as described in the text.



Supplementary Figure S5: *In silico* analysis of the LD block surrounding rs11098403. H3K4me1 activation marks and the unmethylated CpG island near rs11098403 are indicated with solid gray, blue, and red arrows as described in Figure 6. While these features are relatively uncommon in the intergenic region near *NDST3*, an additional hippocampal H3K4me1 mark is observed in the distal end of the LD block (dashed blue arrow), near rs4131119 (which showed a strong effect in the original discovery-cohort GWAS; logistic regression P=4.71E-07). For comparison, the H3K4me1 mark at the *NDST3* promoter is shown (dotted blue arrow).



Supplementary Figure S6: Differences in LD, measured by r^2, across populations (AJ vs CEU) for pairs of SNPs within 500kb of rs11098403. Positive values represent AJ>CEU; negative values represent CEU>AJ. Skew demonstrates enhanced LD in this region in the AJ population relative to the CEU.

Supplementary Note 1: Description of replication cohorts

MGS/GAIN Schizophrenia[49]

Schizophrenia cases were recruited from multiple sites across the United States and Australia. Caucasian subjects and African-American subjects were recruited for two separately analyzed, ethnically homogeneous sub-cohorts, as indicated by separate entries in Table 1. After obtaining written informed consent, all cases were interviewed using the Diagnostic Interview for Genetic Studies v2.0 (DIGS); the Family Interview for Genetic Studies (FIGS) was completed with an informant where possible; and medical records were obtained with the subject's written consent. Almost all cases had at least two sources of information, and two expert diagnosticians then made diagnoses based upon independent review of all available information. Final consensus best-estimate diagnoses were made using DSM-III-R or DSM-IV criteria, depending on the time of ascertainment; schizophrenia or schizoaffective disorder cases were included.

Control subjects were ascertained by two independent contractors (Knowledge Networks, and SSI/Opt-in) using random digit dialing and website banner ads. Using an IRB-approved procedure, individuals who responded were provided with an explanation of the study online (with an opportunity to phone for more information), and provided preliminary consent online. Participation required completion of an online clinical questionnaire. For those who completed the questionnaire, a national phlebotomy company (EMSI) then was asked to contact the individual and obtain a blood specimen. Participants signed written informed consent at that time. All control participants gave written consent for their biological materials and clinical questionnaire

responses to be used for any medical research at the discretion of NIMH. The questionnaire included: The Composite International Diagnostic Interview—Short Form, modified for lifetime (rather than 12-month) common mood, anxiety, and substance use disorders; items for lifetime diagnosis or treatment of schizophrenia or schizoaffective disorder, or hallucinations or delusions, or of bipolar disorder; ethnic origins of each grandparent (permitting multiple responses per grandparent); and a 12-item version of the Eysenck neuroticism and extraversion scales. Individuals were excluded from the control cohort if they endorsed (or failed to answer) any item for a psychotic or bipolar disorder, or if they were outliers in the number of missing items or of "yes" items in the questionnaire as a whole. Note that a subset of the Caucasian controls (n=1377) were fully overlapping with the controls for the GAIN bipolar study described below. This group of control subjects was divided in half randomly between the MGS schizophrenia cohort and the GAIN bipolar cohort. Genotyping of the samples was performed using the Affymetrix Genome-Wide Human SNP Array 6.0, and GWAS analyses utilized the first 5 PCA covariates for European-Americans and two covariates for African-Americans, as described in the original report.

Zucker Hillside Hospital (ZHH) Schizophrenia [50]

Patients with schizophrenia spectrum disorders (n = 178, including schizophrenia, n = 158; schizoaffective disorder, n = 13; or schizophreniform disorder, n = 7) were recruited from the inpatient and outpatient clinical services of The Zucker Hillside Hospital, a division of the North Shore–Long Island Jewish Health System. After providing written informed consent, the Structured Clinical Interview for DSM-IV Axis I disorders (SCID,

version 2.0) was administered by trained raters. Information obtained from the SCID was supplemented by a review of medical records and interviews with family informants when possible; all diagnostic information was compiled into a narrative case summary and presented to a consensus diagnostic committee, consisting of a minimum of three senior faculties. Healthy controls (n = 144) were recruited by use of local newspaper advertisements, flyers and community Internet resources, and underwent initial telephone screening to assess eligibility criteria. The nonpatient SCID (SCID-NP) was administered to subjects who met eligibility criteria, to rule out the presence of an Axis I psychiatric disorder; a urine toxicology screen for drug use and an assessment of the subject's family history of psychiatric disorders were also performed. Exclusion criteria included (current or past) Axis I psychiatric disorder, psychotropic drug treatment, substance abuse, a first-degree family member with an Axis I psychiatric disorder, or the inability to provide written informed consent. Patients (65 female/113 male) and controls (63 F/81 M) did not significantly differ in sex distribution ($P > 0.05$). All self-identified as Caucasian, non-Hispanic. Genotyping was performed using the Affymetrix 500K mapping array.

Japan Schizophrenia [51]

575 patients with schizophrenia (43.5 ±14.8 years) and 564 healthy controls (44.0 ± 14.4 years) were analyzed. All subjects were unrelated, living in the Tokai area of the mainland of Japan, and self-identified as Japanese. Patients were included if they 1) met DSM-IV criteria for schizophrenia, and 2) were physically healthy. Patients were excluded if they had a history of substance abuse, neurodevelopmental disorders, epilepsy or known mental retardation. Consensus diagnoses were made by at least two

experienced psychiatrists according to DSM-IV criteria on the basis of unstructured interviews with patients, their families and review of medical records. Controls were selected from the general population with no history of mental disorders based upon self report. After complete description of the study to the subjects, written informed consent was obtained. This study was approved by the ethics committees of each university participating in this project. Genotyping was performed using the Affymetrix Genome-Wide

Human SNP Array 5.0, and GWAS analyses utilized the first 10 PCA covariates.

Munich and Aberdeen Schizophrenia [52]

The Munich cohort included 439 schizophrenia patients and 418 healthy controls, all self-identifying as of German or central European ancestry. The Aberdeen cohort included 461 schizophrenia patients and 459 controls, all self-identifying as of Scottish or north European ancestry. Patient ascertainment criteria were consistent across both sites, requiring stable inpatient or outpatient status and a diagnosis of schizophrenia according to both DSM-IV and ICD-10 criteria based on detailed medical and psychiatric histories, including SCID interview. Exclusion criteria included a history of head injury or neurological diseases. The study was approved by both local and multiregional academic ethical committees, and all subjects gave informed consent.

Healthy volunteers were randomly ascertained from the local population (recruited by mail for Munich, and by general practitioners for Aberdeen). In both studies, volunteers were screened for absence of psychiatric disorders in themselves and first-degree

relatives. The Munich cohort was genotyped using the Illumina HumanHap300 chip with a total of 317,503 SNPs and the Aberdeen cohort was genotyped using the Illumina Human-Hap550 chip with a total of 555,352 SNPs. However, association analyses to identify schizophrenia risk factors focused on single-marker tests of the 312,565 QC-passed SNPs that were genotyped in both cohorts.

GAIN Bipolar [53]

Cases were drawn from those ascertained by the Bipolar Consortium over the course 18 years, including patients recruited as part of family-based linkage studies and those recruited as unrelated cases for association studies. After providing written informed consent to protocols approved at each participating institution, all subjects were diagnosed with a standard best estimate procedure, including a semi-structured interview (Diagnostic Interview for Genetic Studies). A total of 1,041 European-American (EA) individuals were included, although 40 subjects were ultimately removed due to ineligible or low-confidence diagnosis. EA status was determined based on the subject's self-report that all four grandparents were of EA heritage. The final sample thus included 1,001 EA cases of which 951 had a diagnosis of Bipolar I disorder 50 had a diagnosis of schizoaffective disorder, bipolar type.

Controls were ascertained separately through an NIMH-supported contract mechanism between Dr. Pablo Gejman and Knowledge Networks, Inc.; this mechanism allowed the ascertainment of 4,586 subjects across the U.S. who agreed to donate a blood sample for transformation into lymphoblastoid cell lines and to respond to a medical questionnaire.

All participating subjects, 3,303 EA were given the questionnaire. Only individuals with complete or near-complete psychiatric questionnaire data who did not fulfill diagnostic criteria for major depression, and denied a history of psychosis or bipolar disorder were included as controls for the BiGS analyses. Potential controls were matched for gender and ethnicity with the BiGS EA cases, and the control counts were 1,034 EA. As described above, all control participants gave written consent for their biological materials and clinical questionnaire responses to be used for any medical research at the discretion of NIMH. Genotyping of the samples was performed using the Affymetrix Genome-Wide Human SNP Array 6.0 using the first 4 PCA covariates for GWAS, as in the original report.

German Bipolar [54]

The patients included in the analysis received a lifetime diagnosis of BD according to the DSM-IV criteria on the basis of a consensus best-estimate procedure and structured diagnostic interviews. Protocols and procedures were approved by the local ethics committees, and written informed consent was obtained from all patients and controls. They were recruited from consecutive admissions to psychiatric inpatient units at The Central Institute of Mental Health, Mannheim (n = 1081). All GWAS controls were drawn from three population based epidemiological studies: (1) PopGen13 (n = 490), (2) KORA14 (n = 488), and (3) the Heinz Nixdorf Recall Study (Risk Factors, Evaluation of Coronary Calcification, and Lifestyle) (HNR, n = 383). Ancestry was assigned to patients and controls on the basis of self-reported ancestry. Genotyping was performed on HumanHap550v3 Illumina BeadArrays.

Wellcome Trust Case Control Consortium for Bipolar [55]

Subjects were drawn from a larger study of Caucasian individuals living in Great Britain. An initial cohort of 3000 control samples was drawn equally from two sources: unselected blood donors from the UK Blood Services and epidemiologically ascertained individuals from the 1958 British Birth Cohort study. A total of 2000 Caucasian bipolar cases were ascertained at five recruitment sites across Great Britain. After providing written informed consent, all patients were interviewed by a trained psychologist or psychiatrist using a semi-structured lifetime diagnostic psychiatric interview supplemented by available psychiatric records. Using all available data, the following lifetime psychiatric diagnoses were assigned according to the Research Diagnostic Criteria : bipolar I disorder (71% of all cases), schizoaffective disorder bipolar type (15%), bipolar II disorder (9%) and manic disorder (5%). All samples were genotyped using the Affymetrix 500K mapping array.

Taiwan Bipolar [56]

A total of 1409 unrelated bipolar I patients were consecutively recruited from inpatient units of these psychiatric departments and institutions, including 665 (47.2%) males and 744 females (52.8%). All of them were diagnosed according to DSM-IV criteria for bipolar I disorder with at least one inpatient treatment for manic episode, and nearly all of them (96.3%) had recurrent episodes of mania. Patients with other psychotic and affective disorders were excluded. Clinical phenotype assessment was performed by trained psychiatric nurses and psychiatrists using a cross-culturally validated and reliable

Chinese version of the Schedules for Clinical Assessment in Neuropsychiatry (SCAN), supplemented by available medical records. Diagnosis was generated from the computer diagnostic algorithm for the SCAN (I-Shell). Only the Han-Chinese population, which accounted for 98% of the population in Taiwan, was considered for the recruitment of patients in this study. The control group (N= 2000, with nearly equal proportions of men and women) was randomly selected from the Han-Chinese Cell and Genome Bank in Taiwan. In brief, more than 3300 healthy controls were recruited via a stratified, 3-staged probability clustering sampling scheme through the registry of all the 329 non-aboriginal townships or city districts in Taiwan. The study was approved by the institutional review board of all the participating hospitals and Academia Sinica, Taiwan and written informed consent were obtained from all of the participants. Analysis was conducted among the first 1000 cases and 1000 controls using the Illumina HumanHap550-Duo BeadChip.

Ashkenazi bipolar replication cohort

Bipolar cases (n=214) and controls (n=2531) samples were selected from an Ashkenazi Jewish repository (Hebrew University Genetic Resource, HUGR, http://hugr.huji.ac.il)). All controls were non-overlapping with samples utilized for the discovery GWAS cohort. Patients were recruited from hospitalized inpatients at seven medical centers in Israel. All diagnoses were assigned after direct interview using the structured clinical interview (SCID)[38], a questionnaire with inclusion and exclusion criteria, and cross-references to medical records. The inclusion criteria specified that subjects had to be diagnosed with bipolar I disorder by the Diagnostic and Statistical Manual of Mental Disorders (DSM-

IV), that all four grandparents of each subject were reported by the subject to be of Ashkenazi Jewish ethnic origin, and that each subject or the subject's legal representative has signed the informed-consent form. The exclusion criteria eliminated subjects diagnosed with at least one of the following disorders: schizophrenia or schizoaffective disorder, bipolar II disorder, psychotic disorder due to a general medical condition, substance-induced psychotic disorder, or any Cluster A (schizotypal, schizoid, or paranoid) personality disorder. Samples from healthy Ashkenazi individuals were collected from volunteers at the Israeli Blood Bank; these subjects were not psychiatrically screened but reported no chronic disease and were taking no medication at the time of blood draw. Corresponding institutional review boards and the National Genetic Committee of the Israeli Ministry of Health approved the studies. All samples were fully anonymized immediately after collection and subsequently, genomic DNA was extracted from blood samples through use of the Nucleon kit (Pharmacia).

Genotyping of bipolar cases were performed for ~ 1.4 million genome wide SNPs using Illumina HumanOmni1-Quad arrays according to manufacturers' specifications and under protocols approved by the Institutional Review Board of the North Shore-LIJ Health System.

Controls were genotyped for rs11098403 at Kbioscience (Hoddesdon, UK) using their propietary KASPar chemistry, which is a competitive allele-specific PCR SNP genotyping system using FRET (Fluorescent Resonance Energy Transfer) quencher cassette oligos (http://www.kbioscience.co.uk/reagents/KASP/KASP.html).

Supplementary References:

49. Shi, J. *et al*. Common variants on chromosome 6p22.1 are associated with schizophrenia. *Nature*. **460**,753-757 (2009).

50. Lencz, T. *et al*. Converging evidence for a pseudoautosomal cytokine receptor gene locus in schizophrenia. *Mol Psychiatry*. **12,** 572-580 (2007).

51. Ikeda, M. *et al*. Genome-wide association study of schizophrenia in a Japanese population. *Biol Psychiatry*. **69**, 472-478 (2011).

52. Need, A. C. *et al*. A genome-wide investigation of SNPs and CNVs in schizophrenia. *PLoS Genet*. **5**, e1000373 (2009).

53. Smith, E. N. *et al*. Genome-wide association study of bipolar disorder in European American and African American individuals. *Mol Psychiatry*. **14**, 755-763 (2009).

54. Cichon, S. *et al*. Genome-wide association study identifies genetic variation in neurocan as a susceptibility factor for bipolar disorder. *Am J Hum Genet*. **88**, 372-381 (2011).

55. 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).

56. Lee, M. T. *et al*. Genome-wide association study of bipolar I disorder in the Han Chinese population. *Mol Psychiatry*. **16**, 548-556 (2011).

57. Stefansson, H. *et al*. Common variants conferring risk of schizophrenia. *Nature*. **460**,744-747.

58. Steinberg, S. *et al*. Common variants at VRK2 and TCF4 conferring risk of schizophrenia. *Hum Mol Genet*.**20**,4076-4081 (2011).

59. Bergen, S.E. *et al*. Genome-wide association study in a Swedish population yields support for greater CNV and MHC involvement in schizophrenia compared with bipolar disorder. *Mol Psychiatry*.**17**,880-886 (2012).

60. Irish Schizophrenia Genomics Consortium and the Wellcome Trust Case Control Consortium 2. Genome-wide association study implicates HLA-C*01:02 as a risk factor at the major histocompatibility complex locus in schizophrenia. *Biol Psychiatry*.**72**,620-628 (2012).

61. Chen, D. T. *et al*. Genome-wide association study meta-analysis of European and Asian-ancestry samples identifies three novel loci associated with bipolar disorder. *Mol Psychiatry*.**18**,195-205 (2013).

.

62. Ferreira, M.A. *et al*. Collaborative genome-wide association analysis supports a role for ANK3 and CACNA1C in bipolar disorder. *Nat Genet*.**40**,1056-1058 (2008).

63. Green, E.K. *et al*. Replication of bipolar disorder susceptibility alleles and identification of two novel genome-wide significant associations in a new bipolar disorder case-control sample. *Mol Psychiatry*. 2012 Oct 16. [Epub ahead of print].

64. Vassos, E. *et al.* Replication study and meta-analysis in European samples supports association of the 3p21.1 locus with bipolar disorder. *Biol Psychiatry*.**72**,645-650 (2012).

65. O'Donovan, M.C. *et al*. Identification of loci associated with schizophrenia by genome-wide association and follow-up. *Nat Genet*.**40**,1053-1055 (2008).

66. Steinberg, S. *et al*. Common variant at 16p11.2 conferring risk of psychosis. *Mol Psychiatry*. doi: 10.1038/mp.2012.157 (2012).