HHS Public Access

Author manuscript

Mol Psychiatry. Author manuscript; available in PMC 2019 April 26.

Published in final edited form as:

Mol Psychiatry. 2011 February ; 16(2): 193–201. doi:10.1038/mp.2009.124.

Genome-wide association study of recurrent early-onset major depressive disorder

J Shi[1], **JB Potash**[2], **JA Knowles**[3], **MM Weissman**[4], **W Coryell**[5], **WA Scheftner**[6], **WB Lawson**[7], **JR DePaulo Jr.**[2], **PV Gejman**[8], **AR Sanders**[8], **JK Johnson**[1], **P Adams**[4], **S Chaudhury**[4], **D Jancic**[2], **O Evgrafov**[3], **A Zvinyatskovskiy**[1], **N Ertman**[1], **M Gladis**[1], **K Neimanas**[6], **M Goodell**[2], **N Hale**[5], **N Ney**[1], **R Verma**[2], **D Mirel**[9], **P Holmans**[10], and **DF Levinson**[1]

[1]Department of Psychiatry, Stanford University, Stanford. CA, USA

[2]Department of Psychiatry, Johns Hopkins University, Baltimore, MD, USA

[3]Department of Psychiatry, University of Southern California, Los Angeles, CA, USA

[4]Department of Psychiatry, Columbia University, New York, NY, USA

[5]Department of Psychiatry, University of Iowa, Iowa City, IW, USA

[6]Department of Psychiatry, Rush University Hospital, Chicago, IL, USA

[7]Department of Psychiatry, Howard University, Washington, DC, USA

[8]NorthShore University HealthCare Research Institute and Department of Psychiatry, Northwestern University, Evanston, IL, USA

[9]Genetic Analysis Platform, Broad Institute of Harvard and MIT, Cambridge, MA, USA

[10]Department of Psychological Medicine, Cardiff University, Cardiff, UK

Abstract

A genome-wide association study was carried out in 1,020 case subjects with recurrent early-onset major depressive disorder (MDD) (onset before age 31) and 1,636 control subjects screened to exclude lifetime MDD. Subjects were genotyped with the Affymetrix 6.0 platform. After extensive quality control procedures, 671,424 autosomal SNPs and 25,068 X chromosome SNPs with minor allele frequency greater than 1% were available for analysis. An additional 1,892,186 HapMap II SNPs were analyzed based on imputed genotypic data. Single-SNP logistic regression trend tests were computed, with correction for ancestry-informative principal component scores. No genome-wide significant evidence for association was observed, assuming that nominal $P < 5 \times 10^{-8}$

Correspondence: Douglas F. Levinson, M.D., Department of Psychiatry, Stanford University, 701 Welch Rd., Suite 3325, Palo Alto, CA 94305 (650-724-2827; fax 650-724-3262; dflev@stanford.edu).

Conflicts of Interest

The authors report no competing interests.

Supplementary Information accompanies the paper on the Molecular Psychiatry website (http://www.nature.com/mp). Note that one set of Supplementary Information is provided for this paper and for the companion article “Novel loci for major depression identified by genome-wide association study of STAR*D and meta-analysis of three studies” by Shyn et al.

approximates a 5% genome-wide significance threshold. The strongest evidence for association was observed on chromosome 18q22.1 (rs17077540, $P = 1.83 \times 10^{-7}$) in a region that has produced some evidence for linkage to bipolar I or II disorder in several studies, within an mRNA detected in human brain tissue (BC053410)and approximately 75 kb upstream of *DSEL*. Comparing these results with those of a meta-analysis of three MDD GWAS datasets reported in a companion article, we note that among the strongest signals observed in the GenRED sample, the meta-analysis provided the greatest support (although not at a genome-wide significant level) for association of MDD to SNPs within *SP4*, a brain-specific transcription factor. Larger samples will be required to confirm the hypothesis of association between MDD (and particularly the recurrent early-onset subtype) and common SNPs.

Keywords

major depressive disorder; genetics; GWAS; neuroscience; genotype

Introduction

Major depressive disorder (MDD) is a common psychiatric disorder with a lifetime prevalence of 10–15% in most large studies. Despite the availability of medication and psychotherapeutic treatments, recurrent or chronic course is common (60–80%)[1], often with comorbid anxiety or substance use disorders, substantial impact on family and work life and on physical health, and an approximately 4% risk of eventual suicide (higher in more severe cases).[2] MDD is diagnosed when an individual experiences one or more major depressive episodes in the absence of other diagnoses such as bipolar-I or -II disorder, schizoaffective disorder or schizophrenia. An episode is defined as two or more weeks during which the person experiences impaired functioning and five or more key symptoms (dysphoric mood, loss of enjoyment, suicidal thoughts or acts, agitated or slowed movements, guilty or self-denigrating feelings, fatigue, and disturbances of sleep, appetite, concentration).[3]

The heritability of MDD has been estimated at approximately 40% in population-based twin studies[4], and higher in clinical samples[5] or with repeated assessments.[6] There is an approximately three-fold increase in risk to first-degree relatives.[4] Risk is also increased by severe childhood trauma or parental loss[7], probably interacting with genetic vulnerability.[8] Recurrent episodes and early onset in probands predict greater familial risk, although the size of these effects is controversial.[4, 9–11] Women are at two-fold greater risk of MDD, and there are probably both common and independent genetic factors in men and women, with similar heritability.[12, 13] Although MDD is more frequent in relatives of probands with bipolar disorders and schizophrenia, those disorders are not more frequent in relatives of MDD probands.[14] The degree or nature of overlap in genetic factors underlying these disorders remains unclear.

The Genetics of Recurrent Early-Onset Depression (GenRED) project is creating a large clinical sample, based in the National Institute of Mental Health repository program, for molecular genetic studies of MDD. GenRED I recruited affected sibling pair families for linkage studies[10, 15, 16], and GenRED II is currently recruiting additional cases for

association studies. We have focused on probands with recurrent MDD, early age at onset and positive family history as indices of increased genetic risk.

We present here the results of an initial case-control genome-wide association study (GWAS) of recurrent early-onset MDD using single nucleotide polymorphism (SNP) array technology. A companion article presents a second MDD GWAS in the STAR*D antidepressant effectiveness trial sample and a meta-analysis of the GenRED, STAR*D and publicly-available Genetic Association Information Network (GAIN-MDD) GWAS datasets.[17] We did not observe genome-wide significant evidence for association in the GenRED or combined analyses. The meta-analysis demostrated increased statistical support for one of the top GenRED findings, for SNPs in the gene encoding the Sp4 transcription factor. Many of the most robust GWAS findings to date have been detected in samples of 10,000–20,000 cases (plus controls), with genotypic relative risks typically in the 1.1–1.2 range.[18] Thus, larger samples will be required to clarify whether any of our findings represent true associations, and whether recurrent, early-onset MDD is a uniquely valuable phenotype for association studies.

Methods and procedure

Subjects

Genotyping was attempted for 1,110 MDD cases (655 recruited by GenRED I and 455 by GenRED II). The 1,636 control subjects were selected from the Molecular Genetics of Schizophrenia (MGS)[19] sample. All subjects were of European ancestry. Clinical characteristics are summarized in Table 1 for samples that passed all quality control filters and were included in the analyses presented here.

Cases were recruited from clinical settings and through media and internet announcements and advertisements. After giving written informed consent, participants were interviewed by phone or in person using the Diagnostic Interview for Genetic Studies[20] version 3 (http://nimhgenetics.org). Two independent expert reviewers achieved consensus ratings of DSM-IV mood and comorbid disorder diagnoses and associated course of illness variables, based on the DIGS, a narrative summary, and available treatment records and/or informant reports from the Family Interview for Genetic Studies (FIGS). Eligible probands had an MDD diagnosis, two or more episodes (or one episode lasting at least three years), onset before age 31, at least one sibling or parent with recurrent MDD with onset before age 41, MDD independent of substance dependence (i.e., no lifetime dependence, prior to dependence, or after at least two years of remission from dependence), no diagnosis of bipolar or schizoaffective disorder or schizophrenia, and no suspected bipolar-I disorder in a parent or sibling.[10] In GenRED I, at least one affected sibling was directly interviewed. In GenRED II, MDD in a parent or sibling was documented by FIGS with the proband, supplemented when necessary by a telephone interview with a relative. Additional family history was obtained by FIGS. Because of the excess of female probands, we replaced some female GenRED I probands with a male sibling who met proband eligibility criteria.

MGS control subjects were recruited by Knowledge Networks, Inc. (Menlo Park, CA), a survey research company, from a nationally-representative marketing panel recruited by

random digit dialing methods.[19] Control participants consented to anonymization and deposition of their DNA and clinical information in the NIMH repository for use in any medical research. They completed an online questionnaire including a lifetime version of the Composite International Diagnostic Interview-Short Form (CIDI-SF) for common mood, anxiety and substance use diagnoses[21], supplemented by questions about schizophrenia, psychosis or bipolar disorder. After excluding those who endorsed or failed to answer these latter questions (or who were outliers in total number of items endorsed), MGS selected 2,653 European-ancestry controls who passed SNP QC. We then excluded controls who met CIDI-SF criteria for MDD, or who reported recurrent depression but missed MDD by one criterion, leaving 1,636 controls for analysis.

Genotyping and quality control (QC)

Samples were genotyped with the Affymetrix 6.0 genome-wide SNP array at the Broad Institute Center for Genotyping and Analysis (Cambridge, MA) in three batches: 863 controls, as part of the GAIN schizophrenia project (late 2007); GenRED cases (early 2008); and 773 controls, under MGS grant funding (mid-2008). Genotypes were called with Birdseed (version 2).[22] QC analyses were carried out using PLINK[23] supplemented by local software (see online Supplementary Methods for details). We excluded SNPs and control samples that failed either MGS or GenRED QC criteria, which were selected by determining thresholds that achieved a balance between a low genomic control λ value and inclusion of more data.

Criteria for included SNPs were call rate >97% for autosomes, >98% for chromosome X in females or >99% in males; minor allele frequency >1%; Hardy-Weinberg P-value $>10^{-6}$ in controls; <3 Mendelian errors detected in 30 MGS trios; <2 discordant duplicate genotypes in GenRED duplicates or <3 in 90 MGS specimens genotyped in both the GAIN and NonGAIN experiments; case-control call rate difference <2% for autosomes or <1% for X chromosome; and passing a 1df plate-effect test (no plate differing from all others with $P<10^{-8}$, or <2 plates with $P<10^{-4}$). Each SNP passed these criteria in both MGS and GenRED samples. After QC, 671,424 autosomal and 25,068 X chromosome SNPs were included.

Criteria for DNA samples were call rate >97%; non-outlier for mean heterozygosity across all SNPs and for ancestry principal component scores (EIGENSTRAT[24]); pairwise identity-by-descent estimates not >0.1 with many other samples (pairs of apparent relatives were also inspected and one retained); and non-ambiguous heterozygosity values for X chromosome SNPs in females (these samples were excluded only for X chromosome analyses if they passed autosomal QC). Additional details about population substructure analyses are provided in online Supplementary Methods. Further analyses included 1,020 cases and 1,636 controls.

Statistical analyses

Association between single genotyped SNPs and case-control status was tested with logistic regression (trend test) using PLINK. Genotypic dosages (the estimated number of test alleles) were imputed for all HapMap II SNPs with MACH 1.0 software[25] for autosomal

SNPs and with IMPUTE[26] for X chromosome SNPs, using a Hidden Markov Model algorithm and a training dataset consisting of phased HapMap CEU haplotypes. This provided an additional 1,892,186 SNPs (1,849,062 autosomal and 43,124 X chromosome SNPs) for testing in addition to the genotyped SNPs, after filtering for MAF > 1% and imputation $r^2 > 0.3$ (an estimate of expected agreement between imputed and actual genotypes). This threshold was used in four previous GWAS meta-analyses because it removed most poorly-imputed SNPs but few well-imputed SNPs.[27–30] Association tests for imputed SNPs were carried out with local software using the same logistic regression model. For all tests, ancestry-informative principal components were included as covariates.[24] Each SNP was tested for all subjects, and then separately for males and for females. For the primary analysis of all subjects, a reasonable threshold for 5% genome-wide significance is a nominal *P*-value less than 5×10^{-8}, based on three estimates assuming that all common SNPs have been directly or indirectly tested.[31–33] The analyses of male and of female subjects were considered exploratory.

As described in online Supplementary Results, we also separately examined results for SNPs in or near forty-one mood disorder candidate genes, including single-SNP tests and a permutation-based aggregate test (page S-22) of whether *P*-values in these genes were more significant than expected by chance.

Power analysis

Results of power analyses are shown in Table S3. In the primary analysis, for log additive transmission, power was 78% to detect a locus with MAF of 0.25 conferring a genotypic relative risk (GRR) to heterozygotes of 1.45, or 45% for MAF of 0.4 and GRR of 1.35.

Data sharing

Genotypic and clinical data are available to qualified scientists through controlled-access repository programs: the NIMH repository program (http://nimhgenetics.org). for the GenRED sample; and dbGAP (http://www.ncbi.nlm.nih.gov/gap) for MGS controls.

Results

Figure 1 illustrates results for the primary analysis and the quantile-quantile plot of observed vs. expected chi-square values for all genotyped and imputed SNPs. The genomic control λ value (the observed median chi-square divided by the expected median value of 0.456 under the null hypothesis) was 1.031, indicating that there was no meaningful inflation of test statistics.[34]

Table 2 lists results with $P < 10^{-5}$ for each of the three analyses. (Online file genred_supplementary_data.txt provides data for all results with $P < 0.001$ in any analysis.) Rows report data for the “best” SNP in independent regions (whether gene-containing or not) with $P < 10^{-5}$ for one or more SNPs. In most regions, many SNPs were in strong linkage disequilibrium (LD) and gave similar results. Genes and other functional elements are noted in the table if *P*-values less than 10^{-5} were observed either within the transcribed boundaries of the gene, or within 50 kb upstream or downstream, except that the closest genes are listed for some nongenic regions. Table S13 provides names of these genes

summarizes their known functional roles. Regions in Table 2 for which no gene or element is listed have peaks of similarity to known regulatory sequences by the ESPERR regulatory potential method available as a UCSC browser track.[35]

There were no genome-wide significant findings. Three of the top regions in the primary analysis also produced *P*-values less than 10^{-5} in the meta-analyses that are presented in a companion article.[36] (Low meta-analytic *P*-values required a consistent allelic direction of association across samples.) In the primary meta-analysis (Broad phenotype) that included all MDD cases from GenRED, STAR*D and GAIN-MDD, SNP rs17144465 in SP4 (7p14.3) had $P = 8.38 \times 10^{-7}$, lower than that for GenRED alone ($P = 5.97 \times 10^{-6}$). Another two of top GenRED regions (18q22.1 and 5p13.2) yielded $P < 10^{-5}$ In the Narrow meta-analysis (recurrent early-onset MDD in all samples). For 18q22.1 (rs17077540), the meta-analysis ($P = 7.55 \times 10^{-7}$) did not provide stronger support than GenRED alone (1.83×10^{-7}). For 5p13.2, the meta-analysis *P*-value (1.68×10^{-6}) was slightly smaller than for GenRED alone ($P = 2.49 \times 10^{-6}$). Please see the companion article for further details.

Figure 2 includes genome browser plots showing association *P*-values and relevant genomic information for three regions: 18q22.1 which produced the lowest GenRED *P*-value (all subjects); SP4/7p15.3, the best finding that received increased support in the meta-analysis; and 1p13.3, which was not supported in the meta-analysis, but contained the largest number of SNPs with the lowest *P*-values in two large LD blocks in the GenRED analysis, and which spans a set of interesting candidate genes. Genotyping cluster plots for top SNPs with $P < 10^{-6}$ (Table 2) or tags for those SNPs are provided in online file SNP_intensity_cluster_plots.pdf.

In the 41 mood disorder candidate genes (see Tables S5 and S6), the lowest *P*-value (0.000067) was observed in an intron of *CACNA1C* (calcium channel, voltage-dependent, L type), a gene with strong evidence for association to bipolar disorder.[37] For further details, see Table S6. The aggregate analysis did not yield evidence that the distribution of *P*-values in these genes was more significant than expected by chance.

Discussion

A GWAS of 1,020 recurrent early-onset MDD cases and 1,636 screened controls did not detect genome-wide significant evidence of association. This is consistent with other GWAS results for common, genetically complex diseases[38, 39]: the genotypic relative risks (GRR) of significant findings have typically been in the range of 1.1–1.2, often requiring samples 10,000–20,000 cases obtained by combining multiple samples (in each of which the evidence for association can be quite modest). The odds ratios listed in Table 2 are much higher. They could represent a combination of false positive results and of true associations whose GRRs have been over-estimated by selecting the best results in one study (the "winner's curse" effect), and particularly with an underpowered sample. A companion paper[17] provides details of meta-analyses combining the GenRED, STAR*D and GAIN-MDD samples. We are also participating in an effort to carry out larger meta-analyses of MDD GWAS data through the Psychiatric GWAS Consortium.[39, 40] But because findings

with the strongest statistical evidence for association in each study are most likely to be true positive results, we briefly review here several interesting findings.

Chromosome 18q22.1 (Figure 2) has produced suggestive evidence for linkage to bipolar disorder in several studies[41–47] (although not in the largest combined analysis[48]) and particularly in families with multiple bipolar-II (BP-II) cases[43, 44] (characterized by recurrent depression plus hypomania). Several studies have also reported suggestive linkage to MDD or related personality traits in the same region.[49] Study of a family with diverse mood disorders[46] identified *DSEL* (dermatan sulfate epimerase-like), a brain-expressed gene in which two non-synonymous mutations were observed in cases but not in controls.[50] Dermatan sulfate epimerase is involved in D-glucuronic acid metabolism and tumor rejection. The most strongly associated SNPs in the present study are upstream of *DSEL*, in regions with possible regulatory functions, and within an mRNA (BC053410), identified in pooled human brain tissue, that encodes a hypothetical protein (LOC643542) of unknown function. Familial co-aggregation of MDD and BP-II disorder has been inadequately studied. BP-II was an exclusion criterion for GenRED probands, although in GenRED I, BP-II was diagnosed in 4.2% of the siblings selected for interview because of a history of depression, so the prevalence among all siblings was lower.

On chromosome 1p13.3 (Figure 2), low *P*-values spanned two broad LD blocks that include many genes (Table 3), including those encoding two G proteins (*GNAI3*, *GNAT2*), a G-protein coupled receptor related to biogenic amine receptors (*GPR61*), a transcription factor gene resembling those involved in neurodegenerative syndromes (*ATXN7L2*), and genes involved in neuronal growth and plasticity (*AMIGO1*) and neuronal apoptosis (*SORT1*). Current hypotheses suggest that both G-protein coupled signaling and mechanisms of neuronal plasticity are relevant to the pathophysiology of MDD and the actions of antidepressant drugs.[51]

The signal on chromosome 7p15.3 is within *SP4* (Figure 2), and comes almost entirely from females (*P*=4.44E-05 vs. 0.148 in males). Sp4 is specific to neurons and expressed primarily early in development[52–54], forms complexes with estrogen receptors that influence regulation of many genes[55], and could play a role in the mediation of neuroprotective enzymes and in glutamate-induced neurotoxicity.[56, 57] Zhou et al. have reported that reduced expression of Sp4 in mice leads to hippocampal vacuolization, age-dependent reduced expression of neurotrophin 3 and deficits in sensorimotor gating and contextual memory[54], possibly mediated by defects in the development of dentate gyrus cells.[58] There is also evidence for a possible association between bipolar disorder and SNPs in the promoter of *ADRBK2* (beta adrenergic receptor kinase 2, previously *GRK3*) which disrupt an Sp1/Sp4 binding site.[59] Thus there are several mechanisms by which *SP4* could play a role in psychiatric disorders.

In conclusion, we carried out a GWAS of recurrent early-onset MDD in 1,020 cases and 1,636 controls of European ancestry. No genome-wide significant evidence for association was observed. Of the strongest signals reported in the GenRED sample, the meta-analysis in the companion article[17] provides the greatest support for association to SNPs in *SP4*. Much

larger samples may be needed to determine whether there are true associations between MDD and common SNPs.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The GenRED project is supported by grants from NIMH. We acknowledge the contributions of Dr. George S. Zubenko and Dr. Wendy N. Zubenko, Department of Psychiatry, University of Pittsburgh School of Medicine, to the GenRED I project. The NIMH Cell Repository at Rutgers University and the NIMH Center for Collaborative Genetic Studies on Mental Disorders made essential contributions to this project. Genotyping was carried out by the Broad Institute Center for Genotyping and Analysis with support from grant U54 RR020278 (which partially subsidized the genotyping of the GenRED cases) from the National Center for Research Resources.

GWAS data for the GAIN-MDD dataset were accessed by D.F.L. through the Genetic Association Information Network (GAIN), through dbGaP accession number phs000020.v1.p1 (http://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs000020.v2.p1); samples and associated phenotype data for Major Depression: Stage 1 Genome-wide Association in Population-Based Samples were provided by P. Sullivan.

Data for Molecular Genetics of Schizophrenia (MGS) control subjects was used here by permission of the MGS project. Collection and quality control analyses of the control dataset were supported by grants from NIMH and the National Alliance for Research on Schizophrenia and Depression. Genotyping of the controls was supported by grants from NIMH and by the Genetic Association Information Network (GAIN) (http://www.fnih.org/index.php?option=com_content&task=view&id=338&Itemid=454). Control data are available through dbGAP (http://www.ncbi.nlm.nih.gov/gap). We are grateful to Knowledge Networks,. Inc. (Menlo Park, CA) for assistance in collecting the control dataset.

The authors express their profound appreciation to the families who participated in this project, and to the many clinicians who facilitated the referral of participants to the study. (Additional information is available in online Supplementary Acknowledgements.)

References

1. Pakriev S, Shlik J, Vasar V. Course of depression: findings from cross-sectional survey in rural Udmurtia. Nordic journal of psychiatry. 2001; 55(3):185–189. [PubMed: 11827613]
2. Blair-West GW, Cantor CH, Mellsop GW, Eyeson-Annan ML. Lifetime suicide risk in major depression: sex and age determinants. Journal of affective disorders. 1999 Oct; 55(2–3):171–178. [PubMed: 10628886]
3. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders. 4. American Psychiatric Press; Washington, D. C: 1994.
4. Sullivan PF, Neale MC, Kendler KS. Genetic epidemiology of major depression: review and meta-analysis. Am J Psychiatry. 2000 Oct; 157(10):1552–1562. [PubMed: 11007705]
5. McGuffin P, Katz R, Watkins S, Rutherford J. A hospital-based twin register of the heritability of DSM-IV unipolar depression. Arch Gen Psychiatry. 1996 Feb; 53(2):129–136. [PubMed: 8629888]
6. Kendler KS, Neale MC, Kessler RC, Heath AC, Eaves LJ. The lifetime history of major depression in women. Reliability of diagnosis and heritability. Arch Gen Psychiatry. 1993 Nov; 50(11):863–870. [PubMed: 8215812]
7. Kendler KS, Kuhn JW, Prescott CA. Childhood sexual abuse, stressful life events and risk for major depression in women. Psychol Med. 2004 Nov; 34(8):1475–1482. [PubMed: 15724878]
8. Zimmermann P, Bruckl T, Lieb R, Nocon A, Ising M, Beesdo K, et al. The interplay of familial depression liability and adverse events in predicting the first onset of depression during a 10-year follow-up. Biol Psychiatry. 2008 Feb 15; 63(4):406–414. [PubMed: 17698041]
9. Kendler KS, Gatz M, Gardner CO, Pedersen NL. Clinical indices of familial depression in the Swedish Twin Registry. Acta Psychiatr Scand. 2007 Mar; 115(3):214–220. [PubMed: 17302621]

10. Levinson DF, Zubenko GS, Crowe RR, DePaulo RJ, Scheftner WS, Weissman MM, et al. Genetics of recurrent early-onset depression (GenRED). Am J Med Genet B NeuropsychiatrGenet. 2003 May 15; 119(1):118–130.
11. Weissman MM, Wickramaratne P, Merikangas KR, Leckman JF, Prusoff BA, Caruso KA, et al. Onset of major depression in early adulthood. Increased familial loading and specificity. Arch Gen Psychiatry. 1984 Dec; 41(12):1136–1143. [PubMed: 6508504]
12. Kendler KS, Gardner CO, Neale MC, Prescott CA. Genetic risk factors for major depression in men and women: similar or different heritabilities and same or partly distinct genes? PsycholMed. 2001; 31(4):605–616.
13. Kendler KS, Gatz M, Gardner CO, Pedersen NL. A Swedish National Twin Study of Lifetime Major Depression. American Journal of Psychiatry. 2006 Jan 01; 163(1):109–114. [PubMed: 16390897]
14. Maier W, Lichtermann D, Minges J, Hallmayer J, Heun R, Benkert O, et al. Continuity and discontinuity of affective disorders and schizophrenia. Results of a controlled family study. Arch Gen Psychiatry. 1993 Nov; 50(11):871–883. [PubMed: 8215813]
15. Holmans P, Weissman MM, Zubenko GS, Scheftner WA, Crowe RR, Depaulo JR Jr, et al. Genetics of recurrent early-onset major depression (GenRED): final genome scan report. Am J Psychiatry. 2007 Feb; 164(2):248–258. [PubMed: 17267787]
16. Levinson DF, Evgrafov OV, Knowles JA, Potash JB, Weissman MM, Scheftner WA, et al. Genetics of recurrent early-onset major depression (GenRED): significant linkage on chromosome 15q25-q26 after fine mapping with single nucleotide polymorphism markers. Am J Psychiatry. 2007 Feb; 164(2):259–264. [PubMed: 17267788]
17. Shyn SI, Shi J, Kraft JB, Potash JB, Knowles JA, Weissman MM, et al. Novel loci for major depression identified by genome-wide association study of STAR*D and meta-analysis of three studies. Under review. 2009
18. Psychiatric GWAS Consortium Coordinating Committee. Cichon S, Craddock N, Daly M, Faraone SV, Gejman PV, et al. Genomewide association studies: history, rationale, and prospects for psychiatric disorders. Am J Psychiatry. 2009 May; 166(5):540–556. [PubMed: 19339359]
19. Sanders AR, Duan J, Levinson DF, Shi J, He D, Hou C, et al. No significant association of 14 candidate genes with schizophrenia in a large European ancestry sample: implications for psychiatric genetics. Am J Psychiatry. 2008 Apr; 165(4):497–506. [PubMed: 18198266]
20. Nurnberger JI Jr, Blehar MC, Kaufmann CA, York-Cooler C, Simpson SG, Harkavy-Friedman J, et al. Diagnostic interview for genetic studies. Rationale, unique features, and training. NIMH Genetics Initiative. Arch Gen Psychiatry. 1994 Nov; 51(11):849–859. [PubMed: 7944874]
21. Kessler RC, Andrews G, Mroczek D, Ustun TB, Wittchen H-U. The World Health Organization Composite International Diagnostic Interview Short Form (CIDI-SF). International Journal of Methods in Psychiatric Research. 1998; 7(4):171–185.
22. Korn JM, Kuruvilla FG, McCarroll SA, Wysoker A, Nemesh J, Cawley S, et al. Integrated genotype calling and association analysis of SNPs, common copy number polymorphisms and rare CNVs. Nat Genet. 2008 Oct; 40(10):1253–1260. [PubMed: 18776909]
23. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet. 2007 Sep; 81(3):559–575. [PubMed: 17701901]
24. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. Principal components analysis corrects for stratification in genome-wide association studies. Nat Genet. 2006 Aug; 38(8):904–909. [PubMed: 16862161]
25. Huang L, Li Y, Singleton AB, Hardy JA, Abecasis G, Rosenberg NA, et al. Genotype-imputation accuracy across worldwide human populations. Am J Hum Genet. 2009 Feb; 84(2):235–250. [PubMed: 19215730]
26. 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 Jul; 39(7):906–913. [PubMed: 17572673]

27. Kathiresan S, Willer CJ, Peloso GM, Demissie S, Musunuru K, Schadt EE, et al. Common variants at 30 loci contribute to polygenic dyslipidemia. Nat Genet. 2009 Jan; 41(1):56–65. [PubMed: 19060906]
28. Scott LJ, Mohlke KL, Bonnycastle LL, Willer CJ, Li Y, Duren WL, et al. A genome-wide association study of type 2 diabetes in Finns detects multiple susceptibility variants. Science. 2007 Jun 1; 316(5829):1341–1345. [PubMed: 17463248]
29. Willer CJ, Sanna S, Jackson AU, Scuteri A, Bonnycastle LL, Clarke R, et al. Newly identified loci that influence lipid concentrations and risk of coronary artery disease. Nat Genet. 2008 Feb; 40(2): 161–169. [PubMed: 18193043]
30. Willer CJ, Speliotes EK, Loos RJ, Li S, Lindgren CM, Heid IM, et al. Six new loci associated with body mass index highlight a neuronal influence on body weight regulation. Nat Genet. 2009 Jan; 41(1):25–34. [PubMed: 19079261]
31. Dudbridge F, Gusnanto A. Estimation of significance thresholds for genomewide association scans. Genet Epidemiol. 2008 Apr; 32(3):227–234. [PubMed: 18300295]
32. Hoggart CJ, Clark TG, De Iorio M, Whittaker JC, Balding DJ. Genome-wide significance for dense SNP and resequencing data. Genet Epidemiol. 2008 Feb; 32(2):179–185. [PubMed: 18200594]
33. 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 May; 32(4):381–385. [PubMed: 18348202]
34. Devlin B, Roeder K. Genomic control for association studies. Biometrics. 1999 Dec; 55(4):997–1004. [PubMed: 11315092]
35. Taylor J, Tyekucheva S, King DC, Hardison RC, Miller W, Chiaromonte F. ESPERR: Learning strong and weak signals in genomic sequence alignments to identify functional elements. Genome Res. 2006; 16(12):1596–1604. [PubMed: 17053093]
36. Shyn S. Genomewide association study of major depression (STAR*D). Molecular Psychiatry. 2009
37. Ferreira MAR, O'Donovan MC, Meng YA, Jones IR, Ruderfer DM, Jones L, et al. Collaborative genome-wide association analysis supports a role for ANK3 and CACNA1C in bipolar disorder. Nat Genet. 2008; 40(9):1056–1058. [PubMed: 18711365]
38. Manolio TA, Brooks LD, Collins FS. A HapMap harvest of insights into the genetics of common disease. J Clin Invest. 2008 May; 118(5):1590–1605. [PubMed: 18451988]
39. Psychiatric GWAS Consortium Coordinating Committee. Cichon S, Craddock N, Daly M, Faraone SV, Gejman PV, et al. Genomewide Association Studies: History, Rationale, and Prospects for Psychiatric Disorders. American Journal of Psychiatry. 2009; 166(5):540–556. [PubMed: 19339359]
40. Psychiatric GWAS Consortium. A framework for interpreting genome-wide association studies of psychiatric disorders. Mol Psychiatry. 2009 Jan; 14(1):10–17. [PubMed: 19002139]
41. Fallin MD, Lasseter VK, Wolyniec PS, McGrath JA, Nestadt G, Valle D, et al. Genomewide linkage scan for bipolar-disorder susceptibility loci among Ashkenazi Jewish families. Am J Hum Genet. 2004 Aug; 75(2):204–219. [PubMed: 15208783]
42. McMahon FJ, Hopkins PJ, Xu J, McInnis MG, Shaw S, Cardon L, et al. Linkage of bipolar affective disorder to chromosome 18 markers in a new pedigree series. Am J Hum Genet. 1997 Dec; 61(6):1397–1404. [PubMed: 9399888]
43. McMahon FJ, Simpson SG, McInnis MG, Badner JA, MacKinnon DF, DePaulo JR. Linkage of bipolar disorder to chromosome 18q and the validity of bipolar II disorder. Arch Gen Psychiatry. 2001 Nov; 58(11):1025–1031. [PubMed: 11695948]
44. Nwulia EA, Miao K, Zandi PP, Mackinnon DF, DePaulo JR Jr, McInnis MG. Genome-wide scan of bipolar II disorder. Bipolar disorders. 2007 Sep; 9(6):580–588. [PubMed: 17845272]
45. Stine OC, Xu J, Koskela R, McMahon FJ, Gschwend M, Friddle C, et al. Evidence for linkage of bipolar disorder to chromosome 18 with a parent-of-origin effect. Am J Hum Genet. 1995 Dec; 57(6):1384–1394. [PubMed: 8533768]

46. Verheyen GR, Villafuerte SM, Del-Favero J, Souery D, Mendlewicz J, Van Broeckhoven C, et al. Genetic refinement and physical mapping of a chromosome 18q candidate region for bipolar disorder. Eur J Hum Genet. 1999 May-Jun;7(4):427–434. [PubMed: 10352933]
47. McInnes LA, Escamilla MA, Service SK, Reus VI, Leon P, Silva S, et al. A complete genome screen for genes predisposing to severe bipolar disorder in two Costa Rican pedigrees. Proc Natl Acad Sci U S A. 1996 Nov 12; 93(23):13060–13065. [PubMed: 8917544]
48. McQueen MB, Devlin B, Faraone SV, Nimgaonkar VL, Sklar P, Smoller JW, et al. Combined analysis from eleven linkage studies of bipolar disorder provides strong evidence of susceptibility loci on chromosomes 6q and 8q. Am J Hum Genet. 2005 Oct; 77(4):582–595. [PubMed: 16175504]
49. Levinson DF. The genetics of depression: a review. Biol Psychiatry. 2006 Jul 15; 60(2):84–92. [PubMed: 16300747]
50. Goossens D, Van Gestel S, Claes S, De Rijk P, Souery D, Massat I, et al. A novel CpG-associated brain-expressed candidate gene for chromosome 18q-linked bipolar disorder. Mol Psychiatry. 2003 Jan; 8(1):83–89. [PubMed: 12556911]
51. Krishnan V, Nestler EJ. The molecular neurobiology of depression. Nature. 2008 Oct 16; 455(7215):894–902. [PubMed: 18923511]
52. Supp DM, Witte DP, Branford WW, Smith EP, Potter SS. Sp4, a Member of the Sp1-Family of Zinc Finger Transcription Factors, Is Required for Normal Murine Growth, Viability, and Male Fertility. Developmental Biology. 1996 Jun 15; 176(2):284–299. [PubMed: 8660867]
53. Suske G. The Sp-family of transcription factors. Gene. 1999 Oct 01; 238(2):291–300. [PubMed: 10570957]
54. Zhou X, Long JM, Geyer MA, Masliah E, Kelsoe JR, Wynshaw-Boris A, et al. Reduced expression of the Sp4 gene in mice causes deficits in sensorimotor gating and memory associated with hippocampal vacuolization. Mol Psychiatry. 2004 Nov 23; 10(4):393–406.
55. Safe S, Kim K. Non-classical genomic estrogen receptor (ER)/specificity protein and ER/activating protein-1 signaling pathways. J Mol Endocrinol. 2008; 41(5):263–275. [PubMed: 18772268]
56. Mao X, Moerman-Herzog AM, Wang W, Barger SW. Differential transcriptional control of the superoxide dismutase-2 kappaB element in neurons and astrocytes. J Biol Chem. 2006 Nov 24; 281(47):35863–35872. [PubMed: 17023425]
57. Mao X, Yang SH, Simpkins JW, Barger SW. Glutamate receptor activation evokes calpain-mediated degradation of Sp3 and Sp4, the prominent Sp-family transcription factors in neurons. J Neurochem. 2007; 100(5):1300–1314. [PubMed: 17316402]
58. Zhou X, Qyang Y, Kelsoe JR, Masliah E, Geyer MA. Impaired postnatal development of hippocampal dentate gyrus in Sp4 null mutant mice. Genes Brain Behav. 2007; 6(3):269–276. [PubMed: 16899055]
59. Zhou X, Barrett TB, Kelsoe JR. Promoter variant in the GRK3 gene associated with bipolar disorder alters gene expression. Biol Psychiatry. 2008 Jul 15; 64(2):104–110. [PubMed: 18359007]

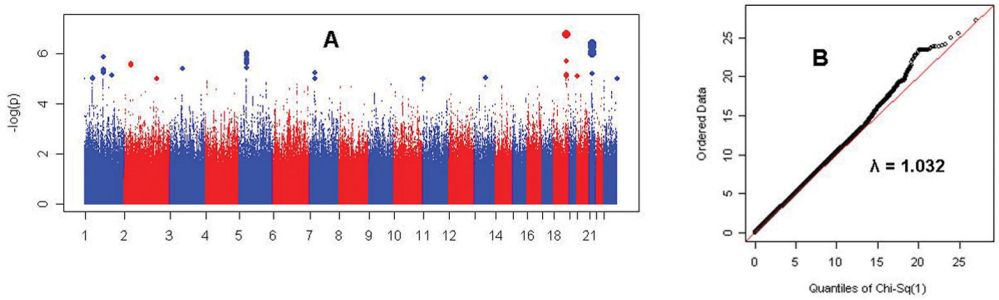

Figure 1. Genome-wide results

Panel A shows the association test result for each SNP (genotyped and imputed) as $-\log_{10}(P\text{-value})$, for the primary analysis of all subjects. The largest symbols represent $P<10^{-6}$, and the intermediate-size symbols represent $P<10^{-5}$. For the same SNPs, Panel B shows the quantile-quantile plot of observed vs. expected X^2 (1df) statistics. The genomic control λ value was 1.032. Manhattan and QQ plots are shown for males and for females separately in online Figures S7 and S14.

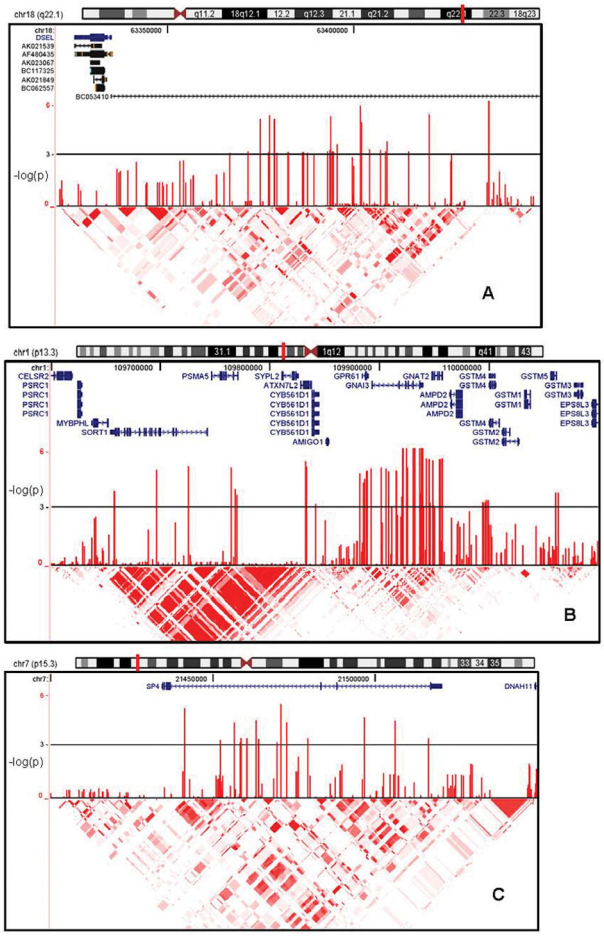

Figure 2. Results for selected chromosomal regions

Shown from top to bottom of each plot are chromosome ideogram (plotted region marked with vertical red bar); genomic information (RefSeq genes with direction of transcription; and in Panel A, mRNAs); association test results ($-\log_{10}[P]$) in the primary analysis of all subjects; and HapMap linkage disequilibrium (r^2) information.

Table 1

Sample characteristics

	Cases	Controls
N	1020	1636
Male	29%	56%
Age at recruitment	40.5 ± 11.9	52.5 ± 17.2
Age at onset	16.85 ± 5.4	
Major Depressive Disorder episodes:		
Recurrent	98%	
Single ≥ 3 years	2%	
Number of episodes	8.4 ± 14.6	
Longest episode (days)	931 ± 1896	
Chronic course (consensus rating)	39%	
Number of 8 MDD criteria met during worst episode (plus dysphoric mood)	6.8 ± 1.1	
Comorbid anxiety disorder diagnosis (panic, agoraphobia, social phobia)	35%	

Table 2

Strongest association findings (All, Male or Female subjects)

Band	SNP	BP	Alleles	Frq	R^2	ALL subjects OR	ALL subjects P	Males OR	Males P	Females OR	Females P	Annotation
P < 10^{-5} in ALL subjects												
18q22.1[1]	rs17077540	63436259	A/**G**	0.11	0.86	1.61	1.83E-07	1.67	6.02E-04	1.57	2.62E-04	mRNA BC053410 (brain; LOC643542); *DSEL* (up). (Figure 2)
21q21.2	rs2828520	24064990	A/**G**	0.31	1.00	1.35	4.33E-07	1.52	2.94E-05	1.28	1.45E-03	
5p13.2[2]	rs270545	38087350	A/**G**	0.69	1.00	1.37	1.03E-06	1.35	5.97E-03	1.42	2.89E-05	*GDNF* (217kb up)
1p13.3	rs6537837	109921255	C/**T**	0.17	0.99	1.43	1.31E-06	1.35	1.55E-02	1.40	5.85E-04	*GNAT2*; *GNAI3*; *AMPD2* (up) (Figure 2)
2p23.2	rs882632	29134265	C/**T**	0.29	1.00	1.34	2.41E-06	1.25	3.07E-02	1.53	4.04E-07	*FAM179A* (LOC165186) (+5665); *C2orf71* (+3796)
3p14.2	rs10514718	61388854	**C**/G	0.94	0.82	2.12	3.96E-06	1.72	4.49E-02	2.46	1.85E-05	*FHIT* (176kb up); *PTPRG* (133kb up)
1p13.3	rs12049330	109832711	**G**/T	0.14	0.95	1.44	5.87E-06	1.56	1.05E-03	1.34	4.46E-03	*ATXN7L2*; *SYPL2* (dwn); *CYB561D1* (up) (Figure 2)
7p15.3	rs17144465	21470952	A/**G**	0.04	1.00	1.82	5.97E-06	1.39	1.48E-01	2.11	4.44E-05	*SP4* (Figure 2)
13q21.33[3]	rs9572423	69744204	A/**G**	0.88	0.96	1.54	9.25E-06	1.94	3.14E-04	1.46	2.53E-03	
1p33	rs1167264	49705632	**C**/T	0.20	0.99	1.35	9.51E-06	1.22	7.53E-02	1.41	1.26E-04	*AGBL4*
11p15.4	rs2898938	3788833	G/**T**	0.72	1.00	1.34	9.60E-06	1.27	3.57E-02	1.40	9.31E-05	*FRAG1*; *NUP98* (up); *STIM1* (up)
11p15.5	rs11024034	2746739	**C**/T	0.09	0.99	1.45	9.64E-06	1.34	3.63E-02	1.53	2.03E-04	*KCNQ1* (including one SNP at exon-intron boundary)
P < 10^{-5} in MALES												
16q12.2	rs2631522	53782584	C/**T**	0.71	1.00	1.13	7.05E-02	1.86	2.50E-07	0.87	7.54E-02	Enhancer element_26 (ends at 53781784 bp)
12p12.3	rs12581840	19725418	*C*/***T***	0.57	1.00	1.18	3.42E-03	1.70	3.47E-07	1.02	8.16E-01	
3p14.2	rs12633494	61120090	**C**/T	0.04	1.00	1.36	2.25E-02	2.75	4.63E-07	0.84	3.65E-01	*FHIT*
1q24.3	rs17657363	170068039	**C**/T	0.18	1.00	1.26	1.05E-03	1.79	5.19E-07	1.01	8.48E-01	*MYOC* (up); *VAMP4* (in, up); *KIAA0859*; *DNM3* (in, up)
P < 10^{-5} in FEMALES												
2p23.2	rs882632	29134265	*C*/***T***	0.29	1.00	1.34	2.41E-06	1.25	3.07E-02	1.53	4.04E-07	*FAM179A* (LOC165186) (dwn); *C2orf71* (dwn)
18q21.2	rs1676089	47448507	*C*/***A***	0.88	1.00	1.38	9.67E-04	0.87	3.56E-01	1.89	1.35E-06	
8p23.2	rs2616976	4465647	**C**/G	0.41	0.97	1.28	2.35E-05	1.06	5.25E-01	1.43	3.60E-06	*CSMD1*
10p14	rs11254981	7148786	*C*/***G***	0.43	0.98	1.28	3.62E-05	1.08	4.57E-01	1.44	4.08E-06	mRNA AK096400 (6.55-7.22 Mb)
11q22.3	rs1430871	102650380	**C**/T	0.76	0.99	1.29	3.07E-04	1.05	6.67E-01	1.53	4.93E-06	*DYNC2H1*
3p14.2	rs11918585	61403222	***C***/*T*	0.93	0.74	1.86	6.12E-05	1.19	4.79E-01	2.58	5.80E-06	*FHIT* (191kb up); *PTPRG* (118kb up)

The SNP with the lowest P-value is listed for all genes or nongenic regions with at least one SNP with $P < 10^{-5}$, separately for the analyses of all subjects (primary analysis), males, and females. The Annotation column lists all genes in the region with one or more SNPs with $P < 10^{-5}$, either within the gene or within 50 kb upstream (up) or downstream (dwn), unless a longer distance is listed. Other functional elements in a region are as noted. Nongenic regions all contain peaks of bioinformatically predicted high homology to known regulatory sequences.[35]

OR=Odds Ratio for the tested allele, indicated in bold font in the Alleles column.

Frq=frequency of the tested allele in Controls. (Case-control frequencies for All Subjects findings are available in online Table S11.)

R^2 indicates the R^2 predicted (by MACH 1.0) between imputed and actual genotypes; $R^2 = 1$ indicates that the SNP was genotyped.

Note that many of these regions contained multiple SNPs with low P-values, see online file genred_supplementary_data.txt.

[1] $P=6.04 \times 10^{-7}$ in the meta-analysis of Narrow cases (GenRED, STAR*D and GAIN) in a companion paper.[17] [2]- $P=1.88 \times 10^{-6}$ in the Narrow meta-analysis.