Published in final edited form as:
Mol Psychiatry. 2009 August ; 14(8): 755–763. doi:10.1038/mp.2009.43.

GENOME-WIDE ASSOCIATION STUDY OF BIPOLAR DISORDER IN EUROPEAN AMERICAN AND AFRICAN AMERICAN INDIVIDUALS

Erin N. Smith[*,1,2], **Cinnamon S. Bloss**[*,1,3], **Judith A. Badner**[4], **Thomas Barrett**[5], **Pamela L. Belmonte**[6], **Wade Berrettini**[7], **William Byerley**[8], **William Coryell**[9], **David Craig**[10], **Howard J. Edenberg**[11,12], **Eleazar Eskin**[13], **Tatiana Foroud**[12], **Elliot Gershon**[4], **Tiffany A. Greenwood**[14], **Maria Hipolito**[15], **Daniel L. Koller**[16], **William B. Lawson**[15], **Chunyu Liu**[4], **Falk Lohoff**[7], **Melvin G. McInnis**[17], **Francis J. McMahon**[18], **Daniel B. Mirel**[19], **Caroline Nievergelt**[14], **John Nurnberger**[16], **Evaristus A. Nwulia**[15], **Justin Paschall**[20], **James B. Potash**[6], **John Rice**[21], **Thomas G. Schulze**[18], **William Scheftner**[22], **Corrie Panganiban**[10], **Noah Zaitlen**[13], **Peter P. Zandi**[6], **Sebastian Zöllner**[17], **Nicholas J. Schork**[1,2], and **John R. Kelsoe**[14,23]

[1]Scripps Genomic Medicine and Scripps Translational Science Institute, La Jolla, CA 92037, USA

[2]Department of Molecular and Experimental Medicine, The Scripps Research Institute, La Jolla, CA 92037, USA

[3]Scripps Health, La Jolla, CA 92037, USA

[4]Department of Psychiatry, University of Chicago, Chicago, IL 60637, USA

[5]Department of Psychiatry, Portland VA Medical Center, Portland, OR, 97239, USA

[6]Department of Psychiatry, Johns Hopkins School of Medicine, Baltimore, MD, 21287, USA

[7]Department of Psychiatry, University of Pennsylvania, Philadelphia, PA, 19104, USA

[8]Department of Psychiatry, University of California, San Francisco, CA, 94143, USA

[9]Department of Psychiatry, University of Iowa, Iowa City, IA 52242, USA

[10]Neurogenomics Division, The Translational Genomics Research Institute, Phoenix, AZ, 85004, USA

[11]Department of Biochemistry and Molecular Biology, Indiana University School of Medicine, Indianapolis, IN, 46202, USA

[12]Department of Medical and Molecular Genetics, Indiana University School of Medicine, Indianapolis, IN, 46202, USA

[13]Department of Computer Science, University of California, Los Angeles, Los Angeles, CA, 90095, USA

[14]Department of Psychiatry, University of California, San Diego, La Jolla, CA 92093, USA

[15]Department of Psychiatry, Howard University, Washington, D.C., 20060, USA

Address Correspondence to: Nicholas J. Schork, Ph.D., The Scripps Translational Science Institute, Department of Molecular and Experimental Medicine, The Scripps Research Institute, 10550 North Torrey Pines Avenue, MEM 275, La Jolla, CA 92037, 858-554-5705 (admin), 858-546-9284 (fax), nschork@scripps.edu.

[*]These authors contributed equally.

[16]Department of Psychiatry, Indiana University School of Medicine, Indianapolis, IN, 46202, USA

[17]Department of Psychiatry, University of Michigan, Ann Arbor, MI, 48109, USA

[18]Genetic Basis of Mood and Anxiety Disorders Unit, National Institute of Mental Health Intramural Research Program, National Institutes of Health, US Dept of Health and Human Services, Bethesda, MD, 20892, USA

[19]Broad Institute of Harvard University and Massachusetts Institute of Technology, Cambridge, MA, 02142, USA

[20]National Center for Biotechnology Information, National Library of Medicine, National Institutes of Health, Bethesda, Maryland 20892, USA

[21]Division of Biostatistics, Washington University, St. Louis, MO, 63130, USA

[22]Department of Psychiatry, Rush University, Chicago, IL, 60612, USA

[23]Department of Psychiatry, VA San Diego Healthcare System, La Jolla, CA, 92151, USA

Abstract

To identify Bipolar Disorder (BD) genetic susceptibility factors, we conducted two genome-wide association (GWA) studies: one involving a sample of individuals of European ancestry (EA; n = 1,001 cases; n = 1,033 controls) and one involving a sample of individuals of African ancestry (AA; n = 345 cases; n = 670 controls). For the EA sample, SNPs with strongest statistical evidence for association included rs5907577 in an intergenic region at Xq27.1 ($p = 1.6 \times 10^{-6}$) and rs10193871 in *NAP5* at 2q21.2 ($p = 9.8 \times 10^{-6}$). For the AA sample, SNPs with strongest statistical evidence for association included rs2111504 in *DPY19L3* at 19q13.11 ($p = 1.5 \times 10^{-6}$) and rs2769605 in *NTRK2* at 9q21.33 ($p = 4.5 \times 10^{-5}$). We also investigated whether we could provide support for three regions previously associated with BD, and we show that the *ANK3* region replicates in our sample, along with some support for *C15Orf53*; other evidence implicates BD candidate genes such as *SLITRK2*. We also tested the hypothesis that BD susceptibility variants exhibit genetic background-dependent effects; SNPs with the strongest statistical evidence for this included rs11208285 in *ROR1* at 1p31.3 ($p = 1.4 \times 10^{-6}$), rs4657247 in *RGS5* at 1q23.3 ($p = 4.1 \times 10^{-6}$), and rs7078071 in *BTBD16* at 10q26.13 ($p = 4.5 \times 10^{-6}$). This study is the first to conduct GWA of BD in individuals of AA and suggests that genetic variations that contribute to BD may vary as a function of ancestry.

Keywords

ANK3; C15Orf53; NAP5; DPY19L3; NTRK2; SLITRK2; ROR1; Bipolar Genome Study; Genetic Information Association Network (GAIN); genetic background; allelic heterogeneity

INTRODUCTION

Bipolar disorder (BD) is a paradigmatic complex phenotype with many genetic and non-genetic determinants. BD is characterized by episodes of mania and depression (1). Onset is usually in late adolescence, although BD typically recurs and relapses throughout life. BD affects approximately 1% of the world's population, and carries a lifetime risk for completed suicide as high as 17%. Family, twin, and adoption studies all support a substantial genetic component in BD (2-4). The sibling recurrence risk is between 7 and 10, and heritability is estimated to be about 80%. While this is consistent with a strong genetic component, the identification of specific genetic variations that influence BD susceptibility has been difficult.

Although some family studies have suggested that BD has an autosomal dominant genetic determinant, the vast majority of genetic studies suggest that BD has a high level of genetic heterogeneity and a substantial polygenic component (5,6). Linkage studies have identified a number of loci with inconsistent replication. While previous linkage studies were highly divergent (7), recent meta-analyses of linkage studies have found consistent supportive evidence for linkage to a few potential BD susceptibility loci (8-10), notably 6q, 8q, 13q, and 22q. A number of polymorphisms in a variety of candidate genes have been tested for association with BD, but most of the polymorphisms have shown no statistically compelling associations with BD; those that appear to be associated show odds ratios of 1.1 to 1.3, which is again consistent with a polygenic basis for BD (11).

In the last several years the development of genome-wide association (GWA) study designs and analysis methods have made it possible to search for multiple genetic variations underlying a condition like BD without *a priori* assumptions about the genes or genomic regions that might harbor susceptibility variations (12). The GWA study approach has revealed a number of genetic variations that are unequivocally associated with traits and diseases (13,14). The US National Institutes of Mental Health–sponsored Genetics Initiative for Bipolar Disorder Consortium (Bipolar Consortium) has collected over 3,500 subjects with BD during 1990-2008. A genetic component of the Bipolar Consortium, termed the 'Bipolar Genome Study (BiGS),' was initiated in 2006 to conduct a GWA study of BD. The initial BiGS GWA study was funded through the Foundation for the National Institutes of Health Genetic Information Association Network (GAIN) initiative (http://www.genome.gov/19518664) and we report the results of this GWA study here. Of note, many investigators have access to and have utilized the DNA samples from individuals with BD that were collected as part of the Bipolar Consortium; thus, it is the case that a portion of the samples on which we report here are not entirely independent of previously published BD GWA studies (15,16) and an ongoing BD GWA study (Scott, Muglia, Upmanyu, Guan, Flickinger, Kong, Tozzi, Li, Burmeister, Absher et al., submitted).

We performed a GWA study of BD separately in individuals of European ancestry (EA) and of African ancestry (AA) using a variety of methods to control for population substructure and admixture among the individuals of AA. Our study is the first GWA study of BD involving individuals of AA. In addition to standard single locus analyses within the EA and AA groups, we also assess the extent to which single nucleotide polymorphisms (SNPs) exhibite evidence of genetic background-dependent effects or allelic heterogeneity across the EA and AA individuals. Although more than 190 GWA studies have been performed to date (12) these studies have virtually all been with EA subjects (NHGRI catalog of published GWA studies: http://www.genome.gov/gwastudies/), raising important questions as to the generalizability of the results to other populations. We assessed the consistency of the SNPs exhibiting the strongest associations with BD in our study with previously published results of BD GWA studies. We tested variation at *ANK3*, which is a candidate gene that was implicated in an earlier GWA study (15) and identified as genome-wide significant in a recent collaborative analysis (17). Finally, we genotyped a subset of SNPs in the regions exhibiting strongest association in a replication case/control group and tracked the co-inheritance of these SNPs and disease in families.

MATERIALS AND METHODS

See Supplemental Material for information regarding study subjects and genotyping and quality control. The final number of BD cases was 1,001 EA subjects and 345 AA subjects; control counts were 1,033 EA subjects and 670 AA subjects. The final dataset consisted of 724,067 SNPs in the EA dataset and 840,730 SNPs in the AA dataset. The two datasets had

702,044 SNPs in common, and these were used to perform analyses addressing SNP and SNPxgenetic background interactions in the combined sample.

Statistical Analysis

Primary analyses—All genetic analyses were conducted using PLINK (18) versions 1.03 and 1.04. Our primary analyses tested the association of each SNP (coded additively) with BD using three distinct methods that control for genetic background heterogeneity (Table 1): (1) single locus, contingency table analysis with genomic control adjustment, (2) logistic regression using a LAMP-derived estimate of ancestry as a covariate, and (3) logistic regression using the top 4 MDS dimensions as covariates. Association analyses using these three methods were conducted within the EA and AA samples separately, as well as within the combined sample (see Table 1). For each analysis, the distribution of expected p-values under the null hypothesis (i.e., no association), and the genomic inflation value (λ) was calculated in PLINK using the adjust command (Figure S1). Genomic inflation values are reported as both uncorrected and corrected for a study size of 1,000 cases and 1,000 controls (19,20).

Haplotype tests—Haplotype analyses were performed in the EA population using the sliding window approach in PLINK (18), and results for 10-SNP windows are highlighted. We used OMNIBUS p-values for the primary analysis and individual haplotype p-values for the analysis of haplotype heterogeneity at *ANK3*. In these analyses, population structure was not taken into account. For the haplotype heterogeneity analysis at *ANK3*, the proportion of haplotypes that reached a p-value of $p < 0.05$ was performed using the 10-SNP window results. For each window, the number of haplotypes with an individual p-value less than 0.05 was counted and a proportion calculated for the window. Smoothing was performed over the results for the rest of the genome, and regions were highlighted that had a smoothed proportion greater than or equal to the maximum found at *ANK3* (0.33). Regions were distinct and were delineated by the first and last marker with a smoothed value greater than or equal to 0.33.

Imputation—Imputation was performed for the cleaned EA dataset using MACH v.1.0.16 (http://www.sph.umich.edu/csg/abecasis/mach/index.html) with HapMap rel21 phased haplotypes as a reference (see Supplemental Material).

Genetic background-dependent effects—We assessed genetic background-specific effects, which could be due to interactions with background-specific variation or alternatively, due to different functional variation at the locus of interest. We combined the EA and AA samples and performed logistic regression analyses with adjustment using a LAMP main effect covariate and a LAMP x SNP interaction term (Table 2). We compared the OR across different categories of admixture. We split the AA individuals into high and low admixture categories by the approximate median %CEU of 15%, and calculated the OR and 95% confidence interval within these and the EA group (see Figure 1).

Power and statistical significance—In the current study, we have employed a significance threshold of 5×10^{-8}, as this threshold has been commonly used in previous GWA studies. As there were no SNPs that exceeded this threshold, we report the most significant SNP for each sample, as well as regions where there are 5 SNPs within 100 kb of each other all with $p < 1 \times 10^{-4}$. In addition, we list all SNPs reported with a $p < 10^{-4}$ (Tables S3-S5). The strongest test of association lies in follow-up replication studies, however, and with the data we have available to us, we have attempted to examine the consistency of our results with previous GWA studies of BD and with our own newly genotyped replication samples (Table 3).

Pritzker study overlap and analysis—A concurrent study of BD is being undertaken by the Pritzker Neuropsychiatric Disorders Research Consortium (PNDRC; Scott, Muglia, Upmanyu, Guan, Flickinger, Kong, Tozzi, Li, Burmeister, Absher et al., submitted). There is some overlap between the cases and controls used in our study and those used in the Pritzker study (407 EA cases and 357 EA controls). As an extension of our analyses we excluded these samples and repeated the association tests (see Table 1) so that the results of our study and the study by the PNDRC could be compared on the basis of independent samples. It is also the case that a portion of the samples on which we report here are not entirely independent of previously published BD GWA studies (15,16). We were not, however, able to perform a similar analysis with the non-overlapping samples from these other studies due to restricted access to the genotype data and time constraints.

Comparison of our results with previous GWA studies—Initially, to assess the extent to which our results corresponded with previous GWA studies of BD, we obtained p-values for genotyped markers in the WTCCC (13) and STEP-BD (16) studies; it was the case, however, that most of the top SNPs that we describe were not genotyped directly within these studies. Therefore, we proceeded by investigating the three regions of interest identified in a recent collaborative GWA study (17) within the EA sample in our study (Figure S3). We evaluated the top SNPs from this study that were directly genotyped in our study, as well as other SNPs in the three regions. We assessed both the direction of associations between the two studies, as well as the strength. For comparison, we include the results from these studies in the plots of regions of interest (Figure S2). Although limited, we further analyzed our results relative to one of the single previously published GWA studies (i.e., Baum et al., 2008) (15), which included about half of the probands and controls used in the present study, in addition to a replication sample that is independent of our sample (i.e., was collected in Germany). We present results in Figure S2, but note that further analysis of the replication within each of the BD GWA studies will require specific delination of the overlapping individuals within all the studies and imputation of all genotypes.

Replication Genotyping

A subset of 85 SNPs from the GWA study was selected for replication genotyping based on several criteria, with a primary focus on the allelic association p-value in the larger EA sample. See Supplemental Material for additional information pertaining to replication genotyping.

RESULTS

Descriptive Statistics

Demographic variables for EA and AA cases were generally similar (Table S1). For the AA population, controls were less likely to be female. In the EA population, gender distributions between cases and controls were similar. Controls were much older than cases in both populations (mean age of cases: 18.0 (AA) and 19.3 (EA) vs. mean age of controls: 45.8 (AA) and 52.2 (EA), which should protect against cryptic disease in the control groups.

Basic Association Analyses

We report the top associated SNP in each ancestry group, as well as the top associated SNP in the combined sample (Table 1; Figure S2). As discussed in the Subjects and Methods section, we utilized different methods to control for genetic background heterogeneity and admixture, but generally obtained similar results for all the methods. As shown in Table 1, the most significant associations were different between the EA and AA subjects. For the EA sample, the most significantly associated SNP was rs1825828 at 3q11.2, $p = 7.0 \times 10^{-7}$.

However, upon inspection of the genotype intensity plot (see Figure S2), rs1825828 appeared to be poorly genotyped. The second-best association was rs5907577 in an intergenic region at Xq27.1, $p = 1.6 \times 10^{-6}$. For the AA sample, the most significantly associated SNP was rs2111504 in *DPY19L3* at 19q13.11, $p = 1.5 \times 10^{-6}$. In addition, we looked for regions where there were multiple SNPs with low p-values ($p < 1 \times 10^{-4}$) with close physical proximity to one another (i.e., within 100kb). We report two regions that contain 5 SNPs that meet this criterion. Among EA, we report *NAP5* (top SNP is rs10193871, $p = 9.8 \times 10^{-6}$), and among AA, we report an intergenic region about 300kb upstream of *NTRK2* (top SNP is rs2769605, $p = 4.5 \times 10^{-5}$). When the EA and AA groups were combined, the most significant SNP was rs4825220 in Xq27.1, $p = 2.6 \times 10^{-7}$, which does not lie near a known gene. We also show Q-Q plots for each analysis (Figure S1). Within AA and EA samples, overall p-value inflation was low. However, we observed high levels of inflation when we combined the EA and AA datasets. This is due to the difference in case-control ratios between the studies, which artificially induces population stratification. Correcting for admixture using either LAMP or MDS covariates effectively removed this stratification. However, using the genomic control method to adjust the p-values overcorrected the association (Table 1).

We investigated the consistency of our top hits across each of the different groups in our study. Table S2 shows to what extent the top hits in each individual sample (e.g., in the EA group) showed evidence of association in the other samples (e.g., in the AA group). Top hits within AA were not significant in EA and vice versa, suggesting that variation that contributes to BD may differ between the two ancestral groups.

Imputation was performed using MACH (http://www.sph.umich.edu/csg/abecasis/mach/index.html) on the EA dataset. Imputed SNPs generally supported, and were of similar strength as the observed associations (see Figures S2 and S3). Near EPHA6, however, there were no imputed SNPs that showed association with BP, supporting the argument that the observed association is due to a genotyping error.

Haplotype Analyses

We performed haplotype-based association tests in the EA population, using the sliding-window approach on the genotyped SNPs in PLINK. In addition to showing 10-SNP haplotype p-values for each region where an individual SNP is highlighted, we also describe the most significant haplotype. This haplotype is located on the X chromosome and is 7 Mb away from the SNP with the highest single locus association strength (10 SNP OMNIBUS $p = 1.9 \times 10^{-11}$). This region is about 1Mb downstream of *SLITRK2* (Figure S2).

Comparison of Our Results with Previous GWA Studies

A recent collaborative GWA study (17) that consisted of 4,387 cases and 6,209 controls pooled from the Wellcome Trust Case Control Consortium (13) Bipolar Analysis, Sklar et al. (STEP-BD) (16), and 2,365 new samples highlighted three regions of interest: *ANK3* (ankyrin G), *CACNA1C* (alpha 1C subunit of the L-type voltage-gated calcium channel), and a region 3.3kb away from *C15ORF53* on chromosome 15q14 (17). This analysis was restricted to individuals of EA; therefore, we investigated these regions within only the EA sample in our study (Figure S3). When we focus on the top SNPs from this study that were directly genotyped in our study, only one of the SNPs reached $p < 0.05$ (*ANK3*, rs1938526/G, $p = 0.036$, OR = 1.31), although the top SNP in the 15q14 region approached $p < 0.05$ (rs2172835, $p = 0.057$, OR = 0.88). In both cases, the association was in the same direction and of a similar strength as previously reported. In both of these regions, we saw additional SNPs that showed low (0.01-0.0001) p-values.

We thought that this pattern could indicate the presence of allelic heterogeneity, which has also been suggested by a separate study focusing on 2 markers in the region (Schulze, Detera-Wadleigh, Akula, Gupta, Kassem, Steele, Pearl, Strohmaier, Breuer, Schwarz, et al., in press), and which might reflect multiple underlying rare variants. In the presence of allelic heterogeneity, we would expect that multiple haplotypes would show association with BD. Using the 10 SNP window result from the haplotype analysis in the EA population, which provide p-values for each haplotype in addition to the OMNIBUS p-values, we investigated the proportion of haplotypes within each window that had p-values less than 0.05 (Figure 2). In the *ANK3* region, there were many 10-SNP windows that had a high proportion of low p-value haplotypes (up to 75%, genome-wide average = 4.7%), with 2-5 haplotypes often being implicated. In order to see if this was unexpected in the genome, we smoothed the proportion of significant haplotypes across the genome and looked for other regions that matched or exceeded the maximum smoothed value found in *ANK3*. We found 10 additional regions covering approximately 1.5 Mb or about 0.05% of the genome that exceeded this level of haplotype heterogeneity, indicating that the proportion of haplotypes that are associated with BD at this locus is relatively high, compared to the rest of the genome (Table S3). This provides additional support that the region shows allelic heterogeneity.

We further analyzed our results relative to one of the single previously published GWA studies (i.e., Baum et al., 2008) (15), which included about half of the probands and controls used in the present study, in addition to a replication sample that is independent of our sample (i.e., was collected in Germany). Several of the 88 SNPs with replicated association signals in the Baum et al. study (15) also show nominal evidence of association in the present study (Table S4), including a SNP in *ANK3* (i.e., rs9804190). The extent to which this can be considered evidence for a consistent finding, however, is limited given the overlap between the samples used in the two studies.

Genetic Background x SNP Interaction Analysis

Table 2 depicts evidence of SNP associations with genetic background-dependent effects. As shown, no SNP showed strong genetic background-dependent effects (i.e., no p-values less than 5×10^{-8}). We do, however, report the top 3 SNPs, which occur in *ROR1*, *RGS5*, and *BTBD16* (all p-values $< 4.5 \times 10^{-6}$). We also show the OR across different categories of admixture to depict these results (Figure 1).

Comparisons of Different Methods for Determining Ancestry

Based on LAMP ancestry estimates using all autosomal SNPs, the EA set showed less than 1% Yoruban admixture (mean= 0.998, range: 0.944-1, STD = 0.005), whereas the AA set showed almost 19% European admixture (mean = 0.188, range: 0.0-1, STD = 0.124). Furthermore, LAMP ancestry estimates including 3053 EA and AA subjects were significantly correlated with estimates generated with the more traditional STRUCTURE method ($r = 0.999$, $p < 0.001$). Including parental allele frequencies of HGDP subjects instead of HapMap subjects did not influence ancestry estimates ($r = 1.0$, $p < 0.001$).

Analysis of Subjects that did not Overlap with the PNDRC

Results of analyses that included the subgroup of EA subjects that did not overlap with the Pritzker Study revealed a different top hit relative to analyses that included the entire sample. Specifically, the most significantly associated SNP in this subsample of EA subjects was rs6046396, which is upstream of *RIN2* at 20p11. 23, $p = 1.43 \times 10^{-6}$ (see Table 1).

Replication Genotyping Analyses

Results from the replication analysis are shown in Table 3, with SNPs shown from genomic regions that demonstrated association ($p < 0.03$) in the family-based replication sample. The analysis of the case-control replication cohort did not detect significance below the multiple-testing threshold ($p < 0.001$) among the 85 SNPs genotyped, although three of the SNPs in the *C15ORF53* region demonstrated some evidence of association in the case-control cohort ($p = 0.03$-0.04) as well as in the family sample ($p = 0.008$-0.015). The preponderance of the association evidence from the family based analysis reported in Table 3 is derived from transmission of SNP alleles to affected individuals. These transmissions are effectively independent of the population association (21). The meta-analysis (Table 3) gave p-values of 0.004-0.01 for SNPs in this region, and reinforced the association evidence for SNP rs13358880 on chromosome 5.

DISCUSSION

We conducted two GWA studies, one in a sample of individuals of EA and the second in a sample of individuals of AA. In order to account for possible genetic background differences, we: 1) considered the analysis of each sample separately; 2) estimated ancestry and genetic background diversity from the genetic data and controlled for it in the association studies; and 3) looked for evidence of genetic background x SNP interactions. In order to qualify our results, we also compared them to previous GWA study results investigating BD, performed replication genotyping of our most strongly associated SNPs on an independent cohort, and conducted analysis of only the non-overlapping subjects in our study (with another study's subjects) in order to tease out independent evidence for associations.

Although no single SNP showed significant association after correction for genome-wide testing in either of our populations, some noteworthy associations were observed with BD candidate genes, as well as with genes known to be expressed in human brain. Of particular interest are *SLITRK2* and *NTRK2*. *SLITRK2* is a member of a family of six genes which are widely expressed in neural tissue (22), producing proteins which are membrane bound. *SLITRK2* regulates neurite outgrowth *in vitro*. Thus, *SLITRK2* is a logical bipolar risk gene. Another member of this gene family, *SLITRK1*, has been reported to be associated with Tourette's syndrome (23). *NTRK2* (also known as TrkB) is a tyrosine kinase receptor which binds brain derived neurotrophic factor (*BDNF*) and possibly other neurotrophins, i.e., (for review see 24). *NTRK2* is a high priority bipolar candidate gene for several reasons. There is abundant evidence from animal models of depression that hippocampal neurogenesis is decreased during the behavioral syndrome, that antidepressants increase neurogenesis, and that *BDNF* has antidepressant-like properties in these animal models (25). *BDNF* expression is increased in animals by treatment with antidepressants or lithium, and *BDNF* SNPs have repeatedly been implicated in genetic risk for BD, although the effect size is quite limited, e.g. odds ratio of 1.1 (26-29).

Our genetic background analysis revealed significant levels of admixture among our AA study subjects. Given this, we explored the comparability of different methods to account for admixture in our analyses and found that three different methods – genomic control adjustment, logistic regression using a LAMP-generated covariate, and logistic regression using MDS-generated covariates – all produced very similar results within the AA population. Regression with either LAMP or MDS based covariates was effective for correcting artificially induced population structure when EA and AA samples were combined, but genomic control overcorrected for this, effectively removing any real association that was not associated with admixture levels.

Different putative associations were observed among individuals of EA and individuals of AA when analyzed separately, and analyses assessing the extent to which some SNPs show genetic background-dependent effects highlighted different areas of potential association. We believe that this is the first report emerging from a GWA to explicitly address the dependency of SNP effects on ancestry, admixture, and/or genetic background.

We sought to replicate our most strongly associated SNPs in two independent sets of subjects, one family-based, and one case-control. Positive results were seen with rs1495186 in *C15ORF53* and some additional SNPs in that region. When considered cumulatively *via* meta-analysis with the primary EA sample, these results demonstrate consistent support for association of SNPs in the *C15ORF53* region with BD, and provide additional supportive evidence of association in other regions.

We also considered the consistency of our results with previous BD GWA studies. There have been three previous GWA studies of BD (13,15,16), in addition to a more recent collaborative GWA meta-analysis study (17), a subset of which were represented in the previous independent GWA studies. The Ferreira et al. (2008) collaborative GWA meta-analysis study (17) identified a region of strong association in *ANK3*, apparently distinct from that detected in Baum et al. (2008). Ferreira et al. also found new evidence for association at 15q14, which is near *C15ORF53*, as well as further support for the previously reported *CACNA1C* gene; these authors concluded that ion channelopathies may be involved in the pathogenesis of BD.

In the current study, we found consistent evidence of both previously-reported *ANK3* findings, and borderline support for replication of a region characterized at 15q14, although we failed to find support for the finding at *CACNA1C*. In *ANK3* and at 15q14, multiple SNPs in weak to no linkage disequilibrium with the previously associated SNP showed stronger association. Investigation of haplotype-based associations in our study provides support for allelic heterogeneity in *ANK3* region. Allelic heterogeneity has the potential to play an important role in genetically influenced disorders, yet can be difficult to detect in population-based samples using common variants, making it a potential explanation of "missing heritability" (30). Of note, and as we have previously indicated however, the Baum et al. (2008) and Sklar et al. (2008) samples both overlap with the current sample (15,16), thus stringent conclusions pertaining to replication are not warranted.

Our GWA study of BD provides some support for previous findings that variation in *ANK3* and at 15q14 influence BD susceptibility. In addition, regions containing *NAP5*, *NTRK2*, *SLITRK2*, and *ROR1* are worthy of follow-up studies. As all of these associated SNPs and regions have small effect size, it is likely, however, that the majority of the genetic variations that influence BD remain yet to be discovered.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank the participants in the study, as without them this work would not have been possible. For best estimate diagnostic work, we thank Vegas Coleman MD, Robert Schweitzer MD, N. Leela Rau MD, and Kelly Rhoadarmer MD. For data management we thank Mariano Erpe and for study coordination Carre Fisher RN. This work was supported by grants from the NIMH and NHGRI to JRK (MH078151, MH081804, MH059567 supplement), and by the Genetic Association Information Network (GAIN). This work was additionally supported by the NIMH Intramural Research Program (FJM and TGS). WHB was supported by a grant from the Tzedakah Foundation, a grant from NIH (R01 MH59553) and a grant from Philip and Marcia Cohen. Falk Lohoff was supported by the Daland Fellowship Award from the American Philosophical Society and by NIH grant K08 MH080372. David

Craig and Corrie Panganiban would like to acknowledge the Stardust foundation. Follow-up genotyping was performed in the laboratory of HE at Indiana University School of Medicine. This research was also supported, in part, by the Intramural Research Program of the NIH, National Library of Medicine. Additionally, this project was supported by NIH/NCRR Grant Number UL1 RR025774. Its contents are the authors(tm) sole responsibility and do not necessarily represent official NIH views. (ENS, CSB, and NJS).

Data and biomaterials were collected in four projects that participated in the National Institute of Mental Health (NIMH) Bipolar Disorder Genetics Initiative. From 1991-98, the Principal Investigators and Co-Investigators were: Indiana University, Indianapolis, IN, U01 MH46282, John Nurnberger, M.D., Ph.D., Marvin Miller, M.D., and Elizabeth Bowman, M.D.; Washington University, St. Louis, MO, U01 MH46280, Theodore Reich, M.D., Allison Goate, Ph.D., and John Rice, Ph.D.; Johns Hopkins University, Baltimore, MD U01 MH46274, J. Raymond DePaulo, Jr., M.D., Sylvia Simpson, M.D., MPH, and Colin Stine, Ph.D.; NIMH Intramural Research Program, Clinical Neurogenetics Branch, Bethesda, MD, Elliot Gershon, M.D., Diane Kazuba, B.A., and Elizabeth Maxwell, M.S.W.

Data and biomaterials were collected as part of ten projects that participated in the National Institute of Mental Health (NIMH) Bipolar Disorder Genetics Initiative. From 1999-07, the Principal Investigators and Co-Investigators were: Indiana University, Indianapolis, IN, R01 MH59545, John Nurnberger, M.D., Ph.D., Marvin J. Miller, M.D., Elizabeth S. Bowman, M.D., N. Leela Rau, M.D., P. Ryan Moe, M.D., Nalini Samavedy, M.D., Rif El-Mallakh, M.D. (at University of Louisville), Husseini Manji, M.D. (at Wayne State University), Debra A. Glitz, M.D. (at Wayne State University), Eric T. Meyer, M.S., Carrie Smiley, R.N., Tatiana Foroud, Ph.D., Leah Flury, M.S., Danielle M. Dick, Ph.D., Howard Edenberg, Ph.D.; Washington University, St. Louis, MO, R01 MH059534, John Rice, Ph.D, Theodore Reich, M.D., Allison Goate, Ph.D., Laura Bierut, M.D.; Johns Hopkins University, Baltimore, MD, R01 MH59533, Melvin McInnis M.D., J. Raymond DePaulo, Jr., M.D., Dean F. MacKinnon, M.D., Francis M. Mondimore, M.D., James B. Potash, M.D., Peter P. Zandi, Ph.D, Dimitrios Avramopoulos, and Jennifer Payne; University of Pennsylvania, PA, R01 MH59553, Wade Berrettini M.D., Ph.D.; University of California at Irvine, CA, R01 MH60068, William Byerley M.D., and Mark Vawter M.D.; University of Iowa, IA, R01 MH059548, William Coryell M.D., and Raymond Crowe M.D.; University of Chicago, IL, R01 MH59535, Elliot Gershon, M.D., Judith Badner Ph.D., Francis McMahon M.D., Chunyu Liu Ph.D., Alan Sanders M.D., Maria Caserta, Steven Dinwiddie M.D., Tu Nguyen, Donna Harakal; University of California at San Diego, CA, R01 MH59567, John Kelsoe, M.D., Rebecca McKinney, B.A.; Rush University, IL, R01 MH059556, William Scheftner M.D., Howard M. Kravitz, D.O., M.P.H., Diana Marta, B.S., Annette Vaughn-Brown, MSN, RN, and Laurie Bederow, MA; NIMH Intramural Research Program, Bethesda, MD, 1Z01MH002810-01, Francis J. McMahon, M.D., Layla Kassem, PsyD, Sevilla Detera-Wadleigh, Ph.D, Lisa Austin, Ph.D, Dennis L. Murphy, M.D.

JRK is a founder and holds equity in Psynomics, Inc.

References

1. Goodwin, FK.; Jamison, KR.; Akiskal, H.; Fawcett, J.; Ghaemi, N.; Hammen, C., et al. Manic-Depressive Illness. 2. Oxford University Press; 2005.
2. Tsuang, MG.; Faraone, SV. The Genetics of Mood Disorders. Johns Hopkins University Press; Baltimore: 1990.
3. Gershon ES, Hamovit J, Guroff JJ, Dibble E, Leckman JF, Sceery W, et al. A family study of schizoaffective, bipolar I, bipolar II, unipolar, and normal control probands. Arch Gen Psychiatry 1982 Oct;39(10):1157–1167. [PubMed: 7125846]
4. Rice J, Reich T, Andreasen NC, Endicott J, Van Eerdewegh M, Fishman R, et al. The familial transmission of bipolar illness. Arch Gen Psychiatry 1987 May;44(5):441–447. [PubMed: 3579495]
5. Spence MA, Flodman PL, Sadovnick AD, Bailey-Wilson JE, Ameli H, Remick RA. Bipolar disorder: evidence for a major locus. Am J Med Genet 1995 Oct 9;60(5):370–376. [PubMed: 8546148]
6. Craddock N, Khodel V, Van Eerdewegh P, Reich T. Mathematical limits of multilocus models: the genetic transmission of bipolar disorder. Am J Hum Genet 1995 Sep;57(3):690–702. [PubMed: 7668299]
7. Risch N, Botstein D. A manic depressive history. Nat Genet 1996 Apr;12(4):351–353. [PubMed: 8630482]
8. Badner JA, Gershon ES. Meta-analysis of whole-genome linkage scans of bipolar disorder and schizophrenia. Mol Psychiatry 2002;7(4):405–411. [PubMed: 11986984]
9. Segurado R, Detera-Wadleigh SD, Levinson DF, Lewis CM, Gill M, Nurnberger JI Jr, et al. Genome scan meta-analysis of schizophrenia and bipolar disorder, part III: Bipolar disorder. Am J Hum Genet 2003 Jul;73(1):49–62. [PubMed: 12802785]

10. 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]
11. Nurnberger, JI.; Berrettini, W.; Niculescu, AB. Genetics of psychiatric disorders. In: Fatemi, SH.; Clayton, PJ., editors. The Medical Basis of Psychiatry. Humana Press; 2008. p. 487-518.
12. 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]
13. WTCCC. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature 2007 Jun 7;447(7145):661–678. [PubMed: 17554300]
14. Altshuler D, Daly MJ, Lander ES. Genetic mapping in human disease. Science 2008 Nov 7;322(5903):881–888. [PubMed: 18988837]
15. Baum AE, Akula N, Cabanero M, Cardona I, Corona W, Klemens B, et al. A genome-wide association study implicates diacylglycerol kinase eta (DGKH) and several other genes in the etiology of bipolar disorder. Mol Psychiatry 2008 Feb;13(2):197–207. [PubMed: 17486107]
16. Sklar P, Smoller JW, Fan J, Ferreira MA, Perlis RH, Chambert K, et al. Whole-genome association study of bipolar disorder. Mol Psychiatry 2008 Jun;13(6):558–569. [PubMed: 18317468]
17. Ferreira MA, 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 Aug 17;
18. 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]
19. Devlin B, Roeder K. Genomic control for association studies. Biometrics 1999 Dec;55(4):997–1004. [PubMed: 11315092]
20. Freedman ML, Reich D, Penney KL, McDonald GJ, Mignault AA, Patterson N, et al. Assessing the impact of population stratification on genetic association studies. Nat Genet 2004 Apr;36(4): 388–393. [PubMed: 15052270]
21. Abecasis GR, Cookson WO, Cardon LR. Pedigree tests of transmission disequilibrium. Eur J Hum Genet 2000 Jul;8(7):545–551. [PubMed: 10909856]
22. Aruga J, Yokota N, Mikoshiba K. Human SLITRK family genes: genomic organization and expression profiling in normal brain and brain tumor tissue. Gene 2003 Oct 2;315:87–94. [PubMed: 14557068]
23. Abelson JF, Kwan KY, O'Roak BJ, Baek DY, Stillman AA, Morgan TM, et al. Sequence variants in SLITRK1 are associated with Tourette's syndrome. Science 2005 Oct 14;310(5746):317–320. [PubMed: 16224024]
24. Reichart LF. Neurotrophin-regulated signalling pathways. Phil Trans R Soc B 2006;(361):1545–1564. [PubMed: 16939974]
25. Schmidt HD, Duman RS. The role of neurotrophic factors in adult hippocampal neurogenesis, antidepressant treatments and animal models of depressive-like behavior. Behav Pharmacol 2007 Sep;18(5-6):391–418. [PubMed: 17762509]
26. Neves-Pereira M, Mundo E, Muglia P, King N, Macciardi F, Kennedy JL. The brain-derived neurotrophic factor gene confers susceptibility to bipolar disorder: evidence from a family-based association study. Am J Hum Genet 2002 Sep;71(3):651–655. [PubMed: 12161822]
27. Sklar P, Gabriel SB, McInnis MG, Bennett P, Lim YM, Tsan G, et al. Family-based association study of 76 candidate genes in bipolar disorder: BDNF is a potential risk locus. Brain-derived neutrophic factor. Mol Psychiatry 2002;7(6):579–593. [PubMed: 12140781]
28. Lohoff FW, Sander T, Ferraro TN, Dahl JP, Gallinat J, Berrettini WH. Confirmation of association between the Val66Met polymorphism in the brain-derived neurotrophic factor (BDNF) gene and bipolar I disorder. Am J Med Genet B Neuropsychiatr Genet 2005 Nov 5;139B(1):51–53. [PubMed: 16152572]
29. Liu L, Foroud T, Xuei X, Berrettini W, Byerley W, Coryell W, et al. Evidence of association between brain-derived neurotrophic factor gene and bipolar disorder. Psychiatr Genet 2008 Dec; 18(6):267–274. [PubMed: 19018231]

30. Maher B. Personal genomes: The case of the missing heritability. Nature 2008 Nov 6;456(7218): 18–21. [PubMed: 18987709]

31. APA. Diagnostic and Statistical Manual of Mental Disorders. 4-Text Revision. Washington DC: 2000.

32. 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. discussion 863-844. [PubMed: 7944874]

33. Feighner JP, Robins E, Guze SB, Woodruff RA Jr, Winokur G, Munoz R. Diagnostic criteria for use in psychiatric research. Arch Gen Psychiatry 1972 Jan;26(1):57–63. [PubMed: 5009428]

34. Akiskal HS, Pinto O. The evolving bipolar spectrum. Prototypes I, II, III, and IV. Psychiatr Clin North Am 1999 Sep;22(3):517–534. vii. [PubMed: 10550853]

35. Horne JA, Ostberg O. A self-assessment questionnaire to determine morningness-eveningness in human circadian rhythms. Int J Chronobiol 1976;4(2):97–110. [PubMed: 1027738]

36. Brown GL, Goodwin FK, Ballenger JC, Goyer PF, Major LF. Aggression in humans correlates with cerebrospinal fluid amine metabolites. Psychiatry Res 1979 Oct;1(2):131–139. [PubMed: 95232]

37. Cloninger, CR.; Przybeck, TR.; Svrakic, DM.; Wetzel, RD. The temperament and character inventory (TCI): A guide to its development and use. Center for Psychobiology of Personality; St. Louis: 1994. p. 1-85.

38. Wender EH. Attention-deficit hyperactivity disorders in adolescence. J Dev Behav Pediatr 1995 Jun;16(3):192–195. [PubMed: 7560123]

39. Zuckerman M, Kuhlman DM, Joireman J, Teta P, Kraft M. A comparison of three structural models of personality: The big three, the big five, and the alternative five. Journal of Personality and Social Psychology 1993;65:757–768.

40. McCarroll SA, Kuruvilla FG, Korn JM, Cawley S, Nemesh J, Wysoker A, et al. Integrated detection and population-genetic analysis of SNPs and copy number variation. Nat Genet 2008 Oct;40(10):1166–1174. [PubMed: 18776908]

41. 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]

42. Sankararaman S, Sridhar S, Kimmel G, Halperin E. Estimating local ancestry in admixed populations. Am J Hum Genet 2008 Feb;82(2):290–303. [PubMed: 18252211]

43. Pritchard JK, Stephens M, Donnelly P. Inference of population structure using multilocus genotype data. Genetics 2000 Jun;155(2):945–959. [PubMed: 10835412]

44. Falush D, Stephens M, Pritchard JK. Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. Genetics 2003 Aug;164(4):1567–1587. [PubMed: 12930761]

45. Li JZ, Absher DM, Tang H, Southwick AM, Casto AM, Ramachandran S, et al. Worldwide human relationships inferred from genome-wide patterns of variation. Science 2008 Feb 22;319(5866): 1100–1104. [PubMed: 18292342]

46. Zeggini E, Scott LJ, Saxena R, Voight BF, Marchini JL, Hu T, et al. Meta-analysis of genome-wide association data and large-scale replication identifies additional susceptibility loci for type 2 diabetes. Nat Genet 2008 May;40(5):638–645. [PubMed: 18372903]

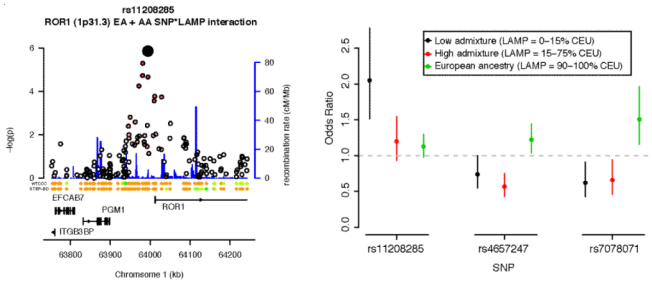

Figure 1. Top genetic background-dependent SNPs

The region around the top background-dependent SNP. The primary SNP is colored in black. Other SNPs are colored according to linkage disequilibrium levels with the primary SNP (r2) calculated from Phase 3 HapMap data using the CEU] population. Recombination rate (HapMap) is shown on the second y-axis in blue. RefSeq genes are shown with all possible exons; arrows indicate transcript direction. In the upper left hand corner of each graph, the genotype intensity plots are shown, with each color indicating the final genotype call (blue and red for homozygotes and purple for the heterozygote). B. Odds ratios for the top three background-dependent SNPs within groups showing different degrees of admixture. Low and High admixture groups are all of African Ancestry, but are split into two groups by the approximate median percent CEU. Lines indicate 95% confidence intervals.

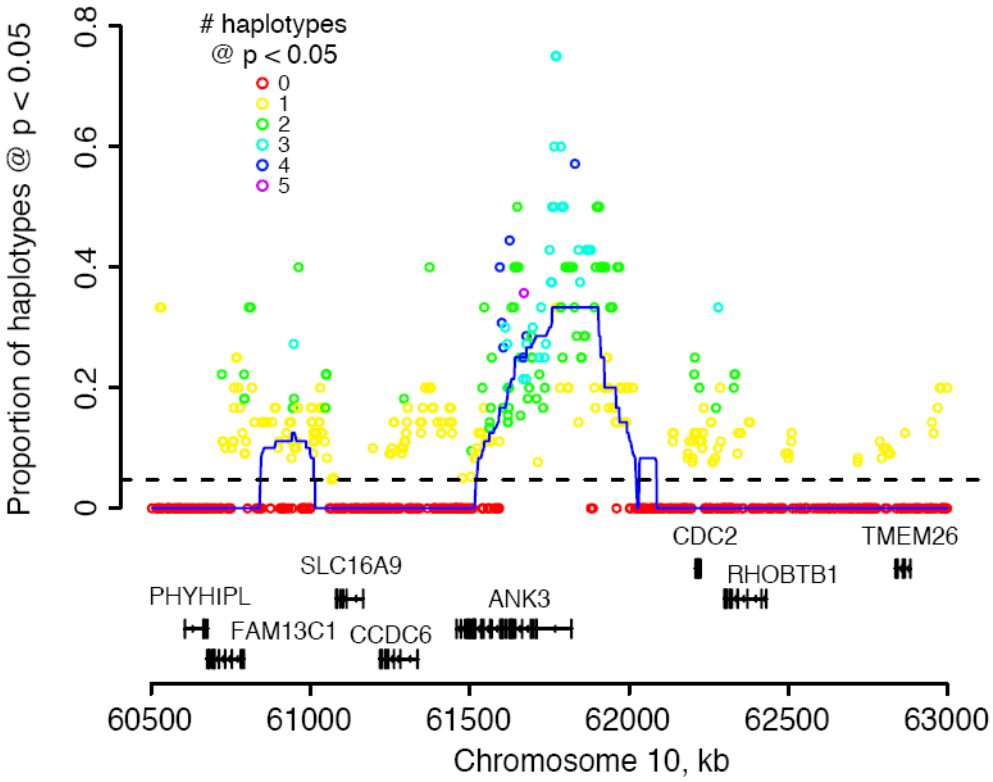

Figure 2. Muliti-haplotype association in the ANK3 region suggests allelic heterogeneity

Each point summarizes the proportion of haplotypes within a given 10 SNP window. The proportion of the haplotypes that show a $p < 0.05$ for a given window is plotted on the y-axis. The x-axis is the physical position of the midpoint of the haplotype. The points are colored according to the number of haplotypes that show $p < 0.05$. The proportion was median smoothed over 101 windows (blue line). The red dotted line indicates the genome-wide mean proportion of haplotypes with a $p < 0.05$. RefSeq genes are shown below the x-axis.

Table 1

Most strongly associated SNPs using different methods to control for admixture and replication p-values.

Analysis	SNP/Allele	Location	MAF	Multidimensional Scaling		LAMP-Adjusted		Genomic Control	
				OR	p-value	OR	p-value	OR	p-value
AA only	rs2111504/T[a]	*DPY19L3* (19q13.11)	0.23	1.74	1.5×10^{-6}	1.73	1.7×10^{-6}	1.66	5.0×10^{-6}
	rs2769605/T[b]	*NTRK2* (9q21.33)	0.22	0.60	4.5×10^{-5}	0.61	4.5×10^{-5}	0.63	1.5×10^{-4}
EA only	rs1825828/C[a,c]	3q11.2	0.26	0.70	7.0×10^{-7}	0.70	7.6×10^{-7}	0.68	2.9×10^{-7}
	rs5907577/T[a]	Xq27.1	0.31	1.48	1.6×10^{-6}	1.47	2.1×10^{-6}	1.444	4.6×10^{-6}
	rs10193871/G[b]	*NAP5* (2q21.2)	0.13	0.65	9.8×10^{-6}	0.65	7.1×10^{-6}	0.65	8.8×10^{-6}
EA/AA combined (main effect)	rs4825220/C[a]	Xq27.1	0.37	0.71	2.6×10^{-7}	0.70	1.9×10^{-7}	0.61	0.8496[d]
EA Pritzker Non-Overlap	rs6046396/G[a]	*RIN2* (20p11.23)	0.29	1.54	1.43×10^{-6}	1.56	6.2×10^{-7}	1.57	4.93×10^{-7}

[a] Overall top hit

[b] Top hit in region with at least 5 SNPs at $P < 1 \times 10^{-4}$

[c] Possibly of poor genotype quality (see text and Figure S2)

[d] This p-value likely reflects the imbalance in the ratio of cases and controls in the EA relative to the AA sample.

Table 2

SNPs showing strongest genetic background-dependent effects (EA/AA combined sample with interaction term).

SNP/Allele	Location	MAF	OR	SNP x Ancestry Interaction p-value
rs11208285/A	*ROR1* (1p31.3)	0.38	2.05	1.4 × 10-6
rs4657247/T	*RGS5* (1q23.3)	0.21	2.11	4.1 × 10-6
rs7078071/T	*BTBD16* (10q26.13)	0.10	2.85	4.5 × 10-6

Table 3

Replication genotyping results.

SNP	Chromosome	Position	Gene	MAF	Family p-value (unaffectedness criteria)		Replication case-control p-value	Meta-analysis p-value
					Permissive	Strict		
rs4671825	2	67762350		0.47	0.186	0.008	0.639	0.099
rs4671826	2	67762453		0.47	0.137	0.008	0.640	0.084
rs13358880	5	56845056		0.14	0.204	0.041	0.196	7.1E-05
rs4389663	5	104451583		0.17	0.166	0.047	0.918	0.130
rs10981870	9	115462097		0.34	0.137	0.047	0.449	0.022
rs4924281	15	36742939		0.27	0.011	0.034	0.363	0.001
rs1495186	15	36778748	*C15ORF53*	0.23	0.015	0.043	0.045	0.009
rs903530	15	36786340		0.23	0.010	0.054	0.033	0.012
rs2643217	15	36791055		0.23	0.008	0.057	0.033	0.004
rs2012353	19	37591133	*DPY19L3*	0.17	0.116	0.030	0.349	0.106

Mol Psychiatry. Author manuscript; available in PMC 2011 February 9.