Supporting Information

Scott et al. 10.1073/pnas.0813386106

SI Methods

Assessment of Stratification. In each set of study samples we performed principal components (PC) analysis based on the sample genotypes (1) in unrelated individuals.

NIMH/Pritzker. We chose every 5th genotyped SNP in the genome for the PC analysis in 1,478 unrelated individuals, choosing 1 person per sibling pair. We assigned the non-included relatives the PC values of the included relative. We excluded 6 individuals with PC >6 SD from the mean of 1 or more of the top 10 PCs.

GSK. We chose SNPs for PC analysis after removing SNPs located in 4 regions of high LD (2) Within windows of 1,500 SNPs, we selected SNPs with pairwise $r^2 < 0.2$. We used the Tracy-Widom test (as described in ref. 2) to assess the significance of the PCs and retained 16 PCs with P value<.05. 26 samples had a PC >6 SD from the mean and were excluded. We repeated the PC analysis with the remaining samples and obtained 7 significant PCs.

WTCCC. We chose every 5th genotyped SNP in the genome for the PC analysis. 120 individuals differed by >6 SD from the mean of a PC (110 from the first PC with 17 case and 93 controls). We retained these individuals to mirror the analysis used by the WTCCC (3).

GWA Analysis. We eliminated 694 SNPs with allele frequency differences >.2 for any pair of studies. We analyzed the observed allele counts or imputed allele dosages using logistic regression assuming an additive genetic model, with covariates as described below, and then repeated the GWA analysis without PCs.

NIMH/Pritzker. We included the 10 most significant PCs as covariates. To account for the presence of case siblings, we used a sandwich estimator (4) to adjust the estimated variances. By clustering on the sibships, the sandwich estimator provides consistent estimates of the variances for the parameters of interest. We also analyzed the data without PC using the sandwich estimator and a method proposed by Bourgain et al. (5). The Bourgain method uses χ^2 test statistic that takes account of the familial relationship in the association test; however, the method cannot adjust for covariates.

GSK. We included recruitment site and the 7 significant PCs as covariates. We analyzed the full sample and we also analyzed a reduced sample removing the 261 London cases in the WTCCC sample.

WTCCC. We compared the BP cases to the extended reference set as our primary analysis. We also tested for marker association comparing the WTCCC NBS controls to the combined 6 sets of non-BP cases. Because of strong signals in this latter analysis in the HLA region on chromosome 6 from 27.2 to 34.0 Mb, we excluded the 5,571 autoimmune disease cases (type 1 diabetes, Crohn's disease, rheumatoid arthritis) from the primary GWA analysis of this region. We included the top 10 PCs as covariates.

Meta-analysis of GWA Samples. We performed a fixed effects meta-analysis using the OR and 95% confidence intervals to combine the association evidence from the study-specific GWA analyses. We used association results for experimentally derived genotypes when available, and for imputed genotypes otherwise. 2,366,197 autosomal SNPs passed QC and had MAF≥.01 in all 3 samples; 75,477 were genotyped in all 3 samples, 412,455 in 2, 312,438 in 1, and 1,565,827 in no samples. Association results were oriented relative to the forward strand of the reference genome (dbSNP125). We adjusted for the genomic control values in each study separately for genotyped and imputed SNPs by increasing the standard error of the OR estimate to correspond to the genomic control P value. Evidence for heterogeneity between ORs was assessed using Cochrans's Q statistic and I^2 (6).

Assessment of Independence of Associated SNPs in Selected Regions. In our 3 regions of strongest association in the 3-study meta-analysis, we tested whether the most strongly associated SNP in the region could account for the association signals at nearby SNPs by including the most strongly associated SNP as a covariate in the logistic regression for each study. Adjusted association results were then combined using fixed effects meta-analysis as described above.

SI Results

Evaluation of the Sandwich Estimator in the Case-Control Analysis of NIMH/Pritzker Sample and Comparison of Results of Analysis with 1 or 2 Siblings per Family. We compared the NIMH/Pritzker BP association results using the sandwich estimator (4) but without covariates to those using Bourgain's method (5). The Pearson correlation coefficient was 0.9994 between the logarithm (base 10) of the *P* values from the 2 approaches. We also compared the analysis of the full NIMH/Pritzker sample to one with a single sibling per family. The correlation coefficient was .87 and the results were randomly distributed around the expected diagonal line. These results suggest that inclusion of siblings in the case-control analysis was appropriately accounted for with the use of the sandwich estimator.

GWA Sample Overlap. The sample sets analyzed in the 3-study meta-analysis overlap those from other BP GWAS. 484 of the NIMH controls were included in the Ferreira et al. study (7). In our meta-analysis, for WTCCC controls we used the NBS sample and the expanded control set that included the WTCCC non-BP cases. WTCCC (3) and Ferreira et al. (7) primary analyses included as controls only the NBS and 1958 Birth Cohort controls, while their secondary analyses included the expanded control set. Our 3-sample meta-analysis set contains 1,833 cases and 992 controls independent of those in the Ferreira et al. (7) analyses. Many of the cases and controls in the pool-based GWAS (8) overlap with our NIMH/Pritzker samples. Finally, 437 cases and 357 controls from our NIMH/Pritzker sample are included in the GAIN GWAS (Nicholas Schork, personal communication).

SI ACKNOWLEDGMENTS. We thank Pablo Gejman for collection of the National Institute of Mental Health (NIMH) Bipolar Disorder Genetics Initiative control samples. Data and biomaterials for bipolar cases were collected as part of 10 projects that participated in the National Institute of Mental Health (NIMH) Bipolar Disorder Genetics Initiative. From 1999–03, 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, M.S.N., R.N., and Lauri Bederow, M.A.; 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. A full list of investigators who contributed to generation of the WTCCC data are available at www.wtccc.org.uk. Funding for the WTCCC was provided by the Wellcome Trust under award 076113.

1. Price AL, *et al* (2006) Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet* 38:904–909.
2. Patterson N, Price AL, Reich D (2006) Population structure and eigenanalysis. *PLoS Genet* 2:2074–2093.
3. Wellcome Trust Case Control Consortium (2007) Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature* 447:661–678.
4. White H (1982) Maximum likelihood estimation of misspecified model. *Econometrica* 50:1–25.
5. Bourgain C, et al. (2003) Novel case-control test in a founder population identifies P-selectin as an atopy-susceptibility locus. *Am J Hum Genet* 73:612–626.
6. Higgins JP, Thompson SG, Deeks JJ, Altman DG (2003) Measuring inconsistency in meta-analyses. *Br Med J* 327:557–560.
7. Ferreira MA, et al. (2008) Collaborative genome-wide association analysis supports a role for *ANK3* and *CACNA1C* in bipolar disorder. *Nat Genet* 40:1056–1058.
8. Baum AE, et al. (2008) A genome-wide association study implicates diacylglycerol kinase eta (*DGKH*) and several other genes in the etiology of bipolar disorder. *Mol Psychiatry* 13:197–207.

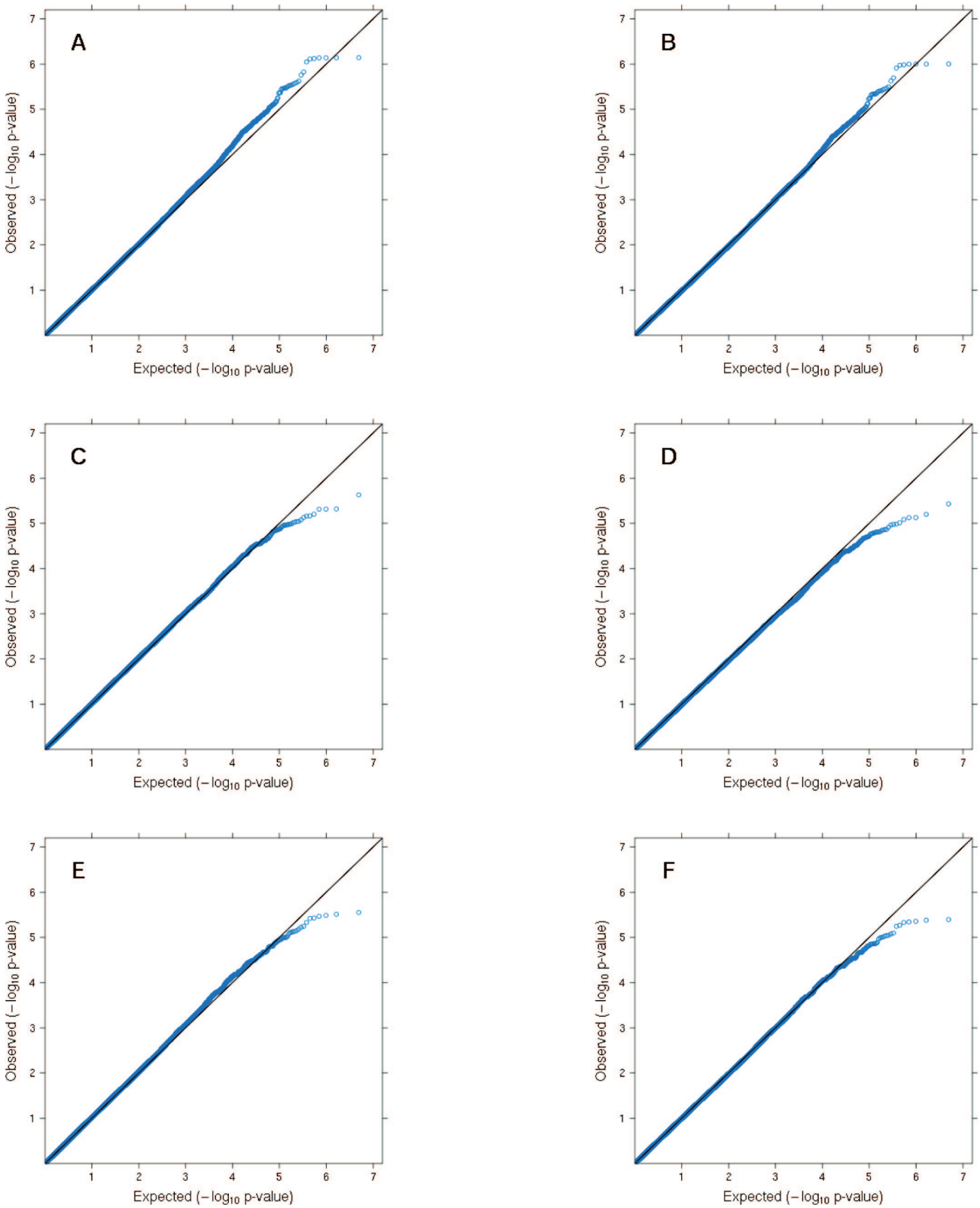

Fig. S1. Quantile-quantile plots of observed vs. expected $-\log_{10} P$ values for BP association including principal components as covariates. (*A*, *C*, *E* and *G*) Without correction for genomic control. (*B*, *D*, *E*, and *G*) With correction for genomic control. (*A* and *B*) NIMH/Pritzker. (*C* and *D*) GSK complete sample. (*E* and *F*) GSK reduced sample. (*G* and *H*) WTCCC (extended reference set).

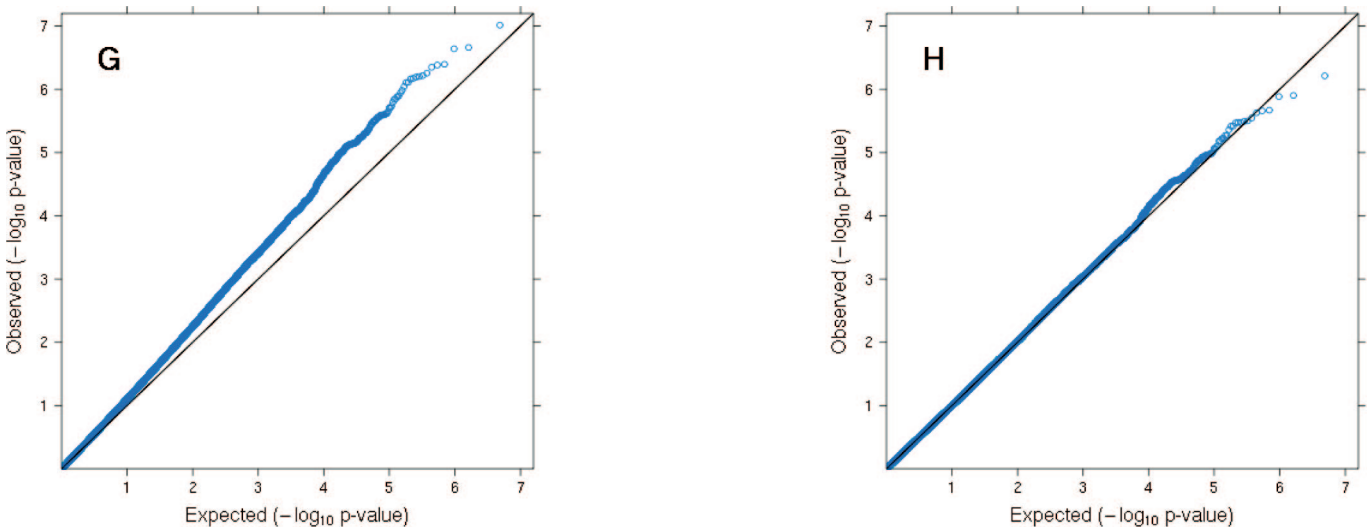

Fig. S1 (continued).

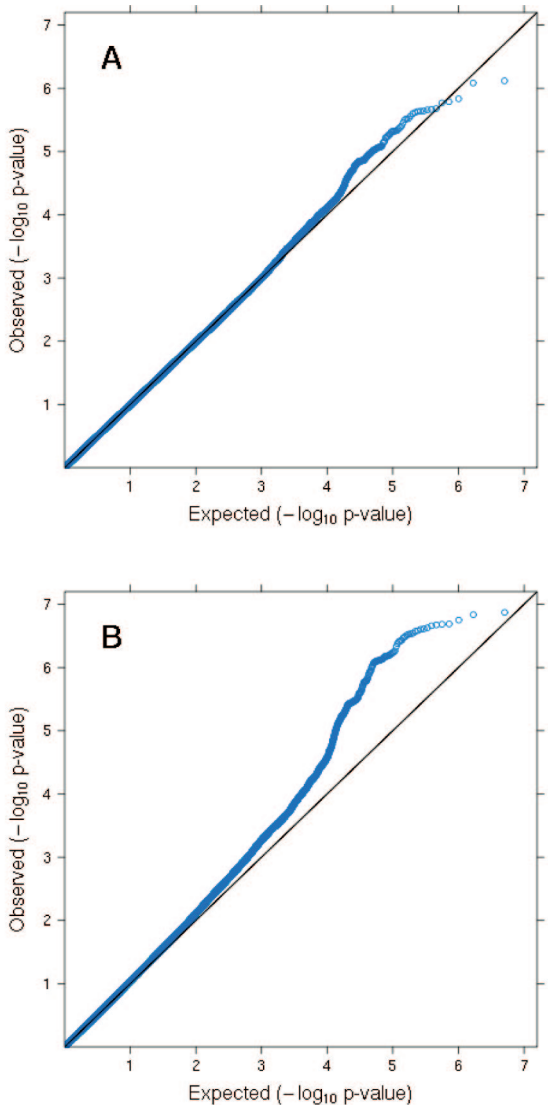

Fig. S2. Quantile-quantile plots of observed vs. expected $-\log_{10} P$ values for BP association meta-analysis including principal components as covariates. (*A*) NIMH/Pritzker and GSK (complete sample). (*B*) NIMH/Pritzker, GSK (reduced sample), and WTCCC.

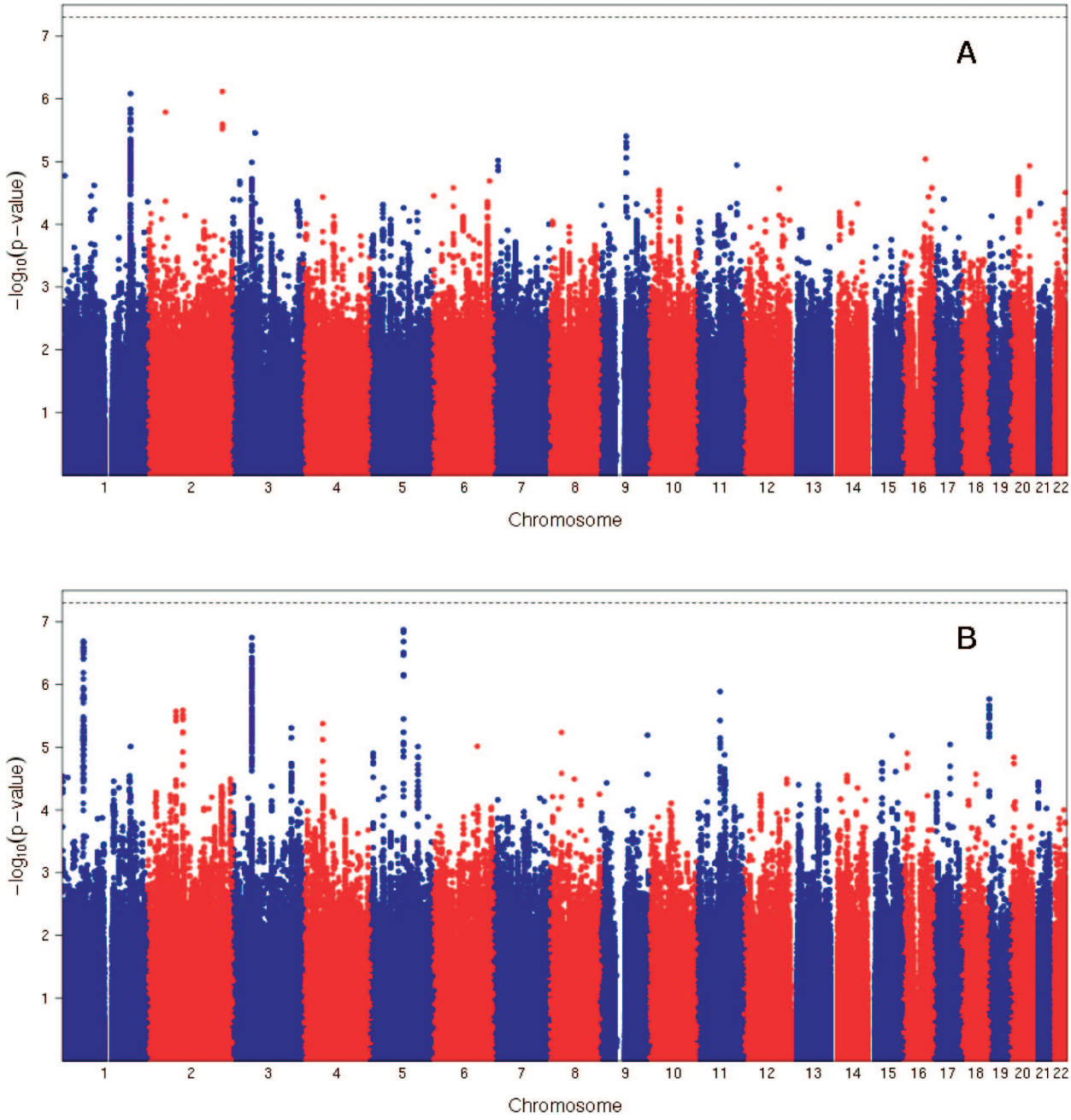

Fig. S3. Plots of -$\log_{10}$ *P* values for BP meta-analysis including principal components as covariates. (*A*) NIMH/Pritzker and GSK (complete sample). (*B*) NIMH/Pritzker, GSK (reduced sample), and WTCCC. The dotted line corresponds to a genome-wide significance threshold $P = 5 \times 10^{-8}$.

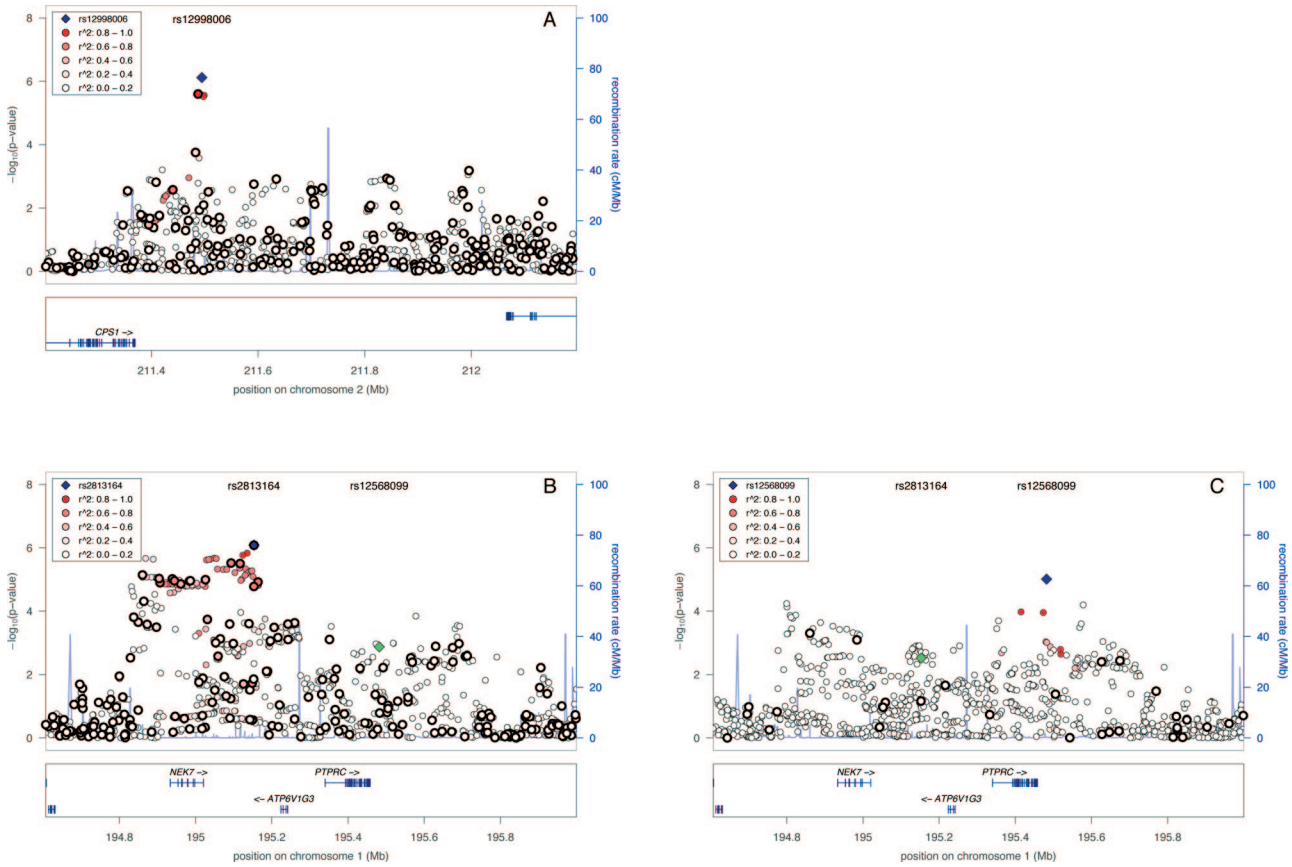

Fig. S4. Plot of $-\log_{10} P$ values for NIMH/Pritzker and GSK (complete sample) (*A* and *B*) and NIMH/Pritzker, GSK (reduced sample), and WTCCC BP-association meta-analysis for chromosomal regions with *P* values $<10^{-6}$ in NIMH/Pritzker and GSK (complete sample) meta-analysis (*C*). *B* and *C* show the same chromosome 1 region. Estimated recombination rates (from Hap Map) are plotted in cyan. Stronger red intensity indicates higher r^2 with the most significant SNP (purple diamond). The most strongly associated SNP in the other panel is shown in green. SNPs genotyped in all 2 or 3 samples are denoted by a thick black circle. refFLAT annotated genes are shown in *A*–*C Lower*.

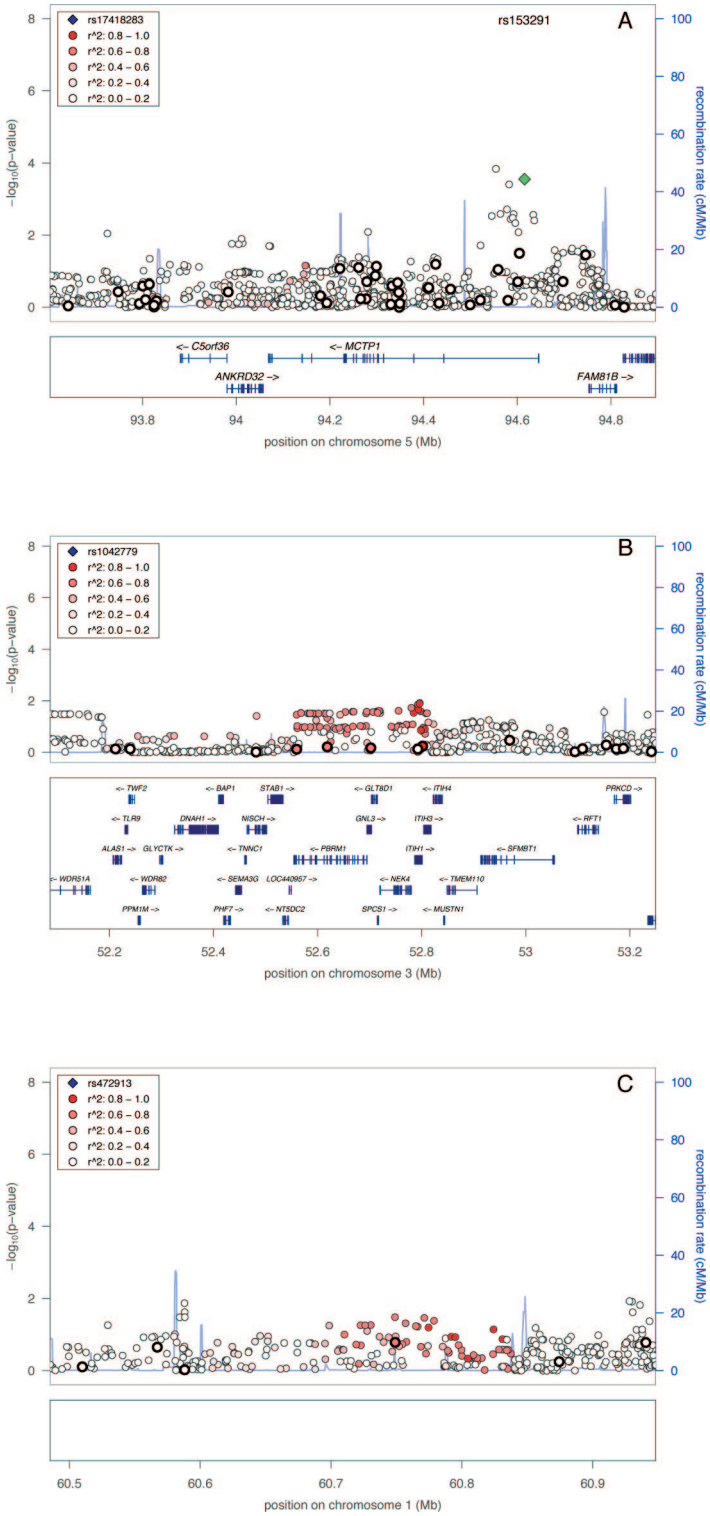

Fig. S5. Plot of $-\log_{10} P$ values for NIMH/Pritzker, GSK (reduced sample), and WTCCC BP association meta-analysis of individual study results conditioned on most strongly associated regional SNP. For each study and region, SNPs were analyzed using a logistic regression model containing the most strongly associated SNP. (*A*) rs17418283. (*B*) rs10426779. (*C*) rs472913. Estimated recombination rates (from Hap Map) are plotted in cyan. Stronger red intensity indicates higher r^2 with the most strongly associated SNP (purple diamond). SNPs genotyped in all 3 samples are denoted by a thick black circle. refFLAT annotated genes are shown in *A–C Lower*. A decrease in the $-\log_{10}$ P value from Fig. 1 indicates that the association signal of the surrounding SNPs can be explained, at least in part, by the most strongly associated SNP. Inclusion of a different nearby strongly associated SNP would have resulted in a similar picture.

Table S1A. NIMH/Pritzker and GSK (complete sample) bipolar meta-analysis association results: loci with $P < 10^{-5}$

SNP	Chr	Position* bp	Risk/ nonrisk allele	Control risk allele freq†	NIMH/Pritzker		GSK		Meta		Heterogeneity	
					OR (95% CI)	*P*	OR (95% CI)	*P*	OR (95% CI)	*P*	I^2, %	*P*
rs2813164	1	195,153,065	G/A	0.28	1.39 (1.19–1.62)	2.1×10^{-5}	1.24 (1.07–1.44)	0.0048	1.31 (1.18–1.46)	8.3×10^{-7}	9	0.30
rs7593459	2	49,607,351	T/A	0.40	1.27 (1.09–1.48)	0.0019	1.31 (1.14–1.51)	0.00014	1.29 (1.16–1.44)	1.6×10^{-6}	0	0.76
rs12998006	2	211,494,512	T/C	0.72	1.27 (1.07–1.5)	0.0061	1.42 (1.21–1.66)	1.3×10^{-5}	1.35 (1.2–1.51)	7.6×10^{-7}	0	0.33
rs11711888	3	62,132,251	A/G	0.96	3.31 (1.84–5.96)	6.7×10^{-5}	2.20 (1.24–3.91)	0.0071	2.69 (1.77–4.08)	3.5×10^{-6}	0	0.33
rs10246960	7	12,478,653	T/C	0.14	1.41 (1.16–1.71)	0.00061	1.33 (1.1–1.61)	0.0034	1.37 (1.19–1.57)	9.6×10^{-6}	0	0.68
rs7867133	9	72,037,564	A/G	0.72	1.23 (1.04–1.44)	0.013	1.39 (1.19–1.63)	3.3×10^{-5}	1.31 (1.17–1.47)	3.9×10^{-6}	20	0.26
rs17498325	16	63,499,388	G/A	0.82	1.42 (1.09–1.87)	0.010	1.46 (1.2–1.79)	0.00021	1.45 (1.23–1.71)	9.1×10^{-6}	0	0.88

*NBCI Build 35-bp position.
†Weighted average of control risk allele frequency for NIMH/Pritzker and GSK (complete sample).

Table S1B. Comparison of NIMH/Pritzker and GSK (complete sample) bipolar meta-analysis to WTCCC association results: loci with $P < 10^{-5}$

SNP	Chr	Position* bp	Risk/ nonrisk allele	Control risk allele freq†	Meta NIMH/Pritzker and GSK (complete sample)		WTCCC		Meta NIMH/Pritzker, GSK (reduced sample), and WTCCC	
					OR (95% CI)	*P*	OR (95% CI)	*P*	OR (95% CI)	*P*
rs2813164	1	195,153,065	G/A	0.28	1.31 (1.18–1.46)	8.3×10^{-7}	1.01 (0.93–1.09)	0.82	1.11 (1.03–1.18)	0.0030
rs7593459	2	49,607,351	T/A	0.40	1.29 (1.16–1.44)	$1. \times 0^{-6}$	1.00 (0.93–1.08)	0.97	1.08 (1.02–1.16)	0.012
rs12998006	2	211,494,512	T/C	0.72	1.35 (1.20–1.51)	7.6×10^{-7}	1.07 (0.97–1.17)	0.16	1.16 (1.08–1.26)	0.00013
rs11711888	3	62,132,251	A/G	0.96	2.69 (1.77–4.08)	3.5×10^{-6}	1.08 (0.82–1.43)	0.58	1.45 (1.13–1.85)	0.0035
rs10246960	7	12,478,653	T/C	0.14	1.37 (1.19–1.57)	9.6×10^{-6}	1.06 (0.97–1.17)	0.21	1.15 (1.06–1.25)	0.0012
rs7867133	9	72,037,564	A/G	0.72	1.31 (1.17–1.47)	3.9×10^{-6}	1.01 (0.93–1.09)	0.84	1.11 (1.04–1.19)	0.0032
rs17498325	16	63,499,388	G/A	0.82	1.45 (1.23–1.71)	9.1×10^{-6}	1.04 (0.94–1.15)	0.41	1.15 (1.05–1.26)	0.0036

GSK reduced sample: Excluding 261 BP cases also present in WTCCC sample.
*NBCI Build 35-bp position.
†Weighted average of control risk allele frequency for NIMH/Pritzker, GSK (reduced sample), and WTCCC.

Table S2. NIMH/Pritzker, GSK (reduced sample), and WTCCC bipolar meta-analysis association results: loci with $P < 10^{-5}$

SNP	Chr	Position* bp	Risk/ nonrisk allele	Control risk allele freq†	NIMH/Pritzker		GSK		WTCCC		Meta		Heterogeneity	
					OR (95% CI)	*P*	OR (95% CI)	*P*	OR (95% CI)	*P*	OR (95% CI)	*P*	I^2, %	*P*
rs472913	1	60,807,579	C/G	0.50	1.12 (0.97–1.29)	0.11	1.17 (1.00–1.36)	0.051	1.20‡ (1.11–1.28)	6.3×10^{-7}	1.18 (1.11–1.25)	2.0×10^{-7}	0	0.72
rs12568099	1	195,482,079	T/C	0.98	1.48 (0.90–2.42)	0.12	2.17 (1.26–3.77)	0.0054	1.49 (1.20–1.85)	0.00031	1.56 (1.28–1.9)	9.8×10^{-6}	0	0.44
rs13409348	2	79,451,643	G/C	0.24	1.12 (0.95–1.32)	0.18	1.20 (1.00–1.43)	0.045	1.22 (1.12–1.33)	7.1×10^{-6}	1.20 (1.11–1.29)	2.7×10^{-6}	0	0.67
rs6733011	2	98,924,020	G/A	0.46	1.11 (0.95–1.29)	0.19	1.13 (0.96–1.32)	0.14	1.20 (1.11–1.29)	1.9×10^{-6}	1.17 (1.1–1.25)	2.6×10^{-6}	0	0.59
rs1042779	3	52,796,051	A/G	0.63	1.20 (1.04–1.38)	0.015	1.31 (1.11–1.54)	0.0012	1.16 (1.07–1.25)	0.00012	1.19 (1.11–1.27)	1.8×10^{-7}	0	0.40
rs7427021	3	165,244,666	G/A	0.56	1.05 (0.91–1.23)	0.49	1.25 (1.06–1.47)	0.0089	1.17 (1.09–1.26)	1.5×10^{-5}	1.16 (1.09–1.24)	4.9×10^{-6}	14	0.31
rs2537859	4	55,323,746	T/C	0.60	1.05 (0.91–1.20)	0.53	1.41 (1.21–1.65)	1.1×10^{-5}	1.14 (1.06–1.23)	0.00071	1.16 (1.09–1.24)	4.2×10^{-6}	76	0.014
rs17418283	5	94,180,344	C/T	0.28	1.09 (0.93–1.28)	0.31	1.19 (1.01–1.41)	0.038	1.25 (1.15–1.36)	9.7×10^{-8}	1.21 (1.13–1.3)	1.3×10^{-7}	15	0.31
rs17169582	5	135,343,067	G/A	0.91	1.21 (0.94–1.55)	0.14	1.11 (0.86–1.43)	0.42	1.30 (1.16–1.46)	7.5×10^{-6}	1.26 (1.14–1.39)	9.8×10^{-6}	0	0.50
rs6901299	6	123,817,025	G/A	0.85	1.05 (0.86–1.27)	0.66	1.18 (0.95–1.46)	0.12	1.26 (1.15–1.39)	2.0×10^{-6}	1.21 (1.11–1.31)	9.7×10^{-6}	35	0.21
rs6990255	8	34,246,490	T/C	0.95	1.41 (1.00–2.00)	0.048	1.28 (0.91–1.81)	0.16	1.33‡ (1.16–1.51)	5.7×10^{-5}	1.33 (1.18–1.51)	5.8×0^{-6}	0	0.92
rs2905072	9	132,874,589	A/G	0.77	1.43 (1.20–1.70)	5.8×10^{-5}	1.14 (0.95–1.38)	0.16	1.16 (1.05–1.28)	0.0043	1.21 (1.11–1.32)	6.4×10^{-6}	56	0.10
rs2242663	11	66,091,884	T/C	0.25	1.29 (1.09–1.53)	0.0028	1.32 (1.10–1.59)	0.0024	1.15 (1.05–1.25)	0.0015	1.20 (1.11–1.29)	1.3×10^{-6}	33	0.23
rs6494849	15	68,267,668	A/C	0.12	1.24 (1.00–1.54)	0.049	1.11 (0.88–1.38)	0.38	1.26 (1.14–1.40)	1.0×10^{-5}	1.23 (1.13–1.35)	6.5×10^{-6}	0	0.57
rs1035050	17	44,919,011	T/C	0.40	1.16 (1.01–1.34)	0.038	1.12 (0.96–1.31)	0.13	1.19 (1.09–1.29)	6.9×10^{-5}	1.17 (1.09–1.25)	9.0×10^{-6}	0	0.84
rs7250872	19	1,762,603	T/C	0.69	1.18 (1.01–1.37)	0.035	1.33 (1.14–1.56)	0.00043	1.18 (1.06–1.30)	0.0024	1.21 (1.12–1.31)	1.7×10^{-6}	0	0.41

GSK reduced sample: Excluding 261 BP cases also present in WTCCC sample.

*NBCI Build 35-bp position.

†Weighted average of control risk allele frequency for NIMH/Pritzker, GSK (reduced sample), and WTCCC.

‡Designated as strong or moderate association in WTCCC (2007) supplementary table 7A and/or 9.

Table S3. Genes in regions with SNPs with $P < 10^{-6}$ in the three study meta-analysis

Chr	Starting position basepair*	Ending position Basepair*	Gene symbol	Gene name
3	52,084,309	52,163,460	*WDR51A*	WD repeat domain 51A
3	52,207,155	52,223,383	*ALAS1*	aminolevulinate, delta-, synthase 1
3	52,230,137	52,235,219	*TLR9*	toll-like receptor 9
3	52,237,666	52,248,223	*TWF2*	twinfilin, actin-binding protein, homolog 2 (Drosophila)
3	52,255,264	52,259,655	*PPM1M*	protein phosphatase 1M (PP2C domain containing)
3	52,263,477	52,287,699	*WDR82*	WD repeat domain 82
3	52,296,911	52,302,532	*GLYCTK*	glycerate kinase
3	52,325,374	52,409,552	*DNAH1*	dynein, axonemal, heavy chain 1
3	52,410,066	52,419,049	*BAP1*	BRCA1 associated protein-1 (ubiquitin carboxy-terminal hydrolase)
3	52,419,566	52,432,696	*PHF7*	PHD finger protein 7
3	52,442,307	52,454,083	*SEMA3G*	sema domain, immunoglobulin domain (Ig), short basic domain, secreted, (semaphorin) 3G
3	52,460,147	52,463,097	*TNNC1*	troponin C type 1 (slow)
3	52,464,563	52,502,128	*NISCH*	nischarin
3	52,504,395	52,533,549	*STAB1*	stabilin 1
3	52,533,424	52,544,133	*NT5DC2*	5′-nucleotidase domain containing 2
3	52,545,660	52,549,626	*LOC440957*	similar to CG32736-PA
3	52,554,407	52,694,906	*PBRM1*	polybromo 1
3	52,694,975	52,703,548	*GNL3*	guanine nucleotide binding protein-like 3 (nucleolar)
3	52,703,544	52,715,088	*GLT8D1*	glycosyltransferase 8 domain containing 1
3	52,714,896	52,717,237	*SPCS1*	signal peptidase complex subunit 1 homolog (S. cerevisiae)
3	52,719,840	52,779,991	*NEK4*	NIMA (never in mitosis gene a)-related kinase 4
3	52,786,647	52,801,117	*ITIH1*	inter-alpha (globulin) inhibitor H1
3	52,803,823	52,818,065	*ITIH3*	inter-alpha (globulin) inhibitor H3
3	52,822,046	52,839,734	*ITIH4*	inter-alpha (globulin) inhibitor H4 (plasma Kallikrein-sensitive glycoprotein)
3	52,842,176	52,844,260	*MUSTN1*	musculoskeletal, embryonic nuclear protein 1
3	52,848,937	52,906,587	*TMEM110*	transmembrane protein 110
3	52,913,667	53,055,110	*SFMBT1*	Scm-like with four mbt domains 1
3	53,099,850	53,139,503	*RFT1*	RFT1 homolog (S. cerevisiae)
3	53,170,262	53,201,771	*PRKCD*	protein kinase C, delta
5	93,880,673	93,980,065	*C5orf36*	chromosome 5 open reading frame 36
5	93,980,146	94,057,329	*ANKRD32*	ankyrin repeat domain 32
5	94,068,956	94,646,035	*MCTP1*	multiple C2 domains, transmembrane 1
5	94,752,803	94,811,900	*FAM81B*	family with sequence similarity 81, member B
5	94,825,879	94,916,438	*TTC37*	tetratricopeptide repeat domain 37

*NBCI Build 35-bp position.

Table S4A. NIMH/Pritzker imputation quality and bipolar association analysis results with and without principal components (PCs)

			PCs		No PCs	
SNP	Chr	Imputation Quality r^2	OR (95% CI)	*P*	OR (95% CI)	*P*
rs472913	1	0.95	1.12 (0.97–1.29)	0.11	1.13 (0.99–1.30)	0.079
rs1042779	3	Genotyped	1.20 (1.04–1.38)	0.015	1.22 (1.06–1.41)	0.0067
rs17418283	5	0.89	1.09 (0.93–1.28)	0.31	1.08 (0.92–1.27)	0.32

Table S4B. GSK (reduced sample) imputation quality and bipolar association analysis results with and without principal components (PCs)

			PCs		No PCs	
SNP	Chr	Imputation quality r^2	OR (95% CI)	*P*	OR (95% CI)	*P*
rs472913	1	0.97	1.17 (1.00–1.36)	0.051	1.19 (1.02–1.38)	0.027
rs1042779	3	Genotyped	1.31 (1.11–1.54)	0.0012	1.35 (1.15–1.58)	0.00026
rs17418283	5	0.98	1.19 (1.01–1.40)	0.038	1.18 (1.00–1.40)	0.044

Reduced sample: Excluding 261 BP cases also present in WTCCC sample.

Table S4C. WTCCC imputation quality and association analysis results with and without principal components (PCs)

			BP vs extended reference set				BP vs NBS		NBS vs non-BP cases	
			PCs		No PCs		PCs		PCs	
SNP	Chr	Imputation quality r^2	OR (95% CI)	*P*	OR (95% CI)	*P*	OR (95% CI)	*P*	OR (95% CI)	*P*
rs472913	1	0.95	1.20* (1.11–1.28)	6.3×10^{-7}	1.20 (1.12–1.29)	3.4×10^{-7}	1.17 (1.06–1.29)	0.0020	1.01 (0.94–1.10)	0.75
rs1042779	3	0.97	1.16 (1.07–1.25)	0.00012	1.16 (1.08–1.25)	6.0×10^{-5}	1.17 (1.06–1.30)	0.0024	1.02 (0.94–1.11)	0.56
rs17418283	5	0.84	1.25 (1.15–1.36)	9.7×10^{-8}	1.24 (1.15–1.35)	1.8×10^{-7}	1.11 (0.99–1.25)	0.076	1.14 (1.04–1.25)	0.0067

*Designated as strong or moderate association in WTCCC (2007) Supplementary Table 7A and 9.

Table S5. Three study meta-analysis with and without adjustment for principal components (PCs)

		PCs		No PCs	
SNP	Chr	OR (95% CI)	*P*	OR (95% CI)	*P*
rs472913	1	1.18 (1.11–1.25)	2.0×10^{-7}	1.18 (1.11–1.26)	7.4×10^{-8}
rs1042779	3	1.19 (1.11–1.27)	1.8×10^{-7}	1.20 (1.13–1.28)	2.7×10^{-8}
rs17418283	5	1.21 (1.13–1.30)	1.3×10^{-7}	1.20 (1.12–1.29)	2.3×10^{-7}