Supplementary Information Appendix

Rare variants in neuronal excitability genes influence risk for bipolar disorder

Seth A. Ament[1], Szabolcs Szelinger[2], Gustavo Glusman[1], Justin Ashworth[1], Liping Hou[3], Nirmala Akula[3], Tanya Shekhtman[4], Judith A. Badner[5], Mary E. Brunkow[1], Denise E. Mauldin[1], Anna-Barbara Stittrich[1], Katherine Rouleau[1], Sevilla Detera-Wadleigh[3], John I. Nurnberger[6,7], Howard J. Edenberg[7,8], Elliot S. Gershon[5], Nicholas J. Schork[9], The Bipolar Genome Study[a], Nathan D. Price[1], Richard Gelinas[1], Leroy Hood[1,b], David W. Craig[2], Francis J. McMahon[3], John R. Kelsoe[4], and Jared C. Roach[1]

[1]Institute for Systems Biology, 401 Terry Avenue North, Seattle, Washington 98109 USA; [2]Neurogenomics Division, The Translational Genomics Research Institute, 445 North Fifth Street, Phoenix, Arizona 85004 USA; [3]Mood and Anxiety Disorders Section, Human Genetics Branch, National Institute of Mental Health Intramural Research Program, National Institutes of Health, US Department of Health and Human Services, 35 Convent Drive, Bethesda, Maryland 20892 USA; [4]Department of Psychiatry, University of California at San Diego, 9500 Gilman Drive, La Jolla, California 92093 USA; [5]Department of Psychiatry, University of Chicago, A27 South Maryland Avenue, Chicago, Illinois 60637 USA; [6]Department of Psychiatry, Indiana University School of Medicine, 355 West 16th Street, Indianapolis, Indiana 46202 USA; [7]Department of Medical and Molecular Genetics, Indiana University School of Medicine, 975 West Walnut Street, Indianapolis, Indiana 46202 USA; [8]Department of Biochemistry and Molecular Biology, Indiana University School of Medicine, 635 Barnhill Drive, Indianapolis, Indiana 46202 USA; and the [9]J. Craig Venter Institute, 4120 Capricorn Lane, La Jolla, California 92037 USA.

Table of Contents

I. Members of the Bipolar Genome Study (BiGS)

University of California, San Diego: John R. Kelsoe, Tiffany A. Greenwood, Caroline M. Nievergelt, Paul D. Shilling, Tatyana Shekhtman; Institute for Systems Biology: Jared C. Roach, Seth A. Ament, Leroy Hood; Indiana University: John I. Nurnberger, Jr., Howard J. Edenberg, Tatiana Foroud, Daniel L. Koller; University of Chicago: Elliot S. Gershon, Judith A. Badner; University of Illinois: Chunyu Liu; Rush University Medical Center: William A. Scheftner; Howard University: William B. Lawson; University of Iowa: William Coryell, James B. Potash; Washington University: John Rice; University of California, San Francisco: William Byerley; National Institute Mental Health: Francis J. McMahon, Liping Hou; University of Pennsylvania: Wade H. Berrettini; Johns Hopkins University: Peter P. Zandi; University of Michigan: Melvin G. McInnis; The Translational Genomics Research Institute: David W. Craig, Szabolcs Szelinger; Georg-August-University Göttingen: Thomas G. Schulze; J. Craig Venter Institute: Nicholas J. Schork, Danjuma Quarless

II. Principal Investigators and Co-Investigators of the National Institute of Mental Health (NIMH) Bipolar Disorder Genetics Initiative and of the NIMH Control Samples Initiative

Data and biomaterials were collected as part of eleven projects (Study 40) that participated in the National Institute of Mental Health (NIMH) Bipolar Disorder Genetics Initiative. From 2003-2007, 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 Johnson and Johnson), Debra A. Glitz, M.D. (at Wayne State University), Eric T. Meyer, Ph.D., M.S. (at Oxford University, UK), Carrie Smiley, R.N., Tatiana Foroud, Ph.D., Leah Flury, M.S., Danielle M. Dick, Ph.D (at Virginia Commonwealth University), 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. K02 DA21237; Johns Hopkins University, Baltimore, M.D., 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 San Francisco, CA, R01 MH60068, William Byerley, M.D., and Sophia Vinogradov, 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 Laurie Bederow, M.A.; NIMH Intramural Research Program, Bethesda, MD, 1Z01MH002810-01, Francis J. McMahon, M.D., Layla Kassem, Psy.D., Sevilla Detera-Wadleigh, Ph.D, Lisa Austin, Ph.D, Dennis L. Murphy, M.D.; Howard University, William B. Lawson, M.D., Ph.D., Evarista Nwulia, M.D., and Maria Hipolito, M.D. This work was supported by the NIH grants P50CA89392 from the National Cancer Institute and 5K02DA021237 from the National Institute of Drug Abuse.

Biomaterials and phenotypic data were obtained from the following projects that participated in the NIMH Control Samples: Control subjects from the National Institute of Mental Health Schizophrenia Genetics Initiative (NIMH-GI), data and biomaterials are being collected by the "Molecular Genetics of Schizophrenia II" (MGS-2) collaboration. The investigators and coinvestigators are: ENH/Northwestern University, Evanston, IL, MH059571, Pablo V. Gejman, M.D. (Collaboration Coordinator; PI), Alan R. Sanders, M.D.; Emory University School of Medicine, Atlanta, GA,MH59587, Farooq Amin, M.D. (PI); Louisiana State University Health Sciences Center; New Orleans, Louisiana, MH067257, Nancy Buccola, APRN, B.C., M.S.N. (PI); University of California-Irvine, Irvine, CA,MH60870, William Byerley, M.D. (PI); Washington University, St. Louis, MO, U01, MH060879, C. Robert Cloninger, M.D. (PI); University of Iowa, Iowa, IA,MH59566, Raymond Crowe, M.D. (PI), Donald Black, M.D.; University of Colorado, Denver, CO, MH059565, Robert Freedman, M.D. (PI); University of Pennsylvania, Philadelphia, PA, MH061675, Douglas Levinson M.D. (PI); University of Queensland, Queensland, Australia, MH059588, Bryan Mowry, M.D. (PI); Mt. Sinai School of Medicine, New York, NY, MH59586, Jeremy Silverman, Ph.D. (PI). The samples were collected by Vishwajit Nimgaonkar's group at the University of Pittsburgh, as part of a multi-institutional collaborative research project with Jordan Smoller, M.D., D.Sc., and Pamela Sklar, M.D., Ph.D., Massachusetts General Hospital (grant MH63420). Data and biomaterials used in Study 23 were collected by the University of Pittsburgh and funded by an NIMH grant (Genetic Susceptibility in Schizophrenia, MH56242) to Vishwajit Nimgaonkar, M.D., Ph.D. Additional Principal Investigators on this grant include Smita Deshpande, M.D., Dr. Ram Moanohar Lohia Hospital, New Delhi, India; and Michael Owen, M.D., Ph.D., University of Wales College of Medicine, Cardiff, UK.

III. Supplementary Methods

1. Whole-genome sequencing of bipolar disorder pedigrees. We sequenced the whole genomes of 200 individuals from 41 pedigrees multiply affected with bipolar disorder (BD). This family-based sequencing strategy is designed to detect effects of transmitted rare variants with moderate to large effect sizes within a pedigree. We hypothesized that transmitted rare variants play an important role in the genetic architecture of BD for the following reasons. First, BD is a strongly familial illness, with 8-fold relative risk in the family members of BD probands. Second, BD has relatively minor effects on evolutionary fitness (1). Third, so far, associations with *de novo* mutations appear relatively weak, compared with schizophrenia and autism (2–4). Fourth, linkage studies have suggested the existence of large-effect loci in some BD pedigrees, although causal variants at these loci have been elusive.

Family-based sequencing also provides more general advantages over sequencing of unrelated cases and controls. Repeated observation of a rare variant across multiple family members provides confidence that it is not a sequencing error. Also, under many disease models we expect stronger or more numerous risk variants in familial cases than in sporadic cases. For instance, if a multiply affected pedigree has a monogenic cause, then the causal variant in that pedigree is predicted to have higher penetrance than risk variants in the general population. If a multiply affected pedigree has a polygenic cause, the affected family members are predicted to inherit an especially large number of risk variants, though each affected family member could inherit a unique combination of risk variants. For all of these reasons, we chose to sequence multiply affected pedigrees as a strategy well suited to find genetic causes of bipolar disorder.

The 41 bipolar disorder pedigrees sequenced in this study were drawn from a set of 972 multiply affected pedigrees collected by the NIMH Genetics Initiative and by sites at the University of California, San Diego, the University of California, San Francisco, and the University of Chicago. This sample has been described previously (5). DNA derived from whole blood or from lymphoblastoid cell lines was obtained from the Rutgers University Cell and DNA Repository and from the Corriell Institute.

We selected pedigrees and individuals for sequencing by considering family structure, per-pedigree LOD scores at 4,500 genome-wide SNPs (5), and the polygenic risk score in each pedigree's proband (6). We chose a subset of individuals within each pedigree so as to maximize power under the most likely inheritance mode, as follows: (i) We sequenced a single affected individual and a single unaffected individual from nine pedigrees with a suggestive per-pedigree linkage peak and predicted dominant inheritance (pedigrees #1-9). (ii) We sequenced a parent-child trio or quartet from eight pedigrees with a single, suggestive per-pedigree linkage peak and either dominant or recessive inheritance (pedigrees #10-17). In 22 pedigrees for which polygenic inheritance was more likely, we sequenced either (iii) a subset of affected individuals and 0-1 unaffected individuals (pedigrees #18-33) or (iv) all available individuals (pedigrees #34-41). Pedigree drawings annotated with the sequenced individuals and segregation criteria used in genetic filtering are shown in Figs. S1-S4 and Table S1.

Whole-genome sequencing was performed to >40x coverage by Complete Genomics, Inc. (Mountain View, CA). We used variant calls for single-nucleotide variants (SNVs) and indels from the Complete Genomics analysis pipeline version 2.0 or 2.2, relative to the human reference genome version GRCh37.

2. Whole-genome sequencing of population controls. We analyzed the 200 genomes from BD pedigrees together with genomes of population controls drawn from a collection of >1,200 genomes in an in-house collection at the Institute for Systems Biology, which were originally ascertained on a variety of diseases. The control pedigrees include a large pedigree segregating a monogenic form of cardiomyopathy (7), 10 pedigrees with Adams-Oliver syndrome (8, 9), and three pedigrees with Fanconi anemia, as well as genome sequences from several ongoing studies. Genome sequences from studies of complex psychiatric and neurological diseases such as epilepsy (10) were excluded from association testing, but are included in some quality control procedures. Individuals in the control pedigrees used for association testing did not undergo diagnostic interviews for psychiatric conditions and are likely to have a rate of BD comparable to that in the broader population (~1-2%). Whole-genome sequencing for all of these individuals was performed by Complete Genomics.

Subsets of control genomes were utilized in several ways throughout our analyses. (i) Allele frequencies and quality-control metrics were calculated across 1,057 non-BD genomes and used as a complement to publicly-available databases for the identification of rare variants (Supplementary Methods Section 4). (ii) For linear mixed model association analysis of single variants, we analyzed the BD genomes together with 254 genomes from individuals and pedigrees of European ancestry. (iii) For gene and pathway burden analyses we used control genomes to calculate an empirical distribution for the rate at which variants in each gene or pathway segregated in non-BD pedigrees. In our analyses of uncommon coding and regulatory variants, we utilized 168 genomes from 34 multigenerational pedigrees, matched to the BD pedigrees by size and structure. (iv) In our analysis of rare, gene-disrupting variants we additionally considered the empirical distribution of transmitted gene-disrupting variants across 242 control parent-offspring trios.

3. Affection status models. For analysis of segregation patterns in single pedigrees, we designated affection status

models by combining phenotypic information with evidence from a published linkage analysis utilizing these same pedigrees (5). We considered all individuals with a diagnosis of bipolar disorder type 1, bipolar disorder type 2, or schizoaffective disorder bipolar type to be affected. Individuals with recurrent or single episodes of major depression were also considered to be affected if (i) the best-supported linkage peak for that pedigree in the analysis by Badner et al. (5) supported an inheritance model that included major depression, and they were not married-in; or (ii) they had offspring with BD and were not married-in to a pedigree. Two individuals with unknown or other phenotypes were considered obligate carriers and treated as affected because pedigree structure and linkage data provided strong evidence that they were carriers of risk alleles. In pedigrees 5407 and 11127 we developed two distinct affection status models, treating individuals with major depression as affected or unaffected, respectively; in these pedigrees, we used the union of candidate variants under the two affection status models for downstream analyses. The designated affection status models for each pedigree are illustrated in Figs. S1-S4.

4. Variant annotation and filtering. SNVs and indels from Complete Genomics masterVar files were annotated on sequence quality, allele frequency, functional annotation, likelihood of false positivity, and genetic segregation. These filters were implemented using a combination of QIAGEN's Ingenuity Variant Analysis software (www.qiagen.com/ingenuity, QIAGEN Redwood City, Redwood City, CA), the Family Genomics Workflow, and PLINK (11), as follows:

4.a. Variant quality. Using Ingenuity, we retained variants with Variable Allele Frequency quality scores >35 in affected individuals or >20 in unaffected individuals. Using the Family Genomics Toolkit, we then filtered out variants at positions with >20% no-call rate across the 200 genomes from BD pedigrees.

4.b. Allele frequency. We annotated allele frequencies using data from the 1,000 Genomes Project, NHLBI Exomes, Complete Genomics Public Genomes, Kaviar (12), and ISB's internal database of 1,057 genome sequences from non-BD pedigrees. We defined "uncommon" variants as those with allele frequencies < 5% in all populations and "rare" variants as those with allele frequencies < 1% in all populations.

4.c. Functional annotation. Using Ingenuity, we categorized gene-disrupting variants, coding variants, and regulatory variants. Gene-disrupting variants were defined by the following categories: (i) stop codon change, (ii) gene fusions, (iii) frameshift indel, (iv) disrupt splice site up to two bases into intron, or (v) a structural variant. Coding variants included gene-disrupting variants, and the following additional categories that alter the amino acid sequence of a protein: (vi) established gain of function in the literature (vii) inferred activating mutations by Ingenuity, (viii) predicted gain of function by BSIFT, (ix) in-frame indel, (x) missense (i.e., all other non-synonymous coding SNVs). Regulatory variants were defined as those (i) in a microRNA binding site predicted by TargetScan, (ii) deleterious to a microRNA, (iii) in a JASPAR cis-motif or ENCODE ChIP-seq transcription factor binding site within 1000 bp upstream of a transcription start site, (iv) within a known or predicted enhancer binding site from the VISTA database, or (v) in a 5' or 3' untranslated region.

4.d. Eliminating likely false positives. We applied four additional filters to remove likely false positives and non-deleterious variants. Coding variants were filtered to exclude 100 bp segments with an unusually high rate of functional variants in Complete Genomics genomes. Gene-disrupting SNVs and indels were further filtered to exclude genes with the highest 1% rate of transmitted gene-disrupting SNVs; genes that contained a segregating, rare gene-disrupting SNV or indel in any of the 34 multigenerational control pedigrees; and variants with CADD (13) scores < 10.

4.e. Genetic segregation. Fully and nearly-fully segregating variants were defined in each pedigree based on allele sharing among the sequenced affected and unaffected individuals. Thresholds for allele-sharing are shown in Table S1. These thresholds were determined based on the number of sequenced individuals and the number of meioses separating the sequenced affected individuals.

4.f. Filter settings for each analysis. We combined annotations in 4.a-4.e to define lists of candidate variants for downstream statistical analysis, as follows. For mixed-model analysis of single variants we used high-quality, uncommon, segregating and non-segregating variants. For gene and pathway burden tests of polygenic variants we used high-quality, rare and uncommon, gene-disrupting, coding, and regulatory variants with low false-positivity scores, which segregated with BD in each pedigree. For pathway analysis of rare, gene-disrupting variants, we used high-quality, rare, gene-disrupting variants with low false-positivity scores, which segregated with BD in each pedigree.

5. Selection of genes and gene sets for analysis. The genetic complexity of BD, combined with the massive number of variants discovered in a WGS study, precluded an unbiased genome-wide analysis. To increase statistical power, we selected for analysis 3,087 candidate genes and 325 candidate gene sets ("pathways"). These genes and pathways were selected based on two hypotheses about BD's genetic causes.

First, we hypothesized that genes influencing risk for BD will be enriched for those with intrinsic neuronal functions. Imaging studies, animal models, and other evidence suggest that mood disorders arise from structural and physiological changes in the brain. It is therefore reasonable to confine an initial search for rare risk variants to genes with predicted functions in neurons. Of course, peripheral tissues and non-neuronal cells in the brain (e.g., glia) could contribute to mood disorders. Also, genes with both neuronal and non-neuronal functions (e.g., metabolic enzymes) may contribute to disease. We anticipate that an unbiased genome-wide search will become possible as sample sizes increase. Since the specific neuronal functions perturbed in mood disorders have yet to be ascertained, we aimed to include a broad spectrum of genes with neuronal functions. Synaptic function appears to be especially relevant, since synapses are highly enriched for proteins that contribute to neuronal activity. We therefore included all the genes that encode proteins that have

been localized to synapses through proteomics studies from the SynaptomeDB database. A second category of neuronal genes of special interest are those with characterized small molecule agonists or antagonists. Genetic findings for these genes may be especially actionable, since they are attractive therapeutic targets. We therefore added to our list any ion channel or G-protein coupled receptor (GPCR) included in the International Union of Pharmacology (IUPHAR) database of drug targets.

Second, we hypothesize that loci with evidence of an association to BD from GWAS are likely to harbor causal rare variants. Although only a handful of genome-wide significant loci have been discovered in BD GWAS, scores of additional loci have sub-threshold associations. These "suggestive" loci are enriched for voltage-gated calcium channels and other pathways (14–16). Also, polygenic risk scores based on the combined genotypes at all nominally significant loci classify cases from controls in independent cohorts (14). These results suggest that some of the genes at loci with sub-threshold associations from GWAS harbor true risk variants. Some of the common SNPs identified by GWAS may tag true causal variants that are more rare. Genes that contain common risk variants may also contain additional rare variants that are not linked to common SNPs. We therefore used data from two recent publications by the Psychiatric Genomics Consortium on Bipolar Disorder to characterize genes with prior evidence from GWAS.

Our final list of candidate genes included:

- 1,887 genes encoding proteins localized to the pre- and post-synaptic densities from the SynaptomeDB database.(17)
- 280 genes encoding ion channels from the IUPHAR database.(18)
- 411 genes encoding G-protein coupled receptors from the IUPHAR database.(18)
- 461 genes located within 100 kb of loci with suggestive associations to BD ($p < 1e\text{-}4$) in the mega-analysis by the Psychiatric Genomics Consortium (PGC-BD1).(16)
- 226 genes with empirical p-values < 0.05 in a meta-analysis of bipolar disorder GWAS by Nurnberger et al.(15)

These 3,087 candidate genes were used in single-variant association tests and gene burden tests.

In addition to variant- and gene-level tests, we conducted pathway-level association and enrichment tests, focusing on gene sets enriched among our 3,087 candidate genes. We downloaded the Gene Ontology, KEGG, and BioCarta pathways from MSigDB. We used Fisher's exact tests to identify pathways that were statistically over-represented among our 3,087 candidate genes (FDR < 0.05). Since the power of pathway-level tests depends on the number of genes in the pathway, we then filtered gene sets by size. For pathway burden association tests (Supplementary Methods Section 7.2), we considered the 269 pathways that contained between 5 and 100 genes. For enrichment analyses of gene-disrupting variants we considered the 325 pathways that contained between 5 and 200 genes (Supplementary Methods Section 7.3). Different size filters reflect the differing power of these tests to detect significance across small vs. large gene sets, respectively.

6. Mixed model association test of uncommon SNVs. We used EMMAX (19) to test for associations between BD and uncommon (1-5% MAF) coding and non-coding SNVs. We integrated the genotypes of the 200 genomes from BD pedigrees with 254 genomes from non-BD pedigrees, also of European descent. For this analysis, we considered individuals with a diagnosis of bipolar disorder type I, bipolar disorder type 2, or schizoaffective disorder to be affected; individuals in BD pedigrees who had no mental illness and all individuals in non-BD pedigrees were considered to be unaffected; individuals with major depression or any other DSM-V diagnosis, and individuals in BD pedigrees for whom a phenotype was not ascertained were coded as unknown. In total, there were 108 affected individuals, 309 unaffected individuals, and 35 unknown. 5,730 uncommon coding and non-coding SNVs were selected for analysis as described in Section 4, above. We used emmax-kin to construct a Balding-Nichols kinship matrix, based on these 5,730 genome-wide SNVs. We then performed association testing with emmax. We evaluated the success of this approach in accounting for pedigree and population structure with quantile-quantile plots and lambda statistics, implemented with the qqman and GenABEL (20) R packages, respectively.

7. Gene and pathway burden tests.

7.1. Overview of the method. Association tests for the effects of individual rare variants (MAF < 1%) have low power because each variant is observed in only a few individuals. Therefore, tests that aggregate the effects of multiple rare variants are commonly used in rare-variant association studies. Typically, these rare-variant tests aggregate the variants in a region into a statistic for association testing, either by collapsing them into a meta-variable ("burden" tests) or by combining them in a variance-component model (e.g., SKAT (21)). At least two challenges have arisen in applying these rare-variant association tests to real data. First, they have low power when only a small fraction of the variants in a region contribute to disease risk. Second, differences in gene size, mutation rate, level of positive and negative selection, sequence depth and rate of sequencing errors at each locus, and other factors lead to biases in the number of transmitted variants observed in each gene or gene set. These biases can lead to uneven power across the genome and to spurious results.

We developed a strategy to test for the effects of rare variants in our family-based sample that is designed to overcome both of the previously mentioned challenges. To enrich for risk-associated variants in each gene, we used segregation patterns within pedigrees in combination with other forms of functional annotation prior to association testing across affected and unaffected pedigrees. This approach may have better power than other methods when only a subset of variants at a locus contribute to risk.

We used an empirical null distribution from control pedigrees to assign p-values that correct for differences in the background rate of transmitted variants. Other groups have used theoretical distributions that account for some of the factors contributing to bias (especially gene length). Empirical distributions are appealing when sufficient

numbers of comparable control genomes are available, since they correct for both known and unknown biases.

We developed three implementations of our approach for testing the effects of (i) single genes, (ii) oligogenic combinations of uncommon variants within a pathway, and (iii) heterogeneous effects of very rare variants within a pathway.

7.2. Gene burden test. The first step of all three burden tests was to identify rare, functional variants that co-segregated with disease in each BD pedigree. Co-segregation was defined by allele sharing among affected and unaffected individuals in a pedigree, according to the affection status models shown in Figures S1-S4 and described in Supplementary Methods Section 3. We allowed slight deviations from perfect co-segregation to account for reduced penetrance and phenocopies. Thresholds for allele sharing in each pedigree are shown in Table S1. Rare and uncommon variants (MAF < 5%) with functional annotations were characterized as described in Section 4. The selection of 3,087 genes for analysis is described in Section 5. We identified a total of 7,783 variants that passed all filtering criteria across all 41 BD pedigrees. These variants are predicted to impact the functions of 2,375 genes, with a median of 146 genes impacted per pedigree.

In order to define an empirical null distribution for the gene burden test, we identified non-bipolar (control) pedigrees from ISB's in-house database that had similar size and structure to the BD pedigrees and which were also of European ancestry. We identified 34 suitable pedigrees in our collection. In each control pedigree, we arbitrarily assigned a subset of individuals as "affected", such that the transmission rate from parents to offspring and the relative risk among siblings matched the values in the sequenced BD pedigrees. We then identified variants that were shared among the "affected" individuals and not the "unaffected" individuals in each control pedigree. A total of 6,686 variants passed all thresholds in control pedigrees, and these variants are predicted to impact 2,283 genes. The distribution of candidate gene counts per pedigree was similar between the BD and control pedigrees (Kolmogorov-Smirnov test: $p = 0.35$). The similarity in the number of genes identified in BD vs. control pedigrees suggests that the structure of the BD vs. control pedigrees was sufficiently similar that the variants in these control pedigrees represent an appropriate null distribution for testing gene enrichments. For each gene, we counted the number of BD pedigrees and control pedigrees in which we observed a variant that co-segregated with affection status. We calculated a p-value for each gene using Fisher's exact test. A quantile-quantile plot indicates that the p-values from this approach are uniformly distributed across all the genes tested (Fig. S11A).

7.3. Pathway burden for uncommon and rare functional variants. Disease risk in an individual may arise from the additive and epistatic effects of rare or uncommon variants across several genes in a protein complex or pathway. Likewise, variants in distinct genes from the same complex or pathway could occur in different pedigrees and have similar effects on disease risk. In these scenarios, the genetic signal from an individual gene may be too weak or too rare to be detected with single-gene burden tests. We developed two approaches to test for an increased burden of variants across multiple genes in a pathway.

In the first pathway burden test, we hypothesize that risk for BD increases with the number of mutated genes in a pathway. To test this hypothesis, we used a score test based on the normalized count of genes per pathway that contained segregating variants in BD vs. control pedigrees. For each of the 269 pathways in our analysis (see Section 5), we counted the number of pathway genes ($N_{pathway}$) in the variant list from each of the 41 BD pedigrees and from each of the 34 control pedigrees. We then calculated a Fractional Pathway Score in each pedigree as $N_{pathway} / N_{total}$, where N_{total} is the total number of genes with a candidate variant in that pedigree. Dividing by N_{total} normalizes for differences in the number of candidate genes arising from segregation analysis in each pedigree (which is itself a function of pedigree structure). We used Wilcoxon's rank-sum test to evaluate the hypothesis that the Fractional Pathway Score for each gene set was larger in BD vs. control pedigrees. A low p-value for a given pathway suggests that segregating variants in BD pedigrees are over-represented for genes within that pathway. Quantile-quantile plots suggest that the distribution from this test is uniform across most pathways, with a single pathway (the BioCarta GABA Pathway) having a large deviation from the null distiribution (Fig. S11B).

7.4. Pathway burden for rare, gene-disrupting variants. A distinct hypothesis is that risk for BD increases if there is a mutation in any of the genes in a pathway. In Boolean terms, the hypothesis in 7.3 proposes an "AND" relationship between the genes in a pathway and risk for disease, whereas the hypothesis in this section proposes an "OR" relationship. We formulated this problem as an enrichment test for pathways that were over-represented for rare, segregating variants in BD pedigrees. This test assumes that each variant has an independent effect on disease risk, and it will be better powered if these effects are large. We therefore focused on rare gene-disrupting variants, which have been shown in other psychiatric diseases to be more strongly enriched than other classes of variants for those with large effects on risk (22).

For each BD pedigree, we identified all likely damaging (CADD score > 10), rare (MAF < 0.01), gene-disrupting variants that co-segregated with disease. 437 genes contained a gene-disrupting variant in one or more of the BD pedigrees. 416 of these 437 genes were disrupted in only one of the 41 BD pedigrees, and 21 genes were disrupted in two pedigrees. This genetic evidence is not strong enough to make inferences to individual genes. We calculated p-values for each pathway by comparing the rate of transmitted gene-disrupting variants in that pathway in BD pedigrees to the empirical null distribution in control pedigrees. Because gene-disrupting variants are rare even in aggregate, we found that a larger set of control genomes was necessary for this analysis than for other burden tests. We therefore calculated the rate of gene-disrupting variants in each gene from a total of 242 non-BD parent-offspring trios in our in-house collection. A small number of genes are common false positives with disrupting variants in a large fraction of trios.

True risk variants in these genes would be indistinguishable from noise. We removed from analysis the top 1% of genes with the highest rate of disrupting variants.

We analyzed 325 pathways that contained between 5 and 200 genes (Section 5). We recorded the overlap of each of these pathways with the set of genes disrupted in BD pedigrees. To test for significance, we then drew a same-sized set of variants at random from the empirical distribution in non-BD trios and recorded whether the overlap of this random set with each of the pathways was greater than or equal to the number observed in BD pedigrees. We repeated this procedure 100,000 times. The empirical p-value for each pathway is calculated as (N + 1) / (100,000 + 1), where N is the number of times out of 100,000 that the overlap to the pathway in variants drawn at random was equal to or greater than the number observed in BD pedigrees. A quantile-quantile plot for the distribution of observed vs. expected p-values across the 325 suggests that this empirical p-value is uniformly distributed except for a handful of pathways with a strong enrichment of disrupting variants in BD pedigrees (Fig. S11C).

7.5. Additional considerations. There are limitations of our approach. Sample size is effectively reduced from the number of sequenced individuals to the number of independent pedigrees. Also, if disease risk within a family is polygenic, then the fully- and nearly-fully segregating variants in a pedigree will represent only a subset of the causal variants. Although the use of control pedigrees accounts for various sources of false positives in sequencing studies, the power of this test may be different in large vs. small genes. Future implementations could explore strategies to weight variants based on the strength of the genetic evidence in each pedigree.

Despite these limitations, our statistical approach has several attractive features. First, because the linkage of individual variants to disease is detected within pedigrees, the test is robust to population stratification. Second, it combines evidence from (within-pedigree) linkage and (between-pedigree) association in an intuitive way and allows us to test the linkage of individual variants to disease within families while collapsing evidence across families into a gene burden test. In contrast to the well-controlled p-value distributions resulting from our method (Fig. S11), evaluation of other rare-variant association tests for family-structured data suggested that they often give anti-conservative results (data not shown).

8. Targeted sequencing of 26 genes in 3,014 cases and 1,717 controls.

8.1. Selection of target regions. We performed targeted resequencing of 185 pools of samples (23) to evaluate associations of BD with rare variants in the coding and non-coding regions of candidate genes (Table S10). Our initial design included 30 candidate genes. Results for 26 of these genes are presented here. Data for 4 genes were not analyzed because WGS analyses conducted after the initial design of the targeted sequencing study revealed poor sequence quality at these loci. 20 of the targeted genes were identified by WGS analysis in this study, and six were identified by GWAS. We targeted all UCSC knownGene exons, as well as the following non-coding regions with putative regulatory functions. 5' and 3' untranslated regions were targeted based on knownGene models, downloaded on December 10, 2012. The core promoter was defined as the region 1-1000 bp upstream of each transcriptional start site (TSS), also based on UCSC knownGene models. We used DNase hypersensitive regions from the ENCODE project (24) to identify putative enhancers. We downloaded the wgEncodeRegDnaseClusteredV2 track from the UCSC Genome Browser and selected hypersensitive regions with quality scores > 300. For the ten voltage-gated channel genes in our study, we targeted the DNase hypersensitive regions starting 5 kb upstream of the first transcriptional start site and spanning the entire gene body. For the remaining genes, we targeted DNase hypersensitive regions 5 kb upstream and downstream of each TSS.

8.2. Amplicon design. We used Design Studio (Illumina, San Diego, CA) to design TruSeq Custom Amplicons for each target region. PCR amplicons passing Design Studio's quality threshold were successfully designed for 90.1% of the targeted regions, spanning a total of 462 kb. Each amplicon was approximately 250 bp and avoided homology to other genomic regions, as well as common SNPs. The TruSeq Custom Amplicon system can target a maximum of 1536 amplicons per assay. We therefore divided these regions into two assays, with 1,536 and 1,525 amplicons, respectively.

8.3. Sample cohort. The data presented in this study are from sequencing of 3,014 BPI cases and 1,717 neurologically cleared controls of European-American ancestry. Samples were primarily from the GAIN (6) and TGEN (25) collections, and also included 169 BD cases from a new prospective sample of lithium responsiveness in BD. These samples were distributed across 149 of the 185 sample pools that were sequenced in parallel. The remaining 36 sample pools contained DNA from African-American BPI cases and controls. Data from the African-American sample will be published elsewhere.

8.4. Sample pooling. DNA was quantified with PicoGreen (Life Technologies, Carlsbad, CA), and equimolar quantities from each individual were combined into 185 pools. Pools contained from a minimum of 16 to a maximum of 37 samples with 152 of the 185 pools containing 32 samples. The samples in each pool were matched for phenotype, ethnicity, and cohort. There were 95 pools with BD cases of European-American (EA) ancestry, 54 pools with EA controls, 14 pools with African-American (AA) cases, and 22 pools with AA controls.

8.5. Library preparation and sequencing. DNA from each pool was aliquoted into multiple 96-well plates to enable parallel library construction from the two sets of amplicons. Each TruSeq Custom Amplicon assay is designed for barcoding of up to 96 samples. We therefore divided the 185 sample pools into two batches, with equal numbers of case/control and EA/AA pools in each batch. As technical controls, we included in each batch a built-in "Illumina QC" pool and two replicate pools of HAPMAP (26) samples. To measure variance between plate preparations, a single EA BD case pool was replicated between the plates.

We generated a separate sequencing library from each of the two batches of samples, for each of the two sets of TruSeq Custom Amplicons (a total of four libraries). We utilized 8 bp dual indexing to provide unique barcodes for each of the 96 sample pools in each batch. We performed paired-end, 2*250 bp Rapid mode sequencing of each library on a HiSeq2500 flowcell. To evaluate replicability, we performed a second round of Rapid mode sequencing of each library on an additional flowcell with paired-end 2*100 bp reads.

8.6. Read alignment, variant calling, and quality control. Sequence reads were aligned to the human reference genome (GRCh37d5) using Burrows-Wheeler Alignment (27), followed by indel realignment with the Genome Analysis Toolkit (28). TruSeq Custom Amplicon library construction incorporates adaptor sequences onto the ends of each amplicon during PCR amplification, and there is no fragmentation step. Therefore, each sequence read starts at either the beginning or end of an amplicon, and reads with an identical start position are not likely to be optical duplicates. Since some of the amplicons were shorter than 250 bp, we clipped reads at the expected length of each amplicon. These steps led to a total of 390 BAM files, one per amplicon set per pool.

We sorted and merged the aligned reads from the two sets of amplicons from each sample pool, using samtools (29). We then used SNVer (30) to call single-nucleotide variants (SNVs) separately in each pool and to estimate per-pool allele frequencies. We used default thresholds of base quality ≥17 and mapping quality ≥20. We excluded variants observed on only one strand. We combined the variant calls from all 185 pools. We retained for further analysis all 3,715 SNVs located at positions with at least 640x total read depth (approximately 20x per diploid), in every pool for which the SNV was called.

The outputs from the steps above were tables of allele frequencies in each of the 185 pools, for 3,715 high-quality SNVs. Statistical association tests require allele counts in each individual (rather than pooled allele frequencies). We constructed a genotype matrix of 3,715 variants x 4,731 individuals by assigning the non-reference alleles from each pool at random to each of the haplotypes in that pool.

8.7. Quality control and normalization. We performed several experiments and analyses to evaluate the accuracy of allele frequency estimates from pooled sequencing. As detailed below, these steps led to the general conclusion that allele frequency estimates were accurate and replicable across most of the allele frequency range. The presence of possible counting errors among very rare variants led us to perform additional normalization steps that improved performance.

In total, targeted sequencing generated 1.28 billion high-quality sequencing reads covering 320 billion base pairs (Table S7). 65-80% of base pairs in each pool reached a read depth threshold of >640x total read depth (~20x coverage per diploid individual; Fig. S7). Across all individuals, per-position read depth was unimodally distributed around a mode of 40x coverage per individual (Fig. S8). These results suggest that sequencing depth was consistent and deep across sample pools and libraries.

We next compared allele frequencies derived from targeted sequencing to the frequencies of these alleles in the Kaviar database (12). Of the 3,715 SNVs identified through targeted sequencing, 1,960 of these SNVs were present in the Kaviar database on January 29, 2014. Allele frequencies for these known SNVs were tightly correlated between targeted sequencing and public data ($r = 0.88$; Fig. S8).

To further evaluate the replicability of allele frequency estimates from targeted sequencing, we performed a second round of sequencing on each of the four sequencing libraries. In the second round of sequencing we used paired 100 bp reads, meaning that only a subset of the positions were targeted. 1,728 alleles were observed at positions with >640x coverage in both sequencing runs. Visual inspection (Fig. S9) indicated that allele frequency estimates were nearly perfectly replicable for common variants and were strongly correlated at allele frequencies as low as 10 copies / 12,002 sequenced haplotypes (MAF = 0.001). At allele frequencies < 0.001 considerably more scatter was present. Still, allele frequencies for the majority of variants at these frequencies (including singleton variants) were estimated identically in the two sequencing runs. These results indicate that results from our targeted sequencing are accurate at MAF > 0.001. Allele frequencies < 0.001 should be treated with caution but provide useful information.

A common source of noise in pooled sequencing arises from the difficulty in estimating an exact copy number for variants present in a small fraction of reads (23). For instance, in a pool of 32 individuals sequenced to 640x depth, each haplotype is represented by 10 reads on average. However, random variation during amplification and sequencing can lead to uneven representation of the haplotypes among the sequencing reads, which in turn can lead to inaccurate estimation of allele counts. These counting errors are likely to be especially important to account for when analyzing very rare variants.

We evaluated possible counting errors by comparing the distribution of estimated counts of each variant in each sample pool to simulated distributions assuming Hardy-Weinberg equilibrium. In simulated distributions, all variants with MAF < 0.001 are expected to be observed in only one haplotype per pool. 97% of variants with MAF < 0.005 were predicted to be observed in only one haplotype per pool. 76% of variants with MAF ~ 0.005-0.01 were predicted to be in only one haplotype per pool. At allele frequencies > 1%, variants are predicted to occur in multiple haplotypes in the majority of pools. Simulated distributions for each allele frequency bin are shown in Figure S11.

Comparison of these simulated distributions to the haplotype count estimates reported by SNVer suggested deviations from the simulated distribution for variants present in 0-2 copies per pool. Specifically, single-copy variants were under-represented at allele frequencies up to 0.05, and 2-copy variants were over-represented at allele frequencies up to 0.005. At allele frequencies > 0.05, the observed distribution approximated simulated values.

Since the vast majority of low-frequency variants are expected to be present in a single haplotype per pool, the most parsimonious explanation for these deviations from

Hardy-Weinberg equilibrium is that some single-haplotype variants are under-called as not present in a pool, and others are over-called as being present in two (or more) haplotypes. These counting errors represent a small fraction of overall genotype calls, but effects on individual rare variants may be significant. We therefore applied two strategies to correct for counting errors. First, we applied an expectation-maximization procedure to normalize allele counts for variants with frequencies < 0.01. Alternatively, we removed from our analysis the extremely rare variants (MAF < 0.001) whose frequency estimates are likely to be less accurate.

The expectation-maximization procedure was applied to all variants with allele frequencies in our data < 0.01. For a variant with a frequency of 0.01 (95/9472 total haplotypes in our study), 70% of pools containing that variant are expected to have only one variant haplotype. For a variant with a frequency of 0.001 (5/9472 haplotypes), >99% of pools containing that variant are expected to have only one variant haplotype. Thus, haplotype counts in a pool that are > 1 are predicted to be incorrect the majority of the time for an allele with a frequency of 0.01, and nearly always incorrect for an allele with a frequency of 0.001. As a corrective, we therefore adjusted allele haplotype counts to 1 for each pool in which the variant was observed. This procedure maximizes the expectation that rare variants will be present in one copy per pool. We applied this normalization only to variants with MAF < 0.01, since the expectation of 1 haplotype per pool does not hold for variants with MAF > 0.01.

Comparison of normalized allele counts to simulations under Hardy-Weinberg equilibrium suggests that our normalization procedure succeeded in correcting haplotype counts per pool to be nearly identical to expected (simulated) values (Fig. S11). Slight deviations from the simulated distribution are still apparent for variants with allele frequencies > 0.01, which were not normalized. However, three considerations suggest that these remaining counting errors will have very little effect on results. First, any counting errors represent only a small fraction of the calls for each common allele. Second, common alleles are given very small weights in the sequence kernel association test (SKAT) used to assess significance. Third, the C-alpha test included only variants with MAF < 0.01 and gave similar results to SKAT (Tables S8, S9).

As an alternative to the normalization procedure described above, we also calculated rare-variant associations while removing all variants with allele frequencies < 0.001. These variants, represented by fewer than 10 haplotypes in our dataset are the most sensitive to counting errors. We chose a cutoff of < 0.001 because the technical reproducibility of allele frequencies for variants above this threshold improved dramatically (Fig. S10).

8.8. Rare-variant association tests. We used SKAT (31), C-alpha (32), and the unidirectional gene burden test implemented in the SKAT R package to test for associations between BD and rare variants in each of the 26 candidate genes. We tested for associations of BD separately for coding vs. regulatory variants. We defined non-synonymous coding variants as those that change the amino-acid sequence of a knownGene gene model, as predicted by ANNOVAR (33). We defined regulatory SNVs as non-coding SNVs with an annotation in RegulomeDB (34).

For the SKAT test, we used default parameters to weight the effects of rare vs. common variants (alpha = 1, beta = 25). These parameters assign strong weights to variants with MAF < 0.01, weak weights to variants with MAF 0.01-0.05, and near-zero weights to variants with MAF > 0.05. P-values reported in Table 1 are based on 100,000 re-sampling permutations. Permutation p-values were nearly identical to the asymptotic p-values reported by SKAT (Tables S8, S9).

We performed a primary SKAT analysis and two secondary analyses. The primary analysis used the allele frequencies with expectation-maximization normalization. The second analysis used “raw” allele frequency estimates but removed variants with MAF < 0.001. The third analysis used raw allele frequency estimates, including all allele frequencies. The p-values from the first and second approaches were nearly identical. The third approach yielded anti-conservative p-values compared to the other approaches. We performed the SKAT gene burden test using normalized allele frequencies.

We performed the C-alpha test using normalized allele frequencies. Since C-alpha assigns an equal weight to each variant regardless of its frequency, we set a hard allele frequency cutoff, MAF < 0.01. We used an implementation of C-alpha in the AssotesteR R package and performed 100,000 permutations.

III. Figures S1-S12

Figure S1. Pedigree diagrams 1-9: WGS from one case and one control. Letters indicate the individuals from whom whole-genome sequences were obtained. Each of these individuals was designated affected (A) or unaffected (U) as described in Supplementary Methods Section 3.

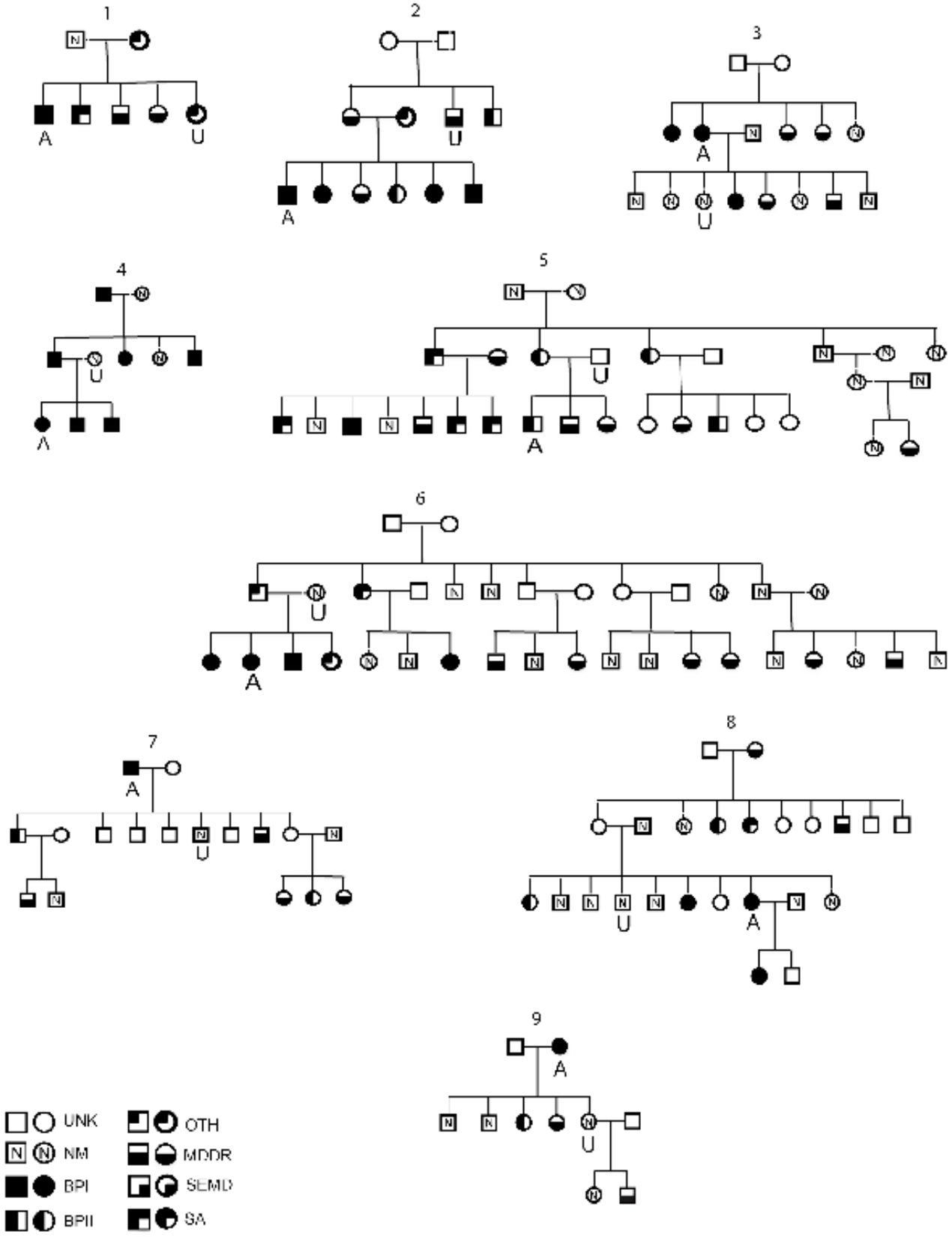


Figure S2. Pedigree diagrams 10-17: WGS from a family quartet or trio. Letters indicate the individuals from whom whole-genome sequences were obtained. Each of these individuals was designated affected (A) or unaffected (U) as described in Supplementary Methods Section 3.

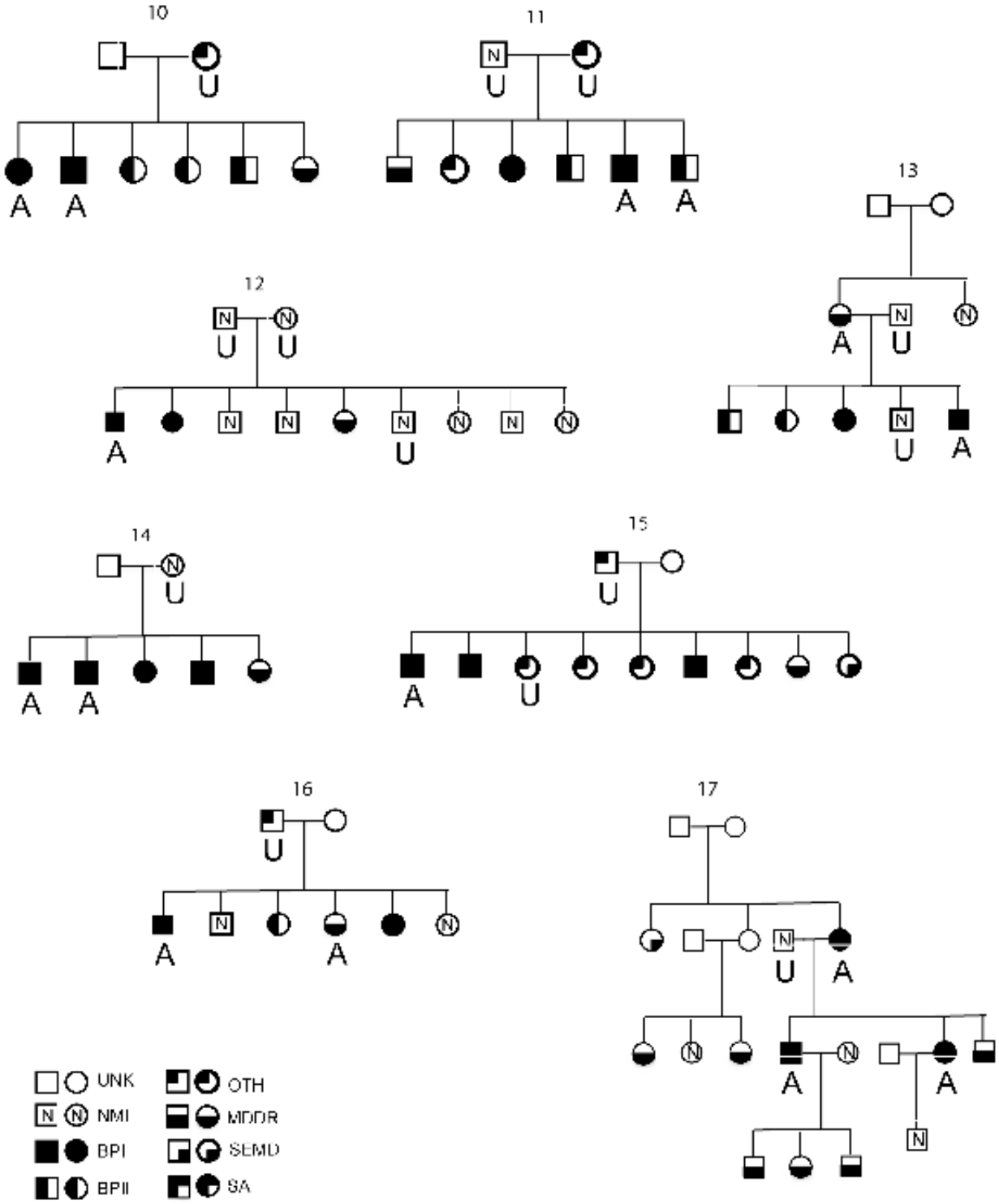


Figure S3. Pedigree diagrams 18-33: WGS from several BD cases. Letters indicate the individuals from whom whole-genome sequences were obtained. Each of these individuals was designated affected (A) or unaffected (U) as described in Supplementary Methods Section 3.

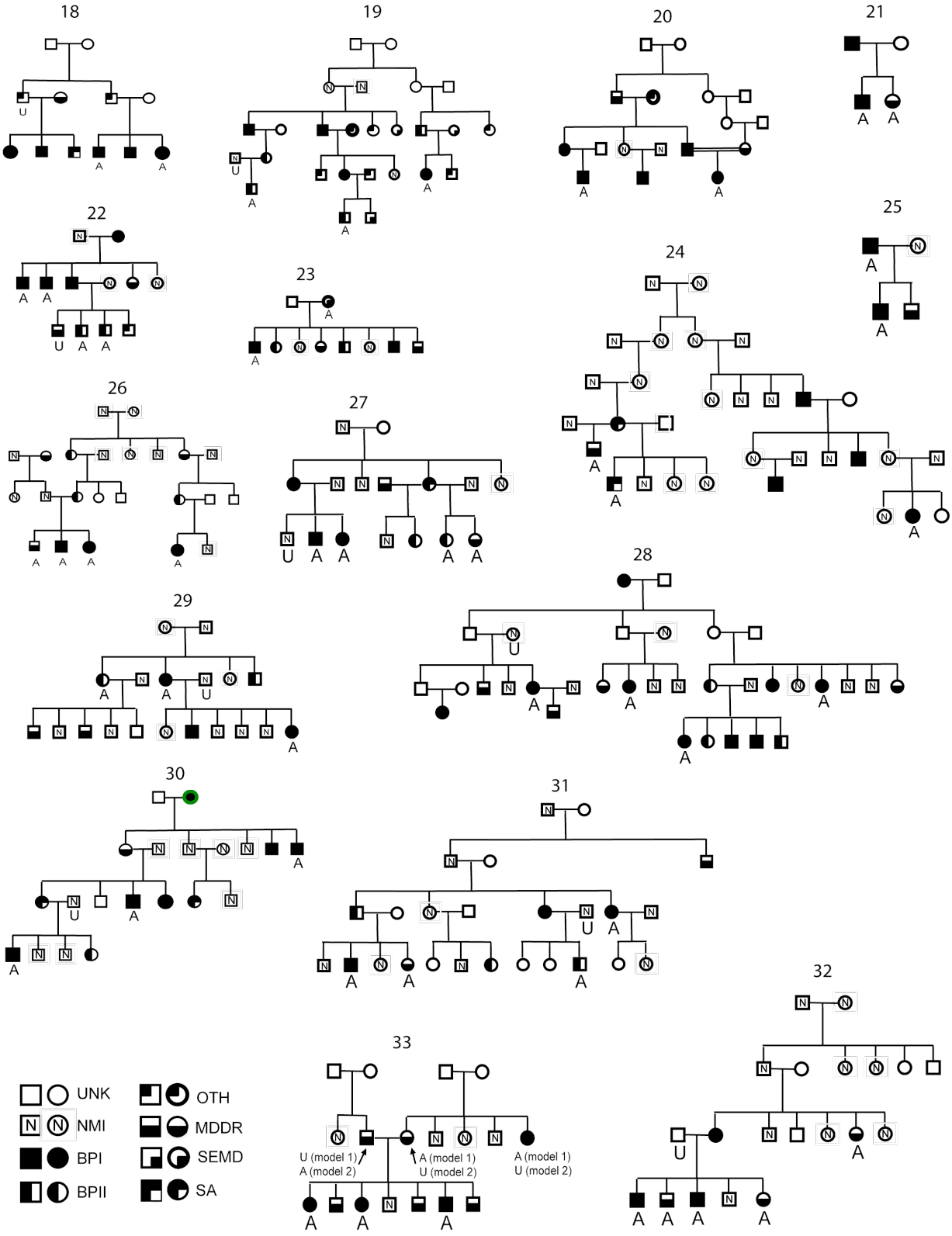


Figure S4. Pedigree diagrams 34-41: WGS from all available individuals. Letters indicate the individuals from whom whole-genome sequences were obtained. Each of these individuals was designated affected (A) or unaffected (U) as described in Supplementary Methods Section 3.

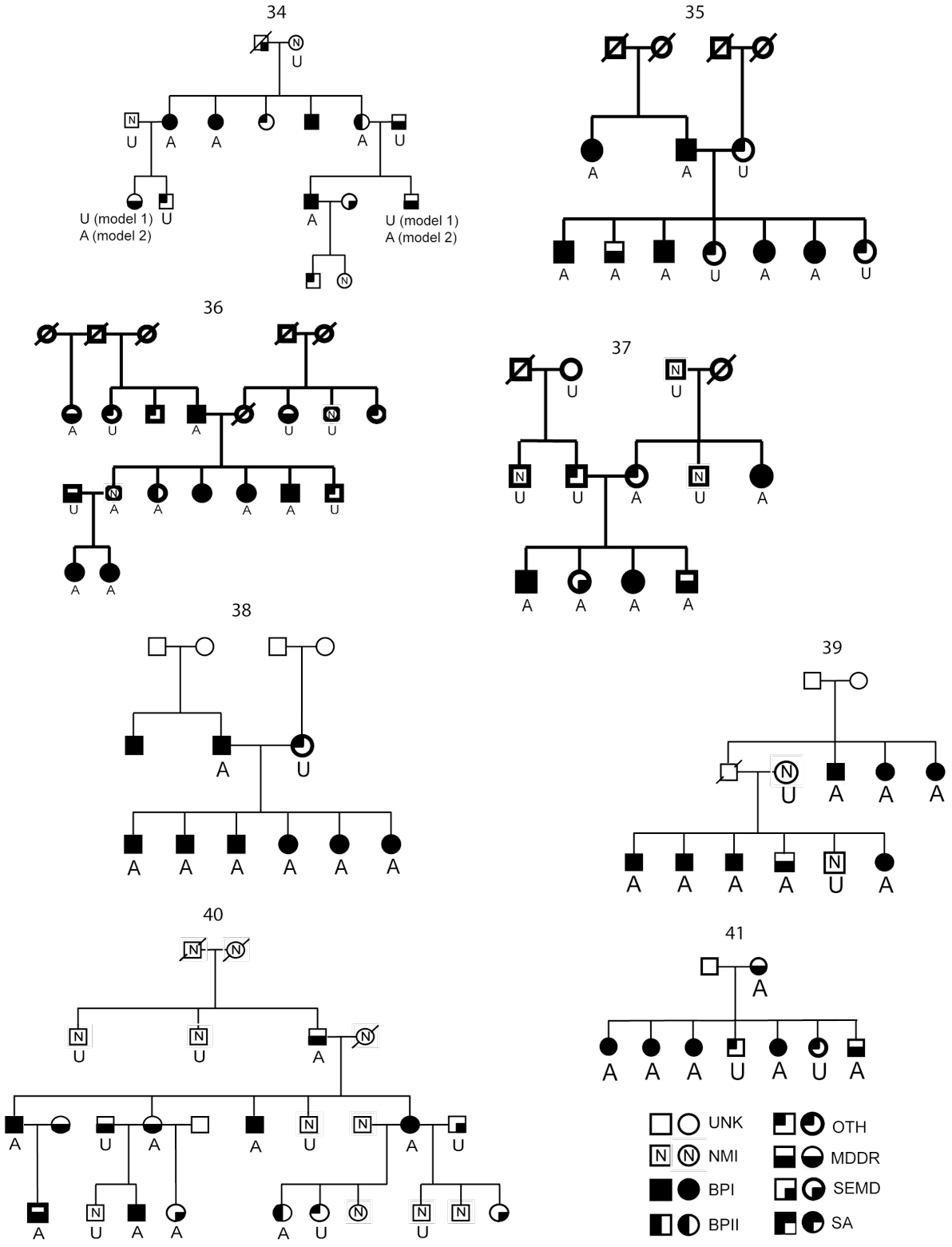


Figure S5. Distribution of coding and non-coding variation in voltage-gated calcium channels and GABA receptors across bipolar disorder cases and controls. A mixed model was implemented in EMMAX to test for associations between BD and uncommon (<5% MAF), bi-allelic SNVs in voltage-gated calcium channels and GABA receptors, using WGS from 108 affected individuals from BD pedigrees ("A"), 57 unaffected individuals from BD pedigrees ("U"), and 254 individuals from control, non-BD pedigrees ("U"). Plots are shown for SNVs in GABA receptors and voltage-gated calcium channels that were associated with risk or protection for BD ($P < 0.05$). Each row represents an SNV, and each column represents an individual person. Individuals are ordered by pedigree along the horizontal axis. Dark blue lines indicate that an individual is heterozygous or homozygous for the SNV in that row. The higher density of SNVs in affected individuals indicates a pathway-wide association of voltage-gated calcium channels and $GABA_A$ receptors with risk for BD. The sparse representation of each SNV suggests that the specific risk variants are heterogeneous among the unaffected individuals. The presence of multiple risk variants in some individuals raises the possibility of additive or epistatic interactions among variants within each pathway.

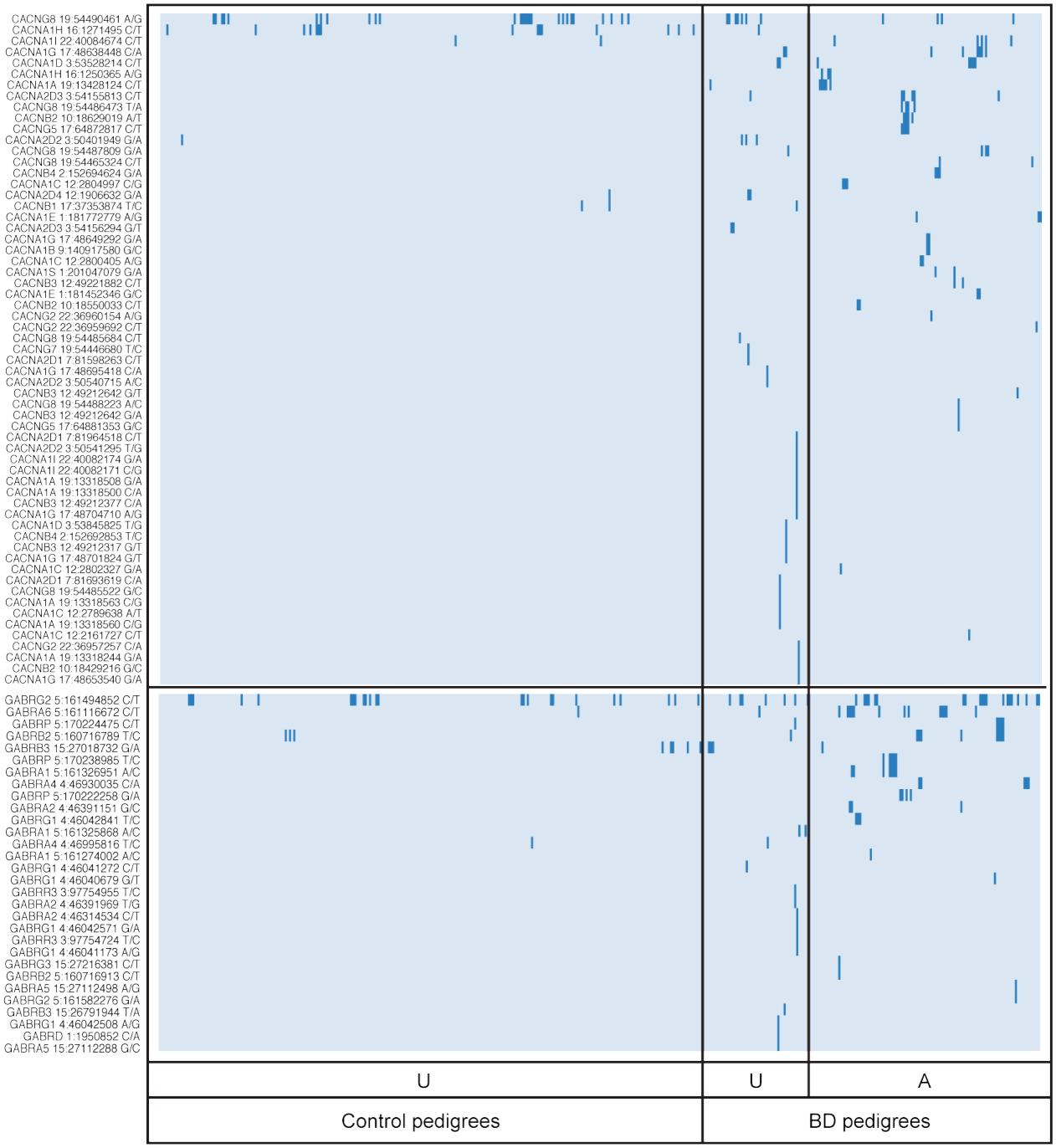


Figure S6. Oligogenic combinations of risk variants in bipolar disorder pedigrees. Combining results from variant- and pathway-level models suggested that most affected individuals inherited multiple risk variants. For this analysis, we defined as risk variants all SNVs with p-values < 0.001 and odds ratios > 1 by mixed model analysis ("EMMAX", 14 SNVs), as well as SNVs with mixed-model p-values < 0.05, odds ratios > 1, and an annotation to one of the following enriched pathways: BioCarta GABA pathway (9 SNVs), calcium channels (38 SNVs), CaM kinases (10 SNVs), GTPases (88 SNVs), and glycolysis / tricarboxylic acid cycle ("Metabolism", 13 SNVs). For each category, we report the number of risk variants identified in each affected BD case ("A", orange), unaffected relative in a BD pedigree ("UR", light blue), or population control ("PC", dark blue). The widths of polygons are proportional to the number of individuals with each variant count. The figure showing combined risk variant scores is reproduced from Fig. 2D.

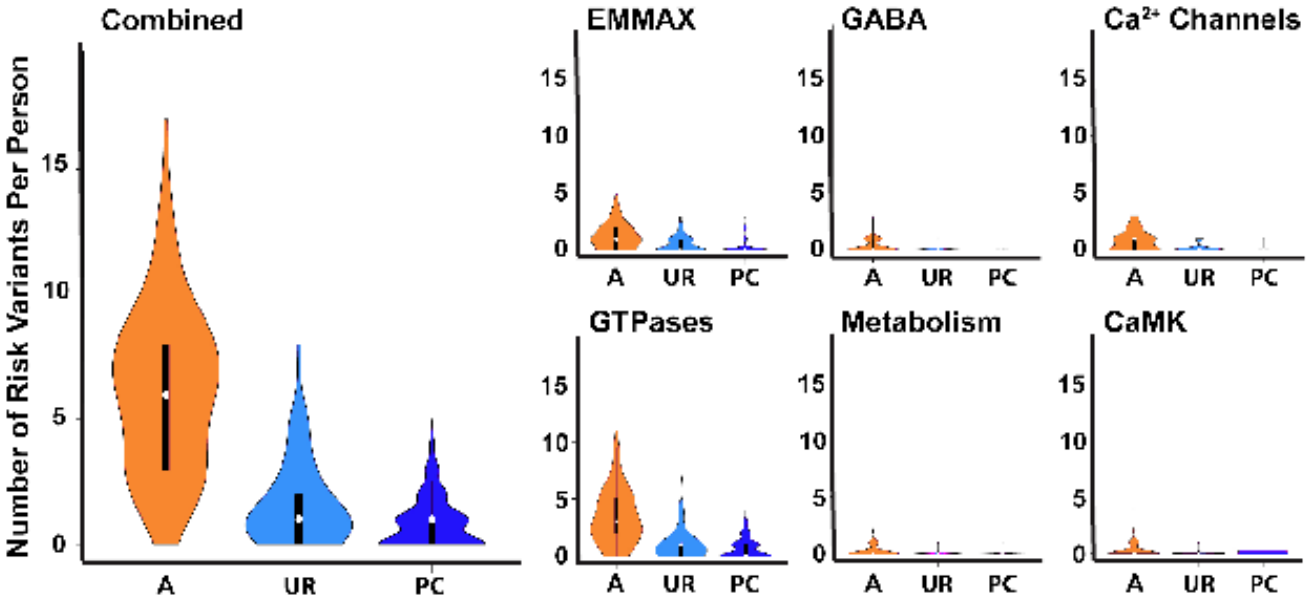


Figure S7. Per-base coverage profiles of the four targeted sequencing libraries. Set1of1 = sample batch #1 + amplicon set 1, Set1of2 = sample batch #1 + amplicon set 2, Set2 of1 = sample batch 2 + amplicon set 1, Set2of2 = sample batch 2 + amplicon set 2.

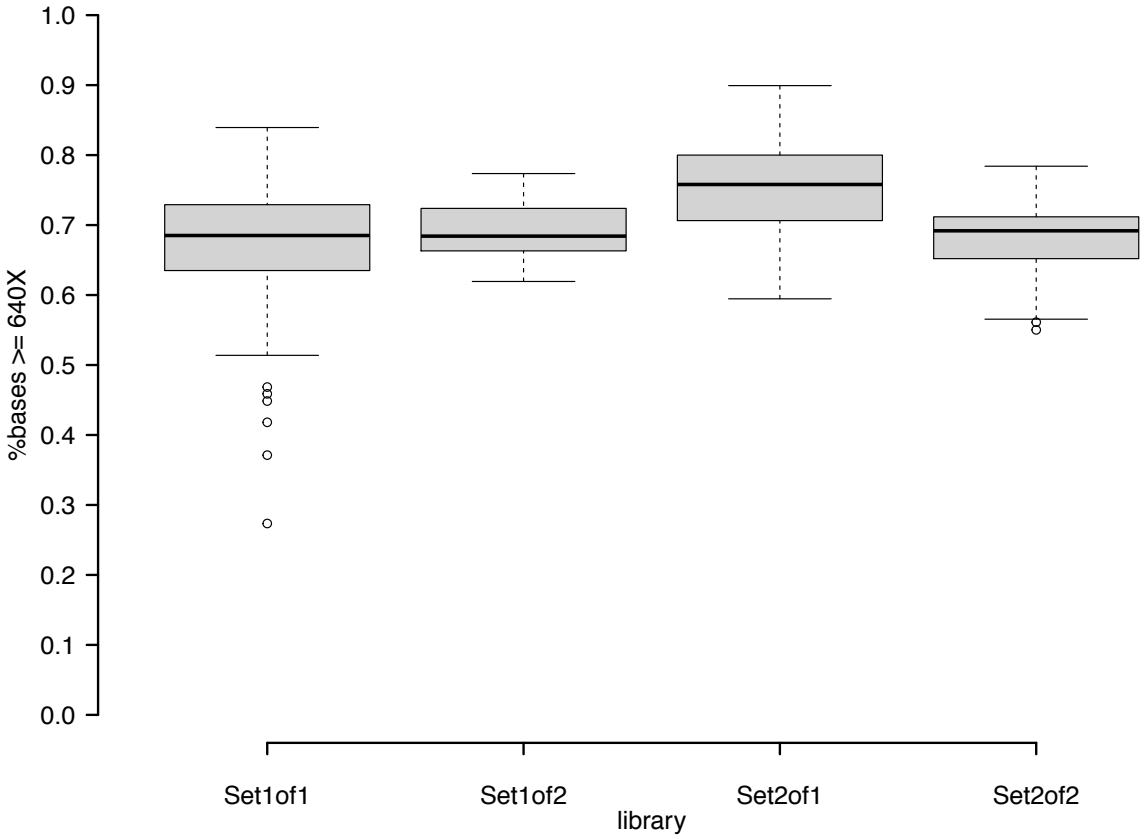


Figure S8. Distribution of average read depth per base in targeted sequencing. Histogram shows all targeted regions, across all sequenced individuals.

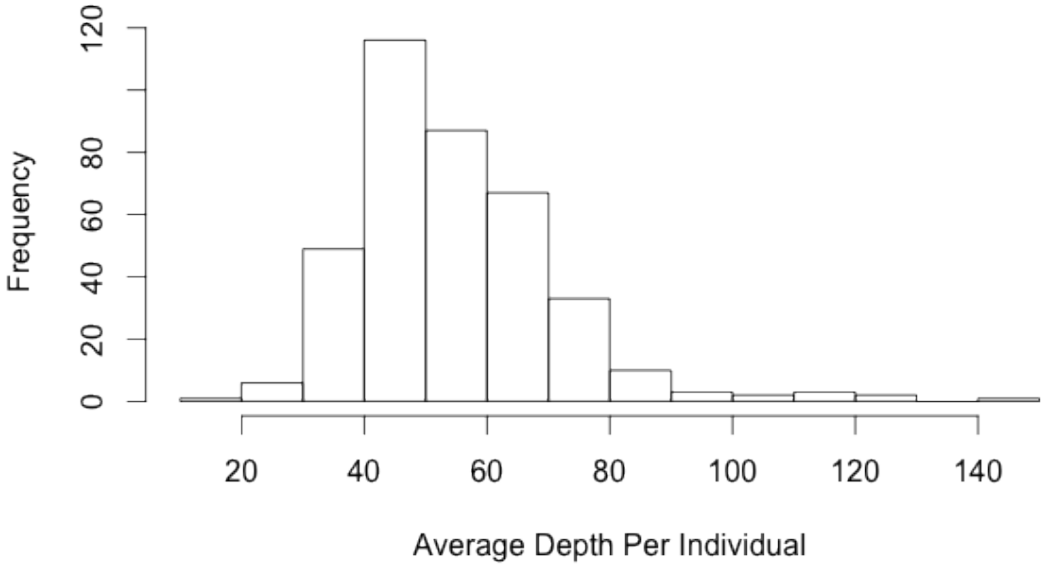


Figure S9. Allele frequencies from pooled sequencing are correlated with allele frequency estimates from public databases. We identified a total of 3,715 distinct SNVs located at positions with at least 20x coverage per individual in every pool in which the variant was observed. 1,960 of these SNVs are present in the Kaviar database (12). Allele frequencies for these known SNVs were tightly correlated between our targeted sequencing and public data (r = 0.88). Allele frequencies in public genomes are based on all data in the Kaviar database on January 29, 2014.

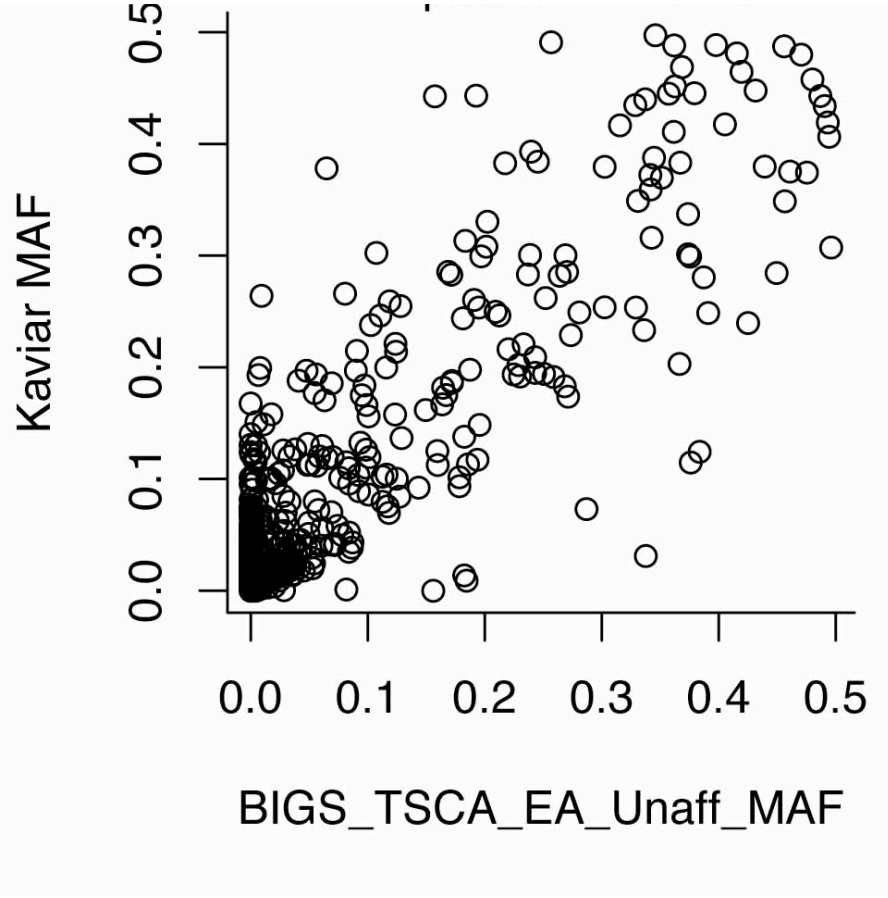


Figure S10. Allele frequency estimates from targeted sequencing are highly replicable across multiple sequencing runs. To evaluate replicability of allele frequency estimates from pooled sequencing, we performed a second round of sequencing on all libraries at a subset of positions. 1,728 alleles were observed at positions with >600x coverage in both sequencing runs. We estimated the number of haplotypes (out of 12,002) containing each minor allele, separately for the primary sequencing run (x-axis) and for the duplicate sequencing run (y-axis). Plots are shown at varying levels of resolution to emphasize both common and rare alleles. In the plot showing the rarest variants, values indicate the number of variants with each haplotype count. MAF, Minor allele frequency.

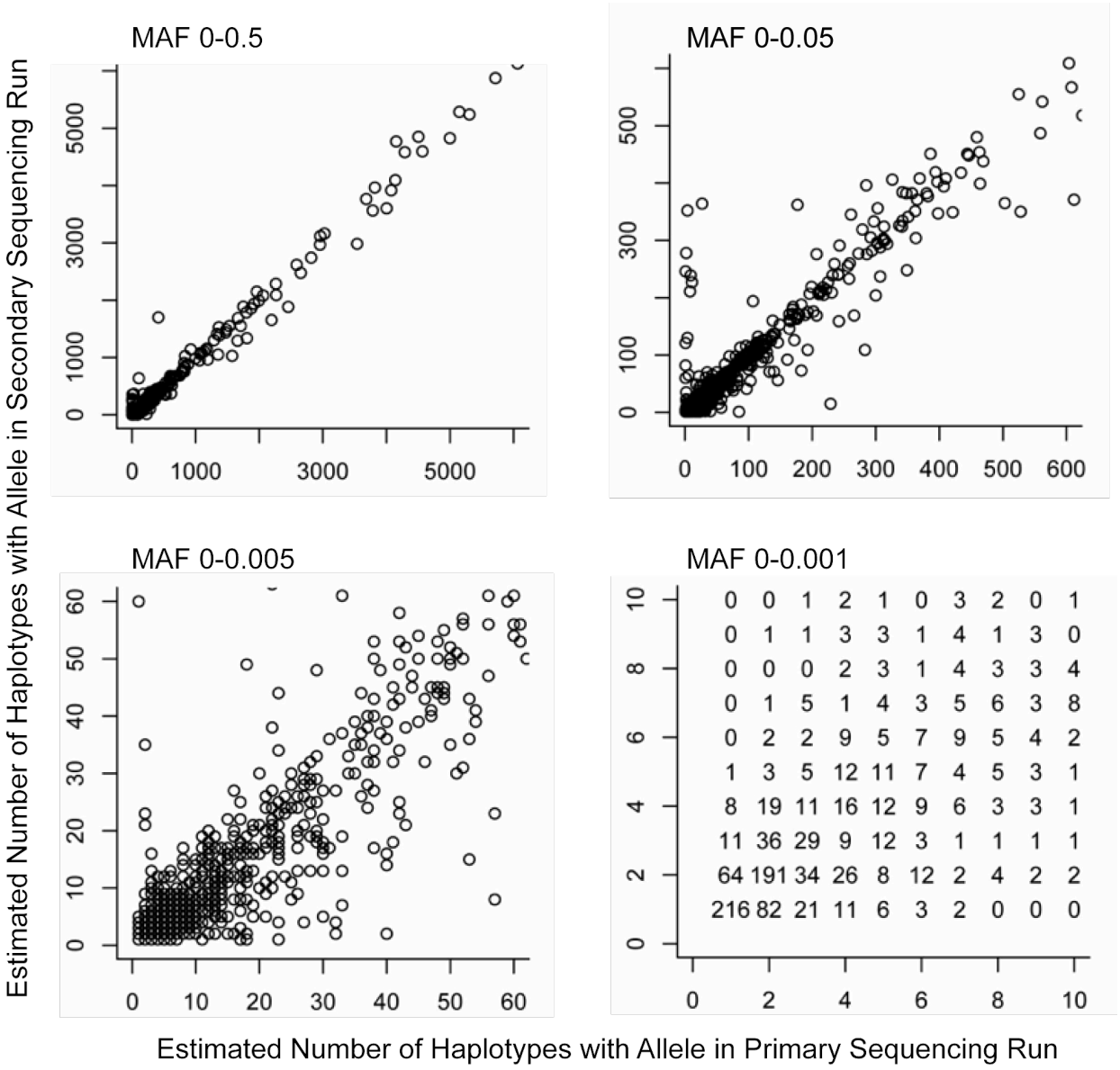


Figure S11. Distribution of alternate allele haplotype counts per pool. To assess the accuracy of allele count estimates within pools, we compared the observed distribution of allele counts to the expected counts if alleles are randomly distributed across pools (dashed red lines). Raw allele counts per pool (black dotted lines) deviated from expected values for counts of 0-2 per pool. Following normalization (solid black lines), allele count estimates were similar to expected values. Association tests reported in the main text and in Table 1 are based on normalized allele counts.

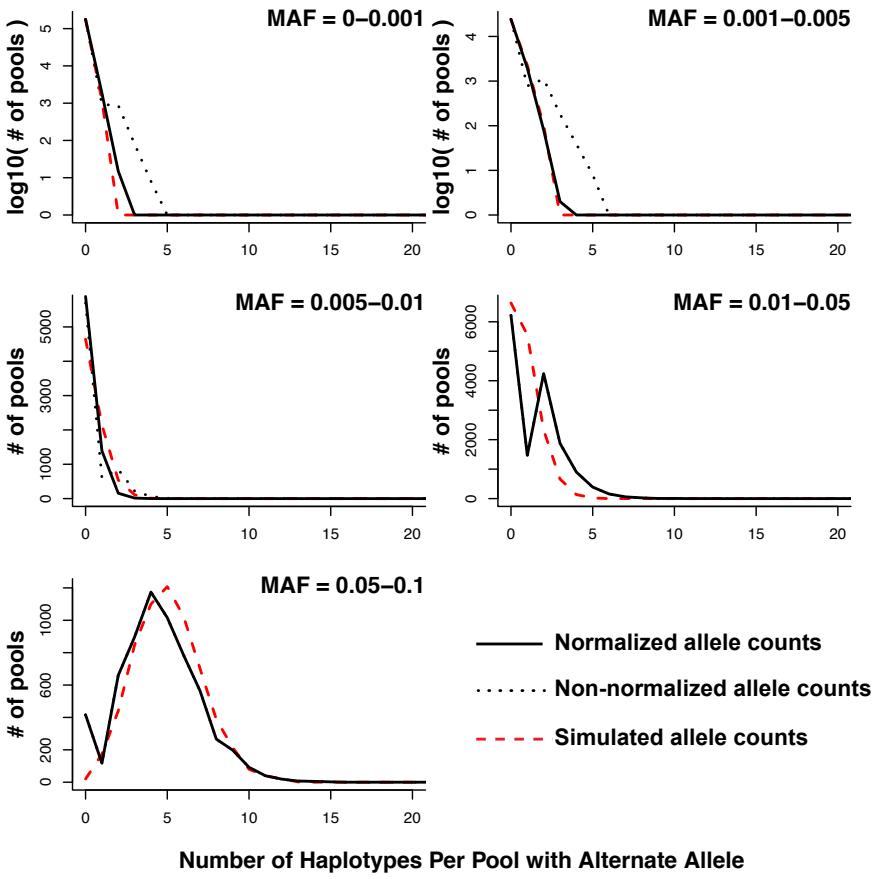
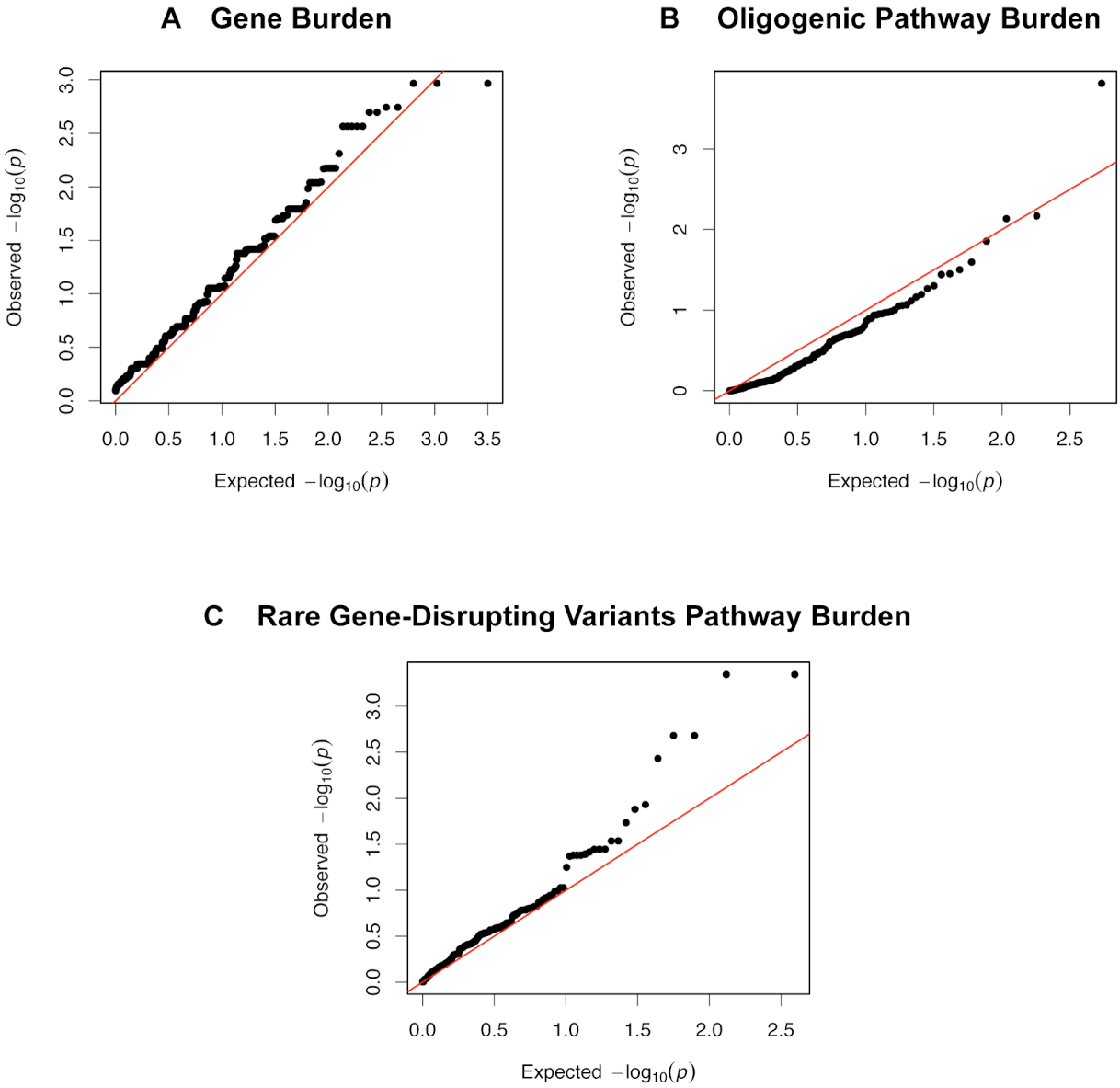


Figure S12. Quantile-quantile plots for gene and pathway burden tests. A. Gene burden p-values were calculated for 3,087 genes. B. Oligogenic pathway burden p-values were calculated for 269 pathways. C. gene-disrupting variant pathway burden p-values were calculated for 325 pathways. We used quantile-quantile plots to compare the observed (y-axis) vs. null expectation (x-axis) p-value distributions for each test. Each point describes a single gene or pathway. The diagonal line on each plot indicates the distribution of p-values that is expected to occur if no genes or pathways have true associations with disease risk.



III. Tables S1-S10

Table S1. Summary of sequenced pedigrees and individuals. The first column lists the numeric identifier for each of the 41 BD pedigrees. The next two columns describe the number of affected and unaffected individuals from each pedigree with WGS. The last two columns describe thresholds for genetic segregation (i.e., allele sharing among affected and unaffected individuals). These thresholds were used to define candidate variants in each pedigree, as described in the supplementary methods. In pedigrees 33 and 34, two distinct affection status models (ASMs) were considered, and both designations of affected vs. unaffected status are shown.

Pedigree ID	# of Affected Individuals Sequenced	# of Unaffected Individuals Sequenced	Minimum # of Affected Individuals with Variant	Maximum # of Unaffected Individuals with Variant
1	1	1	1	0
2	1	1	1	0
3	1	1	1	0
4	1	1	1	0
5	1	1	1	0
6	1	1	1	0
7	1	1	1	0
8	1	1	1	0
9	1	1	1	0
10	2	1	2	0
11	2	2	2	1
12	1	3	1	1
13	2	2	2	0
14	2	1	2	0
15	1	2	1	0
16	2	1	2	0
17	3	1	3	0
18	2	1	2	0
19	3	1	2	0
20	3	0	3	0
21	2	0	2	0
22	4	1	3	0
23	2	0	2	0
24	3	0	2	0
25	2	0	2	0
26	4	0	3	0
27	4	1	3	0
28	4	1	3	0
29	3	1	3	0
30	4	1	2	0
31	4	1	3	0
32	5	1	4	0
33	5 / 4	1 / 2	4 / 3	0 / 0
34	4 / 6	6 / 4	3 / 5	1 / 1
35	7	3	6	1
36	8	5	7	1
37	6	5	5	1
38	7	1	6	0
39	8	2	7	1
40	9	8	7	3
41	6	2	5	1

Table S2. SNVs significantly associated (q < 0.1) with BD in a mixed-model association test of 454 individuals from BD and non-BD pedigrees. We tested for associations of BD with 5,730 SNVs with 1-5% minor allele frequencies and predicted coding or regulatory effects. We compared WGS from 108 cases (BP1, BP2, or SA-BP) within 41 multiply affected pedigrees, to a total of 309 unaffected relatives and population controls. We used linear mixed models implemented with EMMAX(19). These mixed models are designed to test for the associations of single SNVs while accounting for relatedness within pedigrees, population structure, and polygenes. The table provides annotation and association statistics for four SNVs reaching study-wide significance, defined by a False Discovery Rate (q-value) < 10%. Chromosomal position (Chr, Pos) is in 1-based GRCh37 coordinates. Reference (Ref) and alternate (Alt) nucleotides are also based on GRCh37. Gene annotations (Gene, Region, Amino Acid Change) are based on RefSeq gene models. Frequencies of each variant in affected (Freq. Aff.) and unaffected (Freq. Unaff.) individuals, are based on observed genotypes and are not corrected for family structure. No-call rate indicates the fraction of missing genotypes for each SNV. Beta coefficients and p-values are from EMMAX. A beta coefficient < 0 indicates that the SNV is associated with an increased risk for bipolar disorder. Q-values for each SNV were calculated using the method of Benjamini and Hochberg(35), and may be conservative in the presence of non-independence among SNVs.

Chr	Pos	Ref	Alt	Gene	Region	A.A. Change	Freq. Aff.	Freq. Unaff.	No-Call Rate	Beta	P-Value	Q-Value
5	161116672	C	T	*GABRA6*	Exonic	T187M	0.060	0.003	0.002	-0.49	3.0e-5	9.0e-2
10	134902253	C	T	*GPR123*	5'UTR	--	0.097	0.010	0.012	-0.37	4.3e-5	9.0e-2
20	61392681	G	A	*NTSR1*	3'UTR	--	0.097	0.021	0.000	-0.31	6.3e-5	9.0e-2
20	61393859	G	A	*NTSR1*	3'UTR	--	0.097	0.021	0.000	-0.31	6.3e-5	9.0e-2

Table S3. Genes with a suggestive burden of fully and nearly segregating uncommon, coding and non-coding variants in bipolar disorder pedigrees. We identified uncommon and rare functional variants that co-segregated with disease in each BD pedigree and used the empirical distribution of segregating variants in 34 matched control pedigrees to identify genes with an increased burden of rare variants. We calculated a p-value for each gene using Fisher's exact test. Although no gene was significant in this analysis after correcting for multiple testing, there were several genes in which we observed segregating variants in many more BD pedigrees than control pedigrees. We show results for all genes with p-values ≤ 0.01. The top gene, *GABRA4*, is a GABA receptor. This result, together with study-wide significant associations from single-variant tests and pathway burden tests, supports a role for GABA receptors in risk for BD.

Gene	BD Pedigrees (41)	Ctrl Pedigrees (34)	Odds Ratio	P-Value
GABRA4	13	1	14.9	1.1e-3
TFAM	13	1	14.9	1.1e-3
GFAP	11	1	11.8	4.4e-3
TPP1	8	0	Inf	5.7e-3
GRIK2	7	0	Inf	1.1e-2

Table S4. Protein complexes and signaling pathways with an increased burden of fully and nearly-fully segregating variants in bipolar disorder pedigrees. P-values were calculated by comparing the frequency of segregating variants from each pathway in each pedigree, using Wilcoxon's signed rank test (see Methods). Q-values represent a False Discovery Rate calculated using the Benjamini-Hochberg method.

Pathway	BD / Ctrl	Genes in Pathway	Genes with Segregating Variant	P-Value	Q-Value
BioCarta GABA Pathway	4.3	10	8	1.5e-4	4.1e-2
GTPase Activity	1.7	99	88	6.8e-3	> 0.1
BioCarta PGC1A Pathway	2.1	26	20	7.3e-3	> 0.1
BioCarta Ca/CaM Pathway	2.3	16	10	1.4e-2	> 0.1
BioCarta Glycolysis Pathway	2.8	10	9	2.5e-2	> 0.1
Calcium Channel Activity	1.5	33	29	3.1e-2	> 0.1
KEGG Pathogenic *Escherichia coli* Infection	1.6	59	47	3.5e-2	> 0.1
Voltage Gated Calcium Channel Complex	1.6	15	13	3.6e-2	> 0.1
Tricarboxylic Acid Cycle Intermediate Metabolic Process	2.0	11	10	5.0e-2	> 0.1

Table S5. Protein complexes and signaling pathways enriched for fully and nearly-fully segregating, gene-disrupting variants in bipolar disorder pedigrees. For each of 325 functionally-related gene sets from Gene Ontology(36), KEGG(37), and BioCarta, we calculated an empirical p-value by comparison of observed gene-disrupting variants (GDVs) to the genome-wide null distribution of transmitted GDVs in 242 control trios. Results are shown for gene sets with empirical p-values < 0.05.

Gene Set	Genes in Set	Observed GDVs	P-Value	Q-Value	Genes
Nicotinic Acetylcholine Gated Receptor Channel Complex	11	3	4.5e-4	7.3e-2	*CHRNA6, CHRNB4, CHRND*
Nicotinic Acetylcholine Activated Cation Selective Channel Activity	11	3	4.5e-4	7.3e-2	*CHRNA6, CHRNB4, CHRND*
Excitatory Extracellular Ligand Gated Ion Channel Activity	21	3	2.1e-3	> 0.1	*CHRNA6, CHRNB4, CHRND*
Extracellular Ligand Gated Ion Channel Activity	22	3	2.1e-3	> 0.1	*CHRNA6, CHRNB4, CHRND*
Metal Ion Transmembrane Transporter Activity	147	11	3.7e-3	> 0.1	*CACNA1G, CCS, CHRNA6, CHRNB4, CHRND, KCNQ5, KCNS1, MS4A2, SLC39A2, SLC40A1, TRPM1*
Acetylcholine Binding	17	2	1.2e-2	> 0.1	*CHRNA6, CHRNB4*
Amine Binding	23	2	1.3e-2	> 0.1	*CHRNA6, CHRNB4*
Cation Channel Activity	119	8	1.8e-2	> 0.1	*CACNA1G, CHRNA6, CHRNB4, CHRND, KCNQ5, KCNS1, MS4A2, TRPM1*
KEGG Pathogenic Escherichia Coli Infection	59	3	2.9e-2	> 0.1	*NCL, TUBB3, TUBB4B*
Ion Channel Activity	149	9	2.9e-2	> 0.1	*CACNA1G, CHRNA6, CHRNB4, CHRND, CLCA1, KCNQ5, KCNS1, MS4A2, TRPM1*
Substrate Specific Channel Activity	156	9	3.6e-2	> 0.1	*CACNA1G, CHRNA6, CHRNB4, CHRND, CLCA1, KCNQ5, KCNS1, MS4A2, TRPM1*
Response To Nutrient	17	2	3.6e-2	> 0.1	*CCKAR, ENPP1*
Coated Vesicle	47	4	3.6e-2	> 0.1	*COPB1, FURIN, KIAA0368, NECAP2*
Regulation Of MAP Kinase Activity	67	4	3.8e-2	> 0.1	*ERBB2, MAP3K10, MAP3K5, TRIB3*
Positive Regulation Of MAP Kinase Activity	47	3	4.1e-2	> 0.1	*ERBB2, MAP3K10, MAP3K5*
BioCarta Shh Pathway	16	2	4.2e-2	> 0.1	*PRKACB, PTCH1*
Neurotransmitter Binding	53	3	4.2e-2	> 0.1	*CCKAR, CHRNA6, CHRNB4*
Neurotransmitter Receptor Activity	50	3	4.2e-2	> 0.1	*CCKAR, CHRNA6, CHRNB4*
Positive Regulation Of Catalytic Activity	165	7	4.3e-2	> 0.1	*CASP8AP2, ERBB2, MALT1, MAP3K10, MAP3K5, NLRC4, TNNT2*

Table S6. Enrichment of pathways with an increased burden of rare variants in BD pedigrees for differentially expressed genes in dorsolateral prefrontal cortex from BD cases vs. controls. We analyzed RNA-seq gene expression profiles from the dorsolateral prefrontal cortex of 11 BD cases and 11 age- and gender-matched controls (38). We identified 3,059 genes with evidence for differential expression ($p < 0.05$) between cases and controls. Using hypergeometric tests, we evaluated enrichments of these differentially expressed genes for each of the 27 pathways with a significant burden of rare variants in BD pedigrees (Tables S4, S5). DEGs, differentially expressed genes.

Pathway	# of Genes in Pathway	# DEGs	Overlap	P-value
BioCarta SHH Pathway	25	3059	9	1.7e-4
BioCarta PGC1A Pathway	3	3059	11	1.5e-3
Voltage Gated Calcium Channel Complex	8	3059	7	3.0e-3
Metal Ion Transmembrane Transporter Activity	14	3059	39	1.0e-2
Calcium Channel Activity	6	3059	11	1.5e-2
BioCarta Ca/CaM Pathway	4	3059	6	2.0e-2
BioCarta GABA Pathway	1	3059	4	2.7e-2
Cation Channel Activity	17	3059	30	3.0e-2
Ion Channel Activity	19	3059	35	7.3e-2
Substrate Specific Channel Activity	20	3059	36	8.0e-2
Coated Vesicle	22	3059	10	2.7e-1
KEGG Pathogenic *Escherischia coli* Infection	7	3059	11	4.6e-1
GTPase Activity	2	3059	18	5.3e-1
Positive Regulation of MAP Kinase Activity	24	3059	8	5.6e-1
Positive Regulation of Catalytic Activity	28	3059	30	5.7e-1
Nicotinic Acetylcholine Gated Receptor Channel Activity	10	3059	1	6.5e-1
Nicotinic Acetylcholine Activated Cation Selective Channel Activity	11	3059	1	6.5e-1
Excitatory Extracellular Ligand Gated Ion Channel Activity	12	3059	2	8.0e-1
Extracellular Ligand Gated Ion Channel Activity	13	3059	2	8.2e-1
Amine Binding	16	3059	2	8.4e-1
Regulation of MAP Kinase Activity	23	3059	9	8.5e-1
Response to Nutrient	21	3059	1	8.6e-1
Acetylcholine Binding	15	3059	1	8.6e-1
BioCarta Glycolysis Pathway	5	3059	0	8.8e-1
Tricarboxylic Acid Cycle Intermediate Metabolic Process	9	3059	0	9.0e-1
Neurotransmitter Binding	26	3059	3	9.9e-1
Neurotransmitter Receptor Activity	27	3059	3	9.9e-1

Table S7. Genes targeted for sequencing in 3,014 BD cases and 1,717 controls. 10 calcium channels, 5 GABA receptors, and 5 calmodulin-dependent protein kinases were selected for sequencing based on a significant pathway burden in BD pedigrees. An additional six genes were selected based on evidence from GWAS and candidate gene association studies.

Gene	Group	Reference
CACNA1B	Calcium channel	
CACNA1D	Calcium channel	
CACNA1E	Calcium channel	
CACNB2	Calcium channel	
CACNB3	Calcium channel	
CACNB4	Calcium channel	
CACNG1	Calcium channel	
CACNG2	Calcium channel	
RYR3	Calcium channel	
CACNA1C	Calcium channel; GWAS	Ferreira et al. (2008)(39)
CAMK2A	Calmodulin-dependent protein kinase signaling	
CAMK2B	Calmodulin-dependent protein kinase signaling	
CAMK2D	Calmodulin-dependent protein kinase signaling	
CAMKK1	Calmodulin-dependent protein kinase signaling	
CREB1	Calmodulin-dependent protein kinase signaling	
GABRA1	GABA receptor	
GABRA2	GABA receptor	
GABRA4	GABA receptor	
GABRA6	GABA receptor	
GABRG1	GABA receptor	
ANK3	GWAS	Ferreira et al. (2008)(39)
NGF	GWAS	Sklar et al. (2011)(16)
NTRK2	GWAS	Smith et al. (2009)(6)
TENM4	GWAS	Sklar et al. (2011)(16)
BDNF	Candidate gene association studies	Sklar et al. (2002)(40); Neves-Perreira et al. (2002)(41)
NTRK1	Candidate gene association studies	

Table S8. Quality control metrics for targeted re-sequencing data. Targeted sequencing was performed across four lanes of a HiSeq2500 sequencer, so that each of the two sets of amplicons was sequenced in each of the two batches of samples on a single lane. In total, we analyzed 1.28 billion high-quality sequencing reads covering 320 billion base pairs. The table shows quality metrics for individual sequencing lanes, indicating stable performance across the four lanes. Gb, billion base pairs; bp, base pairs.

Lane	Sample Pool	Amplicon Set	Raw Reads (M)	High-Quality Reads (M)	Read Length (bp)	Quality Yield (Gb)	Reads with perfect barcode	Bases > Q30
1	1	1	368.34	314.96	250	78.7	84.9%	70.5%
2	1	2	373.58	326.72	250	81.7	86.7%	74.9%
3	2	1	360.04	320.62	250	80.2	87.9%	76.3%
4	2	2	352.64	318.52	250	79.6	89.1%	81.0%

Table S9. Associations between rare coding variants in 26 candidate genes and risk for BD, based on targeted sequencing of 3,014 cases and 1,717 controls. Rare-variant association tests were used to evaluate associations between regulatory variants at each locus and BD affection status. Coding variants were defined as SNVs with predicted effects on the sequence of a protein produced by any of a gene's knownGene transcripts. We report results from our primary statistical model (Model 1), as well as for four alternate models with distinct variant aggregation statistics and normalization procedures. For further information see Supplementary Methods Section 8. For each test, we report the number of single-nucleotide variants aggregated in that test ("SNVs"), an asymptotic p-value from the test's theoretical distribution ("P-Asym"), an empirical p-value derived from 100,000 permutations ("P-Perm"), and a false discovery rate for P-Perm based on the method of Benjamini and Hochberg ("Q-Val").

	SKAT				SKAT				SKAT				C-Alpha				Burden			
	Expectation Maximization				MAF > 0.001				None				Expectation Maximization				Expectation Maximization			
Gene	**SNVs**	**P-Asym**	**P-Perm**	**Q-Val**	**SNVs**	**P-Asym**	**P-Perm**	**Q-Val**	**SNVs**	**P-Asym**	**P-Perm**	**Q-Val**	**SNVs**	**P-Asym**	**P-Perm**	**Q-Val**	**SNVs**	**P-Asym**	**P-Perm**	**Q-Val**
ANK3	58	4.1e-2	3.4e-2	0.22	15	9.5e-2	9.3e-2	0.10	58	5.3e-2	4.9e-2	0.10	51	0.61	0.60	0.76	58	1.8e-3	1.3e-3	2.2e-2
GABRA6	8	4.6e-2	3.8e-2	0.22	2	1.5e-2	1.2e-2	4.9e-2	8	5.5e-3	4.5e-3	3.8e-2	8	3.9e-2	4.5e-2	0.44	8	0.13	0.16	0.32
CACNA1C	6	5.3e-2	3.8e-2	0.22	1	NA	NA	NA	6	5.2e-2	4.7e-2	0.10	6	5.2e-2	5.2e-2	0.44	6	0.27	0.32	0.52
CACNB4	3	0.12	9.3e-2	0.39	1	NA	NA	NA	3	0.13	0.17	0.26	3	0.29	0.11	0.45	3	0.29	0.39	0.55
CACNA1E	7	0.21	0.14	0.47	0	NA	NA	NA	7	2.3e-2	1.3e-2	7.4e-2	7	0.39	0.29	0.56	7	0.68	0.74	0.79
NTRK2	3	0.17	0.18	0.48	1	NA	NA	NA	3	0.14	0.14	0.24	3	0.54	0.43	0.70	3	4.7e-2	3.6e-2	0.17
CAMK2A	2	0.15	0.22	0.48	0	NA	NA	NA	2	2.4e-2	2.1e-2	7.8e-2	2	0.35	0.20	0.55	2	9.1e-2	6.1e-2	0.17
RYR3	44	0.26	0.23	0.48	6	0.11	0.10	0.10	44	3.1e-2	2.8e-2	7.8e-2	43	0.65	0.63	0.76	44	0.17	0.17	0.32
TENM4	24	0.30	0.34	0.63	5	3.8e-2	3.5e-2	7.0e-2	24	2.8e-2	2.6e-2	7.8e-2	23	0.16	8.8e-2	0.45	24	5.3e-2	5.7e-2	0.17
CACNA1D	19	0.41	0.43	0.74	1	NA	NA	NA	19	3.5e-3	2.4e-3	3.8e-2	19	0.97	0.97	0.97	19	0.41	0.33	0.52
CAMKK1	2	0.38	0.53	0.77	0	NA	NA	NA	2	0.15	0.20	0.28	2	0.95	0.74	0.79	2	1.9e-1	4.5e-2	0.17
CREB1	2	0.38	0.54	0.77	0	NA	NA	NA	2	0.32	0.41	0.50	2	0.95	0.74	0.79	2	1.9e-1	4.5e-2	0.17
CACNB2	7	0.65	0.68	0.88	1	NA	NA	NA	7	0.24	0.24	0.32	7	0.59	0.55	0.76	7	0.15	0.14	0.32
NGF	2	0.72	0.74	0.88	1	NA	NA	NA	2	0.60	0.59	0.62	2	0.53	0.45	0.70	2	0.60	0.62	0.79
NTRK1	3	0.66	0.81	0.88	0	NA	NA	NA	3	0.61	0.67	0.67	3	0.47	0.26	0.55	3	0.86	0.70	0.79
CAMK2D	4	0.86	0.84	0.88	1	NA	NA	NA	4	6.8e-2	6.1e-2	0.12	4	0.48	0.24	0.55	4	0.86	0.96	0.96
GABRA4	4	0.71	0.88	0.88	0	NA	NA	NA	4	0.52	0.49	0.56	4	0.41	0.24	0.55	4	0.88	0.75	0.79
BDNF	0	NA	NA	NA	0	NA	NA	NA	0	NA	NA	NA	0	NA	NA	NA	0	NA	NA	NA
CACNA1B	1	NA	NA	NA	0	NA	NA	NA	1	NA	NA	NA	1	NA	NA	NA	1	NA	NA	NA
CACNB3	0	NA	NA	NA	0	NA	NA	NA	0	NA	NA	NA	0	NA	NA	NA	0	NA	NA	NA
CACNG1	1	NA	NA	NA	1	NA	NA	NA	1	NA	NA	NA	1	NA	NA	NA	1	NA	NA	NA
CACNG2	0	NA	NA	NA	0	NA	NA	NA	0	NA	NA	NA	0	NA	NA	NA	0	NA	NA	NA
CAMK2B	0	NA	NA	NA	0	NA	NA	NA	0	NA	NA	NA	0	NA	NA	NA	0	NA	NA	NA
GABRA1	0	NA	NA	NA	0	NA	NA	NA	0	NA	NA	NA	0	NA	NA	NA	0	NA	NA	NA
GABRA2	0	NA	NA	NA	0	NA	NA	NA	0	NA	NA	NA	0	NA	NA	NA	0	NA	NA	NA
GABRG1	0	NA	NA	NA	0	NA	NA	NA	0	NA	NA	NA	0	NA	NA	NA	0	NA	NA	NA

Table S10. Associations between rare regulatory variants in 26 candidate genes and risk for BD, based on targeted sequencing of 3,014 cases and 1,717 controls. Rare-variant association tests were used to evaluate associations between regulatory variants at each locus and BD affection status. Coding variants were defined as SNVs with predicted effects on the sequence of a protein produced by any of a gene's knownGene transcripts. We report results from our primary statistical model (Model 1), as well as for four alternate models with distinct variant aggregation statistics and normalization procedures. For further information see Supplementary Methods Section 8. For each test, we report the number of single-nucleotide variants aggregated in that test ("SNVs"), an asymptotic p-value from the test's theoretical distribution ("P-Asym"), an empirical p-value derived from 100,000 permutations ("P-Perm"), and a false discovery rate for P-Perm based on the method of Benjamini and Hochberg ("Q-Val").

	Model 1 SKAT Expectation Maximization				Model 2 SKAT MAF > 0.001				Model 3 SKAT None				Model 4 C-Alpha Expectation Maximization				Model 5 Burden Expectation Maximization			
Gene	SNVs	P-Asym	P-Perm	Q-Val	SNVs	P-Asym	P-Perm	Q-Val	SNVs	P-Asym	P-Perm	Q-Val	SNVs	P-Asym	P-Perm	Q-Val	SNVs	P-Asym	P-Perm	Q-Val
CACNA1C	238	2.4e-7	1.0e-5	2.6e-4	52	1.2e-7	1.0e-5	2.6e-4	238	2.4e-9	1.0e-5	2.6e-4	217	0	1.0e-5	2.6e-4	238	6.1e-2	6.3e-2	0.45
CACNA1B	50	4.5e-4	4.3e-4	4.1e-3	16	3.4e-4	3.2e-4	4.2e-3	50	1.5e-4	2.1e-4	1.9e-3	43	2.6E-13	8.0e-5	1.0e-3	50	0.36	0.36	0.61
CACNG2	43	3.8e-4	4.7e-4	4.1e-3	11	5.0e-4	4.8e-4	4.2e-3	43	2.0e-4	2.2e-4	1.9e-3	38	6.6E-10	4.8e-4	4.2e-3	43	0.62	0.62	0.73
CAMK2A	58	1.2e-3	1.4e-3	8.6e-3	19	1.5e-3	1.5e-3	7.8e-3	58	5.1e-4	6.6e-4	3.4e-3	47	9.3e-4	1.0e-2	3.3e-2	58	0.66	0.66	0.74
CACNA1D	111	1.9e-3	1.6e-3	8.6e-3	25	1.6e-3	1.3e-3	7.8e-3	111	7.0e-4	5.6e-4	3.4e-3	100	1.9e-4	3.1e-3	2.0e-2	111	0.22	0.22	0.51
NGF	17	8.9e-3	8.4e-3	3.6e-2	7	8.7e-3	7.6e-3	2.8e-2	17	7.6e-3	7.5e-3	2.8e-2	12	0.30	0.26	0.44	17	0.12	0.12	0.45
CACNG1	18	1.9e-2	1.8e-2	6.5e-2	7	5.1e-3	4.6e-3	2.0e-2	18	3.8e-3	3.6e-3	1.6e-2	16	0.99	1.00	1.00	18	0.11	0.12	0.45
GABRA4	49	2.1e-2	2.1e-2	6.7e-2	15	2.5e-2	2.5e-2	8.2e-2	49	1.2e-2	1.1e-2	3.4e-2	36	0.13	7.5e-2	0.21	49	0.24	0.23	0.51
TENM4	96	4.6e-2	4.3e-2	0.12	30	4.7e-2	4.5e-2	0.13	96	2.3e-2	2.1e-2	6.1e-2	78	1.7e-4	4.7e-3	2.1e-2	96	0.39	0.39	0.61
GABRA1	33	5.1e-2	4.7e-2	0.12	11	6.7e-2	6.2e-2	0.13	33	2.7e-2	2.4e-2	6.3e-2	30	2.8e-4	9.9e-3	3.3e-2	33	2.2e-2	2.2e-2	0.45
CAMK2D	50	5.8e-2	5.6e-2	0.13	11	5.6e-2	5.3e-2	0.13	50	3.4e-2	3.2e-2	6.9e-2	44	0.80	0.80	0.98	50	0.17	0.18	0.51
CAMKK1	42	6.0e-2	6.0e-2	0.13	11	7.0e-2	6.9e-2	0.14	42	2.9e-2	2.9e-2	6.8e-2	34	0.94	0.94	0.98	42	0.13	0.13	0.45
CACNB4	110	7.3e-2	7.1e-2	0.14	35	6.2e-2	6.0e-2	0.13	110	3.6e-2	3.6e-2	7.2e-2	85	1.5e-4	4.8e-3	2.1e-2	110	0.42	0.42	0.61
GABRG1	8	9.4e-2	9.4e-2	0.17	3	9.4e-2	9.5e-2	0.18	8	8.8e-2	8.7e-2	0.13	5	0.27	0.17	0.39	8	0.14	0.14	0.45
CACNB2	87	0.15	0.15	0.25	26	0.12	0.11	0.18	87	6.0e-2	5.8e-2	0.10	75	0.91	0.91	0.98	87	0.57	0.57	0.71
RYR3	83	0.18	0.18	0.29	25	0.11	0.11	0.18	83	6.2e-2	5.9e-2	0.10	73	0.60	0.58	0.78	83	0.83	0.82	0.89
CACNB3	26	0.22	0.22	0.33	3	0.17	0.17	0.26	26	0.14	0.14	0.19	25	0.18	8.2e-2	0.21	26	0.30	0.30	0.57
ANK3	108	0.36	0.36	0.52	30	0.53	0.54	0.66	108	0.23	0.23	0.30	93	0.60	0.58	0.78	108	8.5e-2	8.4e-2	0.45
NTRK1	27	0.44	0.45	0.62	6	0.33	0.33	0.48	27	0.14	0.14	0.19	25	0.90	0.91	0.98	27	0.94	0.94	0.94
NTRK2	118	0.51	0.51	0.66	38	0.50	0.50	0.65	118	0.32	0.32	0.39	100	0.52	0.51	0.78	118	0.46	0.46	0.61
CACNA1E	115	0.58	0.60	0.74	30	0.43	0.44	0.60	115	0.09	0.09	0.13	100	0.56	0.55	0.78	115	8.8e-2	8.9e-2	0.45
CREB1	51	0.64	0.65	0.77	15	0.69	0.70	0.83	51	0.41	0.41	0.49	43	0.61	0.60	0.78	51	0.28	0.28	0.56
GABRA6	13	0.72	0.75	0.80	2	0.84	0.82	0.85	13	0.71	0.74	0.77	12	0.87	0.89	0.98	13	0.46	0.46	0.61
BDNF	16	0.74	0.75	0.80	6	0.79	0.79	0.85	16	0.72	0.74	0.77	13	0.34	0.24	0.44	16	0.20	0.21	0.51
CAMK2B	26	0.76	0.77	0.80	8	0.76	0.76	0.85	26	0.69	0.69	0.77	20	0.34	0.20	0.40	26	0.47	0.47	0.61
GABRA2	34	0.95	0.94	0.94	8	0.94	0.94	0.94	34	0.90	0.90	0.90	29	0.38	0.19	0.40	34	0.92	0.92	0.94

Supplementary References

1. Power RA, et al. (2013) Fecundity of patients with schizophrenia, autism, bipolar disorder, depression, anorexia nervosa, or substance abuse vs their unaffected siblings. *JAMA Psychiatry* 70(1):22–30.
2. Malhotra D, et al. (2011) High frequencies of de novo CNVs in bipolar disorder and schizophrenia. *Neuron* 72(6):951–963.
3. McClellan JM, Susser E, King M-C (2006) Maternal famine, de novo mutations, and schizophrenia. *JAMA* 296(5):582–584.
4. Brown A, Bao Y, McKeague I, Shen L, Schaefer C (2013) Parental age and risk of bipolar disorder in offspring. *Psychiatry Res* 208(3):225–231.
5. Badner JA, et al. (2012) Genome-wide linkage analysis of 972 bipolar pedigrees using single-nucleotide polymorphisms. *Mol Psychiatry* 17(8):818–826.
6. Smith EN, et al. (2009) Genome-wide association study of bipolar disorder in European American and African American individuals. *Mol Psychiatry* 14(8):755–763.
7. Hu H, et al. (2014) A unified test of linkage analysis and rare-variant association for analysis of pedigree sequence data. *Nat Biotechnol*.
8. Stittrich A-B, et al. (2014) Mutations in *NOTCH1* Cause Adams-Oliver Syndrome. *Am J Hum Genet* 95(3):275–284.
9. Lehman A, et al. (2014) Diffuse angiopathy in Adams-Oliver syndrome associated with truncating *DOCK6* mutations. *Am J Med Genet A* 164(10):2656–2662.
10. Schubert J, et al. (2014) Mutations in *STX1B*, encoding a presynaptic protein, cause fever-associated epilepsy syndromes. *Nat Genet* 46(12):1327-1332.
11. Purcell S, et al. (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 81(3):559–575.
12. Glusman G, Caballero J, Mauldin DE, Hood L, Roach JC (2011) Kaviar: an accessible system for testing SNV novelty. *Bioinformatics* 27(22):3216–3217.
13. Kircher M, et al. (2014) A general framework for estimating the relative pathogenicity of human genetic variants. *Nat Genet* 46(3):310–315.
14. Lee SH, et al. (2013) Genetic relationship between five psychiatric disorders estimated from genome-wide SNPs. *Nat Genet* 45(9):984–994.
15. Nurnberger JI, et al. (2014) Identification of pathways for bipolar disorder: a meta-analysis. *JAMA Psychiatry* 71(6):657–664.
16. Large-scale genome-wide association analysis of bipolar disorder identifies a new susceptibility locus near *ODZ4*. (2011) *Nat Genet* 43(10):977–983.
17. Pirooznia M, et al. (2012) SynaptomeDB: an ontology-based knowledgebase for synaptic genes. *Bioinformatics* 28(6):897–899.
18. Pawson AJ, et al. (2014) The IUPHAR/BPS Guide to PHARMACOLOGY: an expert-driven knowledgebase of drug targets and their ligands. *Nucleic Acids Res* 42(Database issue):D1098–D1106.
19. Kang HM, et al. (2010) Variance component model to account for sample structure in genome-wide association studies. *Nat Genet* 42(4):348–354.
20. Aulchenko YS, Ripke S, Isaacs A, van Duijn CM (2007) GenABEL: an R library for genome-wide association analysis. *Bioinformatics* 23(10):1294–1296.
21. Wu MC, et al. (2011) Rare-variant association testing for sequencing data with the sequence kernel association test. *Am J Hum Genet* 89(1):82–93.
22. Iossifov I, et al. (2014) The contribution of de novo coding mutations to autism spectrum disorder. *Nature* 515(7526):216-221.
23. Bansal V, Tewhey R, Leproust EM, Schork NJ (2011) Efficient and cost effective population resequencing by pooling and in-solution hybridization. *PLoS One* 6(3):e18353.
24. Thurman RE, et al. (2012) The accessible chromatin landscape of the human genome. *Nature* 489(7414):75–82.
25. Smith EN, et al. (2011) Genome-wide association of bipolar disorder suggests an enrichment of replicable associations in regions near genes. *PLoS Genet* 7(6):e1002134.
26. Abecasis GR, et al. (2012) An integrated map of genetic variation from 1,092 human genomes. *Nature* 491(7422):56–65.
27. Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25(14):1754–1760.
28. DePristo MA, et al. (2011) A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat Genet* 43(5):491–498.
29. Li H, et al. (2009) The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25(16):2078–2079.
30. Wei Z, Wang W, Hu P, Lyon GJ, Hakonarson H (2011) SNVer: a statistical tool for variant calling in analysis of pooled or individual next-generation sequencing data. *Nucleic Acids Res* 39(19):e132.
31. Lee S, et al. (2012) Optimal unified approach for rare-variant association testing with application to small-sample case-control whole-exome sequencing studies. *Am J Hum Genet* 91(2):224–237.
32. Neale BM, et al. (2011) Testing for an unusual distribution of rare variants. *PLoS Genet* 7(3):e1001322.
33. Wang K, Li M, Hakonarson H (2010) ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res* 38(16):e164.
34. Boyle AP, et al. (2012) Annotation of functional variation in personal genomes using RegulomeDB. *Genome Res* 22(9):1790–1797.
35. Benjamini Y, Drai D, Elmer G (2001) Controlling the false discovery rate in behavior genetics research. *Behav Brain Res* 125(1-2):279-284.
36. Ashburner M, et al. (2000) Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet* 25(1):25–29.
37. Kanehisa M (2000) KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res* 28(1):27–30.
38. Akula N, et al. (2014) RNA-sequencing of the brain transcriptome implicates dysregulation of neuroplasticity, circadian rhythms and GTPase binding in bipolar disorder. *Mol Psychiatry* 19(11):1179–1185.
39. Ferreira MAR, et al. (2008) Collaborative genomed-wide association analysis supports a role for ANK3 and CACNA1C in bipolar disorder. *Nat Genet* 40(9):1056–1058.
40. Sklar P, et al. (2002) Family-based association study of 76 candidate genes in bipolar disorder: BDNF is a potential risk

locus. Brain-derived neutrophic factor. *Mol Psychiatry* 7(6):579–593.

41. Neves-Pereira M, et al. (2002) The brain-derived neurotrophic factor gene confers susceptibility to bipolar disorder: evidence from a family-based association study. *Am J Hum Genet* 71(3):651–655.