Published in final edited form as:

Gene. 2018 March 01; 645: 119–123. doi:10.1016/j.gene.2017.12.025.

Exome sequencing of a large family identifies potential candidate genes contributing risk to bipolar disorder

Tianxiao Zhang¹, Liping Hou², David T. Chen², Francis J. McMahon², Jen-Chyong Wang³, and John P. Rice^{3,*}

¹Department of Epidemiology and Biostatistics, School of Public Health, Xi'an Jiaotong University, Shaanxi, China

²Human Genetics Branch, The National Institute of Mental Health Intramural Research Program, National Institutes of Health, Bethesda, MD, USA

³Department of Psychiatry, Washington University in Saint Louis, MO, USA

Abstract

Bipolar disorder is a mental illness with lifetime prevalence of about 1%. Previous genetic studies have identified multiple chromosomal linkage regions and candidate genes that might be associated with bipolar disorder. The present study aimed to identify potential susceptibility variants for bipolar disorder using 6 related case samples from a four-generation family. A combination of exome sequencing and linkage analysis was performed to identify potential susceptibility variants for bipolar disorder. Our study identified a list of five potential candidate genes for bipolar disorder. Among these five genes, *GRID1* (Glutamate Receptor Delta-1 Subunit), which was previously reported to be associated with several psychiatric disorders and brain related traits, is particularly interesting. Variants with functional significance in this gene were identified from two cousins in our bipolar disorder pedigree. Our findings suggest a potential role for these genes and the related rare variants in the onset and development of bipolar disorder in this one family. Additional research is needed to replicate these findings and evaluate their patho-biological significance.

Keywords

Bipolar disorder; exome sequencing; linkage analysis; *GRID1*

*Corresponding authors: John P. Rice, Department of Psychiatry, Washington University School of Medicine at Washington University Medical Center, Campus Box 8134, 660 South Euclid Avenue, Saint Louis, Missouri, 63110-1093. john@zork.wustl.edu, Tel: +1 314-286-2572.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Conflicts of Interest

The authors declare that they have no conflict of interests.

Contributors

JPR and FJM designed the study. TZ and JPR wrote the main manuscript text. JCW collected the blood samples and extracted DNA. TZ, LH and DTC conducted the statistical analysis and prepared all the tables, figures and supplemental materials for this manuscript. All authors reviewed the manuscript.

Introduction

Bipolar disorder (BPD) is a mental illness with lifetime prevalence of about 1% (Merikangas *et al.*, 2007). BPD is characterized by periods of elevated mood (manic/hypomanic episodes) and periods of depression (depressive episodes) (Anderson, *et al.*, 2012). Currently, there is no cure for BPD, and medications and therapies are used to treat the symptoms. Patients with BPD and their families experience significant losses in functional status and quality of life, placing untoward stress on personal relationships. In addition, BPD is one of the most expensive mental health care diagnosis, both for patients with the illness and for their health insurance plans (Chatterton *et al.*, 2008), and that in turn adds a financial burden on the patients' families, as well as on society as a whole. Biomedical and etiological studies on the onset and development of BPD can throw light on new drug discovery and therapy development.

The etiology of BPD is not clearly understood but extensive research has indicated that both genetic and environmental factors play a role (Barnett *et al.*, 2009). Familial clustering studies have identified a ten-fold higher risk of BPD in people who have affected first degree relatives when compared to the general population (Barnett *et al.*, 2009). The heritability of overall bipolar spectrum disorders is estimated to be 0.71 (Edvardsen *et al.*, 2008). More than 40 linkage scans for BPD have been published and implicate many areas of the genome, although several studies have inconsistent results (Barnett *et al.*, 2009). Genome-wide association studies (GWAS) have identified several susceptibility loci including markers near *PALB2* (Wellcome Trust Case Control Consortium, 2007), *DGKH* (Baum *et al.*, 2008), *ANKK1* & *CACNA1C* (Ferreira *et al.*, 2008), 3p21 (McMahon *et al.*, 2010), *NCAN* (Cichon *et al.*, 2011), *ODZ4* (Psychiatric GWAS Consortium Bipolar Disorder Working Group, 2011), *TRANK1* & *LMAN2L* (Chen *et al.*, 2013), *ADCY2* & 6q16 (Mühleisen *et al.*, 2014), and *SESTD1* (Song *et al.*, 2015). Nevertheless, despite these findings, a recent study has estimated the SNP heritability (the proportion of variation in disease liability that is captured in GWAS by considering all SNPs simultaneously) was ~0.4 for BPD (Lee *et al.*, 2011). This finding indicates that further research is needed to unravel the genetic etiology for BPD.

The traditional microarray chip technology based GWAS focuses on common variants (genetic variants that have minor allele frequency at 5% or higher), and selectively omits the rare variants and structural variants such as short insertion and deletions (indels) due to technological problems (McCarthy *et al.*, 2008). The recent development of “next-generation” sequencing technology has enabled researchers to investigate these variants which are not covered in GWAS at a relatively lower genotyping cost (Mardis *et al.*, 2008; Zhang *et al.*, 2014). Exome sequencing, which sequences the exons of protein coding genes in the genome, is considered to be a powerful tool in genetic association research (Ng *et al.*, 2009). By focusing only on the region of exons, exome sequencing only sequences around 1% of the human genome (far less than whole genome sequencing) while investigating genomic regions of functional significance. A recent research project focusing on lithium-responsive bipolar disorder has identified several rare susceptibility variants by exome sequencing analysis based on 36 familial samples (Cruceanu *et al.*, 2013). This result

indicates that exome sequencing technology combined with proper study design is a promising method for association mapping.

In this project, we have recruited six related BPD cases from a single BPD extended family. The rationale is that a single variant is segregating in this large, unusual family and that this approach will minimize genetic heterogeneity by restricting analysis to a single family. We have performed whole exome sequencing on the six BPD cases, and examined the DNA variants that are shared among all these six cases. Additionally, we have also performed a genome-wide high density linkage analysis based on common SNP data. The linkage peak region has further narrowed down the potential susceptibility variants and genes for BPD.

Material and Methods

Samples

The BPD family was selected via a history of multiple relatives with BPD. An index case was recruited from department of psychiatry at Barnes-Jewish Hospital. Relatives were diagnosed via diagnostic interview (SADS-L) and two independent senior psychiatrists gave best estimated diagnosis made through consensus. Signed consent forms were obtained from all the recruited members. This study was approved by IRB of Washington University in St. Louis. The pedigree structure of the family recruited in this study was shown below (figure 1). Six subjects with BPD were included in this study.

Experimental methods

Genomic DNA was isolated from the peripheral blood leukocytes and the DNA was stored at -80°C for genotyping. The microarray genotyping was done by Illumina OmniExpress 700K SNP array. We applied quality control process to remove singletons and SNPs with missing rate higher than 10%. The whole-exome sequencing was carried out using Agilent SureSelect All Exon 50Mb Target Enrichment kit and on the SOLiD System by EdgeBio. The average read depth for the six bipolar subjects were 57x (supplemental table S1). For the exome-seq data, the alignment was done using novoalignCS (V1.01.15) by EdgeBio, and the data were recalled with HaplotypeCaller (GATK v3.3)(McKenna *et al.*, 2010). The QC was done by GATK's variant quality score recalibration (VQSR). Specifically, in this study variants with VQSLOD score smaller than 1.15 were filtered out (correspondence to filtering tranche 99.0 and above).

Statistical methods

We implemented parametric linkage analysis based on the software Merlin (Abecasis *et al.*, 2002) to identify potential linkage peak regions with the pedigree data. The allele frequency data were extracted from Hapmap 2 CEU samples. SLINK (Weeks *et al.*, 1990) was used to simulate the lineage analysis to obtain a LOD score threshold with acceptable statistical power. The potential effects of exonic SNVs were predicted using SIFT (Choi *et al.*, 2012) and Polyphen2 (Adzhubei *et al.*, 2010).

The redundancy of the microarray SNP chip panel enables us to implement a 10-set replicate analysis strategy in order to reduce the number of potential false positive signals obtained

from the linkage analysis. We 1) randomly selected 10 sets of SNPs with considerations of minor allele frequency (MAF > 0.3), Linkage disequilibrium (LD) structure (to account for LD) and their genomic coverage (each SNPs set contains around 10,000 markers); 2) conducted linkage analysis with each of these 10 replicated SNP sets independently; and 3) identified the peak regions that are repetitively identified in all of the ten rounds of analysis. To account for the potential impact of LD on linkage signals, we only chose one SNP from each LD block which was constructed based on the Hapmap European population data. The LD block was estimated using Plink (Purcell *et al.*, 2007).

We implemented the Perl based software Annovar (Wang *et al.*, 2010) to annotate the exome sequencing data. To investigate the potential susceptibility variants/genes within this large family, we extracted variants that 1) pass the quality control criteria specified in GATK software; 2) are under one of the linkage peak regions; 3) located within genetic region with functional significance (splicing site variants, non-synonymous SNVs, stop gain SNVs, frameshift indels, or non-frameshift indels within exonic regions) and 4) Only variants that were not recorded in 1000 genome database or variants with recorded MAF < 0.05 were included. We also incorporated our filtering results with R package RVsharing (Bureau *et al.*, 2014) to have a statistical estimate of our observed excess sharing among all the related BPD cases of those candidate variants.

Results

Results of linkage analysis

Simulation of the linkage analysis with SLINK showed that a LOD threshold 2.0 will only achieve 22% statistical power. To increase statistical power, we chose 1.8 as the LOD threshold with the cost of increasing the false positive rate. To control the number of potential false positive linkage signals, we implemented a 10-replicate linkage analysis strategy. Genotypes data were released for 733,202 SNPs with our Illumina microarray chip and 537,258 were left after quality control. The redundancy of these markers enabled us to implement a 10-replicate analysis strategy to reduce the number of the potential false positive linkage signals. This marker sets selection and analysis implementation was as indicated above.

The linkage analysis results of 10 replicate SNP sets are shown in table 1. As shown in table 1, four peak regions in total were identified from the linkage analysis. They are 5q33.1-5q33.3, 5q35.1-5q35.2, 10q21.1-10q21.2 and 10q23.1-10q23.33. Two of these 4 linkage peaks, 5q33.1-5q33.3 and 5q35.1-5q35.2 (figure S1) were only identified 3 and 1 times, respectively. This indicates that these two linkage peak signals may be false positive signals. The two consecutive peak regions on chromosome 10, 10q21.1-10q21.2 and 10q23.1-10q23.33 (10q22.3-10q23.33) were identified multiple times in the 10 replicate sets and 10q23.1-10q23.33 (10q22.3-10q23.33) was identified repetitively in all of the 10 replicate SNP sets. This result indicates that the chance that this peak region is a false positive signal is very low (figure S2). Here we provide a brief estimation for the false positive rate of this peak region. The LOD threshold we used here is 1.8, and this is approximately equal to p value < 0.002 (Nyholt *et al.*, 2000). Although we tested around 10,000 SNPs, considering the potential LD among SNPs, the independent number of tests

might be around 400 (this is a reasonable estimate because 400 is the number of tests when using microsatellite markers for linkage analysis). Therefore, the false positive rate for each replicate is around 0.55 ($1-0.998^{400}$). For a genomic region that is proved to be significant in all of the ten replicates, its false positive rate can be as low as 0.0025 (0.55^{10}). Therefore we concentrated those variants under this peak region (10q23.1- 10q23.33) on chromosome 10. We have summarized the results of linkage analyses of each marker set in table 1.

Results of variants filtering

Genotypic data were released for 140,814 variants (supplemental table S2). Among the 140,814 variants, 114,432 (81.26%) variants passed the quality control specified in GATK software. Within these high quality variant calls, 444 variants were located under the linkage peak region of chromosome 10. The number of variants was reduced to 60 if we only considered those variants with potential functional significance (supplemental table S3). Among these 60 variants, a total of 15 variants were not recorded in 1000 genome database or were rare variants with MAF less than 0.05 according to 1000 genome data in Caucasian population. Nine out of these 15 variants were only identified in one of six bipolar subjects sequenced, and 6 variants were shared by 2, 3 or 4 individuals. We have summarized the information of these 6 variants in table 2 (The co-segregation pattern of these 6 mutations across the pedigree was shown in supplemental figure S3). As we can see from the figure of co-segregation pattern, our six BPD cases can be classified into two sub-groups: individuals 142 and 5 are related to mutations (polymorphisms) of *GRID1* and *CDHR1* while the other four individuals, 212,220,356 and 32 seem to be related with mutations of *DYDC2*, *GHITM* and *MINPP1*. Also, from this figure we can identify a clear pattern of interaction (linkage) among these mutations. Mutations of *DYDC2* and *GHITM* were presented together while mutations of *CDHR1* and *GRID1* were presented together in our study subjects. Tests of RVSharing indicated that three out of these six variants (shared among 3 or 4 patients) were statistically significant with *P* value <0.008 (0.05/6). The other three variants that were shared among two patients had a *P* value of 0.0695. These 6 variants come from 5 genes including *DYDC2*, *GHITM*, *MINPP1*, *CDHR1* and *GRID1*. Two SNVs located in the genes *CDHR1* and *GRID1* were predicted to be “damaging” or “possibly damaging” by SIFT and Polyphen-2.

Discussions

Our aim was to identify potential susceptibility variants that contribute to the risk of BPD. In order to minimize genetic heterogeneity, we restricted analysis to a single large family in which several distantly-related individuals suffer from BPD. If a single variant affecting risk for BPD is segregating in this multiply affected family, this approach should identify the variant within the larger set of segregating variants, nominating candidate genes for future studies.

We narrowed the targeted genomic region down to a ~15 MB region by restricting the analysis to a linkage peak identified on chromosome 10. This region (10q22) has been previously reported to be linked with BPD in a large scale linkage study by Fallin et al (Fallin *et al.*, 2004). After filtering for minor allele frequency, sharing among affected

relatives, and functional significance of the potential susceptibility variants, we identified a list of 6 variants within 5 genes. A total of 3 variants in 2 genes, *GRID1* and *CDHR1*, were identified in 2 of the 6 BPD cases. The two cases that shared these 3 variants are also first cousins within this four generation pedigree (individuals #142 and #5 from Figure 1). *GRID1* encodes a subunit of glutamate receptor channels. These channels mediate most of the fast excitatory synaptic transmission in the central nervous system and play key roles in synaptic plasticity (Yamazaki *et al.*, 1992). *GRID1* has been widely investigated in multiple psychiatric disorders and brain related traits (Fallin *et al.*, 2005; Guo *et al.*, 2007; Venken *et al.*, 2008; Nenadic *et al.*, 2012). It was first reported by Fallin *et al.* that *GRID1* was significantly associated with schizophrenia and associated with BPD with suggestive significance among Ashkenazi Jewish case-parent trios (Fallin *et al.*, 2005). A study of mice in which the *GRID1* homologue, *GluD1*, was knocked out reported that the mice were hyperactive, manifested lower anxiety-like and depression-like behavior, and robust aggression (Yadav *et al.*, 2012). The two rare *GRID1* variants we report here (rs2306265 and rs3812645) represent the first evidence that rare variants in *GRID1* may contribute to BPD.

The other gene with a shared rare variant, *CDHR1*, belongs to the cadherin superfamily of calcium-dependent cell adhesion molecules. Its encoded protein is a photoreceptor-specific cadherin that plays a role in outer segment disc morphogenesis (Ba-Abbad *et al.*, 2013). Mutations in this gene have been associated with recessive retinal degeneration (Duncan *et al.*, 2012) and autosomal recessive cone-rod dystrophy (Ostergaard *et al.*, 2010). However, no previous study has linked this gene to any psychiatric disorders. We also identified 3 interesting variants in *DYDC2*, *GHITM* and *MINPP1*. These variants were predicted to be “benign” or “tolerated” and might not affect the protein structure. However, all of these variants were shared by 3 or 4 BPD cases in this pedigree, and their sharing patterns were significant after Bonferroni correction using the RVsharing algorithm.

We note that, given our sample size, these *P* values can only be used as a suggestive guidance when prioritizing variants for further study. One major strength of our study is that as an exome sequencing based study, we can examine both common and rare variants. Previous association studies on BPD mainly focus on common SNPs while ignoring most of the low frequency and rare variants. Restricting to common SNP, may hinder the ability to find susceptibility variants or genes. In our study, we considered both rare and common functional variants via exome sequencing technology. In addition, an advantage of our family based study design is that rare variants that segregate with reasonably high penetrance in an extended pedigree can provide a linkage signal helpful in identifying the susceptibility genes.

There are also several limitations of our study. A main limitation is that we lack familial controls. The variant sharing we utilized as a filtering strategy in our study might generate some false positive signals due to relatives sharing neutral DNA variations. Having exome sequencing data from several healthy family members of this pedigree might further narrow down our candidate gene list. It is still too early for us to make any conclusion on the potential role played by these candidate genes in the onset and development of BPD. Limited by our sample type and study design, further research is needed to replicate our

finding in unrelated individuals. More research is needed to reveal the potential relationship of our candidate gene list and biological mechanisms of BPD.

Conclusion

In summary, our study identified a list of 5 potential candidate genes for BPD based on exome sequencing in a large bipolar disorder pedigree. Among these 5 genes, *GRIDI* has been reported to be associated with several psychiatric and brain related traits in common SNP-based studies. Our results provide evidence linking rare variations in *GRIDI* with BPD. These findings suggest a potential role for these genes in the risk for BPD, but require replication in large, independent studies.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The authors want to thank for the participation of the subjects and their families in this study.

Source of Funding

This study was funded by NIMH Center for Collaborative Genetic Studies-Data Analysis Core (U24 MH068457) and by the NIMH Intramural Research Program.

References

- Abecasis GR, Cherny SS, Cookson WO, Cardon LR. Merlin--rapid analysis of dense genetic maps using sparse gene flow trees. *Nat Genet.* 2002; 30:97–101. [PubMed: 11731797]
- Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, et al. A method and server for predicting damaging missense mutations. *Nat Methods.* 2010; 7(4):248–249. [PubMed: 20354512]
- Anderson IM, Haddad PM, Scott J. Bipolar disorder. *BMJ.* 2012; 345:e8508. [PubMed: 23271744]
- Ba-Abbad R, Sergouniotis PI, Plagnol V, Robson AG, Michaelides M, Holder GE, et al. Clinical characteristics of early retinal disease due to CDHR1 mutation. *Mol Vis.* 2013; 19:2250–2249. [PubMed: 24265541]
- Barnett JH, Smoller JW. The genetics of bipolar disorder. *Neuroscience.* 2009; 164:331–343. [PubMed: 19358880]
- Baum AE, Akula N, Cabanero M, Cardona I, Corona W, Klemens B, et al. A genome-wide association study implicates diacylglycerol kinase eta (DGKH) and several other genes in the etiology of bipolar disorder. *Mol Psychiatry.* 2008; 13(2):197–207. [PubMed: 17486107]
- Bureau A, Younkin SG, Parker MM, Bailey-Wilson JE, Marazita ML, Murray JC, et al. Inferring rare disease risk variants based on exact probabilities of sharing by multiple affected relatives. *Bioinformatics.* 2014; 30(15):2189–2196. [PubMed: 24740360]
- Chatterton ML, Ke X, Lewis BE, Rajagopalan K, Lazarus A. Impact of bipolar disorder on the family: utilization and cost of health care resources. *P T.* 2008; 33(1):15–34. [PubMed: 19749991]
- Chen DT, Jiang X, Akula N, Shugart YY, Wendland JR, Steele CJ, et al. Genome-wide association study meta-analysis of European and Asian-ancestry samples identifies three novel loci associated with bipolar disorder. *Mol Psychiatry.* 2013; 18(2):195–205. [PubMed: 22182935]
- Choi Y, Sims GE, Murphy S, Miller JR, Chan AP. Predicting the functional effect of amino acid substitutions and indels. *PLoS One.* 2012; 7(10):e46688. [PubMed: 23056405]

- Cichon S, Mühleisen TW, Degenhardt FA, Mattheisen M, Miró X, Strohmaier J, et al. Genome-wide association study identifies genetic variation in neurocan as a susceptibility factor for bipolar disorder. *Am J Hum Genet.* 2011; 88(3):372–81. [PubMed: 21353194]
- Cruceanu C, Ambalavanan A, Spiegelman D, Gauthier J, Lafrenière RG, Dion PA, et al. Family-based exome-sequencing approach identifies rare susceptibility variants for lithium-responsive bipolar disorder. *Genome.* 2013; 56:634–640. [PubMed: 24237345]
- Duncan JL, Roorda A, Navani M, Vishweswaraiah S, Syed R, Soudry S, et al. Identification of a novel mutation in the CDHR1 gene in a family with recessive retinal degeneration. *Arch Ophthalmol.* 2012; 130(10):1301–1308. [PubMed: 23044944]
- Edvardsen J, Torgersen S, Røysamb E, Lygren S, Skre I, Onstad S, et al. Heritability of bipolar spectrum disorders. Unity or heterogeneity? *J Affect Disord.* 2008; 106:229–240. [PubMed: 17692389]
- Fallin MD, Lasseter VK, Avramopoulos D, Nicodemus KK, Wolyniec PS, McGrath JA, et al. Bipolar I disorder and schizophrenia: a 440-single-nucleotide polymorphism screen of 64 candidate genes among Ashkenazi Jewish case-parent trios. *Am J Hum Genet.* 2005; 77(6):918–936. [PubMed: 16380905]
- Fallin MD, Lasseter VK, Wolyniec PS, McGrath JA, Nestadt G, Valle D, et al. Genomewide linkage scan for bipolar-disorder susceptibility loci among Ashkenazi Jewish families. *Am J Hum Genet.* 2004; 75(2):204–219. [PubMed: 15208783]
- Ferreira MA, O'Donovan MC, Meng YA, Jones IR, Ruderfer DM, Jones L, et al. Collaborative genome-wide association analysis supports a role for ANK3 and CACNA1C in bipolar disorder. *Nat Genet.* 2008; 40(9):1056–8. [PubMed: 18711365]
- Guo SZ, Huang K, Shi YY, Tang W, Zhou J, Feng GY, et al. A case-control association study between the GRID1 gene and schizophrenia in the Chinese Northern Han population. *Schizophr Res.* 2007; 93(1–3):385–390. [PubMed: 17490860]
- Lee SH, Wray NR, Goddard ME, Visscher PM. Estimating missing heritability for disease from genome-wide association studies. *Am J Hum Genet.* 2011; 88:294–305. [PubMed: 21376301]
- Mardis ER. Next-generation DNA sequencing methods. *Annu Rev Genomics Hum Genet.* 2008; 9:387–402. [PubMed: 18576944]
- McCarthy MI, Hirschhorn JN. Genome-wide association studies: past, present and future. *Hum Mol Genet.* 2008; 17:R100–1. [PubMed: 18852196]
- McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytzky A, et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* 2010; 20:1297–1303. [PubMed: 20644199]
- McMahon FJ, Akula N, Schulze TG, Muglia P, Tozzi F, Detera-Wadleigh SD, et al. Meta-analysis of genome-wide association data identifies a risk locus for major mood disorders on 3p21.1. *Nat Genet.* 2010; 42(2):128–31. [PubMed: 20081856]
- Merikangas KR, Akiskal HS, Angst J, Greenberg PE, Hirschfeld RM, Petukhova M, et al. Lifetime and 12-month prevalence of bipolar spectrum disorder in the National Comorbidity Survey replication. *Arch Gen Psychiatry.* 2007; 64:543–552. [PubMed: 17485606]
- Mühleisen TW, Leber M, Schulze TG, Strohmaier J, Degenhardt F, Treutlein J, et al. Genome-wide association study reveals two new risk loci for bipolar disorder. *Nat Commun.* 2014; 5:3339. [PubMed: 24618891]
- Nenadic I, Maitra R, Scherpiet S, Gaser C, Schultz CC, Schachtzabel C, et al. Glutamate receptor 6 1 (GRID1) genetic variation and brain structure in schizophrenia. *J Psychiatr Res.* 2012; 46(12):1531–1539. [PubMed: 23017809]
- Ng SB, Turner EH, Robertson PD, Flygare SD, Bigham AW, Lee C, et al. Targeted capture and massively parallel sequencing of 12 human exomes. *Nature.* 2009; 461:272–276. [PubMed: 19684571]
- Nyholt DR. All LODs are not created equal. *Am J Hum Genet.* 2000; 67(2):282–288. [PubMed: 10884360]
- Ostergaard E, Batbayli M, Duno M, Vilhelmsen K, Rosenberg T. Mutations in PCDH21 cause autosomal recessive cone-rod dystrophy. *J Med Genet.* 2010; 47(10):665–9. [PubMed: 20805371]

- Psychiatric GWAS Consortium Bipolar Disorder Working Group. Large-scale genome-wide association analysis of bipolar disorder identifies a new susceptibility locus near ODZ4. *Nat Genet.* 2011; 43(10):977–83. [PubMed: 21926972]
- Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, et al. PLINK: a toolset for whole-genome association and population-based linkage analysis. *Am J Hum Genet.* 2007; 81(3): 559–575. [PubMed: 17701901]
- Song J, Bergen SE, Di Florio A, Karlsson R, Charney A, Ruderfer DM, et al. Genome-wide association study identifies SESTD1 as a novel risk gene for lithium-responsive bipolar disorder. *Mol Psychiatry.* 2015 [Epub ahead of print].
- Venken T, Alaerts M, Souery D, Goossens D, Sluijs S, Navon R, et al. Chromosome 10q harbors a susceptibility locus for bipolar disorder in Ashkenazi Jewish families. *Mol Psychiatry.* 2008; 13(4): 442–50. [PubMed: 17579605]
- Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res.* 2010; 38:e164. [PubMed: 20601685]
- Weeks DE, Ott JLG. SLINK: a general simulation program for linkage analysis. *Am J Hum Genet.* 1990; 47:A204.
- Wellcome Trust Case Control Consortium. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature.* 2007; 447(7145):661–678. [PubMed: 17554300]
- Yadav R, Gupta SC, Hillman BG, Bhatt JM, Stairs DJ, Dravid SM. Deletion of glutamate delta-1 receptor in mouse leads to aberrant emotional and social behaviors. *PLoS One.* 2012; 7(3):e32969. [PubMed: 22412961]
- Yamazaki M, Araki K, Shibata A, Mishina M. Molecular cloning of a cDNA encoding a novel member of the mouse glutamate receptor channel family. *Biochem Biophys Res Commun.* 1992; 183(2): 886–892. [PubMed: 1372507]
- Zhang TX, Xie YR, Rice JP. Application of noncollapsing methods to the gene-based association test: a comparison study using Genetic Analysis Workshop 18 data. *BMC Proceedings.* 2014; 8(Suppl 1):S53. [PubMed: 25519333]

Abbreviations

BPD	bipolar disorder
GWAS	genome-wide association study
LD	linkage disequilibrium
MAF	minor allele frequency
QC	quality control
SNPs	single nucleotide polymorphisms
SNVs	single nucleotide variants
VQSR	variant quality score recalibration

Highlights

- Exome sequencing were performed on 6 bipolar cases from a four-generation pedigree
- A suggestive linkage peak region was identified on 10q23.1-10q23.33
- 5 candidate genes including *DYDC2*, *GHITM*, *CDHR1*, *GRID1* and *MINPP1* were identified

Figure 1. Pedigree of the BPD family sequenced in this study

Blood samples of the 6 BPD cases at the bottom of the pedigree were collected.

Table 1

Linkage analysis results of the ten replicate sets.

Set	# of markers	Chromosome Regions	LOD
1	10,396	10q22.3-10q23.33	1.92
2	9,648	10q21.1-10q21.2	1.91
		10q22.3-10q23.33	1.92
3	9,963	10q21.1-10q21.3	1.91
		10q22.3-10q23.33	1.93
4	10,058	5q33.1-5q33.2	1.88
		10q21.1-q22.3	1.91
		10q23.1-10q23.33	1.92
5	10,407	10q21.1-10q21.2	1.91
		10q23.1-10q23.33	1.92
6	9,715	5q33.1- 5q 33.3	1.9
		5q35.1-5q35.2	2.77
		10q23.1-10q23.33	1.92
7	10,418	10q21.1-10q21.2	1.91
		10q23.1-10q23.33	1.92
8	10,402	5q33.1- 5q 33.2	1.9
		10q21.1-10q21.2	1.91
		10q23.1-10q23.33	1.92
9	10,236	10q21.1-10q21.2	1.87
		10q23.1-10q23.33	1.92
10	10,402	10q21.1-10q21.2	1.9
		10q23.1-10q23.33	1.92

The most significant linkage peak region was found on chromosome region 10q22.3- 10q23.33 with LOD score of 1.926. The LOD threshold is 1.8. We have highlighted the 10q23.1-10q23.33 which were identified to be significant in all ten SNP replicates.

Table 2

Summary information for 6 genetic variants identified after applying the filtering strategy.

CHR	SNP	Position	Ref	Alt	Gene	Function	AAChange	SIFT*	Polyphen2**	Sharing	P***
10	rs36027713	82126541	C	G	DYDC2	nonsynonymous SNV	p.P123R	T	B	4	0.0003
10	-	85902497	A	T	GHITM	nonsynonymous SNV	p.E72D	T	B	4	0.0003
10	rs45584033	85974231	C	T	CDHRI	nonsynonymous SNV	p.P812S	D	D	2	0.0695
10	rs2306265	87484382	C	T	GRIDI	nonsynonymous SNV	p.V529I	D	D	2	0.0695
10	rs3812645	87489317	T	C	GRIDI	nonsynonymous SNV	p.M430V	T	B	2	0.0695
10	-	89280872	C	T	MINPPI	nonsynonymous SNV	p.T137I	T	B	3	0.003

* SIFT prediction. T stands for tolerated and D stands for damaging.
** Polyphen2 prediction. B stands for benign, P stands for possibly damaging, and D stands for probably damaging.
*** Pvalues here stands for the probability of observing the sharing pattern in our pedigree. Significant variants after applying Bonferroni correction were shown in bold.