nature COMMUNICATIONS

ARTICLE

Received 14 Aug 2013 | Accepted 29 Jan 2014 | Published 11 Mar 2014

DOI: 10.1038/ncomms4339

Genome-wide association study reveals two new risk loci for bipolar disorder

Thomas W. Mühleisen[1,2,3,*], Markus Leber[4,5,*], Thomas G. Schulze[6,*], Jana Strohmaier[7], Franziska Degenhardt[1,2], Jens Treutlein[7], Manuel Mattheisen[8,9], Andreas J. Forstner[1,2], Johannes Schumacher[1,2], René Breuer[7], Sandra Meier[7,10], Stefan Herms[1,2,11], Per Hoffmann[1,2,3,11], André Lacour[5], Stephanie H. Witt[7], Andreas Reif[12], Bertram Müller-Myhsok[13,14,15], Susanne Lucae[13], Wolfgang Maier[16], Markus Schwarz[17], Helmut Vedder[17], Jutta Kammerer-Ciernioch[17], Andrea Pfennig[18], Michael Bauer[18], Martin Hautzinger[19], Susanne Moebus[20], Lutz Priebe[1,2], Piotr M. Czerski[21], Joanna Hauser[21], Jolanta Lissowska[22], Neonila Szeszenia-Dabrowska[23], Paul Brennan[24], James D. McKay[25], Adam Wright[26,27], Philip B. Mitchell[26,27], Janice M. Fullerton[28,29], Peter R. Schofield[28,29], Grant W. Montgomery[30], Sarah E. Medland[30], Scott D. Gordon[30], Nicholas G. Martin[30], Valery Krasnow[31], Alexander Chuchalin[32], Gulja Babadjanova[32], Galina Pantelejeva[33], Lilia I. Abramova[33], Alexander S. Tiganov[33], Alexey Polonikov[34], Elza Khusnutdinova[35], Martin Alda[36,37], Paul Grof[37,38,39], Guy A. Rouleau[40], Gustavo Turecki[41], Catherine Laprise[42], Fabio Rivas[43], Fermin Mayoral[43], Manolis Kogevinas[44], Maria Grigoroiu-Serbanescu[45], Peter Propping[1], Tim Becker[5,4], Marcella Rietschel[7,*], Markus M. Nöthen[1,2,*] & Sven Cichon[1,2,3,11,*]

Bipolar disorder (BD) is a common and highly heritable mental illness and genome-wide association studies (GWAS) have robustly identified the first common genetic variants involved in disease aetiology. The data also provide strong evidence for the presence of multiple additional risk loci, each contributing a relatively small effect to BD susceptibility. Large samples are necessary to detect these risk loci. Here we present results from the largest BD GWAS to date by investigating 2.3 million single-nucleotide polymorphisms (SNPs) in a sample of 24,025 patients and controls. We detect 56 genome-wide significant SNPs in five chromosomal regions including previously reported risk loci *ANK3*, *ODZ4* and *TRANK1,* as well as the risk locus *ADCY2* (5p15.31) and a region between *MIR2113* and *POU3F2* (6q16.1). *ADCY2* is a key enzyme in cAMP signalling and our finding provides new insights into the biological mechanisms involved in the development of BD.

[1] Institute of Human Genetics, University of Bonn, D-53127 Bonn, Germany. [2] Department of Genomics, Life & Brain Center, University of Bonn, D-53127 Bonn, Germany. [3] Institute of Neuroscience and Medicine (INM-1), Research Centre Jülich, D-52425 Jülich, Germany. [4] Institute for Medical Biometry, Informatics, and Epidemiology, University of Bonn, D-53127 Bonn, Germany. [5] German Center for Neurodegenerative Diseases (DZNE), Bonn, Germany, D-53175 Bonn, Germany. [6] Department of Psychiatry and Psychotherapy, University of Göttingen, D-37075 Göttingen, Germany. [7] Department of Genetic Epidemiology in Psychiatry, Central Institute of Mental Health, Medical Faculty Mannheim/ Heidelberg University, D-68159 Mannheim, Germany. [8] Department of Biomedicine, Aarhus University, DK-8000 Aarhus C, Denmark. [9] Institute for Genomic Mathematics, University of Bonn, D-53127 Bonn, Germany. [10] National Centre Register-Based Research, Aarhus University, DK-8210 Aarhus V, Denmark. [11] Division of Medical Genetics, Department of Biomedicine, University of Basel, Basel CH-4012, Switzerland. [12] Department of Psychiatry, Psychosomatics and Psychotherapy, University of Würzburg, D-97070 Würzburg, Germany. [13] Statistical Genetics, Department of Translational Psychiatry, Max Planck Institute of Psychiatry, D-80804 Munich, Germany. [14] Munich Cluster for Systems Neurology (SyNergy), D-80336 Munich, Germany. [15] Institute of Translational Medicine, University of Liverpool, L69 3BX Liverpool, UK. [16] Department of Psychiatry, University of Bonn, D-53127 Bonn, Germany. [17] Psychiatric Center Nordbaden, D-69168 Wiesloch, Germany. [18] Department of Psychiatry and Psychotherapy, University Hospital, D-01307 Dresden, Germany. [19] Department of Psychology, Clinical Psychology and Psychotherapy, Eberhard Karls University Tübingen, D-72074 Tübingen, Germany. [20] Institute of Medical Informatics, Biometry, and Epidemiology, University Duisburg-Essen, D-45147 Essen, Germany. [21] Department of Psychiatry, Poznan University of Medical Sciences, Poznan PL-60-572, Poland. [22] Department of Cancer Epidemiology and Prevention, Maria Sklodowska-Curie Memorial Cancer Centre and Institute of Oncology, Warsaw PL-02-781, Poland. [23] Department of Epidemiology, Nofer Institute of Occupational Medicine, Lodz PL-91-348, Poland. [24] Genetic Epidemiology Group, International Agency for Research on Cancer (IARC), 69372 Lyon CEDEX 08, France. [25] Genetic Cancer Susceptibility Group, International Agency for Research on Cancer (IARC), 69372 Lyon CEDEX 08, France. [26] School of Psychiatry, University of New South Wales, Randwick, New South Wales 2052, Australia. [27] Black Dog Institute, Prince of Wales Hospital, Randwick, New South Wales 2031, Australia. [28] Neuroscience Research Australia, Randwick, Sydney, New South Wales 2031, Australia. [29] School of Medical Sciences, Faculty of Medicine, University of New South Wales, Sydney, New South Wales 2052, Australia. [30] Queensland Institute of Medical Research (QIMR), Brisbane, Queensland 4006, Australia. [31] Moscow Research Institute of Psychiatry, Moscow 107258, Russian Federation. [32] Institute of Pulmonology, Russian State Medical University, Moscow 105077, Russian Federation. [33] Russian Academy of Medical Sciences, Mental Health Research Center, Moscow 115522, Russian Federation. [34] Department of Biology, Medical Genetics and Ecology, Kursk State Medical University, Kursk 305041, Russian Federation. [35] Institute of Biochemistry and Genetics, Ufa Scientific Center of Russian Academy of Sciences, Ufa 450054, Russian Federation. [36] Department of Psychiatry, Dalhousie University, Halifax, Nova Scotia, Canada B3H 2E2. [37] Mood Disorders Center of Ottawa, Ottawa, Ontario, Canada K1G 4G3. [38] Department of Psychiatry, University of Toronto, Toronto, Ontario, Canada M5T 1R8. [39] The International Group for the Study of Lithium-Treated Patients (IGSLI), Berlin, Germany. [40] Department of Neurology and Neurosurgery, Montreal Neurological Hospital and Institute, McGill University, Montreal, Québec, Canada H3G 1A4. [41] Department of Psychiatry, Douglas Hospital Research Institute, McGill University, Montreal, Quebec, Canada H4H 1R3. [42] Département des sciences fondamentales, Université du Québec à Chicoutimi (UQAC), Saguenay, Canada G7H 2B1. [43] Department of Psychiatry, Hospital Regional Universitario Carlos Haya, Malaga 29009, Spain. [44] Center for Research in Environmental Epidemiology (CREAL), Barcelona 08003, Spain. [45] Biometric Psychiatric Genetics Research Unit, Alexandru Obregia Clinical Psychiatric Hospital, Bucharest RO-041914, Romania. * These authors contributed equally to this work. Correspondence and requests for materials should be addressed to S.C. (email: sven.cichon@unibas.ch).

Bipolar disorder (BD) is a severe disorder of mood, characterized by recurrent episodes of mania and depression, which affect thought, perception, emotion and social behaviour. With a lifetime prevalence of 1% in the general population, BD is a common condition. The World Health Organisation classifies BD as one of the top 10 leading causes of the global burden of disease for the age group of 15–44-year-old people. Formal and molecular genetic data strongly suggest that BD is a multifactorial disease[1]. This means that many genetic and environmental factors influence the risk of disease. The heritability estimates for BD range between 60 and 80%[2]. This suggests a substantial involvement of genetic factors in the development of the disease, but the particular factors underlying the pathophysiology and aetiology of BD are still largely unknown.

Since the first genome-wide association study (GWAS) of BD[3] in 2007, a handful of risk loci have been identified through some larger GWAS, which replicated in adequately sized follow-up studies[4,5], notably *ANK3*, *NCAN*, *CACNA1C* and *ODZ4*. These first genes provide valuable insights into the molecular mechanisms involved in BD and form a basis for future exploration of the molecular pathomechanisms and possibly drug targets. The genetic risk variants at these loci show unprecedented statistical support, but explain only a smaller fraction of BD's heritability. However, this is in line with results of recent studies that provide evidence for a strong polygenic component in BD, suggesting the presence of a large number of additional risk loci, each mediating small disease susceptibility[6,7]. Motivated by this and the successes of GWAS for non-psychiatric phenotypes, a big step forward was taken through a recent large BD GWAS by the multi-national Psychiatric Genomics Consortium Bipolar Disorder Working Group (PGC-BD)[7]. We have previously published a GWAS of BD using a German sample (named Germany I)[8] that was thereafter included in the GWAS of the PGC-BD[7] and another subsequent BD GWAS[9].

Here we generate new and so far unpublished genome-wide SNP data from 2,266 clinically well-characterized patients with BD (~76% BD type I) and 5,028 ethnically matched controls, derived from the MooDS (Systematic Investigation of the Molecular Causes of Major Mood Disorders and Schizophrenia) consortium. The samples originate from four European countries, Canada and Australia. To increase power for detection of risk variants with small genetic effect, we combine and jointly analyse our MooDS samples with the 7,481 patients and 9,250 controls from the aforementioned PGC-BD study (MooDS-PGC). We find strong evidence for common risk variants at three known loci (*ANK3*, *ODZ4* and *TRANK1*) and identify two risk loci (*ADCY2* and the region between *MIR2113* and *POU3F2*), which have not been implicated in BD before.

Results

Association analyses. A total of 2,267,487 imputed SNPs from 9,747 patients and 14,278 controls (Table 1) passed our stringent quality control (QC) and association analysis of autosomal SNPs was performed using a fixed-effects meta-analysis (Methods). Sex chromosomes were not analysed because the published PGC data do not contain the respective information. To adequately correct for inflation of *P* values due to varying sizes of the MooDS and PGC samples, we adjusted the *P* values of the MooDS-PGC analysis (P_{GC}) using a standardized λ factor for genomic control (GC), which was 1.018 (Methods, Fig. 1). A *P* value plot showing an overview of the genome-wide association results is provided in Fig. 2a.

The MooDS-PGC GWAS revealed 56 SNPs reaching genome-wide significance at five genomic loci, three of which had been described before (*ANK3*, *ODZ4* and *TRANK1*)[4,9]. A subset of 18 non-correlated top SNPs, which were selected by a pruning on pairwise linkage disequilibrium (LD) between the associated SNPs, is provided in Table 2. Full lists of SNPs with evidence of strong-to-moderate association ($P_{GC} < 5 \times 10^{-5}$) are given in Supplementary Tables 1–5.

Overall, the most significant association signal and the largest number of genome-wide significant SNPs ($n = 26$) were identified in *ANK3* (ankyrin 3) on chromosome 10q21.2 (top: rs10994415-C, $P_{GC} = 6.88 \times 10^{-11}$, OR = 1.27). SNP rs10994415 is in moderate LD with rs10994397 ($r^2 = 0.62$), which is the most significant finding for *ANK3* in the original PGC-BD study[7]

Table 1 | Phenotypic data for patients with bipolar disorder and controls following QC.

Sample	Ancestry	Patients (*N*)	Controls (*N*)	BD1 (in %)	BD2 (in %)	SAB (in %)	BD-NOS (in %)	MaD (in %)	Diagnosis	Interview	Controls screened (in %)
MooDS											
Australia	European	330	1,811	254 (76.97)	74 (22.42)	1 (0.03)	1 (0.03)	0	DSM-IV	DIGS, FIGS, SCID	No
Germany III	German	490	880	249 (50.82)	123 (24.80)	0	118 (24.08)	0	DSM-IV	AMDP, CID-S, SADS-L, SCID	Yes (100%)
Poland	Polish	411	689	316 (76.89)	95 (23.11)	0	0	0	DSM-IV	SCID	No
Canada	European	285	341	198 (65.26)	82 (27.37)	5 (1.75)	0	0	DSM-IV, RDC	SADS-L	No
Russia	Russian	316	486	315 (91.77)	1 (0.32)	0	0	0	DSM-IV	SCID	No
Germany II	German	181	527	176 (97.24)	5 (2.76)	0	0	0	DSM-IV	AMDP	No
Spain	Spanish	253	294	241 (95.26)	12 (4.72)	0	0	0	DSM-IV	SADS-L	No
Total		2,266	5,028	1,712	388	6	160	0	—	—	—
PGC											
Primary GWAS	European, US-American	7,481	9,250	6,289 (84.01)	824 (11.01)	263 (3.52)	105 (1.40)	0	DSM-IIR, DSM-IV, RDC	DIGS, MINI, SADS, SADS-L, SCAN, SCID	Yes (32.95)
Grand Total	—	9,747	14,278	8,001	1,212	269	265	0	—	—	

AMDP, Association for Methodology and Documentation in Psychiatry[55]; BD1, bipolar disorder type 1; BD2, bipolar disorder type 2; BD-NOS, bipolar disorder not otherwise specified; CID-S, Composite International Diagnostic Screener[56]; DIGS, Diagnostic Interview for Genetic Studies[57]; DSM-IIR/DSM-IV, Diagnostic and Statistical Manual of Mental Disorders[33]; FIGS, Family Interview for Genetic Studies[58]; MaD, Manic disorder according to RDC; MINI, Mini-International Neuropsychiatric Interview[59]; N, number of subjects; PGC, Psychiatric Genomics Consortium; RDC, Research Diagnostic Criteria[60]; SAB, schizoaffective disorder (bipolar type); SADS/SADS-L, Schedule for Affective Disorders and Schizophrenia[61]; SCAN, Schedules for Clinical Assessment in Neuropsychiatry[62]; SCID, Structured Clinical Interview for DSM disorders[35].
Further information about the phenotyping, genome-wide genotyping and QC for patients and controls are included in the Methods section and Supplementary Table 8.

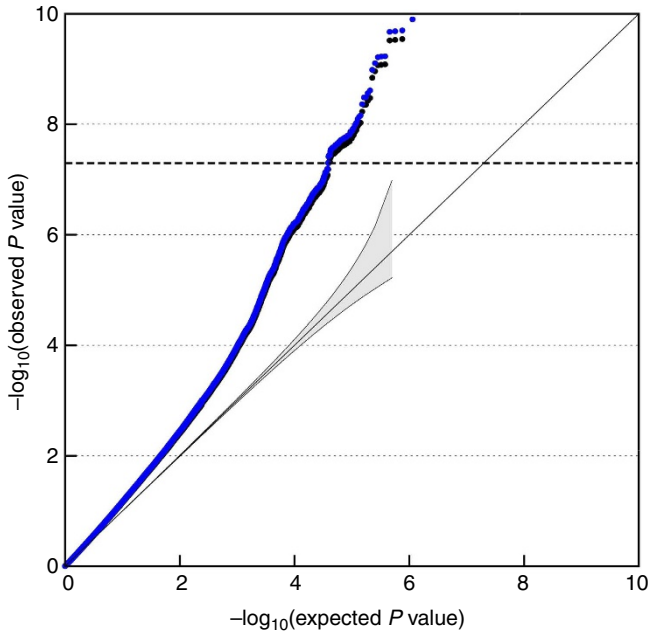

Figure 1 | Quantile-quantile plot of the MooDS-PGC GWAS of BD. A quantile–quantile (QQ) plot displays the relationship between the experimentally observed *P* values (vertical axis) to the expected *P* values of a null distribution (horizontal axis). The grey area represents the 95% concentration band. QQ plots are commonly used for graphical illustration of the effect of a QC on GWAS data. The dashed line (bold) indicates the threshold for genome-wide significance ($P_{GC} = 5 \times 10^{-8}$). Here we show the effect of a standardised λ factor for GC on the 2,267,487 autosomal *P* values from the fixed-effects meta-analysis (blue curve) in the last step of our QC procedure (Methods). After the application, a systematic adjustment of *P* values (P_{GC}, black curve) can be seen, suggesting that this QC step counteracts an inflation of *P* values by varying sample sizes in the MooDS and PGC data.

and the second most significant finding in this GWAS ($P_{GC} = 2.86 \times 10^{-10}$, OR = 1.29). Additionally, rs10994415 is in strong LD with rs1938526 ($r^2 = 0.79$) and in moderate LD with rs10994336 ($r^2 = 0.47$), both genome-wide significant ($P_{GC} = 8.55 \times 10^{-10}$, OR = 1.27; $P_{GC} = 2.34 \times 10^{-8}$, OR = 1.27) and originally found as BD risk variants by Ferreira *et al.*[10] Our second most significant locus was marked by 10 genome-wide significant SNPs in *ODZ4* (odd Oz/ten-m homologue 4 alias teneurin transmembrane protein 4, *TENM4*) on chromosome 11q14.1 (top: rs12290811-A, $P_{GC} = 1.09 \times 10^{-09}$, OR = 1.19). rs12290811 was first described by Ferreira *et al.*[10] with suggestive evidence for association with BD[10]. This SNP is in moderate LD with *ODZ4* rs12576775 ($r^2 = 0.57$, $P_{GC} = 4.46 \times 10^{-9}$, OR = 1.17), the second best finding of the PGC-BD study[7] with genome-wide significance (after their *CACNA1C* rs4765913 finding).

As the third and fourth most significant finding in our study, we discovered two previously undescribed risk loci for BD. On chromosome 5p15.31, two genome-wide significant SNPs are located in *ADCY2* (adenylate cyclase 2, top: rs17826816-G, $P_{GC} = 9.89 \times 10^{-9}$, OR = 1.14) and four were found in a region between the genes *MIR2113* (microRNA 2113) and *POU3F2* (POU class 3 homeobox 2) on chromosome 6q16.1 (top: rs12202969-A, $P_{GC} = 1.08 \times 10^{-8}$, OR = 1.12). *P* value plots that provide a regional overview of the associated SNPs as well as OR plots that show the effect sizes of the top variant in each subsample are provided in Figs 2b, and 3a,b. A comprehensive lookup of previous findings at these loci in the catalogue of published GWAS[11] (as of 4 October 2013) and a literature search did not provide any genome-wide significant GWAS finding for BD in windows of 1 Mb around the top SNPs.

Another known BD risk locus with genome-wide significance was *TRANK1* (tetratricopeptide repeat and ankyrin repeat containing 1 alias lupus Brain antigen 1 homologue, *LBA1*) on chromosome 3p22.2, which was the fifth best-supported locus in our study (top: rs6550435-G, $P_{GC} = 2.05 \times 10^{-8}$, OR = 1.13). rs9834970, a SNP in low LD ($r^2 = 0.13$, $P_{GC} = 4.81 \times 10^{-8}$, OR = 1.12), had already reached genome-wide significance in a GWAS of BD by Chen *et al.*[9], which partially overlaps with the included PGC data. Analysis of rs9834970-C in MooDS showed $P_{GC} = 0.0014$ (OR = 1.15), providing independent evidence for replication.

NCAN (neurocan) on chromosome 19p13.11 was identified as a genome-wide significant risk locus for BD in the first MooDS study[8]. In our present analysis, a SNP in strong LD with the original SNP rs1064395 ($r^2 = 0.96$) showed a sub-genome-wide significant result (rs2011503-C, $P_{GC} = 8.79 \times 10^{-8}$; OR = 0.87). *CACNA1C* (calcium channel, voltage-dependent, L type, alpha 1C subunit) on chromosome 12p13.33 was originally described as BD risk locus in the study by Ferreira *et al.*[10]; in fact, SNPs rs1006737 and rs1024582 were found at the border of genome-wide significance. *CACNA1C* received additional support in the PGC-BD study[7] and a subsequent study by Green *et al.*[12], in which rs4765913, a SNP with moderate LD to rs1006737 and rs1024582 ($r^2 = 0.40$ and $r^2 = 0.49$), was genome-wide significant. In the present MooDS-PGC analysis, rs4765913 was the most significant *CACNA1C* finding but failed to reach genome-wide significance (rs4765913-A, $P_{GC} = 9.69 \times 10^{-6}$, OR = 1.12). None of these three known *CACNA1C* risk variants was supported by the new MooDS data (rs4765913-A, $P_{MooDS} = 0.568$, OR = 1.03; rs1006737-A, $P_{MooDS} = 0.643$, OR = 1.02; rs1024582-A; $P_{MooDS} = 0.614$, OR = 1.02). Regional association plots for the known risk loci (*ANK3*, *ODZ4*, *TRANK1*, *NCAN* and *CACNA1C*) are provided in Supplementary Fig. 1a–e.

Follow-up analyses. To characterize our new findings *in silico*, we have looked up relevant databases for prior knowledge about the SNPs and genes. At both loci, we screened all SNPs with strong-to-moderate evidence for association ($P_{GC} < 5 \times 10^{-5}$, Supplementary Tables 3 and 4) for potentially functional effects such as missense variants with a consequence on the protein's primary structure or variants with an effect on expression in different body tissues including important brain regions. To broaden the search range in data from prior studies or databases that did not contain the original SNP from our analysis, we have also included proxy SNPs that are in strong LD with the analysed GWAS SNPs ($r^2 > 0.8$, Supplementary Tables 6 and 7).

For the 5p15.31 locus, the top SNP rs17826816 is located in intron 2 of the largest protein-coding transcript of *ADCY2* (RefSeq NM_020546.2, Ensembl ENSG00000078295). This transcript (Ensembl ENST00000338316) has a length of 6,575 base pairs and contains 25 exons that encode 1,091 amino-acid residues (UniProt Q08462, Ensembl ENSP00000342952). rs13166360 ($P_{GC} = 1.81 \times 10^{-8}$; OR = 1.14), a proxy for rs17826816 ($r^2 = 0.95$) and the other genome-wide significant SNP in the gene, is located in exon 3 where it mediates an amino-acid change of the nonpolar residue valine to the hydrophobic residue leucine at position 147. rs13166360 impacts on the fourth α-helix of the first transmembrane domain. This missense variant has been predicted as possibly damaging for the protein by a HumDiv data-trained PolyPhen-2 (ref. 13) analysis.

For the 6q16.1 locus, the associated variants that are located closest to the neighbouring genes are rs9385269 ($P_{GC} = 8.74 \times 10^{-7}$,

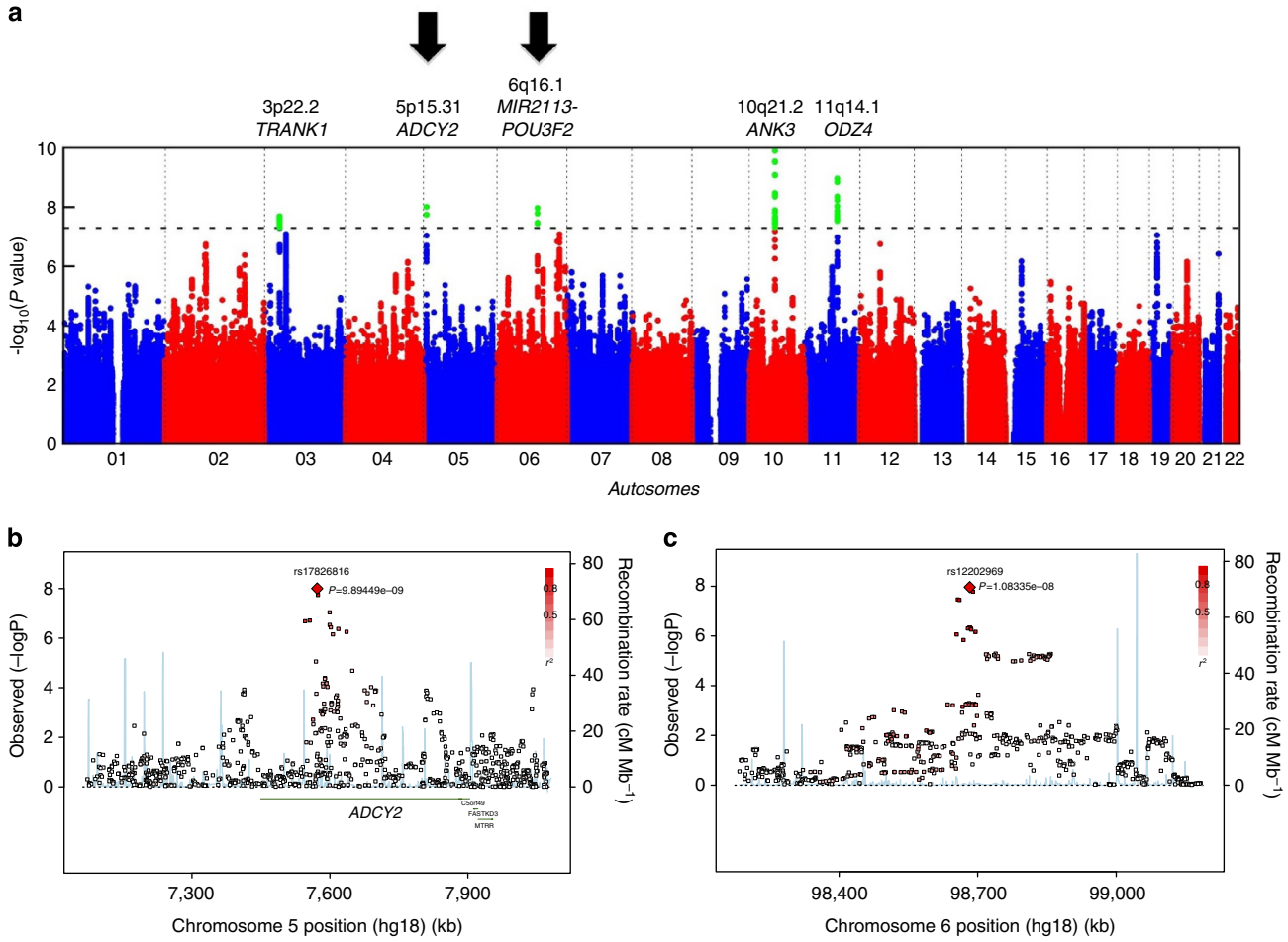

Figure 2 | Association results for the MooDS-PGC GWAS and the two new risk loci for BD. (**a**) Manhattan plot for all analysed SNPs, (**b**,**c**) regional association plots for the SNPs analysed at *ADCY2* (5p15.31) and *MIR2113-POU3F2* (6q16.1). Regional association plots were drawn using SNAP[54] and data for LD (red) and recombination frequency (blue line) from the 1000 Genomes Project.

OR = 1.10) with a distance of 75 kb to *MIR2113* and rs9375371 ($P_{GC} = 5.33 \times 10^{-6}$, OR = 1.11) with a distance of 531 kb to *POU3F2*. We did not find a variant in strong LD ($r^2 > 0.8$) that is directly located in either of the two genes. The genome-wide significant top SNP (rs12202969) is located 104 kb downstream of *MIR2113* and 706 kb upstream of *POU3F2*. To further explore the region between both genes, we investigated ENCODE[14] data at the UCSC Genome Browser[15] as well as data from GTEx[16], PheGenI[17], RegulomeDB[18], seeQTL[19] and an expressed quantitative trait locus (eQTL) browser by the Pritchard lab (http://eqtl.uchicago.edu/cgi-bin/gbrowse/eqtl/). We did not find robust evidence for an eQTL. However, in RegulomeDB, we found suggestive evidence that rs10457441 ($P_{GC} = 4.99E-07$, OR = 1.11) overlaps with regulatory DNA features such as open chromatin, histone modifications and binding sites for the transcription factors (EP300, USF1 and STAT3). Besides that, rs1906252 ($P_{GC} = 3.39 \times 10^{-8}$; OR = 1.12), which is in strong LD with the top SNP (rs12202969), was among the most significant finding in a GWAS of information processing speed (measured by a symbol search test in 991 healthy elderly probands from Scotland)[20].

Since an enrichment of association signals in target genes for *MIR137* has been observed in schizophrenia[4], we also tested whether known or predicted targets of *MIR2113* are enriched for significant *P* values, but did not find support for this hypothesis.

Discussion

The present GWAS is based on 24,025 patients and controls and combines so far unpublished MooDS samples with the published PGC-BD GWAS samples. Our study adds two new risk loci, namely the gene *ADCY2* and the 6q16.1 region, to the list of known BD risk loci identified through GWAS (*ANK3*, *NCAN*, *CACNA1C* and *ODZ4*). The MooDS data provide strong statistical support for three of these loci, except for *CACNA1C*. Several lines of evidence support an involvement of the two new loci in the neurobiology of BD and major psychiatric disorders. For some variants, functional effects on DNA and protein levels are suggested using knowledge from bioinformatical resources.

Our finding in the 6q16.1 region is located within a linkage peak identified by independent genome scans in BD families[21]. The genome-wide significant top SNP from our GWAS (rs12202969) is located downstream of an uncharacterized microRNA gene (*MIR2113*) and upstream of a transcription factor gene (*POU3F2*) that has been found to contribute to neocortex development in mice[22]. Using all available proxies for the SNPs with highly to moderately significant evidence for association, we did not find a variant in a gene in LD. However, one of the four genome-wide significant SNPs in the region (rs1906252) was found to be associated with a cognitive phenotype, speed of information processing, in an independent GWAS[20]. In addition, one SNP below genome-wide significance

Table 2 | Eighteen GWAS SNPs showing genome-wide evidence for association with BD in the combined MooDS and PGC samples.

SNP data				Association data										Gene data
				MooDS (2,266 Pat/ 5,028 Con)		PGC (7,481 Pat/ 9,250 Con)		MooDS-PGC (2,267,487 SNPs) FEM						
Chr	Position	Name	Alleles	*P* value	OR	*P* value	OR	A1	Allele Directions	*P* value	P_{GC} value	OR	Nearest gene or transcript
10q21.2	62,322,034	rs10994415*	C/T	0.0096	1.15	**6.97×10^{-10}**	1.31	C	+ + − + + + + +	**4.69×10^{-11}**	**6.88×10^{-11}**	1.27	*ANK3*, intron
10q21.2	62,279,124	rs10994397†,‡	T/C	0.0315	1.14	**5.54×10^{-10}**	1.35	T	+ + + + + − + +	**2.00×10^{-10}**	**2.86×10^{-10}**	1.29	*ANK3*, intron
10q21.2	62,294,814	rs1938540	T/C	0.0317	1.15	**1.90×10^{-9}**	1.32	T	+ + + + − + + +	**5.85×10^{-10}**	**8.22×10^{-10}**	1.27	*ANK3*, intron
11q14.1	79,083,620	rs12290811§	A/T	0.0023	1.19	9.25×10^{-8}	1.19	A	+ + + − + + + +	**7.81×10^{-11}**	**1.09×10^{-9}**	1.19	*ODZ4*, intron
11q14.1	79,077,193	rs12576775†	G/A	0.0294	1.13	**2.66×10^{-8}**	1.18	G	+ + + − − + + +	**3.27×10^{-9}**	**4.46×10^{-9}**	1.17	*ODZ4*, intron
11q14.1	79,068,394	rs7932890	G/A	0.0263	1.14	7.26×10^{-8}	1.18	G	+ + + − − + + +	**6.94×10^{-9}**	**9.35×10^{-9}**	1.17	*ODZ4*, intron
5p15.31	7,519,298	rs17826816	G/A	1.04×10^{-5}	1.19	3.99×10^{-5}	1.12	G	+ + + + + + + +	**7.35×10^{-9}**	**9.89×10^{-9}**	1.14	*ADCY2*, intron
6q16.1	98,576,223	rs12202969	A/G	**2.16×10^{-8}**	1.22	6.57×10^{-4}	1.08	A	+ + + + + + + +	**8.06×10^{-9}**	**1.08×10^{-8}**	1.12	*MIR2113; POU3F2*
10q21.2	62,136,206	rs10821745†	G/T	0.1691	1.08	**2.33×10^{-9}**	1.35	G	+ + − + + + + +	**9.41×10^{-9}**	**1.26×10^{-8}**	1.27	*ANK3*, intron
11q14.1	79,062,930	rs17138171	C/T	0.0261	1.14	1.06×10^{-7}	1.18	C	+ + + − − + + +	**1.02×10^{-8}**	**1.37×10^{-8}**	1.16	*ODZ4*, intron
10q21.2	62,105,053	rs10821736†	T/C	0.2407	1.06	**1.31×10^{-9}**	1.36	T	+ + − + + − + +	**1.16×10^{-8}**	**1.55×10^{-8}**	1.28	*ANK3*, intron
3p22.2	36,864,489	rs6550435‖	G/T	7.16×10^{-4}	1.15	4.80×10^{-6}	1.12	G	− + + + + + − +	**1.54×10^{-8}**	**2.05×10^{-8}**	1.13	*TRANK1*, intron
10q21.2	62,371,953	rs10994430	T/G	4.96×10^{-4}	1.22	**6.43×10^{-6}**	1.17	T	+ + + + + + + +	**1.61×10^{-8}**	**2.14×10^{-8}**	1.18	*ANK3*, intron
10q21.2	62,179,812	rs10994336†	T/C	0.171	1.11	**4.01×10^{-9}**	1.35	T	+ + + + + − + +	**1.76×10^{-8}**	**2.34×10^{-8}**	1.27	*ANK3*, intron
11q14.1	79,067,472	rs11237799	C/A	0.8125	1.14	2.20×10^{-7}	1.17	C	+ + + − − + + +	**1.91×10^{-8}**	**2.53×10^{-8}**	1.16	*ODZ4*, intron
10q21.2	62,098,952	rs10994308	A/G	0.24	1.10	**3.83×10^{-9}**	1.36	A	+ + − + + − + +	**2.73×10^{-8}**	**3.59×10^{-8}**	1.27	*ANK3*, intron
3p22.16	36,856,030	rs9834970†	C/T	0.0014	1.15	6.19×10^{-6}	1.11	C	− + + + + + + +	**3.67×10^{-8}**	**4.81×10^{-8}**	1.12	*TRANK1*, downstr.
10q21.2	62,071,041	rs10509129	T/G	0.34	1.09	**5.75×10^{-9}**	1.36	T	− + − + + − + +	**3.64×10^{-8}**	**4.77×10^{-8}**	1.29	*ANK3*, intron

Alleles, minor allele first; A1, the allele to which the odds ratio (OR) is predicted; Con, number of control subjects; Chr, cytogenetic band of the chromosome; downstr., downstream; FEM, fixed-effects meta-analysis based on the weighted *z*-score method[51]; MA, minor allele; Pat, number of patient subjects; P_{GC}, *P* value adjusted by a standardized inflation factor for GC.
SNPs are sorted according to their GC *P* values (P_{GC}) from the fixed-effects meta-analysis using the MooDS-PGC samples (Methods). Chromosomal positions refer to genome build GRCh37 (hg19). Allele directions refer to allele A1 in each sample (in the order Germany II, Poland, Australia, Canada, Germany III, Spain, Russia and PGC). Genome-wide significance is defined by the formal threshold of $P_{GC} < 5 \times 10^{-8}$ (bold). In each region, we identified at least two SNPs with genome-wide significance. In each chromosomal region, SNPs in strong LD with the top SNP ($r^2 \geq 0.8$) have been excluded for the purpose of display (LD measures based on HapMap-CEU data from the 1000 Genomes Project, pilot 1, 6 Feb 2012). For each region, a full list of SNPs with $P_{GC} < 5 \times 10^{-5}$ are given in Supplementary Tables 1–5. The numbers of SNPs with $P_{GC} < 5 \times 10^{-5}$ are 35 for the *ANK3* region, 34 for the *ODZ4* region, 14 for the *ADCY2* region, 51 for the region between *MIR2113* and *POU3F2* and 33 for the *TRANK1* region.
*SNP show $r^2 = 0.62$ with *ANK3* rs10994397 (ref. 7), $r^2 = 0.79$ with *ANK3* rs1938526 (ref. 10) and $r^2 = 0.47$ with *ANK3* rs10994336 (ref. 10).
†SNPs reached genome-wide significance in previous studies using a fixed-effects meta-analysis[7,9] or a logistic regression analysis[10].
‡SNP shows $r^2 > 0.8$ with *ANK3* rs10994336 (ref. 10).
§SNP shows $r^2 = 0.57$ with *ODZ4* rs12576775 (ref. 7).
‖SNP shows $r^2 = 0.13$ with *TRANK1* rs9834970 (ref. 9).

(rs10457441) converges with DNA features that are indicative for a potential involvement in gene expression regulation.

The *ADCY2* gene on chromosome 5p15.31 is expressed in the brain and encodes a cell membrane-bound enzyme for the synthesis of the second-messenger molecule cAMP. The ADCY2 protein is primarily regulated by heterotrimeric G proteins and produces cAMP in response to extracellular hormones and neurotransmitters that bind as ligands to G protein-coupled receptors (GPCRs)[23]. ADCY2 is also regulated by the protein kinase C (PKC) and the RAF kinase. ADCY2 is insensitive to calcium ions and calmodulin; both are regulators of other adenylate cyclases. cAMP induces the expression of downstream target genes by activation of the cAMP-dependent protein kinase A (PKA) and subsequent phosphorylation of the transcription factor cAMP response element-binding protein (CREB)[24]. GPCRs represent the majority of neurotransmitter receptors for dopamine, norepinephrine and serotonin, and belong to the earliest functional candidate genes for neuropsychiatric disorders. One of our two genome-wide significant SNPs in *ADCY2* is a missense variant with possibly damaging effect on the protein (rs13166360). We speculate that functional variation in *ADCY2* may have a more pronounced effect on BD susceptibility than functional variation in neurotransmitter receptors where a high degree of redundancy may facilitate functional compensation of one dysfunctional receptor by another.

In a recent co-expression network analysis for the human brain transcriptome, *ADCY2* showed up among the top 50 of 310 genes, which altogether form a network module specific for post-natal telencephalon (neocortex, hippocampus, amygdala and striatum)[25]. Those top (or 'hub') genes show the highest network connectivity within the module and should therefore have the highest potential functional relevance for these brain regions. For SNPs located in 2 of the top 10 hub genes, independent GWAS have found strong association with schizophrenia (neurogranin, *NRGN*)[26] and increasing suicidality in major depression (guanine deaminase, *GDA*)[27], two psychiatric disorders that partially share genetic risk factors with BD[28]. Data from the Stanley Medical Research Institute Online Genomics Database[29] show a 1.5–2.0-fold lower *ADCY2* expression in Brodmann areas 10 and 46 (BA10/46) of patients with BD as well as in patients with major depression disorder compared with controls ($P < 0.01$). Noteworthy, data from the same resource indicate that expression levels of *CACNA1C*, *ODZ4* and *NCAN* in these two areas are also significantly reduced in patients with BD and depression. BA10/46 map to neighbouring frontal and middle frontal regions in the prefrontal cortex and abnormalities in these brain areas have been described in both disorders[30,31].

In summary, our study implicates common variation at two genomic loci as new genetic risk factors for BD. While a specific gene cannot be pinned down at one of the risk loci (6q16.1), it harbours a genome-wide significant SNP that was previously shown to be associated with variation of the 'information processing speed', a cognitive function that has been postulated as a valid and highly specific cognitive endophenotype for BD, differentiating both euthymic BD patients and their healthy first-degree relatives from healthy controls[32]. *ADCY2* plays a key role in cAMP-dependent GPCR pathways. Disturbed neurotransmission at these pathways is a long-standing hypothesis in psychiatric research, which has motivated multiple candidate gene studies. Most of these studies, however, focused on variation in neurotransmitter receptor and transporter genes and less on adenylate cyclase genes, located more downstream and bundling the signals coming in from several neurotransmitter receptor types. Adenylate cyclase may therefore represent a functional bottleneck in signal transduction pathways and genetic variation may have a more direct influence on the phenotype than at

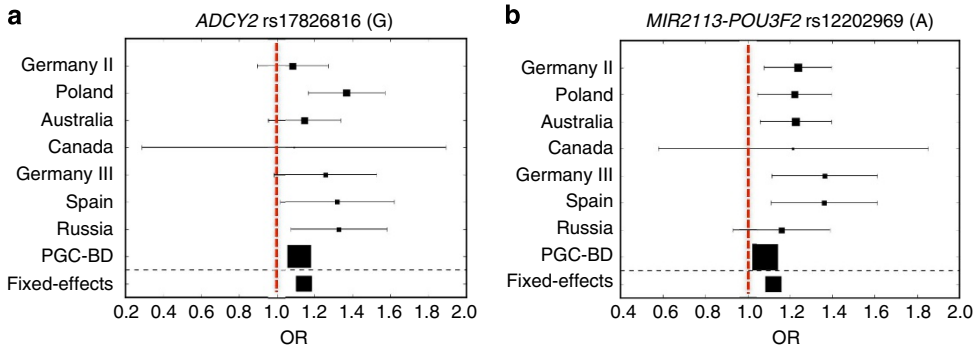

Figure 3 | Genetic effect sizes for the two new risk loci identified through the MooDS-PGC GWAS of BD. (**a**,**b**) Forest plots displaying the most significant SNP's odds ratio (OR, full square) and their 95% confidence interval (horizontal continuous lines) for the gene *ADCY2* (5p15.31) as well as the region between the genes *MIR2113* and *POU3F2* (6q16.1). The overall OR was calculated using a fixed-effects meta-analysis based on the weighted *z*-score method[51]. The effect allele of each SNP is given in brackets. The area of a square reflects the statistical power of the respective study sample. Areas were calculated by the reciprocal value of the standard deviations.

positions with higher functional redundancy (neurotransmitter receptors/transporters). This may also explain the somewhat surprising observation that GWAS in neuropsychiatric disorders performed so far did not identify strong association signals in neurotransmitter/transporter genes.

Methods

Phenotypic data of the study samples. Written informed consent was obtained from all patients and controls before participation in the study. Protocols and procedures were approved by the local Ethics Committees of the hospitals mentioned below. An overview of the phenotyping of the seven new MooDS samples and the PGC-BD GWAS sample are provided in Table 1.

For MooDS patients, affected subjects received a lifetime diagnosis of BD (type I, type II, schizoaffective disorder bipolar type, BD not otherwise specified) according to DSM-IV[33] criteria using a consensus best-estimate procedure[34] and structured interviews[35,36]. Patients were recruited from consecutive admissions to psychiatric in-patient units at the University Hospital Würzburg (Germany II), at the Central Institute of Mental Health in Mannheim and other collaborating psychiatric hospitals in Germany (Germany III), at the Mood Disorder Unit, Prince of Wales Hospital in Sydney (Australia), at the Department of Psychiatry, Poznan University of Medical Sciences (Poland), at specialty mood disorders clinics in Halifax and Ottawa (Canada), at the Russian State Medical University, Moscow (Russia), and at the Civil Hospital Carlos Haya, Málaga (Spain).

For MooDS controls, unaffected subjects were drawn from the population-based Heinz Nixdorf Recall (HNR) Study[37] (Germany II), from a Munich-based community sample by the Max Planck Institute of Psychiatry[38] (Germany III), from parents of adolescent twins by the Brisbane Longitudinal Twin Study and a community sample (Australia), from a hospital-based sample by the University of Szczecin (Poland), from a combined sample by the Saguenay-Lac-Saint-Jean (SLSJ) study (Canada), from a combined sample by the Kursk State Medical University (KSMU) study and the UFA study (Russia), and from parts of the European Community Respiratory Health Survey (ECRHS) (Spain). Controls for the Russian sample were generated by GABRIEL, a multidisciplinary study to identify the genetic and environmental causes of asthma in the European Community[39]. The Canadian and Spanish controls were kindly provided by Catherine Laprise (UQAC, Saguenay, Canada) and Manolis Kogevinas (CREAL, Barcelona, Spain). Polish controls were produced by the International Agency for Research on Cancer (IARC) and the Centre National de Génotypage (CNG) GWAS Initiative for a study of upper aerodigestive tract cancers[40]. Except the controls for Germany III, none of the above-mentioned controls were screened for the presence of a neuropsychiatric phenotype.

Genetic data of the study samples. Over the last 5 years, the MooDS consortium produced two waves of genome-wide data for BD. The first wave of data (Germany I) were generated for the GWAS by Cichon *et al.*[8] and are now part of the PGC-BD GWAS data[7]. The second wave of data were generated for the present GWAS and are described below.

For Polish, Canadian, Russian and Spanish patients, lymphocyte DNA was isolated from EDTA anti-coagulated venous blood by salting-out with saturated sodium chloride solution[41] or by a Chemagic Magnetic Separation Module I (Chemagen, Baesweiler, Germany) according to the manufacturer's recommendations. DNA samples were genome-wide genotyped using Infinium assays (Illumina, San Diego, CA, USA) for BeadChips Human660W-Quad (Polish patients) and HumanOmni1-Quad (Canadian, Russian and Spanish patients). DNA extraction and genotyping was performed at the Department of Genomics, Life & Brain Center, University of Bonn, Germany. Genome-wide genotyped data for Russian and Polish controls were available from previous studies of asthma and cancer[39,40]. Canadian and Spanish data were kindly provided by Catherine Laprise (UQAC, Saguenay, Canada) and Manolis Kogevinas (CREAL, Barcelona, Spain). Genotyping of the samples Australia, Germany II, Germany III, and PGC-BD were described in previous studies[7,8,42–44]. We aimed to genotype the patient and controls from the same population on the same chip type. However, depending on the chip type that was commercially available at the time of our genotyping, respectively, which has been used by other studies from which we have drawn data, some samples were genotyped on different BeadChips. For the seven new MooDS samples, the predominant chip types were the Human610-Quad and the HumanOmni1-Quad. The other chip types were the HumanHap550, the Human660Q-Quad and the HumanHap300. Supplementary Table 8 summarizes the genotyping chips for all new MooDS samples.

In the PGC study[7], SNP genotypes from 11 individual GWAS samples were imputed up to a set of more than two million autosomal markers using reference haplotypes from the HapMap phase 2 CEU sample. To produce a fitting marker set for the seven MooDS samples, the SNPs of each MooDS sample were pruned to the least common set (first and second QC) before the imputation (see below) and subsequent sample-specific association analyses (followed by a third QC). After the alignment of the MooDS and the PGC sets and another filtering step, the final SNP sets were carried on to the meta-analysis. The first and second QC was performed using several tools from INTERSNP[45] (version 1.11) and PLINK[46] (version 1.07) as described below. The third QC consisted of manual filtering of data from IMPUTE[47] (version 2) and SNPTEST[48] (version 1). Supplementary Data 1 illustrate the steps of the QC procedure and the numbers of subjects and SNPs excluded at each individual step.

In particular, the first QC comprised five filtering steps (exclusion criteria are given in brackets) and was performed for each sample separately. We investigated SNP and subject call rates (CR < 95% and CR < 98%), sex status of each subject (inconsistency between data from recruitment centres and X-chromosome heterozygosity rate), deviation of a SNP from the Hardy–Weinberg Equilibrium ($P \leq 1 \times 10^{-6}$) and cryptic relatedness between pairs of subjects (identity-by-state (IBS) > 1.98 or s.d. six times higher than the IBS averaged over all pairs of the sample). If patients and controls have been genotyped on different chip types, the SNP set was reduced to those SNPs available in patients and controls after the first QC. In this way, we avoided potential differential error that have could arised from SNPs that are genotyped in patients but imputed in controls, or *vice versa*.

The second QC consisted of another five filtering steps, that is, SNPs with coding conflicts between SNP sets of different chip types (A/T and G/C SNPs), potential population stratification within and between samples using multi-dimensional scaling analysis (MDS, identification of population outliers by visual assessment, Supplementary Fig. 2), subjects with an excess of heterozygosity (inbreeding coefficient > 4 s.d. from the sample's mean), cryptic relatedness (IBS > 1.98 or s.d. > 6) and monomorphic SNPs (MAF < 0.01%). We assessed the effect of the first and second QC on the genotype data by a calculation of a genomic inflation factor λ. Each sample showed $\lambda \leq 1.05$.

Imputation of the MooDS data sets was carried out using IMPUTE and the February 2012 release of the 1,000 Genomes Project data[49]. Imputed SNPs with an info score < 0.4 or a MAF < 0.03 were removed during the third QC and the sample-specific association analyses (see below). We aligned the imputed SNP sets between MooDS and PGC and excluded SNPs that were absent in three or more MooDS samples and/or were not present in the PGC, resulting in a number of 2,267,487 autosomal SNPs for 2,266 patients and 5,028 controls from MooDS, and 9,747 patients and 9,250 controls from PGC. Finally, we assessed the effect of the QC procedure by further calculations of genomic inflation factors.

The λ values for a standardized sample size of 1,000 patients and 1,000 controls and the non-standardized λ values were 1.027 and 1.084 for MooDS, 1.018 and 1.147 for PGC, and 1.018 and 1.203 for the combination of both.

As ancestry was assigned to patients and controls on the basis of self-reported ancestry, we inspected the ethnical matching between patients and controls within and between the MooDS samples by a MDS analysis. After exclusion of population outliers (described above), we used the significant MDS dimensions used as covariates in the sample-specific association analyses that were conducted using SNPTEST. Using the eight quality-controlled samples for MooDS and PGC, meta-analysis was performed with YAMAS[50] under a fixed-effects model. For each SNP, effect estimates and their s.e. were combined across the eight data sets as described elsewhere[51]. To adequately correct for inflation of P values due to varying sample sizes, we adjusted the P values for the MooDS-PGC analysis (P_{GC}) using the standardized λ factor for GC, which was 1.018 (described above). We considered an association as genome-wide significant if $P_{GC} < 5 \times 10^{-8}$.

Overlap between new samples and previous BD studies. It is noteworthy that none of the seven MooDS samples have been analysed in a genome-wide discovery step for BD before. However, parts of Poland, Spain and Russia have been used for targeted genotyping of 48 SNPs for the first replication step of our previous GWAS of BD[8]. Parts of Australia and Germany III have been used for the second replication step of that study[8], that is, extraction of one SNP from genome-wide data. Thirty eight SNPs from the genome-wide data of Germany II, Germany III (alias BOMA-Rep1 and BOMA–Rep2) and Australia have been used in the replication step of the previous PGC study[7]. The Canadian patients partially overlap with a linkage and expression study of BD[52] and a candidate gene association study of BD[53].

References

1. Lichtenstein, P. *et al.* Common genetic determinants of schizophrenia and bipolar disorder in Swedish families: a population-based study. *Lancet* **373,** 234–239 (2009).
2. Nöthen, M. M., Nieratschker, V., Cichon, S. & Rietschel, M. New findings in the genetics of major psychoses. *Dialogues Clin. Neurosci.* **12,** 85–93 (2010).
3. Baum, A. E. *et al.* A genome-wide association study implicates diacylglycerol kinase eta (DGKH) and several other genes in the etiology of bipolar disorder. *Mol. Psychiatry* **13,** 197–207 (2008).
4. Sullivan, P. F., Daly, M. J. & O'Donovan, M. Genetic architectures of psychiatric disorders: the emerging picture and its implications. *Nat. Rev. Genet.* **13,** 537–551 (2012).
5. Craddock, N. & Sklar, P. Genetics of bipolar disorder. *Lancet* **381,** 1654–1662 (2013).
6. The International Schizophrenia Consortium. Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature* **460,** 748–752 (2009).
7. 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.* **43,** 977–983 (2011).
8. Cichon, S. *et al.* Genome-wide association study identifies genetic variation in neurocan as a susceptibility factor for bipolar disorder. *Am. J. Hum. Genet.* **88,** 372–381 (2011).
9. Chen, D. T. *et al.* Genome-wide association study meta-analysis of European and Asian-ancestry samples identifies three novel loci associated with bipolar disorder. *Mol. Psychiatry* **18,** 195–205 (2013).
10. Ferreira, M. A. R. *et al.* Collaborative genome-wide association analysis supports a role for ANK3 and CACNA1C in bipolar disorder. *Nat. Genet.* **40,** 1056–1058 (2008).
11. Hindorff, L. A. *et al.* Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. *Proc. Natl Acad. Sci. USA* **106,** 9362–9367 (2009).
12. Green, E. K. *et al.* Replication of bipolar disorder susceptibility alleles and identification of two novel genome-wide significant associations in a new bipolar disorder case-control sample. *Mol. Psychiatry* **18,** 1302–1307 (2012).
13. Adzhubei, I. *et al.* A method and server for predicting damaging missense mutations. *Nat. Methods* **7,** 248–249 (2010).
14. Bernstein, B. E. *et al.* An integrated encyclopedia of DNA elements in the human genome. *Nature* **489,** 57–74 (2012).
15. Karolchik, D., Hinrichs, A. S. & Kent, W. J. The UCSC Genome Browser. *Curr. Protoc. Bioinformatics* Chapter 1, 4 (2012).
16. GTEx Consortium. The Genotype-Tissue Expression (GTEx) project. *Nat. Genet.* **45,** 580–585 (2013).
17. Ramos, E. M. *et al.* Phenotype-Genotype Integrator (PheGenI): synthesizing genome-wide association study (GWAS) data with existing genomic resources. *Eur. J. Hum. Genet.* **22,** 144–147 (2013).
18. Boyle, A. P. *et al.* Annotation of functional variation in personal genomes using RegulomeDB. *Genome Res.* **22,** 1790–1797 (2012).
19. Xia, K. *et al.* seeQTL: a searchable database for human eQTLs. *Bioinformatics* **28,** 451–452 (2012).
20. Luciano, M. *et al.* Whole genome association scan for genetic polymorphisms influencing information processing speed. *Biol. Psychol.* **86,** 193–202 (2011).
21. Craddock, N., O'Donovan, M. C. & Owen, M. J. The genetics of schizophrenia and bipolar disorder: dissecting psychosis. *J. Med. Genet.* **42,** 193–204 (2005).
22. Kwan, K. Y., Sestan, N. & Anton, E. S. Transcriptional co-regulation of neuronal migration and laminar identity in the neocortex. *Development* **139,** 1535–1546 (2012).
23. Ostrom, R. S., Bogard, A. S., Gros, R. & Feldman, R. D. Choreographing the adenylyl cyclase signalosome: sorting out the partners and the steps. *Naunyn Schmiedebergs Arch. Pharmacol.* **385,** 5–12 (2012).
24. Delghandi, M. P., Johannessen, M. & Moens, U. The cAMP signalling pathway activates CREB through PKA, p38 and MSK1 in NIH 3T3 cells. *Cell Signal.* **17,** 1343–1351 (2005).
25. Kang, H. J. *et al.* Spatio-temporal transcriptome of the human brain. *Nature* **478,** 483–489 (2011).
26. Stefansson, H. *et al.* Common variants conferring risk of schizophrenia. *Nature* **460,** 744–747 (2009).
27. Perroud, N. *et al.* Genome-wide association study of increasing suicidal ideation during antidepressant treatment in the GENDEP project. *Pharmacogenomics J.* **12,** 68–77 (2012).
28. Lee, S. H. *et al.* Genetic relationship between five psychiatric disorders estimated from genome-wide SNPs. *Nat. Genet.* **45,** 984–994 (2013).
29. Higgs, B. W., Elashoff, M., Richman, S. & Barci, B. An online database for brain disease research. *BMC Genomics* **7,** 70 (2006).
30. Adler, C. M., DelBello, M. P. & Strakowski, S. M. Brain network dysfunction in bipolar disorder. *CNS Spectr.* **11,** 312–320 (2006).
31. Martinowich, K., Schloesser, R. J. & Manji, H. K. Bipolar disorder: from genes to behavior pathways. *J. Clin. Invest.* **119,** 726–736 (2009).
32. Daban, C. *et al.* Is processing speed a valid cognitive endophenotype for bipolar disorder? *J. Affect. Disord.* **139,** 98–101 (2012).
33. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders,* Fourth Edn (DSM-IV) (APA, 1994).
34. Leckman, J. F., Sholomskas, D., Thompson, W. D., Belanger, A. & Weissman, M. M. Best estimate of lifetime psychiatric diagnosis: a methodological study. *Arch. Gen. Psychiatry* **39,** 879–883 (1982).
35. Spitzer, R. L., Williams, J. B., Gibbon, M. & First, M. B. The structured clinical interview for DSM-III-R (SCID). I: history, rationale, and description. *Arch. Gen. Psychiatry* **49,** 624–629 (1992).
36. Farmer, A. E., Wessely, S., Castle, D. & McGuffin, P. Methodological issues in using a polydiagnostic approach to define psychotic illness. *Br. J. Psychiatry* **161,** 824–830 (1992).
37. Schmermund, A. *et al.* Assessment of clinically silent atherosclerotic disease and established and novel risk factors for predicting myocardial infarction and cardiac death in healthy middle-aged subjects: rationale and design of the Heinz Nixdorf RECALL Study. *Am. Heart J.* **144,** 212–218 (2002).
38. Muglia, P. *et al.* Genome-wide association study of recurrent major depressive disorder in two European case-control cohorts. *Mol. Psychiatry* **15,** 589–601 (2010).
39. Moffatt, M. F. *et al.* A large-scale, consortium-based genomewide association study of asthma. *New Engl. J. Med.* **363,** 1211–1221 (2010).
40. McKay, J. D. *et al.* A genome-wide association study of upper aerodigestive tract cancers conducted within the INHANCE consortium. *PLoS Genet.* **7,** e1001333 (2011).
41. Miller, S. A., Dykes, D. D. & Polesky, H. F. A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Res.* **16,** 1215 (1988).
42. McAuley, E. Z. *et al.* A genome screen of 35 bipolar affective disorder pedigrees provides significant evidence for a susceptibility locus on chromosome 15q25-26. *Mol. Psychiatry* **14,** 492–500 (2009).
43. Mitchell, P. B., Johnston, A. K., Corry, J., Ball, J. R. & Malhi, G. S. Characteristics of bipolar disorder in an Australian specialist outpatient clinic: comparison across large datasets. *Aust. N. Z. J. Psychiatry* **43,** 109–117 (2009).
44. Medland, S. E. *et al.* Common variants in the trichohyalin gene are associated with straight hair in Europeans. *Am. J. Hum. Genet.* **85,** 750–755 (2009).
45. Herold, C., Steffens, M., Brockschmidt, F. F., Baur, M. P. & Becker, T. INTERSNP: genome-wide interaction analysis guided by a priori information. *Bioinformatics* **25,** 3275–3281 (2009).
46. Purcell, S. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **81,** 559–575 (2007).
47. Howie, B. N., Donnelly, P. & Marchini, J. A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. *PLoS Genet.* **5,** e1000529 (2009).
48. Marchini, J., Howie, B., Myers, S., McVean, G. & Donnelly, P. A new multipoint method for genome-wide association studies by imputation of genotypes. *Nat. Genet.* **39,** 906–913 (2007).
49. Abecasis, G. R. A map of human genome variation from population-scale sequencing. *Nature* **467,** 1061–1073 (2010).

50. Meesters, C. Quick, 'imputation-free' meta-analysis with proxy-SNPs. *BMC Bioinformatics* **13,** 231 (2012).
51. De Bakker, P. I. W. *et al.* Practical aspects of imputation-driven meta-analysis of genome-wide association studies. *Hum. Mol. Genet.* **17,** R122–R128 (2008).
52. De Lara, C. L. Implication of synapse-related genes in bipolar disorder by linkage and gene expression analyses. *Int. J. Neuropsychopharmacol.* **13,** 1397–1410 (2010).
53. Mamdani, F. *et al.* Lithium response and genetic variation in the CREB family of genes. *Am. J. Med. Genet. B. Neuropsychiatr. Genet.* **147B,** 500–504 (2008).
54. Johnson, A. D. *et al.* SNAP: a web-based tool for identification and annotation of proxy SNPs using HapMap. *Bioinformatics* **24,** 2938–2939 (2008).
55. AMDP. The AMDP-System Association of Methodology and Documentation in Psychiatry. *Manual for the Assessment and Documentation of Psychology* 4th Edn (Springer, 1982).
56. Wittchen, H. U. *et al.* Screening for mental disorders: performance of the Composite International Diagnostic—Screener (CID–S). *Int. J. Methods Psychiatr. Res.* **8,** 59–70 (1999).
57. Nurnberger, J. I. Diagnostic interview for genetic studies. Rationale, unique features, and training. NIMH Genetics Initiative. *Arch. Gen. Psychiatry* **51,** 849–859 (1994).
58. Maxwell, M. E. Family Interview for Genetic Studies. (Clinical Neurogenetic Branch, Intramural Research Program, NIMH, 1992).
59. Sheehan, D. V. *et al.* The Mini-International Neuropsychiatric Interview (M.I.N.I.): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. *J. Clin. Psychiatry* **59**(Suppl 2): 22–33 (quiz 34–57) (1998).
60. Spitzer, R. L., Endicott, J. & Robins, E. Research diagnostic criteria: rationale and reliability. *Arch. Gen. Psychiatry* **35,** 773–782 (1978).
61. Endicott, J. & Spitzer, R. L. A diagnostic interview: the schedule for affective disorders and schizophrenia. *Arch. Gen. Psychiatry* **35,** 837–844 (1978).
62. Wing, J. K. SCAN. Schedules for Clinical Assessment in Neuropsychiatry. *Arch. Gen. Psychiatry* **47,** 589–593 (1990).

Acknowledgements

We are grateful to all patients and control subjects who contributed to this study. This study was supported by the German Federal Ministry of Education and Research (BMBF) through the Integrated Genome Research Network (IG) MooDS (Systematic Investigation of the Molecular Causes of Major Mood Disorders and Schizophrenia; grant 01GS08144 to M.M.N. and S.C., grant 01GS08147 to M.R.), under the auspices of the National Genome Research Network plus (NGFNplus), and through the Integrated Network IntegraMent (Integrated Understanding of Causes and Mechanisms in Mental Disorders), under the auspices of the e:Med Programme. M.M.N. is a member of the DFG-funded Excellence-Cluster ImmunoSensation. M.M.N. also received support from the Alfried Krupp von Bohlen und Halbach-Stiftung. Canadian patients were genotyped within the ConLiGen project (www.ConLiGen.org), supported by a grant from the Deutsche Forschungsgemeinschaft to M.R., M.B., and T.G.S. (RI 908/7-1). Controls for Germany II were drawn from the Heinz Nixdorf Recall Study (HNR) cohort, which was established with the support of the Heinz Nixdorf Foundation. Recruitment of the Australian sample was supported by the Australian NHMRC program grant number 1037196. The collection of the Canadian patients has been supported by a grant from the Canadian Institutes of Health Research #64410 to M.A. Our study uses data generated by the GABRIEL consortium (controls for the sample Russia). Funding for these data was provided by the European Commission as part of GABRIEL contract number 018996 under the Integrated Program LSH-2004-1.2.5-1. Post genomic approaches to understand the molecular basis of asthma aiming at a preventive or therapeutic control and the Wellcome Trust under award 084703. Canadian controls were drawn from the French Canadian study (SLSJ) that was partially supported by the Canada Research Chair, the Canadian Institutes of Health Research (Operating grant No. MOP-13506) and the Quebec Respiratory Network of the Fonds de recherche en Santé du Québec (FRQS). Polish controls were produced by the International Agency for Research on Cancer (IARC)/Centre National de Genotypage (CNG) GWAS Initiative. We thank the Bipolar Disorder Working Group of the Psychiatric Genomics Consortium (PGC-BD) for providing access to the relevant data.

Author Contributions

MooDS management group: M.M.N., M.R., P.P., S.C.; writing group: T.W.M, M.L., T.G.S., M.R., M.M.N. and S.C. coordinated the work and prepared the manuscript, with feedback from the other authors who approved the manuscript; statistical analysis group: M.L., M.M., A.L. and T.B.; bioinformatical analysis group: J.T., S.H. and T.W.M.; sample management, DNA extraction, genotyping, QC in Bonn and Mannheim: F.D., AJ.F., J.S., S.H., P.H., L.P., R.B., S.M. and S.H.W.; recruitment and characterization of patients and controls and generation of control data for Germany I and II: T.G.S., Ja. St., A.R., M.G.-S., W.M., M.S., H.V., J.K.-C., A.P., M.B., M.H., B.M.-M., S.L. and Su.Mo.; recruitment and characterization of patients and controls and generation of control data for Poland: P.M.C., J.H., J.L., N.S.-D., P.B. and J.D.M.; recruitment and characterization of patients and controls and generation of control data for Australia: A.W., P.B.M., J.M.F., P.R.S., G.W.M., S.E.M., S.D.G. and N.G.M.; recruitment and characterization of patients and controls and generation of control data for Russia: V.K., A.C., G.B., G.P., L.I.A., A.S.T., Al.Po., E.K.; recruitment and characterization of patients and controls and generation of control data for Canada: M.A., P.G., G.A.R., G.T. and C.L.; recruitment and characterization of patients and controls and generation of control data for Spain: F.R., F.M. and M.K.

Additional information

Supplementary Information accompanies this paper at http://www.nature.com/naturecommunications

Competing financial interests: The authors declare no competing financial interests.

Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/

How to cite this article: Mühleisen, T. W. *et al.* Genome-wide association study reveals two new risk loci for bipolar disorder. *Nat. Commun.* 5:3339 doi: 10.1038/ncomms4339 (2014).