BJPsych | The British Journal of Psychiatry (2016) 208, 128–137. doi: 10.1192/bjp.bp.114.156976

Impact of a *cis*-associated gene expression SNP on chromosome 20q11.22 on bipolar disorder susceptibility, hippocampal structure and cognitive performance

Ming Li,* Xiong-jian Luo,* Mikael Landén, Sarah E. Bergen, Christina M. Hultman, Xiao Li, Wen Zhang, Yong-Gang Yao, Chen Zhang, Jiewei Liu, Manuel Mattheisen, Sven Cichon, Thomas W. Mühleisen, Franziska A. Degenhardt, Markus M. Nöthen, Thomas G. Schulze, Maria Grigoroiu-Serbanescu, Hao Li, Chris K. Fuller, Chunhui Chen, Qi Dong, Chuansheng Chen, Stéphane Jamain, Marion Leboyer, Frank Bellivier, Bruno Etain, Jean-Pierre Kahn, Chantal Henry, Martin Preisig, Zoltán Kutalik, Enrique Castelao, Adam Wright, Philip B. Mitchell, Janice M. Fullerton, Peter R. Schofield, Grant W. Montgomery, Sarah E. Medland, Scott D. Gordon, Nicholas G. Martin, MooDS Consortium,[a] The Swedish Bipolar Study Group,[a] Marcella Rietschel, Chunyu Liu, Joel E. Kleinman, Thomas M. Hyde, Daniel R. Weinberger and Bing Su

Background
Bipolar disorder is a highly heritable polygenic disorder. Recent enrichment analyses suggest that there may be true risk variants for bipolar disorder in the expression quantitative trait loci (eQTL) in the brain.

Aims
We sought to assess the impact of eQTL variants on bipolar disorder risk by combining data from both bipolar disorder genome-wide association studies (GWAS) and brain eQTL.

Method
To detect single nucleotide polymorphisms (SNPs) that influence expression levels of genes associated with bipolar disorder, we jointly analysed data from a bipolar disorder GWAS (7481 cases and 9250 controls) and a genome-wide brain (cortical) eQTL (193 healthy controls) using a Bayesian statistical method, with independent follow-up replications. The identified risk SNP was then further tested for association with hippocampal volume ($n = 5775$) and cognitive performance ($n = 342$) among healthy individuals.

Results
Integrative analysis revealed a significant association between a brain eQTL rs6088662 on chromosome 20q11.22 and bipolar disorder (log Bayes factor = 5.48; bipolar disorder $P = 5.85 \times 10^{-5}$). Follow-up studies across multiple independent samples confirmed the association of the risk SNP (rs6088662) with gene expression and bipolar disorder susceptibility ($P = 3.54 \times 10^{-8}$). Further exploratory analysis revealed that rs6088662 is also associated with hippocampal volume and cognitive performance in healthy individuals.

Conclusions
Our findings suggest that 20q11.22 is likely a risk region for bipolar disorder; they also highlight the informative value of integrating functional annotation of genetic variants for gene expression in advancing our understanding of the biological basis underlying complex disorders, such as bipolar disorder.

Declaration of interest
None.

Bipolar disorder is a severe, chronic psychiatric disorder with a worldwide lifetime prevalence ranging from 0.5 to 1.5%.[1] Bipolar disorder is characterised by a variety of profound mood symptoms including episodes of mania, hypomania and depression, and is often accompanied by psychotic features and cognitive deficits. To date, there has been a fair amount of data from family and twin studies to highlight a strong genetic predisposition for bipolar disorder.[1] Nevertheless, bipolar disorder is a highly polygenic disorder that can vary substantially from population to population. Although linkage analysis and genetic association studies have yielded numerous candidate variants for bipolar disorder, only a few of these have been satisfactorily replicated across independent samples.[2,3]

With advances in knowledge of human genetic variations, such as data generated by the International HapMap and 1000 Human Genome projects and several subsequent genome-wide association studies (GWAS) by a number of international collaborators, a wealth of novel susceptible variants for bipolar disorder have been reported, particularly single nucleotide polymorphisms (SNPs), in the following genes: calcium channel, voltage-dependent, L type, alpha 1C subunit (*CACNA1C*); ankyrin 3, node of Ranvier (ankyrin G) (*ANK3*); teneurin transmembrane protein 4 (*TENM4*); neurocan (*NCAN*); and tetratricopeptide repeat and ankyrin repeat containing 1 (*TRANK1*).[4–8] These GWAS-identified risk SNPs unfortunately only account for a small portion of the genetic risk for bipolar disorder, which suggests there should be additional loci contributing to the genetic susceptibility. Previous aggregated analyses indicated there might be valid risk loci underlying genetic markers passing only nominal significance in the GWAS,[9] a possibility confirmed by several later studies. For example, a number of schizophrenia and bipolar

*These authors contributed equally to the work.
[a]A list of the additional members is provided in the online Data Supplement to this paper.

https://doi.org/10.1192/bjp.bp.114.156976 Published online by Cambridge University Press

disorder susceptibility SNPs did not reach genome-wide significance in initial GWAS samples, but showed consistent replications in subsequent independent samples, thus implying that these loci might reflect weak but true risk signals.[10]

Genetic loci associated with clinical diagnosis are also expected to be related to so-called intermediate phenotypes implicated in the biology of genetic risk for bipolar disorder. Previous studies have reported hippocampal dysfunction (e.g. memory impairment) in patients with bipolar disorder and their unaffected relatives, implying that variation in hippocampal biology is an intermediate phenotype related to the genetic risk of bipolar disorder.[11] In addition, smaller hippocampal volume has been reported in patients with bipolar disorder.[12,13] Meanwhile, functional neuroimaging studies have revealed that dysfunctions of the hippocampus and its closely related regions underpin abnormal affective responses and dysfunctional emotion regulation in bipolar disorder.[14] Finally, post-mortem studies further support the hypothesis that hippocampal abnormalities are relevant to the altered synaptic plasticity and diminished resilience in bipolar disorder.[15] Therefore, analysis of bipolar disorder-associated SNPs on these hippocampus-related phenotypes may provide a plausible way to uncover their functions in neurodevelopment, and possibly, their involvement in disease susceptibility.

Recent successes in integrating disease GWAS and gene expression data for several other complex diseases have been promising,[16–18] and we wondered whether such an approach may yield novel results for bipolar disorder. Predictably, several lines of evidence have suggested an enrichment of expression quantitative trait loci (eQTL) among bipolar disorder susceptibility SNPs in the brain,[19] further highlighting the importance of integrating the functional annotation of genetic variants for gene expression to advance our understanding of the biological bases of bipolar disorder. In light of these findings, we integrated bipolar disorder GWAS data from 16 731 individuals and genome-wide eQTL data from 193 human cortex samples from healthy individuals, followed by a set of independent replications on both eQTL and disease associations.

Method

Discovery brain eQTL and bipolar disorder GWAS data-sets

The brain eQTL data-set used in this study was reported previously.[20] In brief, after excluding ethnic outliers and samples that were possibly related, a total of 193 independent human cortex samples of European origin from healthy, older individuals (age > 65) were included in the eQTL analysis. Detailed information about genotyping and expression profiling, as well as the statistical methods used, can be found in the online data supplement to this paper or the original publication.[20]

For the bipolar disorder GWAS data, the working group of the Psychiatric Genomics Consortium (PGC) for bipolar disorder recently conducted a meta-analysis of large-scale genome-wide data on bipolar disorder among populations of European descent (PGC1 family GWAS).[6] In this earlier study, the researchers opted to compare patients with bipolar disorder that had experienced pathologically relevant episodes of elevated mood (mania or hypomania) and control patients from the same geographical and ethnic populations. To summarise, we used 2 117 872 SNPs across the genome from the GWAS samples (7481 cases and 9250 controls), and the association significance (*P*-values) for these SNPs was downloaded from the PGC1 data-sharing website (www.med.unc.edu/pgc/downloads). Detailed descriptions of the samples, data quality, genotype imputation, genomic controls and statistical analyses can be found in the original study.[6]

Integrative analysis of eQTL and bipolar disorder GWAS data

We integrated the eQTL and bipolar disorder GWAS data using a Bayesian statistical framework. Statistical analyses for the eQTL and bipolar disorder GWAS were carried out with the *Sherlock* software tool (http://sherlock.ucsf.edu/submit.html), which has been described elsewhere.[17] In brief, *Sherlock* is based on the rationale that a risk gene for the disease may have at least one eQTL, and these eQTLs could alter gene expression, which in turn affects disease susceptibility. Given the probability that this might be true, there should be a significant overlap of the eQTL of a gene and the loci associated with the disorder, which would imply a likely functional role for the gene in that particular disease. At this juncture, *Sherlock* aligns the eQTL and bipolar disorder GWAS and considers only the shared SNPs in both data-sets. *Sherlock*'s scoring rubric both increases the total gene score for overlapping SNPs and provides a penalty in the absence of an overlap, although associations found only in the bipolar disorder GWAS do not alter the score. *Sherlock* computes individual log Bayes factors (LBFs) for each SNP pair in the alignment, and the sum of these constitutes the final LBF score for each gene.

Brain eQTL data for replication analysis

Considering that bipolar disorder is a mental disorder that originates from abnormal brain function, brain samples are presumably appropriate for replication testing of the eQTL results. We first used a brain dorsolateral prefrontal cortex (DLPFC) sample ($n = 320$) consisting of White and African-American healthy controls (labelled as the 'first replication sample'), in which the sample had been previously used to identify psychiatric risk mRNA transcripts.[21–23]

We also used other well-characterised brain expression databases for replication analysis of the eQTL associations. A brief description of the gene expression resources is provided as follows, whereas more detailed information can be found in the original studies:[18,24–26]

(a) BrainCloud: BrainCloud contains genetic information and whole-transcriptome expression data from the post-mortem DLPFC of 261 healthy White and African-American individuals. The data in BrainCloud are aimed at exploring temporal dynamics and genetic control of transcription across the lifespan.[24] Of note, there is partial overlap between BrainCloud data and our 'first replication sample'.

(b) Data from the study by Webster *et al*: Webster and colleagues studied the relationship between the human brain transcriptome and genome in a series of neuropathologically normal post-mortem samples and a confirmed pathological diagnosis of late-onset Alzheimer's disease (final $n = 188$ controls, 176 cases). They suggested that studying the transcriptome as a quantitative endophenotype has greater power for discovering risk SNPs that influence expression than the use of discrete diagnostic categories, such as disease presence or absence.[25] It should be noted that the control sample in this study was the same as our discovery brain eQTL sample.[20]

(c) SNPExpress: The authors, using Affymetrix exon arrays, analysed genome-wide SNPs that are associated with gene expression in human primary cells at the exon level, and

https://doi.org/10.1192/bjp.bp.114.156976 Published online by Cambridge University Press

evaluated 93 autopsy-collected, cortical brain tissue samples with no defined neuropsychiatric conditions.[26]

(d) Data from the study by Zou *et al*: The authors of this study measured the expression levels of 24 526 transcripts in brain samples from the cerebellum and temporal cortex of autopsied individuals with Alzheimer's disease (cerebellar $n = 197$, temporal cortex $n = 202$), and conducted an expression GWAS using 213 528 *cis*-SNPs within 100 kb of the tested transcripts. Their results demonstrated the significant contributions of genetic factors to human brain gene expression, which are reliably detected across different brain regions; they also suggested that the combined assessment of expression and disease GWAS might provide complementary information in the discovery of human disease variants with functional implications.[18]

Bipolar disorder samples for replication analysis

Replication analyses on bipolar disorder samples were conducted in two steps (replication-I and -II), examining a total of 6056 bipolar disorder cases and 46 614 controls from 10 different geographical locations. Detailed information on each sample, including diagnostic assessment, genotyping method and quality control, are shown in the online data supplement and online Table DS1.

Briefly, the bipolar disorder samples used in our replication-I analysis included: (a) Germany II (181 cases and 527 controls);[5] (b) Germany III (490 cases and 880 controls);[5] (c) Australia (330 cases and 1811 controls);[5] (d) France (451 cases and 1631 controls);[2] (e) Sweden I (836 cases and 2093 controls);[6] (f) Sweden II sample (1415 cases and 1271 controls);[6] (g) Iceland (541 cases and 34 546 controls);[6] (h) Romania (244 cases and 174 controls),[5] and (i) China (350 cases and 888 controls).[27] For our replication-II analysis, we used a UK sample (1218 cases and 2913 controls).[28] The 10 samples from the replication-I and II analyses showed no overlap with the PGC1 bipolar disorder samples.[6] Each of the original studies was conducted under appropriate ethical approval. Written informed consent was obtained from all participants.

Samples for analysis of hippocampal volume and cognitive performance

For the analysis of hippocampal volume, we used the data from a recent GWAS conducted by the Enhancing Neuro Imaging Genetics through Meta Analysis (ENIGMA) consortium.[29] The GWAS includes a total of 5775 young and healthy individuals (mean age: 34.8 years). Detailed information on the samples, imaging procedures, genotyping methods and statistical analysis can be found in the original GWAS report.[29]

For analysis of cognitive performance, we used a Chinese sample that included 342 healthy Chinese college students from Beijing Normal University who had self-reported no known history of any neurological or psychiatric disorders (197 women and 145 men, aged 18–23). Cognitive and behavioural measures (shown in online Table DS2) included working memory, executive functions (as assessed with the Attention Network Test, the Wisconsin Card Sorting Task and a reversal-learning test) and personality traits. The institutional review board of the State Key Laboratory of Cognitive Neuroscience and Learning at Beijing Normal University, China approved this experiment. Written informed consent was obtained from all participants following a full explanation of the study procedure.

Statistical analysis

For the replication analysis of bipolar disorder, genomic control was used to correct for relatedness and potential population stratification in each sample;[30] association *P*-values and allele-specific odds ratios (ORs) for each individual sample were calculated with a logistic regression model with an additive effect using a lambda value (genomic control) as a covariate to adjust for potential population stratification. Meta-analyses were then conducted based on *Z*-scores by combining data from different samples in the R software package (www.r-project.org) (meta-module) using the Cochran–Mantel–Haenszel test under the fixed-effects model. As described in a previous GWAS meta-analysis,[6] *P*-values for replication samples are reported as one-tailed tests, whereas *P*-values for all combined samples are shown as two-tailed tests. We used a forest plot to graphically present the individual ORs and their 95% confidence intervals (CIs), i.e. each sample was represented by a square in the forest plot. For the analyses on cognitive performance, two-tailed *t*-tests were conducted with SPSS version 16.0 (IBM Corporation, Armonk, New York, USA).

To explain the logic of the study design, a flow chart summarising the analytical methods and showing how variants were taken forward from one stage of analysis to the next is shown in Fig. 1. All protocols and methods used in this study were approved by the institutional review board of the Kunming Institute of Zoology, Chinese Academy of Sciences and adhere to all relevant national and international regulations.

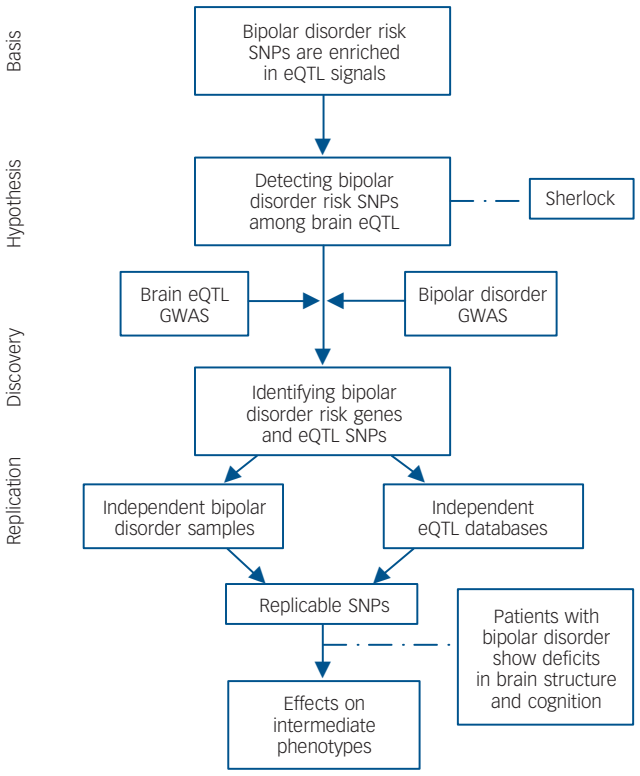

Fig. 1 Flow chart of the present study.

Based on the hypothesis that bipolar disorder risk variants are enriched among eQTL,we systematically integrated bipolar disorder GWAS and genome-wide brain eQTL data with the *Sherlock* software tool. The top genes identified by *Sherlock* were then replicated in independent bipolar disorder samples and eQTL data-sets. Finally, the successfully replicated SNP (rs6088662) was further tested for associations with bipolar disorder phenotypes including hippocampal volume and cognitive performance. SNP, single nucleotide polymorphism; eQTL, expression quantitative trait loci; GWAS, genome-wide association study.

https://doi.org/10.1192/bjp.bp.114.156976 Published online by Cambridge University Press

Results

Integrative analysis of eQTL and bipolar disorder GWAS data

Sherlock identified a total 20 942 SNPs showing significant eQTL effects, and also having bipolar disorder data (e.g. *P*-values), and these SNPs were included for further analyses. Using a Bayesian statistical method to match the 'signature' of genes from the brain eQTL with patterns of association in the bipolar disorder GWAS, we ranked the top candidate genes for bipolar disorder risk according to their LBF scores and *P*-values. Only genes with LBF scores higher than 5.00 were shown and included for further analyses.

The integrative analysis yielded four candidate risk genes (online Table DS3). The first gene was glycosyltransferase 8 domain containing 1 (*GLT8D1*; LBF = 6.78), located on chromosome 3p21.1, which has been repeatedly reported for association with bipolar disorder.[31,32] Detailed analysis revealed that the significant association with this gene was mainly driven by a *cis*-associated SNP (rs2251219). This SNP had already been reported in an earlier GWAS of bipolar disorder,[32] and was replicated in independent bipolar disorder samples. (Their samples overlapped with our replication samples.[33–35]) The second top-ranked gene was chemokine (C-X-C motif) ligand 16 (*CXCL16*; LBF = 6.16), which is located on chromosome 17p13. To the best of our knowledge, this gene has never been reported in genetic association studies on bipolar disorder, and we observed two *trans*-associated SNPs showing moderate associations with bipolar disorder. The third top-ranked gene was transient receptor potential cation channel, subfamily C, member 4 associated protein (*TRPC4AP*; LBF = 5.57) located on chromosome 20q11.22, with the significance mainly driven by a *cis*-associated SNP (rs6088662, $P = 5.85 \times 10^{-5}$ with bipolar disorder). The last top-ranked gene was TAF11 RNA polymerase II, TATA box binding protein (TBP)-associated factor, 28 kDa (*TAF11*; LBF = 5.52) located on chromosome 6p21.31, with a *trans*-associated SNP (rs4482754), which also showed significant association with bipolar disorder.

Replication of eQTL effects in diverse samples

Given the many confounders in a single eQTL database, it is important and necessary to validate the eQTL associations in independent samples. The previously mentioned four candidate genes and their *cis*- or *trans*-associated SNPs were followed up in independent eQTL data-sets.

For the *cis*-associated SNP rs2251219 and *GLT8D1*, we observed significant association in one replication sample of Alzheimer's disease source (online Table DS4),[25] and a marginally significant association in the BrainCloud sample.[24] However, as demonstrated by a previous study,[32] the association of rs2251219 with *GLT8D1* expression in our discovery eQTL sample (Myers *et al* study)[20] may be an artifact, since the probes overlapped with other common SNPs and it could not be replicated in the original cDNA samples of our discovery eQTL data-set by quantitative polymerase chain reaction using probes not overlapping with known SNPs. In addition to *GLT8D1*, we also analysed the expression of other nearby genes around rs2251219; however, no promising findings were observed (Table DS4). For the significant *trans*-eQTL associations in our discovery sample, neither *CXCL16* nor *TAF11* could be validated in any of the replication samples (online Table DS5), implying they might have been generated by chance.

For the *cis*-association between rs6088662 and *TRPC4AP* expression, in the discovery eQTL brain sample,[20] the risk allele G of rs6088662 showed significantly decreased gene expression ($P < 1.0 \times 10^{-8}$, Fig. 2(a)). This pattern was validated in one of the replication samples ($P < 1.0 \times 10^{-8}$),[25] but it should be noted that these replication data include our discovery sample. We therefore re-analysed the result using the non-overlapped Alzheimer's disease patients, and it showed a nominally significant association ($P = 0.023$, Fig. 2(b)). However, rs6088662 showed an opposite effect on *TRPC4AP* expression in our 'first replication sample'. (The risk allele G of rs6088662 showed increased gene expression.) In other replication samples, no significant association between rs6088662 and *TRPC4AP* was observed (online Table DS6).[18,26] These inconsistencies may not be surprising, given a prior report of low-to-moderate overlap between eQTL loci across eQTL studies. (The percentage of overlapped eQTL is from 0 to approximately 35.4% between pairwise brain studies, as shown in Table 4 of the study by McKenzie *et al*[36]). In addition, with the use of several non-brain tissue eQTL databases,[37–39] we also observed significant and consistent associations between rs6088662 and *TRPC4AP* expression. (The *P*-values range from 0.047 to 3.60×10^{-7}; online Figs DS1–DS3.)

To further examine whether rs6088662 is also associated with the expression of other nearby genes, we screened 14 genes in the 20q11.22 region in both discovery and replication eQTL samples (Table DS6). Intriguingly, we observed another gene, gamma-glutamyltransferase 7 (*GGT7*) showing significant association in the discovery sample ($P < 1.0 \times 10^{-7}$; Fig. 3(a)), and it remained significant in the 'first replication sample' with the same direction of effect ($P < 1.0 \times 10^{-8}$; Fig. 3(b)). In other replication samples, the association has also been significant (Webster *et al*[25] and Zou *et al*[18] studies; Fig. 3(c) and Table DS6) or marginally significant (BrainCloud),[24] except for the study by Heinzen *et al*[26] ($P = 0.13$); however, in the sample analysed by Heinzen *et al*, rs6088662 still showed one of the strongest associations with *GGT7* among the genes located on chromosome 20q11.22, and the SNP showed significant or marginally significant associations with the expression of several exons in *GGT7* (online Table DS7), which was not observed in the majority of other nearby genes.

The other genes located on chromosome 20q11.22, *ACSS2*, *MYH7B* and *EDEM2i*, also showed associations in some of the eQTL samples, but the associations were not consistent and these genes are unlikely to be the associated genes (Table DS6). To summarise, from the eQTL analyses in both discovery and replication samples, we have been able to show that rs6088662 is likely to be an authentic eQTL SNP, and we found two potential genes (*GGT7* and *TRPC4AP*) showing an association with this risk SNP.

rs6088662 is associated with bipolar disorder across cohorts

Given the replication of significant associations between rs6088662 and *TRPC4AP* expression, we opted to further analyse this SNP with regard to bipolar disorder risk. In the stage I replication analysis, we analysed rs6088662 in nine independent case–control samples. Although the association between rs6088662 and bipolar disorder did not achieve even nominal significance ($P = 0.05$) in any single cohort, it did show a trend of association in the Germany II and Sweden II samples ($P = 0.08$ and $P = 0.07$ respectively). In the Chinese sample, there was no difference in allele frequencies of this SNP between Han Chinese and Europeans (0.165 *v.* 0.171 for the risk allele G), and the effect size (OR) in the Chinese sample was even higher than in our discovery sample (1.17 *v.* 1.12), the non-significant result being likely due to the limited sample size. When all the replication-I samples were combined, the association *P*-value reached nominal significance level ($P = 4.95 \times 10^{-2}$), with the

https://doi.org/10.1192/bjp.bp.114.156976 Published online by Cambridge University Press

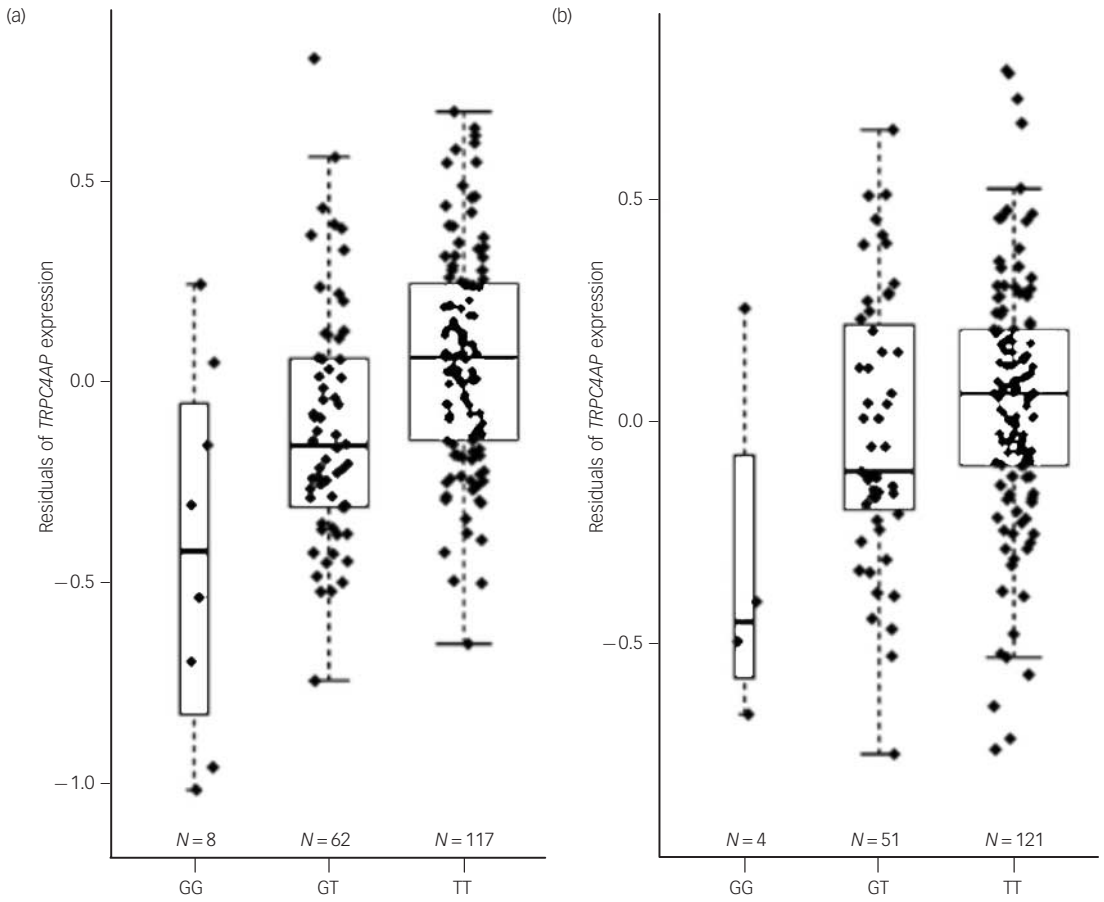

Fig. 2 The risk SNP rs6088662 is significantly associated with *TRPC4AP* mRNA expression. (a) Results in 193 neuropathologically normal human brain (cortical) samples from European individuals. (b) Results in 176 Alzheimer's disease human brain (cortical) samples from European individuals. SNP, single nucleotide polymorphism.

OR being 1.06 (95% CI = 0.99–1.13), consistent with the discovery PGC1 GWAS. There was no significant heterogeneity among the replication-I samples ($P = 0.77$). Detailed results for each individual sample are shown in Table 1. The forest plot of the meta-analysis of all replication-I samples is shown in Fig. 4.

Notably, a previous study[28] reported a significant association of a proxy SNP of rs6088662 (rs13041792, $r^2 = 1.00$ with rs6088662 in Europeans) with bipolar disorder in an independent UK sample (1218 cases and 2913 controls), which is in agreement with our results and was also included in our analysis, denoted as the 'replication-II' sample. Meta-analysis by combining PGC1 GWAS, replication-I and replication-II samples yielded a genome-wide significant association of rs6088662 with bipolar disorder ($P = 3.54 \times 10^{-8}$, OR = 1.12, 95% CI = 1.07–1.16, Table 1). We used the fixed-effects model for meta-analysis because there was no significant heterogeneity among the samples ($P > 0.05$).

Considering the genetic overlap between bipolar disorder and other psychiatric disorders,[1] we also tested the association of rs6088662 with two other mental disorders, schizophrenia and major depressive disorder. It showed a nominally significant association with schizophrenia in the latest PGC2 GWAS ($P = 0.0037$, OR = 1.04, 95% CI = 1.00–1.08, $n = 35\,476/46\,839$);[40] however, it did not show any significant associations with major depressive disorder when using data from the PGC1 major depressive disorder GWAS plus the PsyCoLaus study samples (10 541/11 208) (online Table DS8),[41,42] implying that rs6088662 is likely a psychosis risk SNP rather than a risk SNP for a broader spectrum of mood disorders.

A proxy search for SNPs in high linkage disequilibrium with rs6088662 was performed on the SNP Annotation and Proxy Search (SNAP) database, version 2.2 (www.broadinstitute.org/mpg/snap/ldsearch.php) with the European panel from the 1000 Human Genome project (pilot 1) data-set. This identified 43 SNPs in high linkage disequilibrium ($r^2 > 0.8$) with rs6088662, all of which are located within the *MYH7B* and *TRPC4AP* regions (Fig. 5). Among these, there are one non-synonymous SNP, three synonymous SNPs, one SNP in the 3′ untranslated region (3′ UTR) and one SNP located in the non-coding RNA (ncRNA) region (online Table DS9). However, to identify causal variants for bipolar disorder, further studies are needed.

rs6088662 is associated with hippocampal volume and cognitive performance

To move beyond statistical association with clinical diagnosis and to obtain convergent evidence for an association between rs6088662 and bipolar disorder-related biology, we also performed a series of convergent experiments testing risk-associated SNPs on several intermediate biological phenotypes. The hippocampus is located under the cerebral cortex and it is a region frequently reported to show dysfunction among patients with bipolar disorder.[18,23] We therefore hypothesised that if the identified risk-associated SNP (e.g. rs6088662) affects the anatomy or function of this brain region, then related cognitive deficits, regardless of illness status, should be associated with it. In an exploratory manner, we tested the effects of rs6088662 on the biological phenotypes related to the hippocampus (hippocampal volume and cognitive performance) in healthy individuals.

In the ENIGMA sample, rs6088662 was significantly associated with hippocampal volume across multiple cohorts

https://doi.org/10.1192/bjp.bp.114.156976 Published online by Cambridge University Press

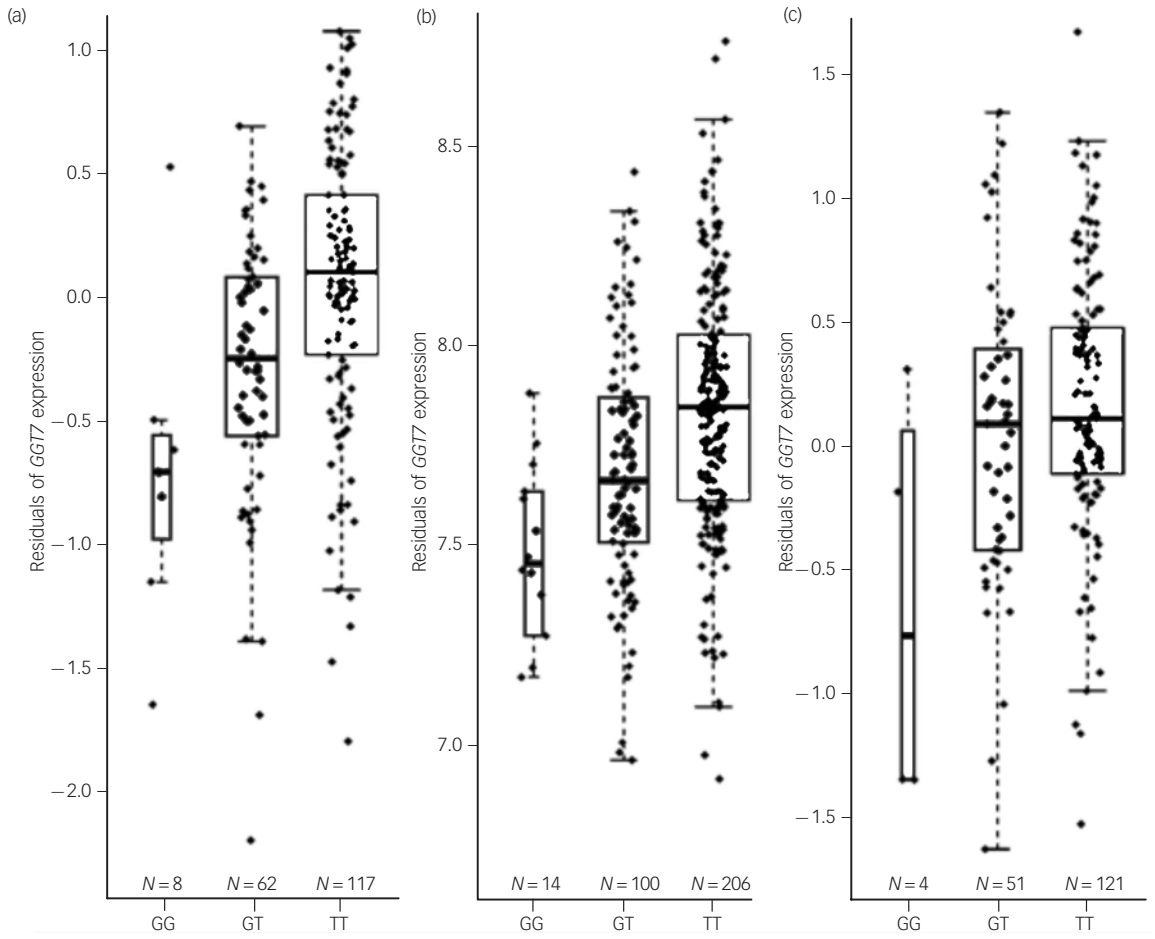

Fig. 3 The risk SNP rs6088662 is significantly associated with *GGT7* mRNA expression. (a) Results in 193 neuropathologically normal human brain (cortical) samples from European individuals. (b) Results in 320 healthy human brain DLPFC samples from White and African-American individuals. (c) Results in 176 Alzheimer's disease human brain (cortical) samples from European individuals. DLPFC, dorsolateral prefrontal cortex.

($P = 0.00063$, $\beta = 27.29\,mm^3$; online Table DS10), supporting the prior hypothesis that bipolar disorder-associated SNPs will likely affect hippocampal structure, although detailed analysis found that the risk allele G led to larger hippocampal volume. As a *post hoc* exploratory test, we then investigated the potential impact of rs6088662 on cognitive performance and found that rs6088662 showed a nominally significant association with executive functions (the alert attention task) ($P = 0.0094$; online Table DS11) and language abilities (visual/auditory) ($P = 0.012$; online Table DS11). Again, the risk allele G indicated a better cognitive performance.

Analysis of bipolar disorder-related phenotypes further confirmed the role of the risk SNPs in bipolar disorder susceptibility and implied it may be functional in the brain. However, as the association results on these intermediate phenotypes (especially for cognitive performance) may not survive multiple correction, further validation in larger samples is needed. In addition, the discrepancy of allelic directionality between clinical diagnosis and intermediate phenotypes suggests that the molecular mechanism at work may be more complicated than we had initially expected when undertaking this study.

Discussion

Findings relating to the 20q11.22 region

In this study, using an integrative analysis that involved both expression and bipolar disorder data, we identified a potential risk locus on chromosome 20q11.22 for bipolar disorder, although it remains unclear which SNPs are actually responsible. This genomic region contains an extensive area of high linkage disequilibrium spanning approximately 276 kb, including at least five protein coding genes (Fig. 5). Of the 43 common SNPs in high linkage disequilibrium ($r^2 > 0.8$) with rs6088662, there is one non-synonymous SNP, three synonymous SNPs, one SNP in the 3′ UTR area and one SNP located in the ncRNA region, all of which are potentially functional but have, as of yet, unknown roles (online Table DS9).

We found a nominally significant association of bipolar disorder risk SNPs with hippocampal volume and cognitive performance, which is consistent with the prevalent perspective that many bipolar disorder-related genes also affect brain structures and cognitive functions. Rather perplexingly though, the rs6088662 risk allele actually seemed to be associated with larger hippocampal volume and better cognition, running entirely opposite to the conventional view that risk alleles generally lead to smaller hippocampal volume and worse cognition. One potential speculative explanation is that the risk genes (*GGT7* or *TRPC4AP*) may play diverse roles in neural development, and the SNP has pleiotropic effects – some detrimental and some beneficial. Another possible explanation is that gene–behaviour association differs by diagnosis status, as previous studies also reported other similar situations: for example, the psychosis risk allele rs1344706 in *ZNF804A* is associated with better cognitive performance in

https://doi.org/10.1192/bjp.bp.114.156976 Published online by Cambridge University Press

Table 1 Summary of logistic regression results for rs6088662 across cohorts

Sample	Ethnicity	Cases	Controls	Effect allele	Additive P[a]	Odds ratio	95% CI	Data source
Discovery								
PGC1	European	7481	9250	G	5.85×10^{-5}	1.12	1.06–1.19	Sklar *et al*[6]
Replication-I								
Germany II	German	181	527	G	0.08	1.23	0.91–1.65	This study
Germany III	German	490	880	G	0.16	1.09	0.90–1.33	This study
Australia	Australian	330	1811	G	0.29	1.07	0.85–1.33	This study
France	French	451	1631	G	0.42	1.02	0.85–1.22	This study
Sweden I	Swedish	836	2093	G	0.37	1.02	0.89–1.18	This study
Sweden II	Swedish	1415	1271	G	0.07	1.12	0.97–1.29	This study
Iceland	Icelandic	541	34 426	G	0.19	0.93	0.79–1.10	This study
Romania	Romanian	244	174	G	0.42	1.04	0.74–1.46	This study
China	Han Chinese	350	888	G	0.11	1.17	0.91–1.50	This study
Total[b]		4838	43 701	G	4.95×10^{-2}	1.06	0.99–1.13	
Replication-II								
UK	British	1218	2913	G	1.06×10^{-6}	1.34	1.19–1.51	Green *et al*[28]
Discovery + replication samples[c]		13 537	55 864	G	3.54×10^{-8}	1.12	1.07–1.16	

CI, confidence interval; PGC, Psychiatric Genomics Consortium.
a. *P*-values are two-sided for the discovery cohort and combined analysis; one-sided *P*-values are listed for the replication-I samples.
b. Heterogeneity test: all replication-I cohorts: $P = 0.77$, $I^2 = 0\%$; meta-analysis was conducted under a fixed-effects model.
c. Heterogeneity test: discovery + replication samples: $P = 0.07$, $I^2 = 41.8\%$; meta-analysis was conducted under a fixed-effects model.

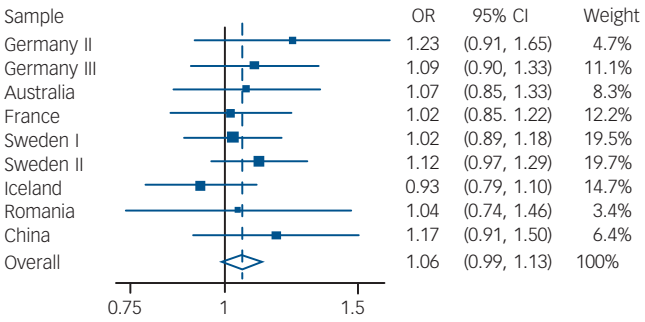

Fig. 4 Forest plot of odds ratios (ORs) with a 95% confidence interval (CI) for total replication-I bipolar disorder samples included in the meta-analysis of rs6088662.

The risk allele G of rs6088662 is overrepresented in bipolar disorder cases in all of the tested cohorts (except for the Icelandic sample).

patients with schizophrenia as seen in two independent samples.[43,44] Likewise, another psychosis risk SNP (rs1006737) in *CACNA1C* was shown to be associated with larger grey matter volume for those with the risk allele.[45,46]

Additional evidence of *GGT7* and *TRPC4AP* in bipolar disorder

TRPC4AP is known to be a substrate-specific adapter of a double-cortin (DCX; DDB1-CUL4-X-box) E3 ubiquitin ligase complex required for cell-cycle control, and *GGT7* is a member of a gene family that encodes enzymes involved in the metabolism of glutathione and in the transpeptidation of amino acids; however, their roles in bipolar disorder susceptibility are still unclear. Here we studied the spatial expression profiling of *GGT7* and *TRPC4AP* in multiple human tissues to see whether their expression was enriched in brain tissues, as bipolar disorder is a mental disorder that mainly originates from abnormal brain function, and if these genes are preferentially expressed in the brain, which would make more sense when considering them as potential risk genes for bipolar disorder. We used the expression data from the Genotype-Tissue Expression project,[47] in which 3797 tissues from 150 post-mortem donors have been collected and subsequently analysed using an RNA sequencing-based gene expression approach. Notably, we found that *GGT7* is abundantly expressed in human brain tissues, such as the cerebellum (online Fig. DS4(a)), whereas the expression level of *GGT7* is generally low in non-neural tissues. However, the expression of *TRPC4AP* in brain tissue is relatively lower than in other tissues (online Fig. DS4(b)), but this gene has been previously reported in association with Alzheimer's disease,[48,49] a neurological disorder showing a high comorbidity with affective disorders, such as bipolar disorder and major depressive disorder, in geriatric populations.[50]

Implications

Alongside our specific findings for genetic susceptibility to bipolar disorder, our results highlight several advantages of convergent analysis using bipolar disorder and eQTL GWAS data-sets (Fig. 1) over conventional analytical strategies aimed at uncovering susceptibility genes. First, analyses such as ours may identify genes that may be missed by traditional univariate analytical approaches, because these genes tend to be authentic risk genes but with small effects. Second, the identification of eQTL effects of the risk SNPs could provide insights for future focused studies, since conventional analyses often observe a large linkage disequilibrium region containing numerous genes showing association with the illness, but actually determining which one is the susceptibility gene is, at best, difficult. Third, significant association between eQTL and illness has been consistently replicated across independent data-sets, providing convergent validity for findings and suggesting potentially higher reproducibility for this kind of system level analysis. Given these advantages, it is likely that further studies using similar methods will strengthen the case for such studies in trying to uncover genetic risk factors for psychiatric diseases.

Study limitations

Although this study offers some interesting observations, it should be noted that the present evidence is limited, and we are therefore cautious in interpreting these results:

(a) In the integrative analysis of bipolar disorder and eQTL GWAS data, we arbitrarily selected genes that were scored higher than 5.0 (LBF score). As such, it is possible that some genes that

https://doi.org/10.1192/bjp.bp.114.156976 Published online by Cambridge University Press

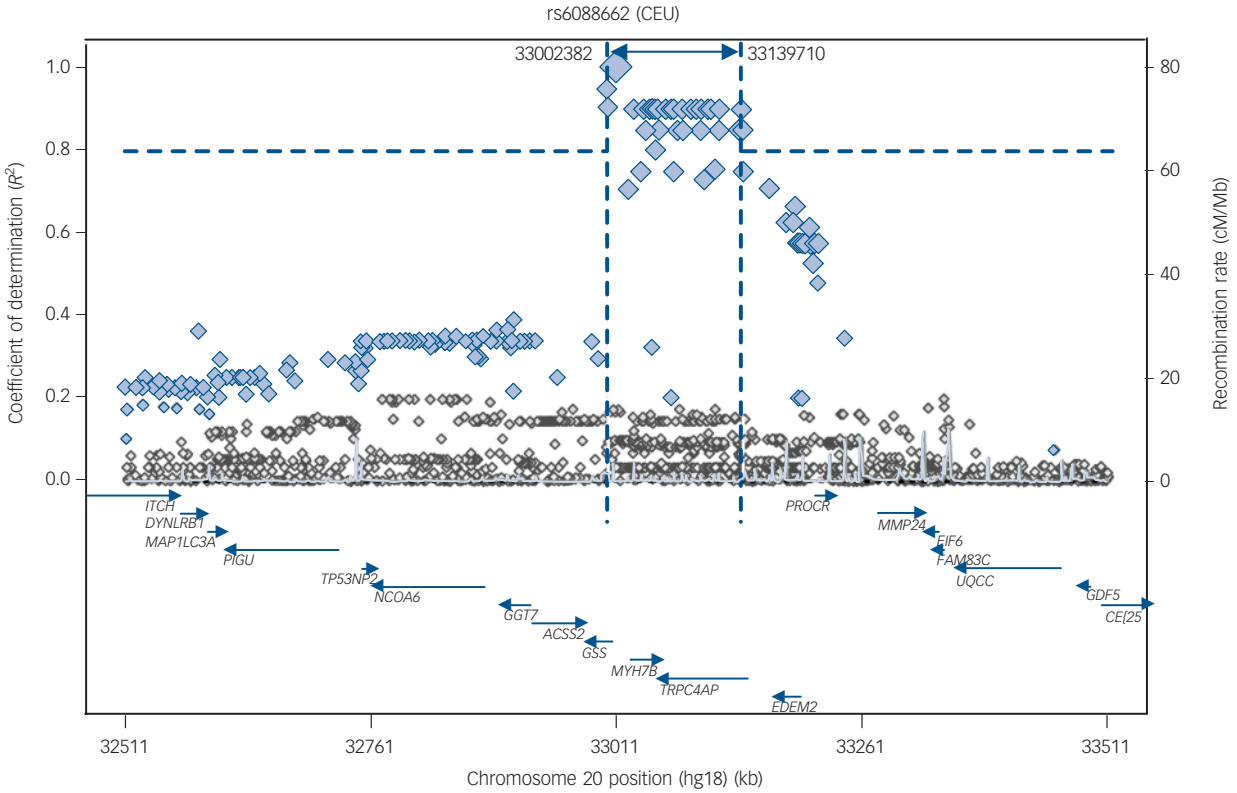

Fig. 5 Plot of chromosome region showing a genomic area of high linkage disequilibrium with rs6088662 in European populations. CEU, Northwestern Europe.

may contribute to bipolar disorder risk but did not meet our selection criteria could have been missed.

(b) Similarly, although we used GWAS data in our analysis, SNP coverage is still relatively low and other true risk SNPs may have been missed. Due to the dearth of functional data, it is difficult to identify the causative variant(s).

(c) Likewise, we cannot exclude the possibility that the positive association signal was actually caused by the hitchhiking effect of rare missense mutations, copy number variations or variants in a distant region. Further focused studies may provide a more complete survey.

(d) The SNPs in the discovery eQTL sample were not imputed, thus reducing the overlap between eQTL and GWAS data-sets and the power of our method, although we believe the obtained results are valuable.

(e) The gene expression coverage in the discovery eQTL data-set is relatively low, and we cannot exclude the possibility of other missing risk genes during the integrative analyses, although we conducted a comprehensive replication and fine mapping analyses to localise the actual risk genes. Further studies using a high coverage array or RNA sequencing are warranted.

(f) It also should be acknowledged that the eQTL databases we used are highly variable, in terms of expression platforms and tissue quality, age and diagnoses. It is highly likely that biological factors mediating eQTL associations, such as epigenetic regulation, transcription factor binding and microRNA dynamics will vary across age and diagnosis.

(g) We would also like to note that our results reached genome-wide significance in the final meta-analysis of our ten new samples added to the public bipolar disorder data-set. Our understanding of the association of rs6088662 with bipolar disorder and with gene expression and hippocampal biology might have started first with the combined GWAS result, but this was not our strategy.

In conclusion, our data from large-scale samples support that SNPs located on chromosome 20q11.22 are significantly associated with bipolar disorder. We observed associations with *GGT7* and *TRPC4AP* mRNA expression, hippocampal volume and cognitive performance. Although the actual risk gene(s) for bipolar disorder in this genomic region are yet to be determined, future studies may give a more compelling picture on the association between these potential risk factors and genetic susceptibility to bipolar disorder.

Funding

This work was supported by grants from the National 973 project of China (2011CBA00401), the National Natural Science Foundation of China (U1202225, 31130051, 31071101 and 31221003), the German Federal Ministry of Education and Research, the National Genome Research Network, Germany, the Integrated Genome Research Network MooDS, Germany (grant 01GS08144 to S.C. and M.M.N., grant 01GS08147 to M.R. and T.G.S.), the 111 Project (B07008) of the Ministry of Education of China, the Strategic Priority Research Program (B) of the Chinese Academy of Sciences (XDB02020000), the National Authority for Scientific Research, Bucharest, Romania (UEFISCDI – PN-II-89/2012) and Personal Genetics SRL, Bucharest, Romania.

Acknowledgements

We would like to acknowledge the efforts of the Psychiatric Genomics Consortium Bipolar Disorder Working Group for their contribution to this study. We are deeply grateful to Stacy Steinberg, Hreinn Stefansson, Kari Stefansson, Thorgeir Thorgeirsson (deCODE Genetics, Reykjavík, Iceland) and Engilbert Sigurdsson (Landspitali University Hospital, Reykjavík, Iceland) for their results from the Icelandic samples. We are also grateful to Andrew Willden (Kunming Institute of Zoology, Kunming, Yunnan, China) for the language editing of the manuscript.

https://doi.org/10.1192/bjp.bp.114.156976 Published online by Cambridge University Press

Ming Li, PhD, State Key Laboratory of Genetic Resources and Evolution, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming, Yunnan, China, and Lieber Institute for Brain Development, Johns Hopkins University, Baltimore, Maryland, USA; **Xiong-jian Luo**, PhD, Key Laboratory of Animal Models and Human Disease Mechanisms of the Chinese Academy of Sciences and Yunnan Province, Kunming Institute of Zoology, Kunming, Yunnan, China; **Mikael Landén**, MD, PhD, Department of Medical Epidemiology and Biostatistics, Karolinska Institute, Stockholm, and Section of Psychiatry and Neurochemistry, Sahlgrenska Academy at Gothenburg University, Gothenburg, Sweden; **Sarah E. Bergen**, PhD, Department of Medical Epidemiology and Biostatistics, Karolinska Institute, Stockholm, Sweden, and Stanley Center for Psychiatric Research, Broad Institute of Harvard and MIT, Cambridge, Massachusetts, USA; **Christina M. Hultman**, PhD, Department of Medical Epidemiology and Biostatistics, Karolinska Institute, Stockholm, Sweden; **Xiao Li**, MSc, **Wen Zhang**, PhD, **Yong-Gang Yao**, PhD, Key Laboratory of Animal Models and Human Disease Mechanisms of the Chinese Academy of Sciences and Yunnan Province, Kunming Institute of Zoology, Kunming, Yunnan, China; **Chen Zhang**, PhD, Schizophrenia Program, Shanghai Mental Health Center, and Shanghai Jiao Tong University School of Medicine, Shanghai, China; **Jiewei Liu**, MSc, State Key Laboratory of Genetic Resources and Evolution, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming, Yunnan, and Kunming College of Life Science, University of Chinese Academy of Sciences, Kunming, Yunnan, China; **Manuel Mattheisen**, MD, Department of Biomedicine, Aarhus University, Aarhus C, Denmark; **Sven Cichon**, PhD, Division of Medical Genetics, University of Basel, Basel, Switzerland, Institute of Human Genetics and Department of Genomics, Life and Brain Center, University of Bonn, Bonn, and Institute of Neuroscience and Medicine (INM-1), Structural and Functional Organization of the Brain, Genomic Imaging, Research Centre Jülich, Germany; **Thomas W. Mühleisen**, PhD, Institute of Human Genetics and Department of Genomics, Life and Brain Center, University of Bonn, Bonn, and Institute of Neuroscience and Medicine (INM-1), Structural and Functional Organization of the Brain, Genomic Imaging, Research Centre Jülich, Germany; **Franziska A. Degenhardt**, MD, Institute of Human Genetics and Department of Genomics, Life and Brain Center, University of Bonn, Bonn, Germany; **Markus M. Nöthen**, MD, Institute of Human Genetics and Department of Genomics, Life and Brain Center, University of Bonn, Bonn, and German Center for Neurodegenerative Diseases (DZNE), Bonn, Germany; **Thomas G. Schulze**, MD, Department of Genetic Epidemiology in Psychiatry, Central Institute of Mental Health, Medical Faculty Mannheim/University of Heidelberg, Mannheim, Section on Psychiatric Genetics, Department of Psychiatry and Psychotherapy, University Medical Center, Georg-August-University, Göttingen, and Institute of Psychiatric Phenomics and Genomics (IPPG), Ludwig-Maximilians-University Munich, Munich, Germany; **Maria Grigoroiu-Serbanescu**, PhD, Biometric Psychiatric Genetics Research Unit, Alexandru Obregia Clinical Psychiatric Hospital, Bucharest, Romania; **Hao Li**, PhD, **Chris K. Fuller**, MSc, Department of Biochemistry and Biophysics, University of California at San Francisco, San Francisco, California, USA; **Chunhui Chen**, PhD, **Qi Dong**, PhD, State Key Laboratory of Cognitive Neuroscience and Learning & IDG/McGovern Institute for Brain Research, Beijing Normal University, Beijing, and Center for Collaboration and Innovation in Brain and Learning Sciences, Beijing Normal University, Beijing, China; **Chuansheng Chen**, PhD, Department of Psychology and Social Behavior, University of California, Irvine, California, USA; **Stéphane Jamain**, PhD, Inserm U 955, IMRB, Psychiatrie Génétique, Créteil, Université Paris Est, Faculté de Médecine, and Fondation Fondamental, Créteil, France; **Marion Leboyer**, MD, PhD, Inserm U 955, IMRB, Psychiatrie Génétique, Créteil, Université Paris Est, Faculté de Médecine, Fondation Fondamental, Créteil, and AP-HP, Hôpital A. Chenevier – H. Mondor, Pôle de Psychiatrie, Créteil, France; **Frank Bellivier**, MD, PhD, Inserm U 955, IMRB, Psychiatrie Génétique, and Fondation Fondamental, Créteil, AP-HP, Groupe hospitalier Lariboisière – F. Widal, Pôle de Psychiatrie, Paris, and Université Paris Diderot, Paris, France; **Bruno Etain**, MD, PhD, Inserm U 955, IMRB, Psychiatrie Génétique, Créteil, Université Paris Est, Faculté de Médecine, Créteil, Fondation Fondamental, Créteil, and AP-HP, Hôpital A. Chenevier – H. Mondor, Pôle de Psychiatrie, Créteil, France; **Jean-Pierre Kahn**, MD, PhD, Fondation Fondamental, Créteil, and Département de Psychiatrie et de Psychologie Clinique, CHU de Nancy, Hôpital Jeanne d'Arc, Toul, France; **Chantal Henry**, MD, PhD, Inserm U 955, IMRB, Psychiatrie Génétique, Créteil, Université Paris Est, Faculté de Médecine, Créteil, Fondation Fondamental, Créteil, and AP-HP, Hôpital A. Chenevier – H. Mondor, Pôle de Psychiatrie, Créteil, France; **Martin Preisig**, MD, MPH, Department of Psychiatry, Centre Hospitalier Universitaire Vaudois, Prilly, Switzerland; **Zoltán Kutalik**, PhD, Institute of Social and Preventive Medicine, Centre Hospitalier Universitaire Vaudois, Lausanne, and Swiss Institute of Bioinformatics, Lausanne, Switzerland; **Enrique Castelao**, MSc, Department of Psychiatry, Centre Hospitalier Universitaire Vaudois, Prilly, Switzerland; **Adam Wright**, PhD, **Philip B. Mitchell**, MD, School of Psychiatry, University of New South Wales, Sydney, and Black Dog Institute, Prince of Wales Hospital, Sydney, Australia; **Janice M. Fullerton**, PhD, **Peter R. Schofield**, PhD, Neuroscience Research Australia, Sydney and School of Medical Sciences, Faculty of Medicine, University of New South Wales, Sydney, Australia; **Grant W. Montgomery**, PhD, **Sarah E. Medland**, PhD, **Scott D. Gordon**, PhD, **Nicholas G. Martin**, PhD, MooDS Consortium and The Swedish Bipolar Study Group, Queensland Institute of Medical Research, Brisbane, Australia; **Marcella Rietschel**, MD, PhD, Department of Genetic Epidemiology in Psychiatry, Central Institute of Mental Health, Medical Faculty Mannheim/University of Heidelberg, Mannheim, Germany; **Chunyu Liu**, PhD, Department of Psychiatry, University of Illinois at Chicago, Chicago, Illinois, USA; **Joel E. Kleinman**, MD, **Thomas M. Hyde**, MD, **Daniel R. Weinberger**, MD, Lieber Institute for Brain Development, Johns Hopkins University, Baltimore, Maryland, USA; **Bing Su**, PhD, State Key Laboratory of Genetic Resources and Evolution, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming, Yunnan, China

Correspondence: Bing Su, Kunming Institute of Zoology, Chinese Academy of Sciences, 32 East Jiao-chang Road, Kunming 650223, Yunnan, China. Email: sub@mail.kiz.ac.cn; Ming Li, Lieber Institute for Brain Development, Johns Hopkins University, Baltimore, MD, USA. Email: limingkiz@gmail.com

First received 20 Apr 2014, final revision 1 Sep 2014, accepted 21 Oct 2014

References

1 Craddock N, Jones I. Genetics of bipolar disorder. *J Med Genet* 1999; **36**: 585–94.

2 Etain B, Dumaine A, Mathieu F, Chevalier F, Henry C, Kahn JP, et al. A SNAP25 promoter variant is associated with early-onset bipolar disorder and a high expression level in brain. *Mol Psychiatry* 2010; **15**: 748–55.

3 Li M, Luo XJ, Rietschel M, Lewis CM, Mattheisen M, Müller-Myhsok B, et al. Allelic differences between Europeans and Chinese for CREB1 SNPs and their implications in gene expression regulation, hippocampal structure and function, and bipolar disorder susceptibility. *Mol Psychiatry* 2014; **19**: 452–61.

4 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**: 195–205.

5 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**: 372–81.

6 Sklar P, Ripke S, Scott LJ, Andreassen OA, Cichon S, Craddock N, et al. Large-scale genome-wide association analysis of bipolar disorder identifies a new susceptibility locus near ODZ4. *Nat Genet* 2011; **43**: 977–83.

7 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**: 1056–8.

8 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.

9 Purcell SM, Wray NR, Stone JL, Visscher PM, O'Donovan MC, Sullivan PF, et al. Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature* 2009; **460**: 748–52.

10 Steinberg S, de Jong S, Irish Schizophrenia Genomics Consortium, Andreassen OA, Werge T, Børglum AD, et al. Common variants at VRK2 and TCF4 conferring risk of schizophrenia. *Hum Mol Genet* 2011; **20**: 4076–81.

11 Quraishi S, Walshe M, McDonald C, Schulze K, Kravariti E, Bramon E, et al. Memory functioning in familial bipolar I disorder patients and their relatives. *Bipolar Disord* 2009; **11**: 209–14.

12 Rimol LM, Hartberg CB, Nesvåg R, Fennema-Notestine C, Hagler Jr DJ, Pung CJ, et al. Cortical thickness and subcortical volumes in schizophrenia and bipolar disorder. *Biol Psychiatry* 2010; **68**: 41–50.

13 Haukvik UK, Westlye LT, Mørch-Johnsen L, Jørgensen KN, Lange EH, Dale AM, et al. In vivo hippocampal subfield volumes in schizophrenia and bipolar disorder. *Biol Psychiatry* 2015; **77**: 581–8.

14 Phillips ML, Ladouceur CD, Drevets WC. A neural model of voluntary and automatic emotion regulation: implications for understanding the pathophysiology and neurodevelopment of bipolar disorder. *Mol Psychiatry* 2008; **13**: 829, 833–57.

15 Frey BN, Andreazza AC, Nery FG, Martins MR, Quevedo J, Soares JC, et al. The role of hippocampus in the pathophysiology of bipolar disorder. *Behav Pharmacol* 2007; **18**: 419–30.

16 Conde L, Bracci PM, Richardson R, Montgomery SB, Skibola CF. Integrating GWAS and expression data for functional characterization of disease-associated SNPs: an application to follicular lymphoma. *Am J Hum Genet* 2013; **92**: 126–30.

17 He X, Fuller CK, Song Y, Meng Q, Zhang B, Yang X, et al. Sherlock: detecting gene-disease associations by matching patterns of expression QTL and GWAS. *Am J Hum Genet* 2013; **92**: 667–80.

18 Zou F, Chai HS, Younkin CS, Allen M, Crook J, Pankratz VS, et al. Brain expression genome-wide association study (eGWAS) identifies human disease-associated variants. *PLoS Genet* 2012; **8**: e1002707.

19 Gamazon ER, Badner JA, Cheng L, Zhang C, Zhang D, Cox NJ, et al. Enrichment of cis-regulatory gene expression SNPs and methylation quantitative trait loci among bipolar disorder susceptibility variants. *Mol Psychiatry* 2013; **18**: 340–6.

20 Myers AJ, Gibbs JR, Webster JA, Rohrer K, Zhao A, Marlowe L, et al. A survey of genetic human cortical gene expression. *Nat Genet* 2007; **39**: 1494–9.

21 Morita Y, Callicott JH, Testa LR, Mighdoll MI, Dickinson D, Chen Q, et al. Characteristics of the cation cotransporter NKCC1 in human brain: alternate transcripts, expression in development, and potential relationships to brain function and schizophrenia. *J Neurosci* 2014; **34**: 4929–40.

22 Nakata K, Lipska BK, Hyde TM, Ye T, Newburn EN, Morita Y, et al. DISC1 splice variants are upregulated in schizophrenia and associated with risk polymorphisms. *Proc Natl Acad Sci U S A* 2009; **106**: 15873–8.

https://doi.org/10.1192/bjp.bp.114.156976 Published online by Cambridge University Press

23 Tao R, Li C, Newburn EN, Ye T, Lipska BK, Herman MM, et al. Transcript-specific associations of SLC12A5 (KCC2) in human prefrontal cortex with development, schizophrenia, and affective disorders. *J Neurosci* 2012; **32**: 5216–22.

24 Colantuoni C, Lipska BK, Ye T, Hyde TM, Tao R, Leek JT, et al. Temporal dynamics and genetic control of transcription in the human prefrontal cortex. *Nature* 2011; **478**: 519–23.

25 Webster JA, Gibbs JR, Clarke J, Ray M, Zhang W, Holmans P, et al. Genetic control of human brain transcript expression in Alzheimer disease. *Am J Hum Genet* 2009; **84**: 445–58.

26 Heinzen EL, Ge D, Cronin KD, Maia JM, Shianna KV, Gabriel WN, et al. Tissue-specific genetic control of splicing: implications for the study of complex traits. *PLoS Biol* 2008; **6**: e1.

27 Zhang X, Zhang C, Wu Z, Wang Z, Peng D, Chen J, et al. Association of genetic variation in CACNA1C with bipolar disorder in Han Chinese. *J Affect Disord* 2013; **150**: 261–5.

28 Green EK, Hamshere M, Forty L, Gordon-Smith K, Fraser C, Russell E, 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* 2013; **18**: 1302–7.

29 Stein JL, Medland SE, Vasquez AA, Hibar DP, Senstad RE, Winkler AM, et al. Identification of common variants associated with human hippocampal and intracranial volumes. *Nat Genet* 2012; **44**: 552–61.

30 Devlin B, Roeder K. Genomic control for association studies. *Biometrics* 1999; **55**: 997–1004.

31 Scott LJ, Muglia P, Kong XQ, Guan W, Flickinger M, Upmanyu R, et al. Genome-wide association and meta-analysis of bipolar disorder in individuals of European ancestry. *Proc Natl Acad Sci U S A* 2009; **106**: 7501–6.

32 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**: 128–31.

33 Breen G, Lewis CM, Vassos E, Pergadia ML, Blackwood DH, Boomsma DI, et al. Replication of association of 3p21.1 with susceptibility to bipolar disorder but not major depression. *Nat Genet* 2011; **43**: 3–5.

34 Kondo K, Ikeda M, Kajio Y, Saito T, Iwayama Y, Aleksic B, et al. Genetic variants on 3q21 and in the Sp8 transcription factor gene (SP8) as susceptibility loci for psychotic disorders: a genetic association study. *PLoS One* 2013; **8**: e70964.

35 Vassos E, Steinberg S, Cichon S, Breen G, Sigurdsson E, Andreassen OA, et al. Replication study and meta-analysis in European samples supports association of the 3p21.1 locus with bipolar disorder. *Biol Psychiatry* 2012; **72**: 645–50.

36 McKenzie M, Henders AK, Caracella A, Wray NR, Powell JE. Overlap of expression quantitative trait loci (eQTL) in human brain and blood. *BMC Med Genomics* 2014; **7**: 31.

37 Dimas AS, Deutsch S, Stranger BE, Montgomery SB, Borel C, Attar-Cohen H, et al. Common regulatory variation impacts gene expression in a cell type-dependent manner. *Science* 2009; **325**: 1246–50.

38 Nica AC, Parts L, Glass D, Nisbet J, Barrett A, Sekowska M, et al. The architecture of gene regulatory variation across multiple human tissues: the MuTHER study. *PLoS Genet* 2011; **7**: e1002003.

39 Stranger BE, Montgomery SB, Dimas AS, Parts L, Stegle O, Ingle CE, et al. Patterns of cis regulatory variation in diverse human populations. *PLoS Genet* 2012; **8**: e1002639.

40 Ripke S, Neale BM, Corvin A, Walters JT, Farh KH, Holmans PA, et al. Biological insights from 108 schizophrenia-associated genetic loci. *Nature* 2014; **511**: 421–7.

41 Ripke S, Wray NR, Lewis CM, Hamilton SP, Weissman MM, Breen G, et al. A mega-analysis of genome-wide association studies for major depressive disorder. *Mol Psychiatry* 2013; **18**: 497–511.

42 Preisig M, Waeber G, Vollenweider P, Bovet P, Rothen S, Vandeleur C, et al. The PsyCoLaus study: methodology and characteristics of the sample of a population-based survey on psychiatric disorders and their association with genetic and cardiovascular risk factors. *BMC Psychiatry* 2009; **9**: 9.

43 Walters JT, Corvin A, Owen MJ, Williams H, Dragovic M, Quinn EM, et al. Psychosis susceptibility gene *ZNF804A* and cognitive performance in schizophrenia. *Arch Gen Psychiatry* 2010; **67**: 692–700.

44 Chen M, Xu Z, Zhai J, Bao X, Zhang Q, Gu H, et al. Evidence of IQ-modulated association between *ZNF804A* gene polymorphism and cognitive function in schizophrenia patients. *Neuropsychopharmacology* 2012; **37**: 1572–8.

45 Wang F, McIntosh AM, He Y, Gelernter J, Blumberg HP. The association of genetic variation in CACNA1C with structure and function of a frontotemporal system. *Bipolar Disord* 2011; **13**: 696–700.

46 Perrier E, Pompei F, Ruberto G, Vassos E, Collier D, Frangou S. Initial evidence for the role of CACNA1C on subcortical brain morphology in patients with bipolar disorder. *Eur Psychiatry* 2011; **26**: 135–7.

47 Lonsdale J, Thomas J, Salvatore M, Phillips R, Lo E, Shad S, et al. The Genotype-Tissue Expression (GTEx) project. *Nat Genet* 2013; **45**: 580–5.

48 Poduslo SE, Huang R, Huang J. The frequency of the TRPC4AP haplotype in Alzheimer's patients. *Neurosci Lett* 2009; **450**: 344–6.

49 Poduslo SE, Huang R, Huang J, Smith S. Genome screen of late-onset Alzheimer's extended pedigrees identifies TRPC4AP by haplotype analysis. *Am J Med Genet B Neuropsychiatr Genet* 2009; **150B**: 50–5.

50 Teipel SJ, Walter M, Likitjaroen Y, Schönknecht P, Gruber O. Diffusion tensor imaging in Alzheimer's disease and affective disorders. *Eur Arch Psychiatry Clin Neurosci* 2014; **264**: 467–83.

EXTRA CONTENT ONLINE

https://doi.org/10.1192/bjp.bp.114.156976 Published online by Cambridge University Press