Molecular Psychiatry (2018) 23, 1957–1967

www.nature.com/mp

ORIGINAL ARTICLE

The cAMP responsive element-binding (CREB)-1 gene increases risk of major psychiatric disorders

X Xiao[1], C Zhang[2], M Grigoroiu-Serbanescu[3], L Wang[1], L Li[1], D Zhou[4], T-F Yuan[5], C Wang[6], H Chang[1], Y Wu[1], Y Li[7], D-D Wu[8], Y-G Yao[1,9] and M Li[1,9]

Bipolar disorder (BPD), schizophrenia (SCZ) and unipolar major depressive disorder (MDD) are primary psychiatric disorders sharing substantial genetic risk factors. We previously reported that two single-nucleotide polymorphisms (SNPs) rs2709370 and rs6785 in the cAMP responsive element-binding (CREB)-1 gene (*CREB1*) were associated with the risk of BPD and abnormal hippocampal function in populations of European ancestry. In the present study, we further expanded our analyses of rs2709370 and rs6785 in multiple BPD, SCZ and MDD data sets, including the published Psychiatric Genomics Consortium (PGC) genome-wide association study, the samples used in our previous *CREB1* study, and six additional cohorts (three new BPD samples, two new SCZ samples and one new MDD sample). Although the associations of both *CREB1* SNPs with each illness were not replicated in the new cohorts (BPD analysis in 871 cases and 1089 controls (rs2709370, $P = 0.0611$; rs6785, $P = 0.0544$); SCZ analysis in 1273 cases and 1072 controls (rs2709370, $P = 0.230$; rs6785, $P = 0.661$); and MDD analysis in 129 cases and 100 controls (rs2709370, $P = 0.114$; rs6785, $P = 0.188$)), an overall meta-analysis of all included samples suggested that both SNPs were significantly associated with increased risk of BPD (11 105 cases and 51 331 controls; rs2709370, $P = 2.33 \times 10^{-4}$; rs6785, $P = 6.33 \times 10^{-5}$), SCZ (34 913 cases and 44 528 controls; rs2709370, $P = 3.96 \times 10^{-5}$; rs6785, $P = 2.44 \times 10^{-5}$) and MDD (9369 cases and 9619 controls; rs2709370, $P = 0.0144$; rs6785, $P = 0.0314$), with the same direction of allelic effects across diagnostic categories. We then examined the impact of diagnostic status on *CREB1* mRNA expression using data obtained from independent brain tissue samples, and observed that the mRNA expression of *CREB1* was significantly downregulated in psychiatric patients compared with healthy controls. The protein–protein interaction analyses showed that the protein encoded by *CREB1* directly interacted with several risk genes of psychiatric disorders identified by GWAS. In conclusion, the current study suggests that *CREB1* might be a common risk gene for major psychiatric disorders, and further investigations are necessary.

Molecular Psychiatry advance online publication, 21 November 2017; doi:10.1038/mp.2017.243

INTRODUCTION

The putative similarities between affective and non-affective major psychoses have been proposed since the birth of the research field.[1–3] Shared risk components among the major psychiatric disorders have been initially reported by epidemiological studies. For example, the prevalence of unipolar major depressive disorder (MDD) in family members of an individual diagnosed with bipolar disorder (BPD) is usually higher than that in the general population.[4] Twin studies also suggest that shared genetic susceptibility factors exist between BPD and MDD.[5] Several family studies have also shown that the morbid risk of schizophrenia (SCZ) or all major psychiatric disorders (BPD, MDD, SCZ and schizoaffective disorders) within the families of bipolar probands is greater than expected,[6–8] which is supported by a recent cross-disorder genome-wide association study (GWAS).[9]

In addition to these lines of initial evidence of shared genetic components among major psychiatric disorders, recent molecular genetic analyses have further shown that particular genes, biological processes and systems may contribute to the susceptibility to these illnesses across the diagnostic categories.[10–12] For instance, single-nucleotide polymorphisms (SNPs) in certain genes (for example, *PCLO*, *PCDH17* and *CACNA1C*) as well as the genetic loci at chromosome 3p21.1 have been shown to confer susceptibility to both BPD and MDD.[13–18] Similarly, genes and genomic regions at specific loci, including *CACNA1C*, *ZNF804A*, *NRGN* and the extended major histocompatibility complex region likely have important roles in the pathogenesis of both SCZ and BPD.[19–24]

Indeed, numerous genetic association analyses using a large number of subjects with various psychiatric disorders have been conducted to examine the potential shared genetic risk factors

[1]Key Laboratory of Animal Models and Human Disease Mechanisms of the Chinese Academy of Sciences and Yunnan Province, Kunming Institute of Zoology, Kunming, China; [2]Schizophrenia Program, Shanghai Mental Health Center, Shanghai Jiao Tong University School of Medicine, Shanghai, China; [3]Biometric Psychiatric Genetics Research Unit, Alexandru Obregia Clinical Psychiatric Hospital, Bucharest, Romania; [4]Ningbo Kangning Hospital, Ningbo, China; [5]Shanghai Key Laboratory of Psychotic Disorders, Shanghai Mental Health Center, Shanghai Jiao Tong University School of Medicine, Shanghai, China; [6]Department of Pharmacology, and Provincial Key Laboratory of Pathophysiology in Ningbo University School of Medicine, Ningbo, China; [7]Laboratory for Conservation and Utilization of Bio-Resource, Yunnan University, Kunming, China; [8]State Key Laboratory of Genetic Resources and Evolution, Chinese Academy of Sciences, Kunming, China and [9]CAS Center for Excellence in Brain Science and Intelligence Technology, Chinese Academy of Sciences, Shanghai, China. Correspondence: Dr C Zhang, Schizophrenia Program, Shanghai Mental Health Center, Shanghai Jiao Tong University School of Medicine, 600 Wan Ping Nan Road, Shanghai 200030, China or Professor M Grigoroiu-Serbanescu, Biometric Psychiatric Genetics Research Unit, Alexandru Obregia Clinical Psychiatric Hospital, 10, Sos. Berceni, R-041914 Bucharest, Romania or Professor M Li, Key Laboratory of Animal Models and Human Disease Mechanisms of the Chinese Academy of Sciences and Yunnan Province, Kunming Institute of Zoology, NO 32 Jiao-Chang Donglu, Kunming 650223, China.
E-mail: zhangchen645@gmail.com or maria.serbanescu@googlemail.com or limingkiz@mail.kiz.ac.cn
Received 4 May 2017; revised 25 August 2017; accepted 14 September 2017

among different psychiatric conditions.[9,11,17] However, many of such studies failed to provide compelling evidence of genetic risk factors reaching the genome-wide significance, partially due to the polygenic nature of psychiatric disorders, in which case numerous risk genes may individually confer only limited effects.[25] However, lack of compelling statistical evidence does not deny the possibility that particular genes are true risk factors for these illnesses. In fact, previous aggregated analyses indicated that markers reaching nominal significance in the GWAS of psychiatric disorders were still bona fide risk factors if they showed consistent significant associations (though may not be on the genome-wide level) with psychiatric illnesses in later replication studies using independent samples.[26] Examples of such markers include *CMYA5*, *VRK2* and *PCDH17*,[13,27–33] and functional investigations have already revealed potential biological mechanisms underlying their effects on the onset of the diseases.[34]

In line with this theory, we previously identified a BPD susceptibility gene *CREB1* through large-scale meta-analytic studies of clinical samples followed by additional analyses of related neuroimaging phenotypes.[35] The protein encoded by *CREB1* gene, the transcription factor cyclic adenosine monophosphate (cAMP) responsive element-binding protein 1, has a critical role in the cAMP signaling pathway, which is found to be impaired in most patients with psychiatric disorders.[36,37] Consistent with previous animal behavioral studies, researchers have shown that *CREB1* is also involved in anxiety and MDD.[38] Moreover, this gene is implicated in early-onset familial MDD,[39] antidepressant response,[40] anger,[41] neuronal plasticity and memory formation.[37] Therefore, *CREB1* is theoretically a potential common risk gene for major psychiatric disorders.

To gain initial insights into this hypothesis, we further investigated the associations between *CREB1* SNPs and BPD in larger clinical samples, and examined whether the BPD risk alleles in *CREB1* were also associated with SCZ and MDD. In the meantime, we also examined the expression levels of *CREB1* mRNA in patients with BPD, SCZ and MDD, and compared *CREB1* expression levels between those patients and healthy controls. Finally, we analyzed CREB1-participated protein–protein interaction (PPI) network to understand its interactions with other psychiatric disease-related proteins. Our results suggest that *CREB1* might affect the risk of major psychiatric disorders.

MATERIALS AND METHODS

BPD clinical association samples

The current BPD meta-analysis utilized 10 independent case–control samples comprising 11 105 patients and 51 331 controls. In brief, the samples were collected either from published GWAS data sets,[42] from unpublished data provided by corresponding researchers or from publicly available data sets such as dbGaP (https://www.ncbi.nlm.nih.gov/gap). All the subjects included in the studies provided written informed consents, and all procedures were performed following relevant legal and ethical guidelines in the respective areas. Most of these samples were published in previous large-scale collaborative studies,[42,43] and have been proven to be reliable in detection of genetic risk variants for BPD, warranting the appropriate design of the current meta-analysis. Briefly, the samples in the current study are from the following: (1) Psychiatric Genomics Consortium 1 (PGC1) GWAS (7481 cases and 9250 controls);[42] (2) France (451 cases and 1631 controls);[35] (3) Germany (181 cases and 527 controls);[35] (4) Iceland (544 cases and 34 426 controls);[35] (5) Poland (411 cases and 504 controls);[35] (6) Sweden (836 cases and 2093 controls);[35] (7) Australia (330 cases and 1811 controls);[35] (8) Romania (451 cases and 318 controls);[13] (9) GAIN AA (362 cases and 671 controls);[44] and (10) USA EA (58 cases and 100 controls).[13] Among these 10 cohorts, Romania, GAIN AA and USA EA were new samples in which the associations of *CREB1* SNPs with BPD were not individually reported previously.[35] Clinical information of these subjects (such as the diagnostic history of psychiatric conditions, age of onset and so on) was collected through standardized semi-structured interviews, and operational criteria were applied in making lifetime diagnoses. Briefly, all cases were diagnosed according to the DSM-IV, ICD-9 or ICD-10 criteria. Controls were populations of matched ethnicities from the same geographical areas as the cases, and were proven to have low probabilities of diagnosis with BPD. Among these samples, GAIN AA sample included subjects of African American ancestry, and all the other samples were of European origin. Detailed information of individual samples are shown in the Supplementary Data.

SCZ clinical association samples

The samples for SCZ meta-analysis included three independent case–control data sets comprising 34 913 patients and 44 528 controls. The samples were collected from a published GWAS[45] and dbGaP. The origins and sizes of the samples were as follows: (1) PGC2 GWAS[45] (33 640 cases and 43 456 controls, European ancestry) (The original PGC2 GWAS[45] included 46 cohorts from Europeans and 3 cohorts from East Asians, with a total size of 34 241 cases and 45 604 controls. Considering that rs2709370 and rs6785 are not polymorphic in East Asians, we retrieved only the 46 European cohorts from the PGC2 GWAS[45] for the current analyses, which yielded 33 640 cases and 43 456 controls.); (2) GAIN AA (1195 cases and 954 controls, African American ancestry);[46] and (3) USA AA (78 cases and 118 controls, European ancestry). GAIN AA and USA AA cohorts for SCZ were newly included samples in which *CREB1* was not individually examined. All participants were recruited according to the relevant guidelines, and written informed consents were collected before their participation. Notably, there was a substantial overlap of the control subjects in SCZ GAIN AA sample[46] and BPD GAIN AA sample,[44] although their total control sizes were different. To avoid potential bias in the current analyses, we did not normalize the control subjects in SCZ GAIN AA and BPD GAIN AA into the same number. Detailed information regarding sample description, diagnosis, genotyping and statistical analyses methods can be found in the Supplementary Data.

MDD clinical association samples

The MDD meta-analysis included two independent case–control data sets of 9369 patients and 9619 controls. The samples were collected from the published PGC1 GWAS[47] and dbGaP. Briefly, the samples used in the current analysis were 9240 cases and 9519 controls from PGC1 GWAS,[47] as well as 129 cases and 110 controls from USA EA sample in dbGaP (the new sample not used in PGC GWAS or the previous *CREB1* study[35]). Cases were diagnosed with lifetime MDD establishments by trained individuals according to the DSM-IV criteria. The primary sources of these two samples included health-care providers and randomly selected controls screened to be free of lifetime MDD history from the population. All the included subjects were of European ancestry. Detailed descriptions of the samples, data and statistical approaches are listed in the Supplementary Data.

SNP selection, genotyping and statistical analysis

In the current meta-analysis, we focused on two previously reported BPD susceptibility SNPs rs2709370 and rs6785.[35] In the samples re-analyzed in the current study, the genotyping was mainly done with the Illumina (San Diego, CA, USA) and Affymetrix platforms, yielding at least 98% call in all cases and controls. The two tested SNPs (rs2709370 and rs6785) were either directly genotyped or imputed using standard methods. Specifically, in the PGC2 SCZ GWAS,[45] prephasing imputation stepwise approaches implemented in IMPUTE2 (ref. 48) and SHAPEIT[49] were used for the genotype imputation, with the low quality and rare variants removed before the analyses. The imputation reference was used from the 1000 Human Genomes Project data set,[50] separately by platform; while in the PGC1 BPD and MDD GWAS,[42,47] BEAGLE 3.0 was used to impute the genotype data,[51] and phased HapMap data[52] was set as a reference. Only SNPs with very high imputation quality were subject to further analyses. We retrieved the odds ratio (OR) and s.e. to estimate heterogeneity among individual samples. Specifically, we performed the standard Cochran's (Q) χ^2-test as previously described.[13] In the event that no heterogeneity exists among individual samples, all the samples were combined under a fixed-effect model; if heterogeneity was seen, the samples were combined using a random-effect model. Following the heterogeneity analyses, the OR and 95% confidence interval were calculated for the combined total sample pool. The meta-analysis was performed with the classical inverse variance weighted methods ('*metafor*' package in R; http://www.R-project.org) as described in our recent studies.[53,54] A forest plot presenting the pooled ORs and 95% confidence intervals was generated, with each square (the sizes of the squares represent the weight of each respective study) in the

figure denoting a particular study. All the protocols and methods used in this study were approved by the institutional review board of the Kunming Institute of Zoology, Chinese Academy of Sciences.

Comparison of brain *CREB1* expression between psychiatric patients and healthy controls

To compare the mRNA expression levels of *CREB1* between psychiatric patients and healthy subjects, we obtained the microarray data in the frontal cortex of 64 adult BPD patients, 97 adult SCZ patients and 131 adult MDD patients, as well as 103 adult controls from dbGaP (accession number phs000979.v1.p1). All the subjects were of European ancestry. The expression data were normalized with $\log_2$ ratios of florescent intensities. The $\log_2$ ratios obtained through normalization were then adjusted using surrogate variable analysis[55] to alleviate the impact of systematic noise on gene expression results. All of these microarray data analyses were conducted using codes and tools from the Bioconductor project (http://www.bioconductor.org/) in R.

We also collected hippocampal RNA-sequencing data of 15 adult BPD cases, 15 adult SCZ cases and 15 adult MDD cases, as well as 15 adult controls from the Stanley Medical Research Institute (SMRI) data set (http://sncid.stanleyresearch.org/) in the FASTQ file format. The RNA-sequencing reads underwent adaptation and low-quality filtering using btrim64,[56] and were then aligned to human reference genome (Human GRCh38 (hg38), http://asia.ensembl.org/index.html) by splice-read mapper (Tophat2 v2.0.14).[57] The known transcript maps were extracted from Ensembl Build GRCh38. Cufflinks v2.2.1 (ref. 58) was applied to call new transcripts as well as to assemble and quantify both the novel and known transcripts with default parameters. For each subject, accepted hits bam files from Tophat2 alignment were merged by Samtools v0.1.18 (ref. 59) for the following Cufflinks quantification: (1) reads that were uniquely mapped to particular genes were used to quantify the respective gene expression levels; (2) fragments per kilobase per million mapped reads ($FPKM = F \times 10^3/L \times 10^6/N$, where F is the number of fragments mapping to the gene annotation, L is the length of the gene structure in nucleotides and N is the total number of sequence reads mapped to the genome) was calculated to quantify mRNA expression levels as described elsewhere.[13] The FPKM were log2-transformed, and then underwent surrogate variable analysis[60,61] normalization to remove covariate effects before the diagnostic analyses.

We then calculated the differences of *CREB1* mRNA levels between different diagnostic categories (BPD, SCZ, MDD and healthy controls) using an analyses of covariance model. The diagnostic statuses were applied as independent variables, while surrogate variables, age, gender, RNA integrity number, brain pH and postmortem interval were set as covariates. Following these analyses, pairwise comparisons adjusted with the Tukey's method were performed. In addition, we assessed the effects of medication treatment with antipsychotics, antidepressants and benzodiazepines on *CREB1* mRNA level within the patients using two-sided *t*-test. The same procedure was undertaken for daily use of nicotine and alcohol (yes/no), as this information would provide valuable insights but was not available in the control group.

PPI analysis

It is generally accepted that proteins interact with each other to facilitate a wide array of cellular processes, and disruption of one member or subunit of a particular protein complex will most likely leads to cascades of functional consequences. This theory has been proven true in a variety of illnesses and physiological conditions with convincing evidence.[62] Intriguingly, it is believed that for a defined illness, the frequency of interaction between proteins encoded by susceptibility genes of this disease tends to be higher than that between random proteins, and PPI network analysis is a robust approach to investigate such phenomenon.[62] In fact, our recent studies revealed that proteins encoded by SCZ susceptibility genes formed a highly interconnected PPI network.[62] Therefore, we also performed the PPI analyses to understand whether CREB1 was involved in the PPI network containing proteins products of the top susceptibility genes of psychiatric disorders. Specifically, we utilized a well-characterized data set containing promising psychiatric susceptibility gene candidates identified by recent GWAS (for example, *ZNF804A*, *CACNA1C* and so on).[42,45,63] The PPI network consists of nodes and edges, denoting proteins and physical interactions, respectively. Physically interacting protein molecules are therefore visualized as nodes connected by edges. In the current study, proteins encoded by psychiatric susceptibility genes were used as seed proteins, and STRING (http://string-db.org/)[64] was used for exaction and reconstruction of the PPI network.

Table 1. Meta-analysis of *CREB1* SNPs with BPD, SCZ and MDD in general populations

				rs2709370 (chr2:208090847, C/A)			*rs6785 (chr2:208176242, A/G)*		
Disease	*Sample*	*Case*	*Control*	*OR*	*S.e.*	P-*value*	*OR*	*S.e.*	P-*value*
BPD	PGC1 GWAS	7481	9250	1.094	0.029	1.80×10^{-3}	1.111	0.029	3.38×10^{-4}
	France	451	1631	1.083	0.097	0.411	1.066	0.096	0.510
	Germany II	181	527	1.180	0.153	0.241	1.235	0.152	0.145
	Iceland	544	34 426	1.012	0.074	0.862	0.987	0.077	0.865
	Poland	411	504	1.208	0.108	0.0536	1.186	0.108	0.0891
	Sweden I	836	2093	1.050	0.072	0.501	1.051	0.075	0.508
	Australia	330	1811	0.927	0.106	0.467	0.956	0.108	0.662
	Romania	451	318	1.228	0.132	0.118	1.241	0.132	0.103
	GAIN AA	362	671	1.101	0.137	0.482	1.138	0.134	0.334
	USA EA	58	100	1.406	0.305	0.264	1.196	0.323	0.579
	Meta-analysis	11 105	51 331	1.085	—	2.33×10^{-4}	1.095	—	6.33×10^{-5}
SCZ	PGC2 GWAS	33 640	43 456	1.055	0.013	7.30×10^{-5}	1.059	0.014	2.90×10^{-5}
	GAIN AA	1195	954	1.100	0.092	0.300	1.047	0.094	0.625
	USA AA	78	118	1.224	0.289	0.484	0.979	0.279	0.939
	Meta-analysis	34 913	44 528	1.056	—	3.96×10^{-5}	1.059	—	2.44×10^{-5}
MDD	PGC1 GWAS	9240	9519	1.063	0.027	0.0224	1.056	0.027	0.0422
	USA EA	129	100	1.462	0.240	0.114	1.374	0.241	0.188
	Meta-analysis	9369	9619	1.067	—	0.0144	1.059	—	0.0314

Abbreviations: AA, African American; BPD, bipolar disorder; *CREB1*, cAMP responsive element-binding-1; EA, European American; GWAS, genome-wide association study; MDD, major depressive disorder; OR, odds ratio; PGC1, Psychiatric Genomics Consortium 1; SCZ, schizophrenia; SNP, single-nucleotide polymorphism.

RESULTS

Genetic associations of *CREB1* polymorphisms with major psychiatric disorders

We herein performed meta-analyses of two *CREB1* SNPs (rs2709370 and rs6785), which showed associations with BPD in our previous report.[35] While these SNPs have been covered in the earlier PGC GWAS[42,45,47] and our previous study,[35] several additional cohorts (for example, BPD Romania, BPD GAIN AA, BPD USA EA, SCZ GAIN AA, SCZ USA AA and MDD USA EA; sample descriptions are shown in Supplementary Data) are also available for further characterization of their involvements in major psychiatric disorders. Combining all these samples, we have in total collected data from 11 05 patients and 51 331 controls for BPD analysis, 34 913 patients and 44 528 controls for SCZ analysis, as well as 9369 patients and 9619 controls for MDD analysis. Table 1 lists the information of the samples involved in the SNPs of interest in this meta-analysis together with the results of each individual sample.

SNP rs2709370 is a flanking variant in the upstream region of the *CREB1* gene. Our previous study showed that this SNP was associated with the onset of BPD, as well as hippocampal volume and activity during the record memory task.[35] With the aim to further analyze its associations with BPD in greater samples, and to examine its roles in the genetic risk of SCZ and MDD, we performed analyses of all these three diagnostic categories. A fixed-effect model was used as there was no evidence for significant heterogeneity between different samples (BPD, rs2709370, $P=0.710$, $I^2=0$; BPD, rs6785, $P=0.718$, $I^2=0$; SCZ, rs2709370, $P=0.791$, $I^2=0$; SCZ, rs6785, $P=0.954$, $I^2=0$; MDD, rs2709370, $P=0.188$, $I^2=42.4\%$; MDD, rs6785, $P=0.278$, $I^2=15.0\%$). Briefly, the current meta-analysis of BPD included the samples used in our previous *CREB1* study[35] and three new independent samples from Romania (451 cases and 318 controls), GAIN AA (362 cases and 671 controls) and USA EA (58 cases and 100 controls).[35] In each of the three new BPD samples, rs2709370 did not show significant associations (Romania, $P=0.118$, OR = 1.228; GAIN AA, $P=0.482$, OR = 1.101; USA EA, $P=0.264$, OR = 1.406), but its effect sizes (that is, OR) were all larger than 1.10 (Table 1). Meta-analysis combining these three new samples (871 cases and 1089 controls) revealed a marginal association between rs2709370 and BPD ($P=0.0611$, OR = 1.185). Notably, the OR of rs2709370 in the African American population (GAIN AA sample) was in line with that in the Europeans (Table 1), and the allelic frequencies of this SNP were also similar between different populations (minor allele frequency was 0.129 in Africans vs 0.181 in Europeans). To increase the statistical power, we then analyzed the association of rs2709370 in the combination of the 10 independent BPD data sets from European and African American populations (11 105 BPD patients and 51 331 controls), and found that rs2709370 was significantly associated with BPD ($P=2.33\times10^{-4}$, OR = 1.085; Figure 1). Given that the majority of these BPD cohorts have already been included in our previous study,[35] this result is in line with our expectation. Meanwhile, we also observed an interesting phenomenon that 9 of the 10 included samples (except for the Australia sample) showed the same direction of allelic effects at rs2709370. Furthermore, although this SNP did not achieve the conventional nominal significance in the new BPD samples, its effect sizes in all these independent samples were similar. After BPD, we then performed the meta-analysis of rs2709370 in SCZ samples using three data sets comprising of 34 913 cases and 44 528 controls in total, and observed a significant association ($P=3.96\times10^{-5}$, OR = 1.056; Figure 1). The meta-analysis using two MDD data sets with 9369 cases and 9619 controls also yielded a significant association of rs2709370 with this disease ($P=0.0144$, OR = 1.067; Figure 1). It should be noted that the meta-analytic result of MDD was less significant than those of BPD or SCZ, which was likely due to the smaller sample size used in the analysis for MDD compared with

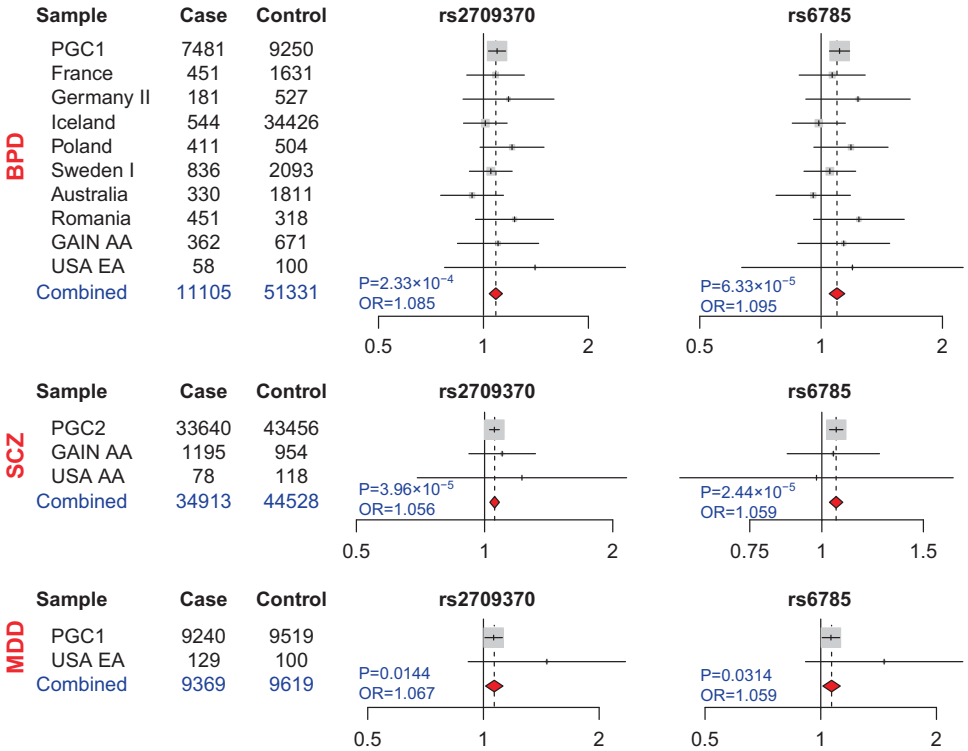

Figure 1. Forest plots of *CREB1* single-nucleotide polymorphisms with bipolar disorder (BPD), schizophrenia (SCZ) and unipolar major depressive disorder (MDD). The sample information was included in the Supplementary Data.

the other two disorders, as the effect sizes of rs2709370 were similar across all three diagnostic groups (Table 1). As the commonly observed associations of rs2709370 suggested potential shared genetic factors between these psychiatric disorders, we also performed a meta-analysis of this SNP using the combined samples of the three illnesses. As there were duplicated control subjects among the controls of the samples with different diagnosis categories from PGC GWAS (PGC1 BPD, PGC2 SCZ and PGC1 MDD in Table 1), we utilized the 'smaller PGC GWAS' samples (6990 cases and 4820 controls for BPD, 9379 cases and 7736 controls for SCZ, and 9227 cases and 7383 controls for MDD; Supplementary Table S1), which have excluded the duplicated individuals as previously described.[9] Although the overall sample size after exclusion was significantly reduced (a total of 25 596 cases and 19 939 controls), meta-analysis in these PGC samples yielded a nominal significant association ($P = 1.15 \times 10^{-3}$, OR = 1.059; Supplementary Table S1) between rs2709370 and the three major psychiatric disorders. We then combined these PGC GWAS samples and 'non-PGC samples' to conduct an overall meta-analysis (Supplementary Table S1). Notably, given the substantial overlap in control subjects between BPD GAIN AA and SCZ GAIN AA samples, we only used the SCZ GAIN AA sample in this meta-analysis; besides, the BPD USA EA and MDD USA EA samples were merged as the 'Mood USA EA' sample as they used the same control individuals. Finally, we observed a stronger association between rs2709370 and the three major psychiatric disorders (30 260 cases and 62 421 controls, $P = 7.82 \times 10^{-5}$, OR = 1063; Supplementary Table S1). As our analyses were carried out using both published and new samples, we then analyzed the data only in the new psychiatric samples (BPD Romania, Mood USA EA, SCZ GAIN AA and SCZ USA AA, including a total of 1911 cases and 1490 controls) that were not included in either the former PGC GWAS[9] or our previous *CREB1* study.[35] Indeed, the meta-analysis of these new psychiatric samples revealed a nominal association at rs2709370 ($P = 0.0235$, OR = 1.170; Supplementary Table S1).

The other SNP we examined in this meta-analysis, rs6785, is a substitution in the 3′-untranslated region of *CREB1*. We previously showed that this SNP was associated with BPD and hippocampal volume.[35] After meta-analytic investigation in the present BPD (11 105 cases and 51 331 controls), SCZ (34 913 cases and 44 528 controls) and MDD (9369 cases and 9619 controls) groups, we observed significant associations between this SNP with these psychiatric disorders, respectively ($P = 6.33 \times 10^{-5}$, OR = 1.095 for BPD; $P = 2.44 \times 10^{-5}$, OR = 1.059 for SCZ; and $P = 0.0314$, OR = 1.059 for MDD; Figure 1). In the three new BPD samples (Romania, GAIN AA and USA EA), including 871 cases and 1089 controls, rs6785 showed a marginal association with the illness via meta-analysis ($P = 0.0544$, OR = 1.190). Subsequently, using the same sample set as rs2709370, we performed an overall meta-analysis of rs6785 in the 30 260 cases with psychiatric disorders and 62 421 controls from the previous PGC GWAS,[9] our *CREB1* study[35] and new samples as described above. Consistent with our expectation, rs6785 was significantly associated with the three major psychiatric disorders ($P = 6.61 \times 10^{-5}$, OR = 1065; Supplementary Table S1). Collectively, the present results expanded the knowledge of rs2709370 and rs6785 in the genetic risk of psychosis and mood disorders to a broader diagnostic spectrum.

The above associations did not reach genome-wide statistical significance, which was likely explained by the limited statistical power of the present sample sizes. To understand whether the current sample sizes for the tested psychiatric disorders allowed detection of significant associations, we conducted a power analysis using the Power and Sample Size Programme software.[65] As the SCZ sample had the largest size (34 913 SCZ patients and 44 528 controls) among samples of all three diagnostic groups, analyzing the statistical power of SCZ cohorts should answer this question. Because the frequencies of risk alleles for rs6785 (A-allele 0.17) and rs2709370 (C-allele 0.18) were similar, we used rs6785 as an example. Assuming that two-tailed α of 5.00×10^{-8} denoted a genome-wide significant association, we assumed the OR to be 1.10 (which was commonly observed in psychiatric genetic association studies, and was similar to the effect size of rs6785 and rs2709370 for psychiatric disorders) for the analysis. The present sample size demonstrated only 36.2% power to capture a genome-wide significant association for rs6785. On the other hand, to observe a genome-wide significant association for rs6785 assuming the effect size of 1.10, the sample should contain at least 59 909 cases and 59 909 controls to achieve an 80% statistical power. Nevertheless, the effect sizes of the risk SNPs were similar with other genome-wide significant variants highlighted in the large-scale GWAS.[45] More importantly, the allelic effects across all diagnostic groups were in the same direction, and the effect sizes were also similar across the investigated illnesses. Therefore, our results provided initial evidence for a potential role of *CREB1* in the genetic risk of major psychiatric disorders.

CREB1 mRNA levels across diagnostic categories

We first performed *CREB1* mRNA expression analysis in the dbGaP brain samples (discovery sample, $n = 395$ subjects) and found that *CREB1* mRNA levels in the dorsolateral prefrontal cortex (PFC) tissues differed significantly between the BPD, SCZ, MDD and control groups (analyses of covariance, $P = 2.96 \times 10^{-3}$). Pairwise comparisons revealed that the expression levels of *CREB1* in patients with SCZ, BPD and MDD were similar (all $P > 0.95$), and were all significantly lower than that in healthy controls (BPD vs control, $P = 0.0288$; SCZ vs control, $P = 9.23 \times 10^{-3}$; MDD vs control, $P = 0.0101$; Figure 2a). These data suggested that decreased expression of *CREB1* might be a susceptibility factor for these

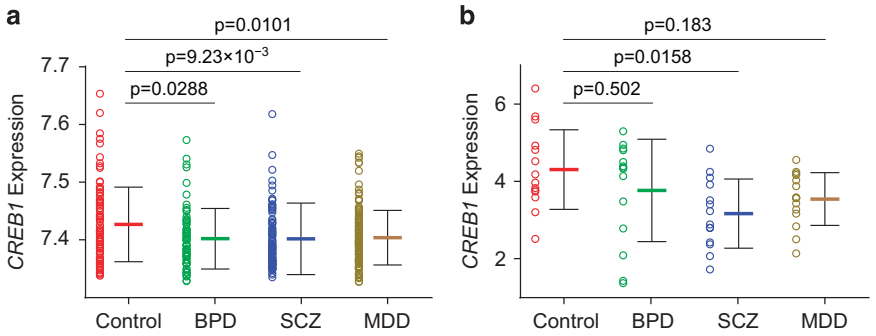

Figure 2. Differences in *CREB1* expression across diagnostic categories in two independent brain samples from dbGaP (**a**) and Stanley Medical Research Institute data set (**b**). The plot was made using the florescent intensities (**a**) and raw fragments per kilobase per million mapped reads (**b**). BPD, bipolar disorder; MDD, unipolar major depressive disorder; SCZ, schizophrenia.

psychiatric disorders. We then assessed the effects of medication status on *CREB1* expression in these samples. However, no significant effects of antipsychotics, antidepressants or benzodiazepines on *CREB1* mRNA level were found (antipsychotics: BPD, $P=0.366$; SCZ, $P=0.763$; MDD, $P=0.487$. Antidepressants: BPD, $P=0.103$; SCZ, $P=0.394$; MDD, $P=0.752$. Benzodiazepines: BPD, $P=0.757$; SCZ, $P=0.413$; MDD, $P=0.051$; Figure 3). Smoking or alcohol abuse status did not influence *CREB1* mRNA levels within the patients either (nicotine: BPD, $P=0.962$; SCZ, $P=0.369$; MDD, $P=0.784$. Ethanol: BPD, $P=0.534$; SCZ, $P=0.486$; MDD, $P=0.191$; Figure 3).

We then examined *CREB1* expression in the SMRI data set (replication data set, $n=60$ subjects) with a smaller sample size of hippocampal tissues. Again, we saw that *CREB1* mRNA levels differed significantly between BPD, SCZ, MDD and controls (analyses of covariance, $P=0.0253$). Pairwise comparisons revealed that the expression level of *CREB1* in patients with SCZ was significantly decreased compared with that in healthy controls ($P=0.0158$; Figure 2b); though *CREB1* expression did not exhibit significant differences between BPD patients and controls or between MDD patients and respective controls (BPD vs control, $P=0.502$; MDD vs control, $P=0.183$; Figure 2b), the mRNA levels were still overall lower in the patients, which was in line with the results obtained with the discovery sample, and the nonsignificant results during replication analyses were likely due to the smaller sample sizes.

To further replicate these results in expanded samples, we also analyzed the *CREB1* mRNA expression in multiple samples comprising patients with psychiatric disorders from the Gene Expression Omnibus (GEO) data sets (https://www.ncbi.nlm.nih.gov/gds/). Intriguingly, in the GSE53987 (ref. 66) data set, *CREB1* expression was significantly decreased in patients with SCZ and in patients with MDD compared with healthy controls in both PFC and hippocampus tissues (PFC, SCZ vs control, $P=5.00\times10^{-3}$; MDD vs control, $P=0.036$; Supplementary Figure S1A; hippocampus, SCZ vs control, $P=0.028$; MDD vs control, $P=0.036$; Supplementary Figure S1B), while in this data set *CREB1* expression was lower in patients with BPD than in healthy control in hippocampus only ($P=0.044$, Supplementary Figure S1B). The reduction of *CREB1* expression in patients with psychiatric disorders was also observed in other GEO data sets. For example, *CREB1* expression was lower in BPD patients than in healthy controls in the GSE5392 (ref. 67) data set ($P=0.012$; Supplementary Figure S1C). In addition, *CREB1* expression was significantly decreased in patients with SCZ compared with healthy controls in the GSE21138 (ref. 68) data set ($P=8.00\times10^{-3}$; Supplementary Figure S1D). In the GSE12654 (ref. 69) data set, *CREB1* expression was reduced in MDD patients compared with healthy controls ($P=0.033$; Supplementary Figure S1E). These data further confirmed our findings in the dbGaP brain samples (dorsolateral PFC tissue) and in the SMRI data set (hippocampal tissue). Notably, the samples of GEO data sets may not be completely independent from those in the SMRI data set, as a certain proportion of the GEO subjects was also provided by the Stanley Foundation Brain Collection. To avoid potential duplication, we analyzed the hippocampal RNA-sequencing results in the SMRI data set, and investigated microarray-based gene expression data in the PFC and hippocampus in the GEO data sets. Therefore, these analyses can still be considered as technical replications of the earlier results.

To investigate whether rs2709370 and rs6785 contributed to the lower *CREB1* expression level in patients with psychiatric disorders, we also analyzed the associations of these two SNPs with *CREB1* mRNA expression in the dbGaP brain samples as well as in public expression quantitative trait locus (eQTL) data sets from GTEx (https://gtexportal.org/home/)[70,71] and Braineac (http://www.braineac.org/).[72] In the dbGaP samples, the two risk SNPs were not associated with *CREB1* mRNA in either controls or SCZ patients (for rs2709370: control, $P=0.940$; SCZ, $P=0.051$. For rs6785: controls, $P=0.594$; SCZ, $P=0.108$; Supplementary Figure S2), and were in nominal (or marginal) association in BPD and MDD patients (for rs2709370: BPD, $P=0.0027$; MDD, $P=0.089$. For rs6785: BPD, $P=0.011$; MDD, $P=0.025$; Supplementary Figure S2); it should be noted that in the diagnostic analysis of mRNA expression, the effect size for downregulation of *CREB1* mRNA in patients with SCZ was the most prominent across all three illnesses. To validate the above results, we then examined the associations between these two SNPs and *CREB1* expression in GTEx and Braineac eQTL data sets, but no significant associations were observed (Supplementary Figures S3 and S4). These data suggested that though *CREB1* expression was lower in patients with psychiatric disorders (especially SCZ) compared with healthy controls, rs2709370 and rs6785 were unlikely the potential reasons. We also investigated the possibility that the decreased *CREB1* expression in patients was caused by other SNPs in high or low linkage disequilibrium with rs2709370 and rs6785, by

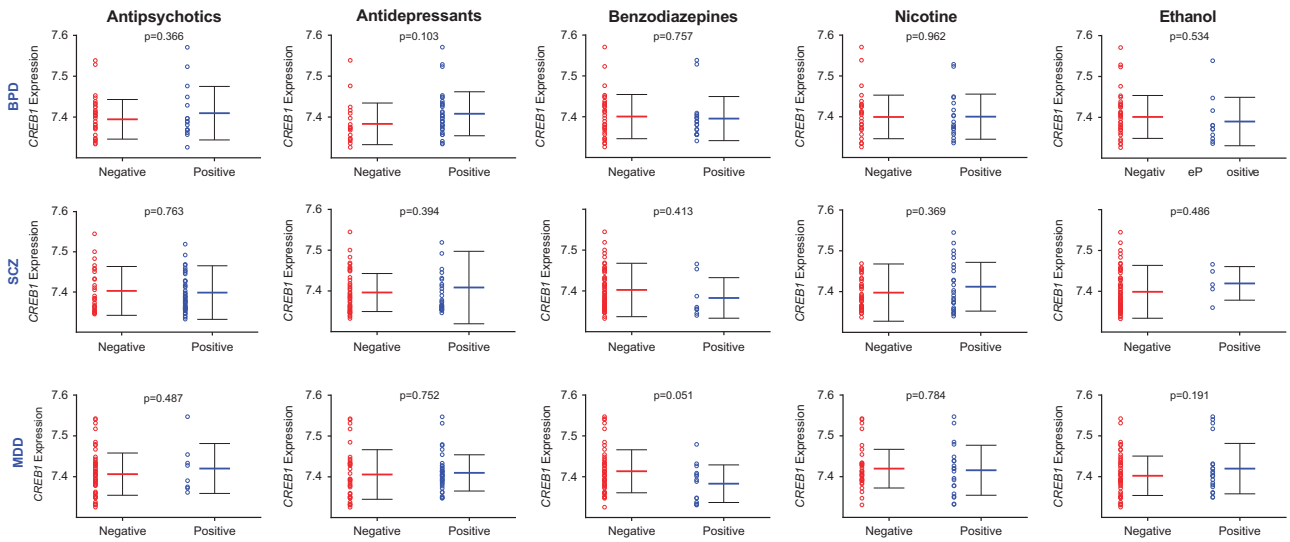

Figure 3. Effects of medication and certain substrates on *CREB1* expression in patients with bipolar disorder (BPD), schizophrenia (SCZ) and unipolar major depressive disorder (MDD). 'Positive' means 'yes', while 'negative' means 'no'.

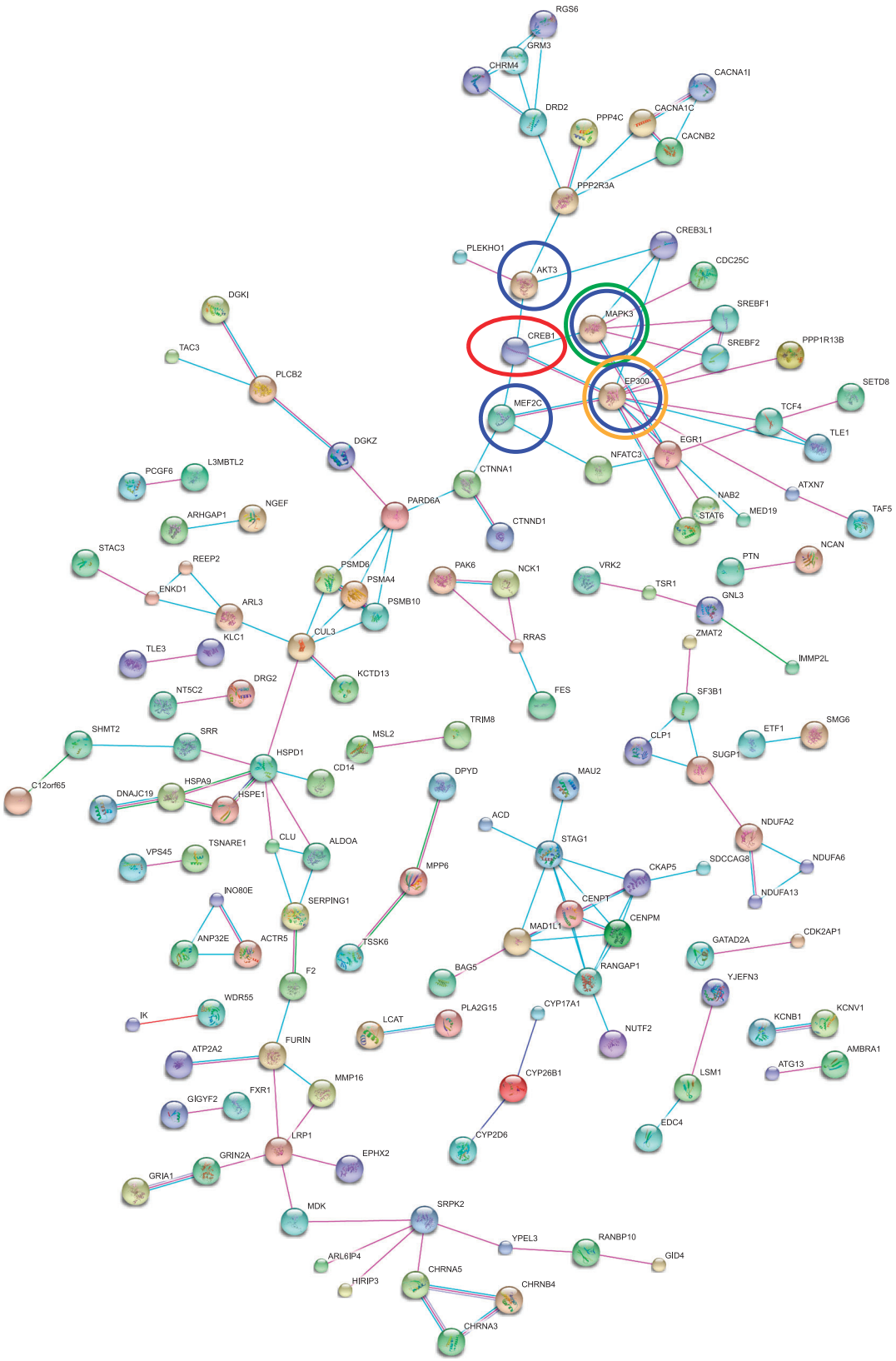

Figure 4. Protein encoded by *CREB1* in a densely interconnected protein–protein interaction (PPI) network formed by top susceptibility genes of psychiatric disorders. The amaranthine line indicates known PPIs that have been experimentally verified, and the light blue line indicates known PPIs from curated databases. The green line indicates predicted interactions through gene neighborhood, the dark blue line indicates predicted interactions through gene co-occurrence and the red line indicates predicted interactions through gene fusions. The CREB1 protein was highlighted with red circle. For the four proteins that have direct interactions with CREB1, blue circle suggests that the protein has been reported in susceptibility to schizophrenia, green circle means the protein has been reported in susceptibility to bipolar disorder and orange circle means the protein has been reported in susceptibility to unipolar major depressive disorder.

performing genome-wide eQTL analyses based on *CREB1* expression in the frontal cortex and hippocampus tissues in GTEx and Braineac data sets. However, no significant eQTLs that could survive multiple testing correction ($P < 0.001$) were observed in either data set (data not shown). Therefore, the decreased *CREB1* expression in patients with psychiatric disorders is unlikely caused by genetic factors that are detectable in these data sets. Future studies analyzing other potential factors (for example, epigenetic components, miRNA, lncRNA and so on) affecting *CREB1* mRNA expression in psychiatric patients are therefore needed. In sum, *CREB1* is likely involved in the susceptibility of the three major psychiatric disorders (SCZ, BPD and MDD), and the regulatory mechanisms underlying its reduced expression in patients with these illnesses are yet to be determined.

CREB1 participates in a highly interconnected PPI network involving risk genes of psychiatric disorders

The idea that disturbances of certain cellular processes or pathways contributes to the risk of psychiatric disorders has been emerging and gained support from accumulating evidence.[62,73,74] The well-characterized databases have also provided essential information for constructing highly interconnected PPI networks, which may reveal the underlying biological mechanisms. Therefore, we examined the interaction between CREB1 protein and protein products of multiple defined risk genes for psychiatric disorders. We found that the top susceptibility genes of major psychiatric disorders identified by GWAS[42,45,63] encoded a densely interconnected PPI network (Figure 4). Intriguingly, CREB1 also participates in this network ($P < 0.05$), indicating its potential involvement in the common molecular network modulating the pathogenesis of psychiatric disorders. It is noteworthy that CREB1 directly interacts with proteins encoded by *AKT3* (AKT serine/threonine kinase 3), *MAPK3* (mitogen-activated protein kinase 3), *MEF2C* (myocyte enhancer factor 2C) and *EP300* (E1A binding protein p300; Figure 4), the genes repeatedly highlighted in genetic risk studies of major psychiatric disorders in distinct samples.[45,75–81] *CREB1* encodes a transcription factor that has central roles in the cAMP signaling and in neuronal signal transduction,[82] primarily by mediating the induction of cAMP signal transduction pathway following activation of a bunch of G-protein-coupled receptors. In fact, several proteins highlighted in our PPI network have been implicated in previous studies for their interaction with CREB1 in cAMP pathways. For example, CREB protein has been demonstrated to be phosphorylated by ERK,[83] a member of the MAP kinase family that encoded by *MAPK3*; the protein encoded by *EP300* functions as a histone acetyltransferase and regulates cAMP gene by specifically binding to phosphorylated CREB protein, and therefore facilitates important processes in cell proliferation and differentiation.[84–86] Intriguingly, both the *EP300* and *MAPK3* genes locate in the most significant genomic regions identified in the latest PGC2 SCZ GWAS (*EP300* locates in 22q13.2, the most significant SNP was rs9607782, $P = 2.07 \times 10^{-11}$; *MAPK3* locates in 16p11.2, the most significant SNP was rs12691307, $P = 4.55 \times 10^{-11}$).[45] Therefore, CRBE1 may contribute to the risk of psychiatric disorders via interactive effects on these top-risk loci. Further studies are needed to dissect the mechanisms by which CREB1 interacts with the proteins encoded by *AKT3* and by *MEF2C*, as well as the roles of these protein interactions in brain function and development, as well as and susceptibility to psychiatric disorders.

DISCUSSION

In the current study, we performed meta-analyses of *CREB1* SNPs followed by expression comparisons, and highlighted potential involvement of *CREB1* in the risk of major psychiatric disorders, including SCZ, BPD and MDD. Although a genome-wide significance level was not reached, we have presented converging evidence supporting a putative role of this gene in a broader spectrum of psychiatric conditions in agreement with several recent studies.[87–90] These findings further imply that different psychiatric disorders likely share common genetic risk factors, and support the contention that psychiatric diagnostic entities may have less clear-cut boundaries.

While we found significant associations between *CREB1* SNPs and psychiatric conditions, the *P*-values reported here would not be significant after correction for all SNPs across the genome. This probably reflects the fact that this genomic region has not been highlighted as a major locus in the recently published GWASs of psychiatric disorders to date.[42,45] However, although GWAS analysis is considered a persuasive approach in identifying genetic risk factors for certain diseases, its rigid statistical correction and its nature as a discovery analysis without any *prior* hypothesis may overshadow certain critical genetic variants and create false negatives. It is also accepted that the relevance of a particular gene to psychiatric disorders should be determined based on the overall base of genetic and biological evidence rather than solely relying on statistical significance. In the case of *CREB1*, which is of sufficient biological interest as indicated by numerous studies,[36,40,82,91] the less significant statistics should not be taken as a denial of its potentially important roles in psychiatric disorders. Future investigations on its biological basis in such illnesses are still of great interest.

Moreover, although the risk SNPs in *CREB1* likely have only moderate roles in the major psychiatric disorders, and the associations did not achieve genome-wide level significance ($P < 5.0 \times 10^{-8}$), their ORs were similar with the those of susceptibility loci confirmed in large-scale genetic association studies of major psychiatric disorders,[45] suggesting that *CREB1* SNPs might confer comparable risk for diseases with those genome-wide significant risk loci. Given the polygenic characteristics of psychiatric disorders, the *CREB1* SNPs are likely within the complex array of risk loci across the allelic frequency spectrum. In addition, the possible involvement of *CREB1* in the pathogenesis of SCZ, BPD and MDD was supported by our diagnostic expression analyses results. The PPI analysis also demonstrated direct interactions between CREB1 and GWAS top-risk genes for major psychiatric disorders, providing essential insights into the connection between *CREB1* gene and those illnesses.

Although the exact roles of *CREB1* in psychiatric disorders are unknown, the present study provides several perspectives for future investigations. The decreased mRNA expression of *CREB1* in psychiatric patients compared with healthy controls in the PFC and hippocampus is in line with the previous studies showing reductions of total and phosphorylated CREB1 protein levels in the orbitofrontal cortex of antidepressant-free MDD patients.[92] In addition, CREB1 is also a potentially crucial mediator for effects of antidepressant targeting the hippocampal function and activity, as multiple standard antidepressant treatments (for example, noradrenaline-reuptake inhibitors, selective serotonin-reuptake inhibitors and electroconvulsive seizures) have been shown to cause elevated CREB1 activity (for example, phosphorylation level) within the hippocampus.[93,94] On the other hand, genetic modification resulting in increased CREB1 protein levels in the murine hippocampal tissues can produce antidepressant-like effects.[95] Combining these previous discoveries with our observation in the present study that antidepressants did not affect the mRNA expression of *CREB1*, we speculate that antidepressants likely modulate the translation or post-translational modification (for example, phosphorylation) of the CREB1 protein rather than affecting its mRNA levels. Consistent with this notion, researchers have revealed that systemic administration of PDE4 inhibitors that stimulate cAMP signaling, which could lead to phosphorylation and activation of CREB1, also results in antidepressant effects in both animals[96] and humans,[97] probably through altering CREB1

activities in hippocampus. Therefore, these lines of evidence and our data suggest potential beneficial effects of hippocampal CREB1 induction on the therapeutic efficacy of antidepressant treatments. However, it should be noted that CREB1 activation could be sometimes detrimental depending on the brain regions involved. For example, elevated CREB1 activity in the nucleus accumbens leads to various depression-like responses in rodents.[91,98] While the molecular mechanisms underlying these complications remain undetermined, our finding and these previous studies together highlight the pivotal roles of CREB1 in the central nervous system. To obtain information regarding possible therapeutic targeting of this transcription factor in psychiatric disorders, future studies dissecting the nonuniform effects of CREB1 throughout the brain, and the molecular elements regulating specialized roles of this protein in different brain regions and neural circuits, are needed. In addition, varied activities or regulating mechanisms of CREB1 downstream targets may have also contributed to this phenomenon, and future work focusing on revealing such mechanisms could also provide valuable hints for the pathogenesis of psychiatric disorders.[99]

In conclusion, our meta-analyses suggest that SNPs located in *CREB1* are associated with major psychiatric disorders, and the decreased expression of *CREB1* might be a risk factor for the illnesses. However, the exact risk structures of the *CREB1* gene in these psychiatric disorders remain largely unknown. Future studies depicting the overall picture of the genetic and biological basis of *CREB1* and its protein product in susceptibility to psychiatric disorders are urgently needed.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGMENTS

This work was supported by grants from the Strategic Priority Research Program of the Chinese Academy of Sciences (Grant No., XDB13000000); Chinese Academy of Sciences (CAS) Pioneer Hundred Talents Program (to ML); National Natural Science Foundation of China (81471358); and Shanghai Municipal Education Commission—Gaofeng Clinical Medicine Grant Support (20152530). The Romanian sample recruitment and genotyping was funded by UEFISCDI, Bucharest, Romania (grant no. 89/2012 to MG-S) and by the German Federal Ministry of Education and Research (BMBF) through the Integrated Network IntegraMent (grant 01ZX1314A). The data of USA samples used for the analyses described in this manuscript were obtained from dbGaP accession number phs000979.v1.p1; the data of GAIN AA sample used for the analyses described in this manuscript were obtained from dbGaP accession number phs000021.v3.p2 (schizophrenia) and phs000017.v3.p1 (bipolar disorder). This research was supported by the Intramural Research Program of the NIMH (NCT00001260, 900142).

REFERENCES

1 Barch DM. Neuropsychological abnormalities in schizophrenia and major mood disorders: similarities and differences. *Curr Psychiatry Rep* 2009; **11**: 313–319.
2 Benabarre A, Vieta E, Colom F, Martinez-Aran A, Reinares M, Gasto C. Bipolar disorder, schizoaffective disorder and schizophrenia: epidemiologic, clinical and prognostic differences. *Eur Psychiatry* 2001; **16**: 167–172.
3 Greenwood TA. Positive traits in the bipolar spectrum: the space between madness and genius. *Mol Neuropsychiatry* 2017; **2**: 198–212.
4 McGuffin P, Katz R. The genetics of depression and manic-depressive disorder. *Br J Psychiatry* 1989; **155**: 294–304.
5 McGuffin P, Rijsdijk F, Andrew M, Sham P, Katz R, Cardno A. The heritability of bipolar affective disorder and the genetic relationship to unipolar depression. *Arch Gen Psychiatry* 2003; **60**: 497–502.
6 Berrettini WH. Are schizophrenic and bipolar disorders related? A review of family and molecular studies. *Biol Psychiatry* 2000; **48**: 531–538.
7 Grigoroiu-Serbanescu M, Rietschel M, Hauser J, Czerski PM, Herms S, Sun X *et al.* Commingling analysis of age-of-onset in bipolar I disorder and the morbid risk for major psychoses in first degree relatives of bipolar I probands. *J Affect Disord* 2014; **168**: 197–204.
8 Lichtenstein P, Yip BH, Bjork C, Pawitan Y, Cannon TD, Sullivan PF *et al.* Common genetic determinants of schizophrenia and bipolar disorder in Swedish families: a population-based study. *Lancet* 2009; **373**: 234–239.
9 Cross-Disorder Group of the Psychiatric Genomics Consortium. Identification of risk loci with shared effects on five major psychiatric disorders: a genome-wide analysis. *Lancet* 2013; **381**: 1371–1379.
10 Craddock N, Owen MJ. The Kraepelinian dichotomy - going, going... but still not gone. *Br J Psychiatry* 2010; **196**: 92–95.
11 Moskvina V, Craddock N, Holmans P, Nikolov I, Pahwa JS, Green E *et al.* Gene-wide analyses of genome-wide association data sets: evidence for multiple common risk alleles for schizophrenia and bipolar disorder and for overlap in genetic risk. *Mol Psychiatry* 2009; **14**: 252–260.
12 Ding Y, Chang LC, Wang X, Guilloux JP, Parrish J, Oh H *et al.* Molecular and genetic characterization of depression: overlap with other psychiatric disorders and aging. *Mol Neuropsychiatry* 2015; **1**: 1–12.
13 Chang H, Hoshina N, Zhang C, Ma Y, Cao H, Wang Y *et al.* The protocadherin 17 gene affects cognition, personality, amygdala structure and function, synapse development and risk of major mood disorders. *Mol Psychiatry* 2017; DOI:10.1038/mp.2016.231.
14 Choi KH, Higgs BW, Wendland JR, Song J, McMahon FJ, Webster MJ. Gene expression and genetic variation data implicate PCLO in bipolar disorder. *Biol Psychiatry* 2011; **69**: 353–359.
15 Green EK, Grozeva D, Jones I, Jones L, Kirov G, Caesar S *et al.* The bipolar disorder risk allele at CACNA1C also confers risk of recurrent major depression and of schizophrenia. *Mol Psychiatry* 2010; **15**: 1016–1022.
16 Liu Y, Blackwood DH, Caesar S, de Geus EJ, Farmer A, Ferreira MA *et al.* Meta-analysis of genome-wide association data of bipolar disorder and major depressive disorder. *Mol Psychiatry* 2011; **16**: 2–4.
17 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**: 128–131.
18 Li L, Chang H, Peng T, Li M, Xiao X. Evidence of AS3MTd2d3-associated variants within 10q24.32-33 in the genetic risk of major affective disorders. *Mol Neuropsychiatry* 2017; **2**: 213–218.
19 O'Donovan MC, Craddock N, Norton N, Williams H, Peirce T, Moskvina V *et al.* Identification of loci associated with schizophrenia by genome-wide association and follow-up. *Nat Genet* 2008; **40**: 1053–1055.
20 Ou J, Li M, Xiao X. The schizophrenia susceptibility gene ZNF804A confers risk of major mood disorders. *World J Biol Psychiatry* 2017; **18**: 557–562.
21 Ripke S, O'Dushlaine C, Chambert K, Moran JL, Kahler AK, Akterin S *et al.* Genome-wide association analysis identifies 13 new risk loci for schizophrenia. *Nat Genet* 2013; **45**: 1150–1159.
22 Williams HJ, Craddock N, Russo G, Hamshere ML, Moskvina V, Dwyer S *et al.* Most genome-wide significant susceptibility loci for schizophrenia and bipolar disorder reported to date cross-traditional diagnostic boundaries. *Hum Mol Genet* 2011; **20**: 387–391.
23 Chang H, Xiao X, Li M. The schizophrenia risk gene ZNF804A: clinical associations, biological mechanisms and neuronal functions. *Mol Psychiatry* 2017; **22**: 944–953.
24 Liu Z, Huang L, Luo XJ, Wu L, Li M. MAOA variants and genetic susceptibility to major psychiatric disorders. *Mol Neurobiol* 2016; **53**: 4319–4327.
25 Tang J, Fan Y, Li H, Xiang Q, Zhang D-F, Li Z *et al.* Whole-genome sequencing of monozygotic twins discordant for schizophrenia indicates multiple genetic risk factors for schizophrenia. *J Genet Genomics* 2017; **44**: 295–306.
26 International Schizophrenia Consortium, Purcell SM, Wray NR, Stone JL, Visscher PM, O'Donovan MC *et al.* Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature* 2009; **460**: 748–752.
27 Steinberg S, de Jong S. Irish Schizophrenia Genomics Consortium, Andreassen OA, Werge T, Borglum AD *et al.* Common variants at VRK2 and TCF4 conferring risk of schizophrenia. *Hum Mol Genet* 2011; **20**: 4076–4081.
28 Li M, Wang Y, Zheng XB, Ikeda M, Iwata N, Luo XJ *et al.* Meta-analysis and brain imaging data support the involvement of VRK2 (rs2312147) in schizophrenia susceptibility. *Schizophr Res* 2012; **142**: 200–205.
29 Chang H, Zhang C, Xiao X, Pu X, Liu Z, Wu L *et al.* Further evidence of VRK2 rs2312147 associated with schizophrenia. *World J Biol Psychiatry* 2016; **17**: 457–466.
30 Chen X, Lee G, Maher BS, Fanous AH, Chen J, Zhao Z *et al.* GWA study data mining and independent replication identify cardiomyopathy-associated 5 (CMYA5) as a risk gene for schizophrenia. *Mol Psychiatry* 2011; **16**: 1117–1129.
31 Zhang R, Zhang H, Li M, Li H, Li Y, Valenzuela RK *et al.* Genetic analysis of common variants in the CMYA5 (cardiomyopathy-associated 5) gene with schizophrenia. *Prog Neuropsychopharmacol Biol Psychiatry* 2013; **46**: 64–69.
32 Li M, Luo XJ, Zhang X, Yang ZH, Xiang K, Xiao X *et al.* A common variant of the cardiomyopathy associated 5 gene (CMYA5) is associated with schizophrenia in Chinese population. *Schizophr Res* 2011; **129**: 217–219.

33 Wang Q, He K, Li Z, Chen J, Li W, Wen Z *et al.* The CMYA5 gene confers risk for both schizophrenia and major depressive disorder in the Han Chinese population. *World J Biol Psychiatry* 2014; **15**: 553–560.
34 Hoshina N, Tanimura A, Yamasaki M, Inoue T, Fukabori R, Kuroda T *et al.* Protocadherin 17 regulates presynaptic assembly in topographic corticobasal Ganglia circuits. *Neuron* 2013; **78**: 839–854.
35 Li M, Luo XJ, Rietschel M, Lewis CM, Mattheisen M, Muller-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**: 452–461.
36 Carlezon WA Jr., Duman RS, Nestler EJ. The many faces of CREB. *Trends Neurosci* 2005; **28**: 436–445.
37 Weeber EJ, Levenson JM, Sweatt JD. Molecular genetics of human cognition. *Mol Interv* 2002; **2**: 376–91, 39.
38 Wallace TL, Stellitano KE, Neve RL, Duman RS. Effects of cyclic adenosine monophosphate response element binding protein overexpression in the basolateral amygdala on behavioral models of depression and anxiety. *Biol Psychiatry* 2004; **56**: 151–160.
39 Zubenko GS, Hughes HB 3rd, Maher BS, Stiffler JS, Zubenko WN, Marazita ML. Genetic linkage of region containing the CREB1 gene to depressive disorders in women from families with recurrent, early-onset, major depression. *Am J Med Genet* 2002; **114**: 980–987.
40 Blendy JA. The role of CREB in depression and antidepressant treatment. *Biol Psychiatry* 2006; **59**: 1144–1150.
41 Perlis RH, Purcell S, Fagerness J, Cusin C, Yamaki L, Fava M *et al.* Clinical and genetic dissection of anger expression and CREB1 polymorphisms in major depressive disorder. *Biol Psychiatry* 2007; **62**: 536–540.
42 Psychiatric Gwas Consortium Bipolar Disorder Working Group. Large-scale genome-wide association analysis of bipolar disorder identifies a new susceptibility locus near ODZ4. *Nat Genet* 2011; **43**: 977–983.
43 Muhleisen 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.
44 Smith EN, Bloss CS, Badner JA, Barrett T, Belmonte PL, Berrettini W *et al.* Genome-wide association study of bipolar disorder in European American and African American individuals. *Mol Psychiatry* 2009; **14**: 755–763.
45 Schizophrenia Working Group of the Psychiatric Genomics Consortium. Biological insights from 108 schizophrenia-associated genetic loci. *Nature* 2014; **511**: 421–427.
46 Shi J, Levinson DF, Duan J, Sanders AR, Zheng Y, Pe'er I *et al.* Common variants on chromosome 6p22.1 are associated with schizophrenia. *Nature* 2009; **460**: 753–757.
47 Major Depressive Disorder Working Group of the Psychiatric Gwas Consortium, Ripke S, Wray NR, Lewis CM, Hamilton SP, Weissman MM *et al.* A mega-analysis of genome-wide association studies for major depressive disorder. *Mol Psychiatry* 2013; **18**: 497–511.
48 Howie B, Marchini J, Stephens M. Genotype imputation with thousands of genomes. *G3 (Bethesda)* 2011; **1**: 457–470.
49 Delaneau O, Marchini J, Zagury JF. A linear complexity phasing method for thousands of genomes. *Nat Methods* 2011; **9**: 179–181.
50 Genomes Project Consortium, Abecasis GR, Auton A, Brooks LD, DePristo MA, Durbin RM *et al.* An integrated map of genetic variation from 1,092 human genomes. *Nature* 2012; **491**: 56–65.
51 Browning SR, Browning BL. Rapid and accurate haplotype phasing and missing-data inference for whole-genome association studies by use of localized haplotype clustering. *Am J Hum Genet* 2007; **81**: 1084–1097.
52 International HapMap Consortium, Altshuler DM, Gibbs RA, Peltonen L, Altshuler DM, Gibbs RA *et al.* Integrating common and rare genetic variation in diverse human populations. *Nature* 2010; **467**: 52–58.
53 Li M, Chang H, Xiao X. BDNF Val66Met polymorphism and bipolar disorder in European populations: a risk association in case-control, family-based and GWAS studies. *Neurosci Biobehav Rev* 2016; **68**: 218–233.
54 Xiao X, Luo XJ, Chang H, Liu Z, Li M. Evaluation of European Schizophrenia GWAS loci in Asian Populations via comprehensive meta-analyses. *Mol Neurobiol* 2016; **54**: 4071–4080.
55 Leek JT, Storey JD. Capturing heterogeneity in gene expression studies by surrogate variable analysis. *PLoS Genet* 2007; **3**: 1724–1735.
56 Kong Y. Btrim: a fast, lightweight adapter and quality trimming program for next-generation sequencing technologies. *Genomics* 2011; **98**: 152–153.
57 Kim D, Pertea G, Trapnell C, Pimentel H, Kelley R, Salzberg SL. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol* 2013; **14**: R36.
58 Trapnell C, Roberts A, Goff L, Pertea G, Kim D, Kelley DR *et al.* Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nat Protoc* 2012; **7**: 562–578.
59 Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 2009; **25**: 2078–2079.
60 Leek JT, Johnson WE, Parker HS, Jaffe AE, Storey JD. The sva package for removing batch effects and other unwanted variation in high-throughput experiments. *Bioinformatics* 2012; **28**: 882–883.
61 Johnson WE, Li C, Rabinovic A. Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics* 2007; **8**: 118–127.
62 Luo X, Huang L, Jia P, Li M, Su B, Zhao Z *et al.* Protein-protein interaction and pathway analyses of top schizophrenia genes reveal schizophrenia susceptibility genes converge on common molecular networks and enrichment of nucleosome (chromatin) assembly genes in schizophrenia susceptibility loci. *Schizophr Bull* 2014; **40**: 39–49.
63 Wu Y, Yao YG, Luo XJ. SZDB: a database for schizophrenia genetic research. *Schizophr Bull* 2017; **43**: 459–471.
64 Szklarczyk D, Franceschini A, Wyder S, Forslund K, Heller D, Huerta-Cepas J *et al.* STRING v10: protein-protein interaction networks, integrated over the tree of life. *Nucleic Acids Res* 2015; **43**(Database issue): D447–D452.
65 Dupont WD, Plummer WD Jr. Power and sample size calculations. A review and computer program. *Control Clin Trials* 1990; **11**: 116–128.
66 Lanz TA, Joshi JJ, Reinhart V, Johnson K, Grantham LE 2nd, Volfson D. STEP levels are unchanged in pre-frontal cortex and associative striatum in post-mortem human brain samples from subjects with schizophrenia, bipolar disorder and major depressive disorder. *PLoS ONE* 2015; **10**: e0121744.
67 Ryan MM, Lockstone HE, Huffaker SJ, Wayland MT, Webster MJ, Bahn S. Gene expression analysis of bipolar disorder reveals downregulation of the ubiquitin cycle and alterations in synaptic genes. *Mol Psychiatry* 2006; **11**: 965–978.
68 Narayan S, Tang B, Head SR, Gilmartin TJ, Sutcliffe JG, Dean B *et al.* Molecular profiles of schizophrenia in the CNS at different stages of illness. *Brain Res* 2008; **1239**: 235–248.
69 Iwamoto K, Kakiuchi C, Bundo M, Ikeda K, Kato T. Molecular characterization of bipolar disorder by comparing gene expression profiles of postmortem brains of major mental disorders. *Mol Psychiatry* 2004; **9**: 406–416.
70 GTEx Consortium. The Genotype-Tissue Expression (GTEx) project. *Nat Genet* 2013; **45**: 580–585.
71 GTEx Consortium. Human genomics. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans. *Science* 2015; **348**: 648–660.
72 Ramasamy A, Trabzuni D, Guelfi S, Varghese V, Smith C, Walker R *et al.* Genetic variability in the regulation of gene expression in ten regions of the human brain. *Nat Neurosci* 2014; **17**: 1418–1428.
73 Gilman SR, Chang J, Xu B, Bawa TS, Gogos JA, Karayiorgou M *et al.* Diverse types of genetic variation converge on functional gene networks involved in schizophrenia. *Nat Neurosci* 2012; **15**: 1723–1728.
74 Network Pathway Analysis Subgroup of Psychiatric Genomics Consortium. Psychiatric genome-wide association study analyses implicate neuronal, immune and histone pathways. *Nat Neurosci* 2015; **18**: 199–209.
75 Mitchell AC, Javidfar B, Pothula V, Ibi D, Shen EY, Peter CJ *et al.* MEF2C transcription factor is associated with the genetic and epigenetic risk architecture of schizophrenia and improves cognition in mice. *Mol Psychiatry* 2017.
76 Blizinsky KD, Diaz-Castro B, Forrest MP, Schurmann B, Bach AP, Martin-de-Saavedra MD *et al.* Reversal of dendritic phenotypes in 16p11.2 microduplication mouse model neurons by pharmacological targeting of a network hub. *Proc Natl Acad Sci USA* 2016; **113**: 8520–8525.
77 Harrington AJ, Raissi A, Rajkovich K, Berto S, Kumar J, Molinaro G *et al.* MEF2C regulates cortical inhibitory and excitatory synapses and behaviors relevant to neurodevelopmental disorders. *Elife* 2016; doi: 10.7554/eLife.20059.
78 Focking M, Lopez LM, English JA, Dicker P, Wolff A, Brindley E *et al.* Proteomic and genomic evidence implicates the postsynaptic density in schizophrenia. *Mol Psychiatry* 2015; **20**: 424–432.
79 Erk S, Mohnke S, Ripke S, Lett TA, Veer IM, Wackerhagen C *et al.* Functional neuroimaging effects of recently discovered genetic risk loci for schizophrenia and polygenic risk profile in five RDoC subdomains. *Transl Psychiatry* 2017; **7**: e997.
80 Kalkman HO. Potential opposite roles of the extracellular signal-regulated kinase (ERK) pathway in autism spectrum and bipolar disorders. *Neurosci Biobehav Rev* 2012; **36**: 2206–2213.
81 Melas PA, Lennartsson A, Vakifahmetoglu-Norberg H, Wei Y, Aberg E, Werme M *et al.* Allele-specific programming of Npy and epigenetic effects of physical activity in a genetic model of depression. *Transl Psychiatry* 2013; **3**: e255.
82 Lonze BE, Ginty DD. Function and regulation of CREB family transcription factors in the nervous system. *Neuron* 2002; **35**: 605–623.
83 Impey S, Obrietan K, Wong ST, Poser S, Yano S, Wayman G *et al.* Cross talk between ERK and PKA is required for Ca2+ stimulation of CREB-dependent transcription and ERK nuclear translocation. *Neuron* 1998; **21**: 869–883.
84 Li J, Sutter C, Parker DS, Blauwkamp T, Fang M, Cadigan KM. CBP/p300 are bimodal regulators of Wnt signaling. *EMBO J* 2007; **26**: 2284–2294.

85 Dyson HJ, Wright PE. Role of intrinsic protein disorder in the function and interactions of the transcriptional coactivators CREB-binding protein (CBP) and p300. *J Biol Chem* 2016; **291**: 6714–6722.
86 Blobel GA. CREB-binding protein and p300: molecular integrators of hematopoietic transcription. *Blood* 2000; **95**: 745–755.
87 Kumar G, Clark SL, McClay JL, Shabalin AA, Adkins DE, Xie L *et al.* Refinement of schizophrenia GWAS loci using methylome-wide association data. *Hum Genet* 2015; **134**: 77–87.
88 Juhasz G, Dunham JS, McKie S, Thomas E, Downey D, Chase D *et al.* The CREB1-BDNF-NTRK2 pathway in depression: multiple gene-cognition-environment interactions. *Biol Psychiatry* 2011; **69**: 762–771.
89 Ma L, Wu DD, Ma SL, Tan L, Chen X, Tang NL *et al.* Molecular evolution in the CREB1 signal pathway and a rare haplotype in CREB1 with genetic predisposition to schizophrenia. *J Psychiatr Res* 2014; **57**: 84–89.
90 Sokolowski M, Wasserman J, Wasserman D. Polygenic associations of neurodevelopmental genes in suicide attempt. *Mol Psychiatry* 2016; **21**: 1381–1390.
91 Barrot M, Olivier JD, Perrotti LI, DiLeone RJ, Berton O, Eisch AJ *et al.* CREB activity in the nucleus accumbens shell controls gating of behavioral responses to emotional stimuli. *Proc Natl Acad Sci USA* 2002; **99**: 11435–11440.
92 Yamada S, Yamamoto M, Ozawa H, Riederer P, Saito T. Reduced phosphorylation of cyclic AMP-responsive element binding protein in the postmortem orbitofrontal cortex of patients with major depressive disorder. *J Neural Transm (Vienna)* 2003; **110**: 671–680.
93 Nibuya M, Nestler EJ, Duman RS. Chronic antidepressant administration increases the expression of cAMP response element binding protein (CREB) in rat hippocampus. *J Neurosci* 1996; **16**: 2365–2372.
94 Jeon SH, Seong YS, Juhnn YS, Kang UG, Ha KS, Kim YS *et al.* Electroconvulsive shock increases the phosphorylation of cyclic AMP response element binding protein at Ser-133 in rat hippocampus but not in cerebellum. *Neuropharmacology* 1997; **36**: 411–414.
95 Chen AC, Shirayama Y, Shin KH, Neve RL, Duman RS. Expression of the cAMP response element binding protein (CREB) in hippocampus produces an antidepressant effect. *Biol Psychiatry* 2001; **49**: 753–762.
96 O'Donnell JM, Zhang HT. Antidepressant effects of inhibitors of cAMP phosphodiesterase (PDE4). *Trends Pharmacol Sci* 2004; **25**: 158–163.
97 Fleischhacker WW, Hinterhuber H, Bauer H, Pflug B, Berner P, Simhandl C *et al.* A multicenter double-blind study of three different doses of the new cAMP-phosphodiesterase inhibitor rolipram in patients with major depressive disorder. *Neuropsychobiology* 1992; **26**: 59–64.
98 Pliakas AM, Carlson RR, Neve RL, Konradi C, Nestler EJ, Carlezon WA Jr. Altered responsiveness to cocaine and increased immobility in the forced swim test associated with elevated cAMP response element-binding protein expression in nucleus accumbens. *J Neurosci* 2001; **21**: 7397–7403.
99 Barco A, Pittenger C, CREB Kandel ER. memory enhancement and the treatment of memory disorders: promises, pitfalls and prospects. *Expert Opin Ther Targets* 2003; **7**: 101–114.

Supplementary Information accompanies the paper on the Molecular Psychiatry website (http://www.nature.com/mp)

Bipolar disorder sample

PGC1 sample

The Psychiatric Genomics Consortium (PGC1) BPD group recently conducted a meta-analysis of large-scale genome-wide data on BPD in populations of European ancestry. In this study, all patients had experienced pathologically relevant episodes of elevated mood (mania or hypomania) and met the established criteria for BPD within the primary study classification system, and the subjects with a low probability of having BPD from the same geographic and ethnic populations were selected to use as a control. We extracted the results of candidate SNPs from the primary GWAS samples (7,481 cases/9,250 controls). Detailed descriptions of the samples, data quality, genomic controls and statistical analyses can be found in the original GWAS (1).

France sample

Patients with BPD and controls were recruited as part of a large study on genetics of BPD in France (Paris-Creteil, Bordeaux, Nancy) with a protocol approved by relevant IRBs and with written informed consent. Cases were of French descent for more than three generations and were all been assessed by a well-trained psychiatrist or psychologist with the DIGS (2) and the FIGS. Diagnoses were based on structured interviews supplemented by medical case notes, mood scales and a self rating questionnaire assessing dimensions. Genotyping of controls were provided by the Centre National de Génotypage (M Lathrop, Evry). Patients and controls were genotyped on the Illumina platform (HumanHap300, HumanHap550, HumanHap 610-quad).

Germany II sample

Cases were ascertained from consecutive admissions to the inpatient units of the Department of Psychiatry and Psychotherapy at the University of Bonn and at the Central Institute for Mental Health in Mannheim, University of Heidelberg, as well as at other collaborating psychiatric university hospitals in Germany. DSM-IV lifetime diagnoses of BPD were assigned using a consensus best-estimate procedure, based on all available information, including semi-structured interviews (AMDP), medical records, and family history. In addition, the OPCRIT system (3) was used for the detailed poly-diagnostic documentation of symptoms.

Controls were ascertained from the population-based Heinz Nixdorf Recall Study.(4) Study protocols were reviewed and approved in advance by Institutional Review Boards of the participating institutions. All subjects provided written informed consent and were genotyped using the Illumina platform.

Detailed descriptions of the samples, data quality, genomic controls and statistical analyses can be found in the original GWAS (5).

Iceland sample

The Iceland sample consisted of 541 subjects with BPD and 34,546 population controls. Patients and controls were Icelandic and were recruited throughout the country. Diagnoses were assigned according to RDC through the use of the SADS-L for 303 subjects. DSM-IV BPD diagnoses were obtained through the use of the Composite International Diagnostic Interview (CIDI-Auto) for 82 subjects. In addition, there were 150 subjects with ICD-9 or ICD-10 BPD diagnoses and nine subjects with DSM-III BPD diagnoses.

The 34,546 controls were recruited as a part of various genetic programs at

deCODE and were not screened for psychiatric disorders. Approval for the study was granted by the National Bioethics Committee of Iceland and the Icelandic Data Protection Authority and written informed consent was obtained for all participants.

Poland sample

All patients were recruited from consecutive hospital admissions and were directly interviewed with the Structured Clinical Interview for DSM-IV-TR-Axis I Disorders (Patient Edition). Information provided via medical records and interviews of family members was also used in a best estimate procedure of diagnosis on the basis of DSM-IV-TR criteria. The control samples were population-based, drawn from the same population as the patients. The ethnicity of the patients and control subjects was determined by genealogical investigation up to the grandparental generation.

Sweden sample

SBP Bipolar cases were recruited from St. Göran's Hospital in Stockholm, Sweden. All participants provided written informed consent to participate in a genetic study of BPD, and the study was approved by the Regional Ethics Committee of Stockholm. Diagnoses were based on physician administered ADE (6) and MINI (7).

BPD cases were identified from the Swedish Bipolar Quality Assurance Registry. Patient information within the registry includes disease sub-classification, psychosis, age at onset, number of manic and depressive episodes, number of hospitalizations and family history. Participants provided written informed consent to participate in a genetic study of psychiatric disease, and the study was approved by the Regional Ethics Committee of Stockholm.

Hospital Discharge Registry (HDR) Bipolar cases were identified from the Swedish Hospital Discharge Registry if they a) have at least two admissions with discharge diagnoses of BPD) were born in Sweden or another Nordic country. The register contains a nearly complete record of all individuals hospitalized in Sweden since 1973. Diagnoses were established by an attending physician and were shown to have high sensitivity and specificity (8). The study was approved by the Regional Ethics Committee of Stockholm. All participants provided written informed consent to participate in genetic studies of psychotic disorders and were interviewed by a research nurse about other medical conditions.

The SBP BPD cases were recruited from the Stockholm County catchment area. All patients provided written informed consent to participate in a genetic study of BPD, and the study was approved by the Regional Ethics Committee of Stockholm. Diagnoses were made according to the DSM-IV criteria.

Sweden control samples were obtained from the Swedish Hospital Discharge Registry on the condition they had never received discharge diagnoses of BPD, schizophrenia and/or schizoaffective disorder.

Australia sample

Subjects were ascertained through two studies: 1) a BPD pedigree sample (described in McAuley et al. (9)) and 2) a specialized Sydney Black Dog Institute BPD clinic sample (described in Mitchell et al. 2009 (10)). All subjects were interviewed by trained research staff using the DIGS or SCID, using best-estimate DSM-IV diagnoses derived from those instruments, medical records and FIGS. First, for the pedigree sample, only one BPD subject per family was included in the case sample. Pedigrees were only included in the original genetic study if there was unilineal inheritance, and at least two

BPD subjects including at least one with bipolar I disorder. Subjects were ascertained through clinical presentations to the Mood Disorders Unit at the Prince of Wales Hospital in Sydney, direct referrals from Australian clinicians, and BPD consumer organizations. Second, for the clinic sample, subjects comprised consecutive subjects referred by psychiatrists or general practitioners for specialized clinical review. All patients provided written informed consent to participate in this study and the study was approved by the local ethics committee. Patients were included in the BOMA study and genotyped at the Life & Brain Centre in Bonn.

Australian controls were drawn from families participating in the Brisbane Longitudinal Twin Study, an unselected community sample recruited to take part in studies of melanoma risk factors, cognition, and other phenotypes. Subjects were not screened for any phenotype relevant to BPD. The study was approved by the ethic committee and all proband gave written informed consent. All subjects were genotyped as a single project by deCODE and have been through an extensive QC process including exclusion for non-European ancestry. The sample is overwhelmingly of northern European origin, predominately from the British Isles.

Romania sample

The Romania sample consisted of 451 BPD patients and 318 healthy controls. All patients were recruited from consecutive hospital admissions and directly interviewed with the Structured Clinical Interview for DSM-IV-TR-Axis I Disorders - Patient Version (SCID-I, 1994) and the Diagnostic Interview for Genetic Studies (DIGS) version 3.0 (1999). Information provided by medical records and interviews of family members was also used in a best estimate procedure of diagnosis on the basis of DSM-IV-TR criteria. The control sample was population-based, drawn from the same population as the patients, and was screened for major psychiatric disorders prior to inclusion. The ethnicity of the patients and control subjects was determined by genealogical investigation to the grandparental generation. Only the patient sample was previously reported in other collaborative studies. The controls were genotyped on Illumina Omni-Express chips at the Life & Brain Center in Bonn, and the patients were also genotyped on Illumina chips (partly on Omni1-Quad) at the Life & Brain Center in Bonn.

GAIN AA sample

Cases were selected from those collected and characterized by the Bipolar Consortium over the past 18 years. All subjects were diagnosed with a standard best estimate (BEFD) procedure. For the BiGS GWA study we selected unrelated Diagnostic and Statistical Manual (DSM) IV-defined BPI subjects. Among those participants, African American (AA) status was based on self-report of at least one grandparent being of AA. A total of 362 of these AA BPD subjects were ultimately included in the BiGS analyses after review of best estimate diagnoses. Controls were ascertained separately through a NIMH-supported contract mechanism between Dr. Pablo Gejman and Knowledge Networks, Inc.; this mechanism allowed the ascertainment of 4,586 subjects across the U.S. who agreed to donate a blood sample for transformation into lymphoblastoid cell lines and to respond to a medical questionnaire. Only individuals with complete or near-complete psychiatric questionnaire data who did not fulfill diagnostic criteria for major depression and denied a history of psychosis or BPD were included as controls for the BiGS analyses. The control groups included 716 AA subjects.

All case subjects were interviewed with the Diagnostic Interview for Genetic Studies. These included diagnosis by DSM-IV, DSM-IIIR, and the Research Diagnostic

Criteria (RDC), as well as age of onset, number of episodes for depression, hypomania and mania, temporal relationship of mood disorder to substance abuse and psychosis, evidence of mixed episodes and rapid cycling, and a summary of the family history information. All of these indicators were scored independently by a senior clinician (generally a psychiatrist) based on all available information, including medical records, interviewer observations, the coded DIGS, and the Family Instrument for Genetic Studies ('FIGS,' developed for the NIMH Genetics Initiative). The FIGS incorporates clinician judgment on family patterns of illness, including presence or absence of BPD, unipolar disorder, and/or other psychiatric disorders in first and second-degree relatives. Genotyping of the AA samples was carried out using the Affymetrix Genome-Wide Human SNP Array 6.0. Detailed descriptions of the samples, data quality, genomic controls and statistical analyses can be found in the original GWAS (11).

USA EA sample

The genotype data in USA EA BPD sample was obtained from dbGaP accession number phs000979.v1.p1. In brief, postmortem brains of BPD patients and healthy controls were collected at the Human Brain Collection Core, NIMH with informed consent from the legal next of kin (NIMH protocol 90-M-0142), and at the Brain and Tissue Bank for Developmental Disorders of the NICHD (contracts NO1-HD-4-3368 and NO1-HD-4-3383) and through the Stanley Medical Research Institute. All BPD patients met DSM-IV criteria for a lifetime Axis I diagnosis of BPD, and controls had no history of psychological or psychiatric problems. Genotyping was performed using HumanHap650Yv3.0 and Human1M-Duov3_B.

Schizophrenia sample

PGC2 sample

Recently, the Psychiatric Genomics Consortium (PGC) reported the largest schizophrenia GWAS so far (PGC2 release). In brief, genome-wide genotypes from 49 independent samples (46 of European and 3 of Asian ancestry, with a final sample size up to 35,476 schizophrenia cases and 46,839 controls) were combined and systematically meta-analyzed. The study identified 128 independent associations that reached genome-wide significant level and most of the significant associations (83 of 128) are newly reported in that study. Summary statistics (genome-wide SNP associations, i.e., P values) from PGC2 were used as input in this study. Detailed information about sample description, diagnosis, genotyping, and statistical analyses can be found in the original study and PGC website (http://www.med.unc.edu/pgc/) (12).

GAIN AA sample

Cases. Self-reported European ancestry (EA) and African American (AA) unrelated adult cases with DSM-IIIR (SGI study) or DSM-IV (MGS1, MGS2 studies) schizophrenia (SZ) or schizoaffective disorder (SA) were collected under institutional review board-approved protocols in three studies, Schizophrenia Genetics Initiative (SGI), Molecular Genetics of Schizophrenia Part 1 (MGS1), and MGS2. Briefly: SGI subjects (2.4%) were recruited by three research groups in the United States in a study designed to collect families with affected sibling pairs (and other affected members when available), ascertaining through probands with schizophrenia recruited through clinical settings. One member per family (proband) SGI family was selected for the GWAS sample. MGS1 subjects (10.4%) were collected by ten sites (see acknowledgements) in the United States and Australia in a study designed to collect affected sibling pairs (and other affected members when available), ascertaining through probands with schizophrenia recruited through clinical settings and community residences. One member per MGS1 family (proband) was selected for the GWAS sample. MGS2 (87.2%) subjects were collected by the same ten sites as MGS1. Single cases were selected. Probands had DSM-IV schizophrenia, or DSM-IV schizoaffective disorder but with a history of meeting Criteria A for schizophrenia for at least six months of the course of illness. The three studies used the same clinical assessments (with minor modifications for MGS1 and 2 compared with SGI) and diagnostic procedures. Interviews were conducted in person by trained research interviewers using the Diagnostic Interview for Genetic Studies v2.0 (DIGS); the Family Interview for Genetic Studies (FIGS) was completed with an informant where possible; and medical records were obtained with the subject's written consent.

Controls. EA and AA unrelated adult control subjects were collected under MGS2. Briefly, random digit dialing to area codes were used and selected to represent the national population, to recruit individuals to join a nationwide survey panel. Communication with panel members is by email and web interaction, but initial recruitment does not require internet connection, and those who agree to join but have no internet access are given a web TV to facilitate participation. Panel members are then contacted to invite participation in specific surveys. All EA control subjects and 41% of AA controls were recruited using these methods. There was an insufficient number of AA individuals in the panel who were interested in participating to complete the collection, thus a second survey research company (SSI Opt-In) was contracted. SSI/Opt-in uses banner ads on websites to recruit participants. In total, 3,364 EA and 1,301 AA control subjects were recruited, of which 772 AA subjects (59%) were recruited by SSI/Opt-in.

Detailed descriptions of the samples, data quality, genomic controls and statistical analyses can be found in the original GWAS (13).

USA AA sample

The genotype data in USA AA SCZ sample was obtained from dbGaP accession number phs000979.v1.p1. In brief, postmortem brains of SCZ patients and healthy controls were collected at the Human Brain Collection Core, NIMH with informed consent from the legal next of kin (NIMH protocol 90-M-0142), and at the Brain and Tissue Bank for Developmental Disorders of the NICHD (contracts NO1-HD-4-3368 and NO1-HD-4-3383) and through the Stanley Medical Research Institute. All SCZ patients met DSM-IV criteria, and controls had no history of psychological or psychiatric problems. Genotyping was performed using HumanHap650Yv3.0 and Human1M-Duov3_B.

Major depressive disorder sample

PGC1 sample

The MDD GWAS included 9,240 patients and 9,519 controls. The cases were defined by having lifetime diagnoses of MDD according to DSM-IV criteria by trained interviewers, or based on clinician-administered DSM-IV checklists using structured diagnostic instruments. For most of these participants, cases were obtained from clinical sources, and controls were randomly selected from the population. We determined the relatedness of all pairs of individuals using genotypes of SNPs present on all platforms, and excluded one of each duplicate or closely related pair. Detailed descriptions of the samples, data quality, genomic controls and statistical analyses can be found in the original GWAS (14).

USA EA sample

The genotype data in USA EA MDD sample was obtained from dbGaP accession number phs000979.v1.p1. In brief, postmortem brains of MDD patients and healthy controls were collected at the Human Brain Collection Core, NIMH with informed consent from the legal next of kin (NIMH protocol 90-M-0142), and at the Brain and Tissue Bank for Developmental Disorders of the NICHD (contracts NO1-HD-4-3368 and NO1-HD-4-3383) and through the Stanley Medical Research Institute. All MDD patients met DSM-IV criteria, and controls had no history of psychological or psychiatric problems. Genotyping was performed using HumanHap650Yv3.0 and Human1M-Duov3_B.

REFERENCES

1. Psychiatric Gwas Consortium Bipolar Disorder Working Group (2011): Large-scale genome-wide association analysis of bipolar disorder identifies a new susceptibility locus near ODZ4. *Nat Genet*. 43:977-983.

2. Nurnberger JI, Jr., Blehar MC, Kaufmann CA, York-Cooler C, Simpson SG, Harkavy-Friedman J, et al. (1994): Diagnostic interview for genetic studies. Rationale, unique features, and training. NIMH Genetics Initiative. *Arch Gen Psychiatry*. 51:849-859; discussion 863-864.

3. McGuffin P, Farmer A, Harvey I (1991): A polydiagnostic application of operational criteria in studies of psychotic illness. Development and reliability of the OPCRIT system. *Arch Gen Psychiatry*. 48:764-770.

4. Schmermund A, Mohlenkamp S, Stang A, Gronemeyer D, Seibel R, Hirche H, et al. (2002): 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. Risk Factors, Evaluation of Coronary Calcium and Lifestyle. *Am Heart J*. 144:212-218.

5. Muhleisen TW, Leber M, Schulze TG, Strohmaier J, Degenhardt F, Treutlein J, et al. (2014): Genome-wide association study reveals two new risk loci for bipolar disorder. *Nat Commun*. 5:3339.

6. Spitzer RL, Williams JB, Gibbon M, First MB (1992): The Structured Clinical Interview for DSM-III-R (SCID). I: History, rationale, and description. *Arch Gen Psychiatry*. 49:624-629.

7. Sheehan DV, Lecrubier Y, Sheehan KH, Amorim P, Janavs J, Weiller E, et al. (1998): 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 20:22-33;quiz 34-57.

8. Sellgren C, Landen M, Lichtenstein P, Hultman CM, Langstrom N (2011): Validity of bipolar disorder hospital discharge diagnoses: file review and multiple register linkage in Sweden. *Acta Psychiatr Scand*. 124:447-453.

9. McAuley EZ, Fullerton JM, Blair IP, Donald JA, Mitchell PB, Schofield PR (2009): Association between the serotonin 2A receptor gene and bipolar affective disorder in an Australian cohort. *Psychiatr Genet*. 19:244-252.

10. Mitchell PB, Johnston AK, Corry J, Ball JR, Malhi GS (2009): Characteristics of bipolar disorder in an Australian specialist outpatient clinic: comparison across large datasets. *Aust N Z J Psychiatry*. 43:109-117.

11. Smith EN, Bloss CS, Badner JA, Barrett T, Belmonte PL, Berrettini W, et al. (2009): Genome-wide association study of bipolar disorder in European American and African American individuals. *Mol Psychiatry*. 14:755-763.

12. Schizophrenia Working Group of the Psychiatric Genomics Consortium (2014): Biological insights from 108 schizophrenia-associated genetic loci. *Nature*. 511:421-427.

13. Shi J, Levinson DF, Duan J, Sanders AR, Zheng Y, Pe'er I, et al. (2009): Common variants on chromosome 6p22.1 are associated with schizophrenia. *Nature*. 460:753-757.

14. Major Depressive Disorder Working Group of the Psychiatric Gwas Consortium, Ripke S, Wray NR, Lewis CM, Hamilton SP, Weissman MM, et al. (2013): A mega-analysis of genome-wide association studies for major depressive disorder. *Mol Psychiatry*. 18:497-511.