Journal of Psychiatric Research 57 (2014) 84–89

Contents lists available at ScienceDirect

Journal of Psychiatric Research

journal homepage: www.elsevier.com/locate/psychires

Molecular evolution in the CREB1 signal pathway and a rare haplotype in *CREB1* with genetic predisposition to schizophrenia

Liang Ma [a, b], Dong-Dong Wu [c], Suk Ling Ma [d], Liwen Tan [e], Xiaogang Chen [e], Nelson L.S. Tang [f, g, h, **], Yong-Gang Yao [a, b, f, *]

[a] Key Laboratory of Animal Models and Human Disease Mechanisms of the Chinese Academy of Sciences & Yunnan Province, Kunming Institute of Zoology, Kunming, Yunnan 650223, China
[b] Kunming College of Life Science, University of Chinese Academy of Sciences, Kunming, Yunnan 650201, China
[c] State Key Laboratory of Genetic Resources and Evolution, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming, Yunnan 650223, China
[d] Department of Psychiatry, The Chinese University of Hong Kong, Hong Kong, China
[e] Institute of Mental Health, The Second Xiangya Hospital, Central South University, Changsha, Hunan 410011, China
[f] KIZ/CUHK Joint Laboratory of Bioresources and Molecular Research in Common Diseases, Kunming, Yunnan, China
[g] Department of Chemical Pathology and Laboratory of Genetics of Disease Susceptibility, Li Ka Shing Institute of Health Sciences, The Chinese University of Hong Kong, Hong Kong, China
[h] Functional Genomics and Biostatistical Computing Laboratory, Shenzhen Research Institute, The Chinese University of Hong Kong, Shenzhen, China

ARTICLE INFO

Article history:
Received 28 April 2014
Received in revised form
8 June 2014
Accepted 16 June 2014

Keywords:
CREB1 signal pathway
Schizophrenia
Positive selection
Association

ABSTRACT

CREB1 is a cAMP responsive transcriptional factor which plays a key role in neural development. CREB1 signal pathway (CSP) has been implicated repeatedly in studies of predisposition for schizophrenia. We speculated that CSP has undergone positive selection during evolution of modern human and some genes that have undergone natural selection in the past may predispose to schizophrenia (SCZ) in modern time. Positive selection and association analysis were employed to explore the molecular evolution of CSP and association with schizophrenia. Our results showed a pan-ethnic selection event on *NRG1* and *CREB1*, as confirmed in all 14 ethnic populations studied, which also suggested a selection process occurred before the "Out of Africa" scenario. Analysis of 62 SNPs covering 6 CSP genes in 2019 Han Chinese (976 SCZ patients and 1043 healthy individuals) showed an association of two SNPs (rs4379857, $P = 0.009$, OR [95% CI]: 1.200 [1.379–1.046]; rs2238751, $P = 0.023$, OR [95% CI]: 1.253 [1.522–1.032]) with SCZ. However, none of these significances survived after multiple testing corrections. Nonetheless, we observed an association of a rare *CREB1* haplotype CCGGC (Bonferroni corrected $P = 1.74 \times 10^{-5}$) with SCZ. Our study showed that there was substantial population heterogeneity in genetic predisposition to SCZ, and different genes in the CSP pathway may predispose to SCZ in different populations.

© 2014 Elsevier Ltd. All rights reserved.

1. Introduction

Schizophrenia (SCZ) [OMIM: 181500] is a severe and chronic neuropsychiatric disorder, with a lifetime prevalence of approximately 1%. The heritability was estimated to be 81% (Sullivan et al., 2003). Despite of a consistent prevalence and high heritability globally, the genetic underpinnings of schizophrenia have not been sufficiently resolved. Most genetic association studies were not successfully replicated in different ethnic population. For example, despite the strong support of putative mechanism, some studies did not replicate the association between *NRG1* (Neuregulin 1) and SCZ (Doi et al., 2009; Gong et al., 2009; Li et al., 2006; Munafo et al., 2006). Previous association studies of *DGCR2* (DiGeorge syndrome critical region gene 2) with SCZ had yielded conflicting results (Betcheva et al., 2009; Georgi et al., 2009; Ishiguro et al., 2008; Shifman et al., 2006). We failed to replicate the positive signals in two genome-wide association studies (GWAS) of Han Chinese patients with SCZ (Shi et al., 2011; Yue et al., 2011) in an independent Han Chinese population from Hunan Province of China (Ma et al., 2013). SCZ might be caused by the disorder of neural circuits in brain and

* Corresponding author. Key Laboratory of Animal Models and Human Disease Mechanisms, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming, Yunnan, 650223, China. Tel./fax: +86 871 65180085.
** Corresponding author. Department of Chemical Pathology, The Chinese University of Hong Kong, Hong Kong, China.
E-mail addresses: nelsontang@cuhk.edu.hk (N.L.S. Tang), yaoyg@mail.kiz.ac.cn, ygyaozh@gmail.com (Y.-G. Yao).

http://dx.doi.org/10.1016/j.jpsychires.2014.06.008
0022-3956/© 2014 Elsevier Ltd. All rights reserved.

many predisposition genes reported so far play roles in neural development (Lewis and Sweet, 2009). Therefore, it will be worthy to analyze the key genes and their pathway regulating neural development, and CSP is a good candidate.

SCZ may be related to de-coupling of molecular components underlying intelligence and language, some of which might have an important role during the evolution of human brain, which favors a selective advantage for higher cognitive function (Keller and Miller, 2006). This situation of "advantageous" gene in ancient time turned into disease gene in modern time could be found in diabetes mellitus (Neel, 1962). Here, we hypothesized genes and pathways that have been subjected to adaptive evolution since the divergence of human populations might be involved in the pathogenesis of SCZ and other mental disorders. Therefore, we studied the evolutionary signals of CREB1 signaling pathway (CSP).

CREB1 (cAMP responsive element binding protein 1) is a transcriptional factor in response to cAMP level and, in central nervous system, it plays crucial roles in mediating neurotrophin-induced axon growth (Lonze et al., 2002). Lacking CREB and CREM (cAMP response element modulator) in neurons of the developing central nervous system could lead to apoptosis and neurodegeneration in hippocampus and dorsolateral striatum (Mantamadiotis et al., 2002). Animal and behavioral studies have demonstrated that CREB1 signaling pathway can influence mood, addiction, learning and memory by activating an array of proteins (Carlezon et al., 2005). Two large European studies found that multiple single nucleotide polymorphisms (SNPs) of *CREB1* were significantly associated with SCZ and bipolar disorder (Li et al., 2014; Ripke et al., 2013). Previous studies of candidate genes in SCZ also focused on CSP genes (Kvajo et al., 2010). In addition, CREB1 could regulate gene expression of a group of neuronal proteins, and ultimately affect the function of neurons, brain circuits and complex behavior (Carlezon et al., 2005).

Recently, the catalog of downstream genes regulated by transactivation of CREB1 was systematically characterized (Zhang et al., 2005). Among the gene list, *NRG1*, *CNNM2* (cyclin M2), *NT5C2* (5′-nucleotidase, cytosolic II), *DGCR2* and *OPCML* (opioid binding protein/cell adhesion molecule-like) are related to neural development and have been implicated in SCZ (Aberg et al., 2013; Bergen et al., 2012; Law et al., 2012; Mistry et al., 2013; Ripke et al., 2011; Xu et al., 2011). This downstream gene set of CSP play essential roles in neural development and synaptic plasticity (Fazzari et al., 2010; Kato et al., 2011; Li et al., 2007). Disease-associated SNPs in *NRG1* could affect expression of specific NRG1 isoforms in human brain (Law et al., 2006; Tan et al., 2007), and this abnormal expression was associated with human brain activation and psychotic symptoms (Hall et al., 2006; Hatzimanolis et al., 2013; Law et al., 2012). *CNNM2-NT5C2* was reported to reach a genome-wide significance in a large European SCZ-GWAS (Ripke et al., 2011) and in a meta-analysis with a large Swedish sample (Bergen et al., 2012). This positive association was consistently validated in a family-based study (Aberg et al., 2013) and a comprehensive replication study (Bergen et al., 2012). *DGCR2* encodes a potential adhesion receptor protein with unclear function (Kajiwara et al., 1996). Deletion of this region accounted for ~2% of sporadic schizophrenia (Xu et al., 2011). *OPCML* is highly expressed in hippocampus and cortex and was implicated in neuronal connection (Struyk et al., 1995). Genome-wide expression study of a large SCZ cohort showed that *OPCML* was significantly down-regulated in prefrontal cortex of SCZ patients (Mistry et al., 2013). Recent association studies supported its genetic effects on SCZ (O'Donovan et al., 2008).

In light of the above findings, we speculated that historical molecular evolution of CSP genes may contribute to emergence of SCZ pathogenesis in modern society (Fig. 1). Functional

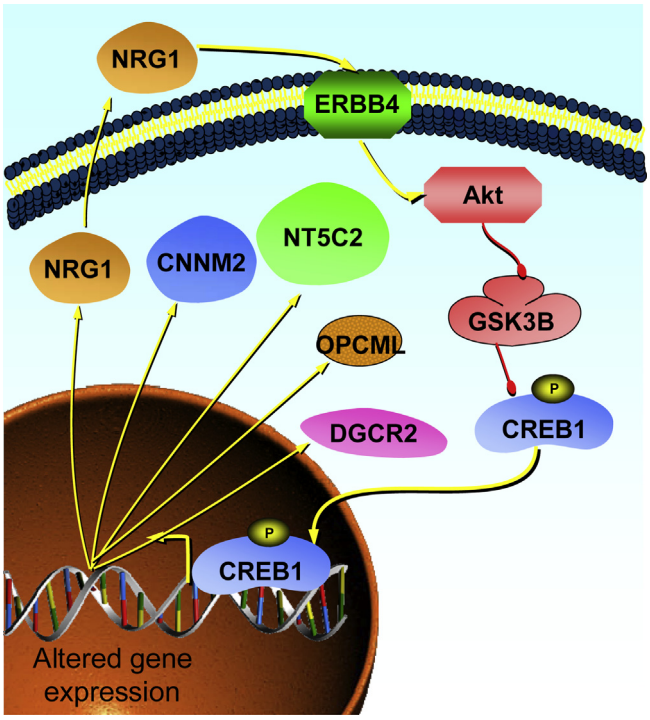


Fig. 1. CREB1 signal pathway. Depicted is the interaction of intracellular signal pathways involving *CREB1* and other susceptibility genes. Yellow lines indicate activation, whereas red lines indicate inhibition or deactivation. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

adaptations that drive species differentiation could be found in genes that have undergone positive selection (Vamathevan et al., 2008). Therefore we investigated whether there were signatures of selection effect on the six candidates CSP genes related to neural development during the origin of modern human and determined whether CSP genetic variants contribute to SCZ risk in Han Chinese.

2. Materials and methods

2.1. Molecular evolution analysis of CSP genes using the 1000 genome project data

Imputation data for *CREB1* and its target genes (*NRG1*, *CNNM2*, *NT5C2*, *DGCR2* and *OPCML*) in the 1000 Genomes Project were obtained from web (http://mathgen.stats.ox.ac.uk/impute/). This data set includes 97 Han Chinese in Beijing [CHB], 100 Han Chinese South [CHS] and 89 Japanese in Tokyo [JPT], 85 Utah residents (CEPH) with Northern and Western European ancestry [CEU], 93 Finnish from Finland [FIN], 89 British from England and Scotland [GBR], 14 Iberian populations in Spain [IBS] and 98 Toscani in Italy [TSI], 61 African Ancestry in Southwest US [ASW], 88 Yoruba in Ibadan, Nigeria [YRI] and 97 Luhya in Webuye, Kenya [LWK], 60 Colombian in Medellin, Colombia [CLM], 66 Mexican Ancestry in Los Angeles, CA [MXL] and 55 Puerto Rican in Puerto Rico [PUR]. Nucleotide diversity (π) and Tajima's *D* (Tajima, 1989) were used to measure the deviation from neutrality in each population. These two statistics were estimated on the basis of 10 kb windows. In addition, integrated haplotype score (iHS) were used to detect long haplotypes that have not been broken by recombination (Voight et al., 2006).

2.2. Genetic analysis of CSP genes in a Han Chinese SCZ patient cohort

A total of 976 unrelated SCZ patients (mean age ± SD 24.92 ± 8.33) and 1043 matched healthy controls (mean age ± SD 37.41 ± 14.22), all of Han Chinese, were recruited from Hunan Province in South Central China, which were described in our previous studies (Ma et al., 2013). In brief, patients were diagnosed independently by two experienced psychiatrists as SCZ according to DSM-IV criteria. The controls were clinically diagnosed as no family history of psychiatric disorders. All participants gave written informed consent. This study was approved by the ethics committee of the Kunming Institute of Zoology.

Genomic DNA was extracted from whole blood by using the AxyPrep™ Blood Genomic DNA Miniprep Kit (Axygen, USA). 68 SNPs of six CSP genes (*CREB1*, *NRG1*, *CNNM2-NT5C2*, *DGCR2* and *OPCML*) were selected for genotyping. These SNPs included susceptible SNPs reported in previous genome-wide association studies for SCZ (Aberg et al., 2013; Bergen et al., 2012; Ripke et al., 2011) and tagSNPs that covering the entire gene region. SNP location, gene structure and linkage disequilibrium pattern of these SNPs were shown in Supplementary Fig. 1. SNPs were genotyped using SNaPshot (Supplementary Table 1). GeneMarker software (Holland and Parson, 2011) was used to read the genotyping results. The detailed procedures were described in the supplementary file (Supplementary methods).

Calculations of Hardy–Weinberg equilibrium (HWE) test, allele and genotype comparisons were performed using PLINK (http://pngu.mgh.harvard.edu/~purcell/plink/). Haplotype test was calculated by R Haplo Stats Package (http://www.r-project.org). *P*-value was calculated by logistic model in which gender was used as the covariate in allele and genotype tests. We allocated the SNPs and reconstructed the haplotypes according to the chromosome that they were located. There were five blocks in chromosome 2 (*CREB1*), chromosome 8 (*NRG1*), chromosome 10 (*CNNM2-NT5C2*), chromosome 11 (*OPCML*) and chromosome 22 (*DGCR2*), respectively. Linkage disequilibrium patterns (r^2 algorithm) of these five blocks were performed by using Haploview (Barrett et al., 2005). Power calculations were performed using the G-power software (http://www.gpower.hhu.de/).

3. Results

3.1. Positive selection on CSP during the divergence of modern human populations

We used three detection methods with different estimates of time scale (π and Tajima's *D*, <250,000 years; iHS, <30,000 years) (Sabeti et al., 2006) to properly indicate the different ages of selection events. We first detected whether genetic diversity was reduced in human populations during their dispersal across the world by using π and Tajima's *D* (Tajima, 1989). The sliding window fashion calculations showed that π and Tajima's *D* values of *CREB1* region were significantly lower than those of its upper and lower region in all 14 populations of 1000 Genomes Project (Fig. 2). This finding implied that *CREB1* might have undergone positive selection in human populations and the selection might appear before the divergence of human populations, as all populations showed the signal. The positive selection was unlikely caused by population history such as population expansion because it should have the same effect on the genome. Note that negative selection might also lead to the deviation from the neutral test. We failed to detect any significant signal in other CSP genes.

Large positive iHS (iHS > 3) indicates that haplotypes on the ancestral allele background are longer than derived allele background (Voight et al., 2006). In contrast, an extreme negative iHS score (iHS < −3) indicates that haplotypes on the derived allele background are longer than haplotypes containing the ancestral allele. We found a total of 22 SNPs within the *NRG1* region had iHS values either >3 or <−3 in all 14 populations of 1000 Genomes Project (Supplementary Table 2).

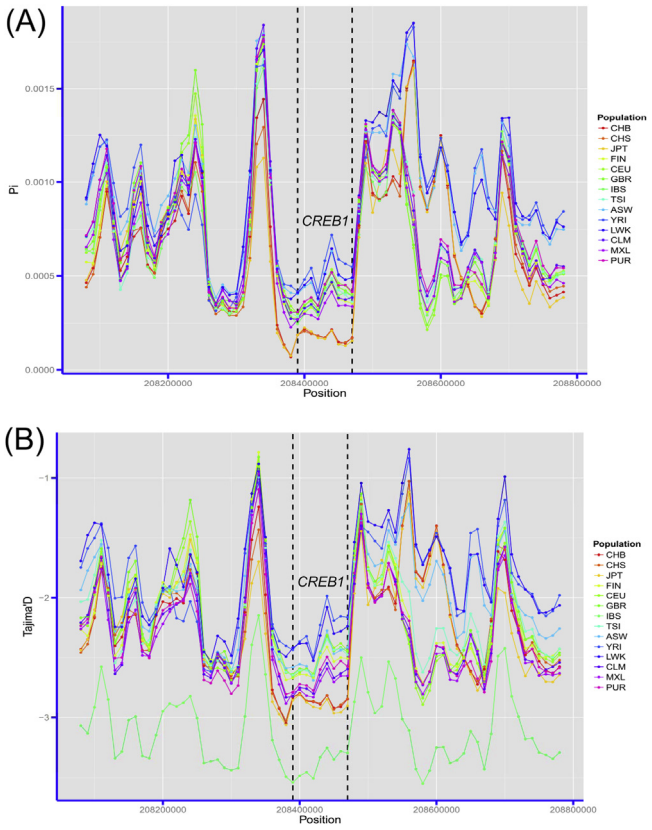


Fig. 2. Nucleotide diversity (π, Pi) (A) and Tajima's *D* test (B) for *CREB1* gene in 14 human populations of the 1000 Genomes Project. CHB, Han Chinese in Beijing, China; CHS, Han Chinese South; JPT, Japanese in Tokyo, Japan; CEU, Utah residents (CEPH) with Northern and Western European ancestry; FIN, Finnish from Finland; GBR, British from England and Scotland; IBS, Iberian populations in Spain; TSI, Toscani in Italia; ASW, African Ancestry in Southwest US; YRI, Yoruba in Ibadan, Nigeria; LWK, Luhya in Webuye, Kenya; CLM, Colombian in Medellin, Colombia; MXL, Mexican Ancestry in Los Angeles, CA; PUR, Puerto Rican in Puerto Rico. The imputation data were obtained from web (http://mathgen.stats.ox.ac.uk/impute/).

3.2. CSP genetic variants and SCZ susceptibility in Han Chinese

To detect whether CSP is associated with SCZ, we genotyped 68 SNPs of CSP genes in a total of 976 SCZ patients and 1043 normal controls from South Central China. This sample had a statistical power of 98.59% to detect a genetic effect with an effect size index of 0.1 under an assumption of weak gene effect. Our recent analysis of the matrilineal components of the patient and control populations showed no evidence of population stratification (Zhang et al., 2014). We excluded five SNPs with a successful calling rate less than 95%, and one SNP that was failed to be genotyped from the subsequent analyses. The genotype and allele frequencies of the remaining 62 SNPs were shown in Fig. 3 and Supplementary Tables S3 and S4. The strongest signal was located in intron 1 of *OPCML* (rs4379857, $P = 0.009$, odds ratio (OR) [95% confidence interval (CI)]: 1.200 [1.379–1.046]) and intron 2 of *DGCR2* (rs2238751, $P = 0.023$, OR [95% CI]: 1.253 [1.522–1.032]) at the allele frequency level (Fig. 3 and Supplementary Table 3).

We further evaluated potential haplotype association between SCZ patients and healthy controls. Haplotype test identified a strong association signal: a rare *CREB1* haplotype CCGGC surpassed stringent Bonferroni correction and showed an association with SCZ ($P = 1.74 \times 10^{-5}$) (Table 1). To further confirm this result, we downloaded and re-analyzed genotype data of these SNPs for CHB from HapMap database (http://hapmap.ncbi.nlm.nih.gov/). The result showed that haplotype frequencies of the four common haplotypes in our data were similar to those of the CHB population from HapMap (Table 1). Linkage disequilibrium pattern of *CREB1* SNPs composing the significant haplotype was also similar between our control sample and CHB population (Supplementary Fig. 2).

Table 1
Haplotype frequency and association results of *CREB1*.

Haplotype[a]	Frequency			*P*-value[b]	OR (95% CI)
	Case	Control	HapMap CHB		
CCGTC	0.437	0.429	0.423	0.659	Reference
CTAGC	0.227	0.236	0.238	0.642	0.960 (0.809–1.140)
CCGTA	0.150	0.177	0.214	0.049	0.846 (0.694–1.030)
TCAGC	0.111	0.125	0.119	0.219	0.882 (0.708–1.100)
CCGGC	0.025	0.004	NA	1.74×10^{-5}	4.442 (1.909–10.34)

The global *P*-value of the entire haplotype comparison between the case and control samples is 3.39×10^{-5}.
[a] SNP order: rs2360969, rs16839849, rs2253206, rs11904814 and rs4234080.
[b] *P*-value was obtained by comparing the frequency of certain haplotype in the case and control populations.

4. Discussion

SCZ is a complex neural development disorder and had an involvement of a series of genes, especially genes related to neural circuit development (Lewis and Sweet, 2009). The tremendous shifts experienced by human populations in diet, climate, population densities, social communications and mental pressure have surely led to direct selection pressures on medically relevant phenotypes (Vamathevan et al., 2008). Genes in certain pathway driven by selective pressure are required for functional convergence and mental diseases may be regarded as by-products of human brain evolution (Vamathevan et al., 2008). In this study, we focused on *CREB1* and its five targeted genes, which have been reported as SCZ susceptibility genes in recent studies (Aberg et al., 2013; Bergen et al., 2012; Law et al., 2012; Li et al., 2014; Mistry et al., 2013; Ripke et al., 2011; Xu et al., 2011). We initially detected whether there are footprints of positive selection on these CSP genes in human populations by using a variety of evolutionary approaches, which are sensitive to show different age estimates of selection events.

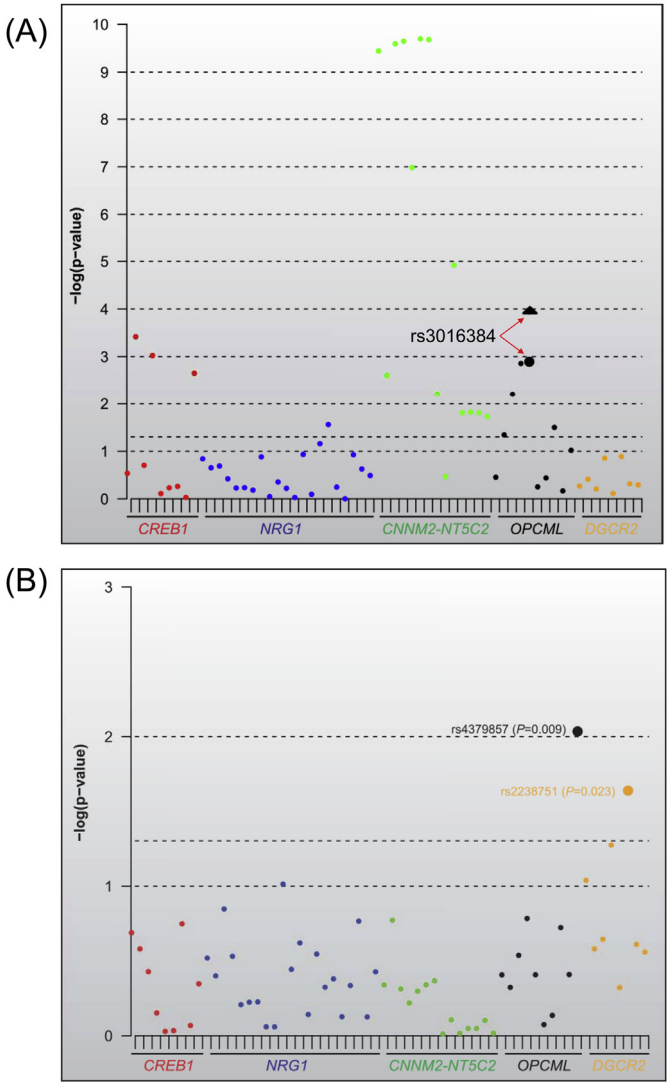


Fig. 3. Plot of *P*-values of 62 single nucleotide polymorphisms selected from *CREB1* (red), *NRG1* (blue), *CNNM2-NT5C2* (green), *OPCML* (black) and *DGCR2* (orange) genes in Psychiatric Genomics Consortium (PGC) data (A) and Han Chinese from Hunan Province, China (B). *P* values in the case-control study are calculated on a $-\log_{10}$ scale. The risk SNP rs3016384 (marked with a triangle) from O'Donovan et al. (2008) was also included in the PGC plot for comparison. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Consistent with the finding that *NRG1* fall into the top 1% of iHS signals in East Asians and the cross population extended haplotype homozygosity (XP-EHH) signals observed in Africans (Pickrell et al., 2009), we found an abundance of extreme iHS signals (around 30,000 years ago) that were located within *NRG1* (Supplementary Table 2). We also found that *NRG1* and *CREB1* had an adaptation pressure around 250,000–300,000 years ago. All 14 populations in 1000 Genomes Project were consistently experienced selection pressure at *NRG1* and *CREB1*, which implied that these regions had been selected before the dispersal of modern human populations across the world. This selection advantage lasted for many generations in global populations and was hardly affected by inheritance (Voight et al., 2006). These results confirmed our speculation and previous findings that CREB1 signal pathway, especially *NRG1* and *CREB1*, had undergone fast evolution in modern human evolution and this effect may affect brain development of our species.

To further determine whether genetic variants of the CSP genes contributed to SCZ risk, we analyzed 62 SNPs in six genes of CSP in Han Chinese. Unfortunately, we found no robust association between these SNPs and SCZ in Han Chinese populations at the allele and genotype levels, albeit we observed two SNPs showing a positive association with SCZ (Fig. 3), but the significance disappeared after multiple testing corrections. Nonetheless, a rare *CREB1* haplotype CCGGC was found to be significantly associated with SCZ in our population. This association remained significant after multiple testing corrections, and this haplotype was not found in CHB data set from HapMap but was enriched in our SCZ patients (Table 1), which further supported that CREB1 is a risk for SCZ. This pattern seemed to support our initial speculation that the CSP genes (*CREB1* in particular) that have undergone fast evolution during human evolution would harbor susceptible variants for SCZ.

There are several limitations in this study. First, our sample size is relatively modest, which might not be sufficient to detect SCZ

susceptible SNPs with very small effects. Second, the significant haplotype found in this study should be viewed with caution, because only one homogenous Chinese population was used in this study. The result remains to be confirmed in independent samples. Finally, we did not perform functional assay to further characterize the association between the significant rare haplotype and SCZ.

In summary, we found that *CREB1* and *NRG1* underwent fast evolution in recent human history and the selection appeared before the dispersal of modern human out of Africa. None of the 62 SNPs of the six CSP genes showed any significant association with SCZ after multiple testing corrections. However, we found a rare haplotype of *CREB1* confers susceptibility to SCZ. Taken all these results and previous reports together, we concluded that different genes in the CSP pathway may predispose to SCZ in different populations and substantial population heterogeneity in genetic predisposition to SCZ would account for the inconsistency. *CREB1*, the core component of CSP, might be a risk gene for SCZ in Han Chinese, albeit the risk effect was contributed by a rare haplotype.

Role of funding source

This study was supported by the Strategic Priority Research Program (B) of the Chinese Academy of Sciences (XDB02020300).

Contributors

LM and YGY designed the study; XC and LT collected the samples and clinical information; LM performed laboratory work; LM, DDW, SLM, NLST and YGY analyzed data; LM, NLST and YGY wrote the manuscript. All the authors read and revised the manuscript.

Conflict of interest

There is no conflict of interest to declare.

Acknowledgments

We thank Dr. Zhiguang Zhao and Mr. Xiao-Feng Liu for valuable suggestions. We are grateful to the support of the Large-scale Instrument Regional Center of Biodiversity in Kunming Institute of Zoology.

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.jpsychires.2014.06.008.

References

Aberg KA, Liu Y, Bukszar J, McClay JL, Khachane AN, Andreassen OA, et al. A comprehensive family-based replication study of schizophrenia genes. JAMA Psychiatry 2013;70:1–9.
Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics 2005;21:263–5.
Bergen SE, O'Dushlaine CT, Ripke S, Lee PH, Ruderfer DM, Akterin S, et al. Genome-wide association study in a Swedish population yields support for greater CNV and MHC involvement in schizophrenia compared with bipolar disorder. Mol Psychiatry 2012;17:880–6.
Betcheva ET, Mushiroda T, Takahashi A, Kubo M, Karachanak SK, Zaharieva IT, et al. Case-control association study of 59 candidate genes reveals the DRD2 SNP rs6277 (C957T) as the only susceptibility factor for schizophrenia in the Bulgarian population. J Hum Genet 2009;54:98–107.
Carlezon Jr WA, Duman RS, Nestler EJ. The many faces of CREB. Trends Neurosci 2005;28:436–45.
Doi N, Hoshi Y, Itokawa M, Usui C, Yoshikawa T, Tachikawa H. Persistence criteria for susceptibility genes for schizophrenia: a discussion from an evolutionary viewpoint. PLoS One 2009;4:e7799.
Fazzari P, Paternain AV, Valiente M, Pla R, Lujan R, Lloyd K, et al. Control of cortical GABA circuitry development by Nrg1 and ErbB4 signalling. Nature 2010;464:1376–80.
Georgi A, Schumacher J, Leon CA, Wolf AV, Klein K, Bohenz KV, et al. No association between genetic variants at the DGCR2 gene and schizophrenia in a German sample. Psychiatr Genet 2009;19:104.
Gong YG, Wu CN, Xing QH, Zhao XZ, Zhu J, He L. A two-method meta-analysis of Neuregulin 1(NRG1) association and heterogeneity in schizophrenia. Schizophr Res 2009;111:109–14.
Hall J, Whalley HC, Job DE, Baig BJ, McIntosh AM, Evans KL, et al. A neuregulin 1 variant associated with abnormal cortical function and psychotic symptoms. Nat Neurosci 2006;9:1477–8.
Hatzimanolis A, McGrath JA, Wang R, Li T, Wong PC, Nestadt G, et al. Multiple variants aggregate in the neuregulin signaling pathway in a subset of schizophrenia patients. Transl Psychiatry 2013;3:e264.
Holland MM, Parson W. GeneMarker® HID: a reliable software tool for the analysis of forensic STR data. J Forensic Sci 2011;56:29–35.
Ishiguro H, Imai K, Koga M, Horiuchi Y, Inada T, Iwata N, et al. Replication study for associations between polymorphisms in the CLDN5 and DGCR2 genes in the 22q11 deletion syndrome region and schizophrenia. Psychiatr Genet 2008;18:255–6.
Kajiwara K, Nagasawa H, Shimizu-Nishikawa K, Ookura T, Kimura M, Sugaya E. Cloning of SEZ-12 encoding seizure-related and membrane-bound adhesion protein. Biochem Biophys Res Commun 1996;222:144–8.
Kato T, Abe Y, Sotoyama H, Kakita A, Kominami R, Hirokawa S, et al. Transient exposure of neonatal mice to neuregulin-1 results in hyperdopaminergic states in adulthood: implication in neurodevelopmental hypothesis for schizophrenia. Mol Psychiatry 2011;16:307–20.
Keller MC, Miller G. Resolving the paradox of common, harmful, heritable mental disorders: which evolutionary genetic models work best? Behav Brain Sci 2006;29:385–404. discussion 405–352.
Kvajo M, McKellar H, Gogos JA. Molecules, signaling, and schizophrenia. Curr Top Behav Neurosci 2010;4:629–56.
Law AJ, Lipska BK, Weickert CS, Hyde TM, Straub RE, Hashimoto R, et al. Neuregulin 1 transcripts are differentially expressed in schizophrenia and regulated by 5′ SNPs associated with the disease. Proc Natl Acad Sci U S A 2006;103:6747–52.
Law AJ, Wang Y, Sei Y, O'Donnell P, Piantadosi P, Papaleo F, et al. Neuregulin 1-ErbB4-PI3K signaling in schizophrenia and phosphoinositide 3-kinase-p110delta inhibition as a potential therapeutic strategy. Proc Natl Acad Sci U S A 2012;109:12165–70.
Lewis DA, Sweet RA. Schizophrenia from a neural circuitry perspective: advancing toward rational pharmacological therapies. J Clin Invest 2009;119:706–16.
Li B, Woo RS, Mei L, Malinow R. The neuregulin-1 receptor erbB4 controls glutamatergic synapse maturation and plasticity. Neuron 2007;54:583–97.
Li D, Collier DA, He L. Meta-analysis shows strong positive association of the neuregulin 1 (NRG1) gene with schizophrenia. Hum Mol Genet 2006;15:1995–2002.
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;19:452–61.
Lonze BE, Riccio A, Cohen S, Ginty DD. Apoptosis, axonal growth defects, and degeneration of peripheral neurons in mice lacking CREB. Neuron 2002;34:371–85.
Ma L, Tang J, Wang D, Zhang W, Liu W, Liu XH, et al. Evaluating risk loci for schizophrenia distilled from genome-wide association studies in Han Chinese from Central China. Mol Psychiatry 2013;18:638–9.
Mantamadiotis T, Lemberger T, Bleckmann SC, Kern H, Kretz O, Martin Villalba A, et al. Disruption of CREB function in brain leads to neurodegeneration. Nat Genet 2002;31:47–54.
Mistry M, Gillis J, Pavlidis P. Genome-wide expression profiling of schizophrenia using a large combined cohort. Mol Psychiatry 2013;18:215–25.
Munafo MR, Thiselton DL, Clark TG, Flint J. Association of the NRG1 gene and schizophrenia: a meta-analysis. Mol Psychiatry 2006;11:539–46.
Neel JV. Diabetes mellitus: a "thrifty" genotype rendered detrimental by "progress"? Am J Hum Genet 1962;14:353–62.
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;40:1053–5.
Pickrell JK, Coop G, Novembre J, Kudaravalli S, Li JZ, Absher D, et al. Signals of recent positive selection in a worldwide sample of human populations. Genome Res 2009;19:826–37.
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;45:1150–9.
Ripke S, Sanders AR, Kendler KS, Levinson DF, Sklar P, Holmans PA, et al. Genome-wide association study identifies five new schizophrenia loci. Nat Genet 2011;43:969–76.
Sabeti PC, Schaffner SF, Fry B, Lohmueller J, Varilly P, Shamovsky O, et al. Positive natural selection in the human lineage. Science 2006;312:1614–20.
Shi Y, Li Z, Xu Q, Wang T, Li T, Shen J, et al. Common variants on 8p12 and 1q24.2 confer risk of schizophrenia. Nat Genet 2011;43:1224–7.
Shifman S, Levit A, Chen ML, Chen CH, Bronstein M, Weizman A, et al. A complete genetic association scan of the 22q11 deletion region and functional evidence reveal an association between DGCR2 and schizophrenia. Hum Genet 2006;120:160–70.
Struyk AF, Canoll PD, Wolfgang MJ, Rosen CL, D'Eustachio P, Salzer JL. Cloning of neurotrimin defines a new subfamily of differentially expressed neural cell adhesion molecules. J Neurosci 1995;15:2141–56.

Sullivan PF, Kendler KS, Neale MC. Schizophrenia as a complex trait: evidence from a meta-analysis of twin studies. Arch Gen Psychiatry 2003;60:1187–92.
Tajima F. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics 1989;123:585–95.
Tan W, Wang Y, Gold B, Chen J, Dean M, Harrison PJ, et al. Molecular cloning of a brain-specific, developmentally regulated neuregulin 1 (NRG1) isoform and identification of a functional promoter variant associated with schizophrenia. J Biol Chem 2007;282:24343–51.
Vamathevan JJ, Hasan S, Emes RD, Amrine-Madsen H, Rajagopalan D, Topp SD, et al. The role of positive selection in determining the molecular cause of species differences in disease. BMC Evol Biol 2008;8:273.
Voight BF, Kudaravalli S, Wen X, Pritchard JK. A map of recent positive selection in the human genome. PLoS Biol 2006;4:e72.
Xu B, Roos JL, Dexheimer P, Boone B, Plummer B, Levy S, et al. Exome sequencing supports a de novo mutational paradigm for schizophrenia. Nat Genet 2011;43:864–8.
Yue WH, Wang HF, Sun LD, Tang FL, Liu ZH, Zhang HX, et al. Genome-wide association study identifies a susceptibility locus for schizophrenia in Han Chinese at 11p11.2. Nat Genet 2011;43:1228–31.
Zhang X, Odom DT, Koo SH, Conkright MD, Canettieri G, Best J, et al. Genome-wide analysis of cAMP-response element binding protein occupancy, phosphorylation, and target gene activation in human tissues. Proc Natl Acad Sci U S A 2005;102:4459–64.
Zhang W, Tang J, Zhang AM, Peng MS, Xie HB, Tan L, et al. A matrilineal genetic legacy from the Last Glacial Maximum confers susceptibility to schizophrenia in Han Chinese. J Genet Genomics 2014.

Molecular evolution in the CREB1 signal pathway and a rare haplotype in *CREB1* with genetic predisposition to schizophrenia

Running title: CREB1 signal pathway and schizophrenia

Liang Ma[1,2], Dong-Dong Wu[3], Suk Ling Ma[4], Liwen Tan[5], Xiaogang Chen[5], Nelson L.S. Tang[6,7,8*], Yong-Gang Yao[1,2,6*]

[1]Key Laboratory of Animal Models and Human Disease Mechanisms of the Chinese Academy of Sciences & Yunnan Province, Kunming Institute of Zoology, Kunming, Yunnan 650223, China
[2]Kunming College of Life Science, University of Chinese Academy of Sciences, Kunming, Yunnan 650201, China
[3]State Key Laboratory of Genetic Resources and Evolution, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming, Yunnan 650223, China
[4]Department of Psychiatry, The Chinese University of Hong Kong, Hong Kong, China
[5]Institute of Mental Health, the Second Xiangya Hospital, Central South University, Changsha, Hunan 410011, China
[6]KIZ/CUHK Joint Laboratory of Bioresources and Molecular Research in Common Diseases, Kunming, Yunnan, China
[7]Department of Chemical Pathology and Laboratory of Genetics of Disease Susceptibility, Li Ka Shing Institute of Health Sciences, The Chinese University of Hong Kong, Hong Kong, China
[8]Functional Genomics and Biostatistical Computing Laboratory, Shenzhen Research Institute, The Chinese University of Hong Kong, Shenzhen, China.

*Corresponding authors:
Dr. Yong-Gang Yao, Tel/Fax: 86-871-65180085; E-mail: ygyaozh@gmail.com or yaoyg@mail.kiz.ac.cn or Dr. Nelson Tang, E-mail: nelsontang@cuhk.edu.hk

Supplementary material

Supplementary methods

Selection of SNPs

A total of 68 SNPs were selected for the six CSP genes (CREB1, NRG1, CNNM2-NT5C2, DGCR2 and OPCML). These SNPs included susceptible SNPs reported in previous association studies for SCZ, tagSNPs that covering the entire gene region, and eQTLs (Table S4). We followed the following selection criteria to choose tagSNP: MAF > 0.05 and linkage disequilibrium r2 < 0.8. As there are many SNPs met the criteria, we chose the minimum number of tagSNPs that could tag the studied gene region.

Genotyping

Genomic DNA was extracted from whole blood by using the AxyPrep™ Blood Genomic DNA Miniprep Kit (Axygen, USA). The 68 SNPs were detected by multiplex PCR and single-base extension method based on SNaPshot assay (Table S1). Briefly, all PCR reactions were carried out in a volume of 8 μL reaction mixture containing 10-20 ng templates DNA, 0.4 mM dNTPs, 0.2-0.5 μM of each primer, 2.0 mM $MgCl_2$ and 1.0 U of Fast Star DNA polymerase (Roche Ltd, Basel, Switzerland). The amplification program is consisted of a first denaturation at 94°C for 2 min, followed by 40 cycles at 94°C for 30 s, 55°C for 30 s, 72°C for 1 min, and then ended with incubation at 4°C. PCR products were cleaned up with 1.0 U of shrimp alkaline phosphatase (SAP) and 0.5 U of Exonuclease I (TaKaRa Biotechnology Co. Ltd. [Dalian, China]) at 37°C for 40 min, followed by a final incubation at 90°C for 10 min. The single-base extension was performed according to the protocol of the ABI PRISM® SNaPshot® Multiplex Kit (Applied Biosystems) in a total of 10 μL reaction solution containing 4 μL of the above treated-PCR products, 5 μL SNaPshot Multiplex Ready Reaction Mix and 0.4-0.8 μM pooled SNP-specific oligonucleotide primers (Table S1; SNP-specific oligonucleotide primers contained a special complementary sequence at 5' end with poly [GACT]). The thermal cycling program included 25 cycles of 96 °C for 10 s, 50 °C for 5 s, and 60 °C for 30 s. The final products were purified by treatment with 1.0 U of SAP at 37°C for 40 min and a deactivation at 75°C for 20 min. We loaded 0.5 μL purified product, 9 μL of Hi-Di^{TM} formamide and 0.5 μL of GeneScan^{TM} 120 LIZ^{TM} size standard (Applied Biosystems) for capillary electrophoresis on ABI PRISM ^{TM} 3730xl DNA analyzer (Applied Biosystems). The GeneMarker software (Holland and Parson 2011) was used to read the genotyping result. To validate the results, about one percent genotypes were randomly selected for sequencing using Sanger sequencing chemistry on an ABI3730 sequencer at the Kunming Biodiversity Large-Apparatus Regional Center, Kunming Institute of Zoology.

Supplementary Table 1 Primers used for SNapShot assay

Panel	SNP	Gene	Position	Region	Source [a]	Forward primer	Reverse primer	Extension primer	Extension primer tail
Panel 1	rs10786719	*CNNM2*	104637992	5' region	eQTL	TGAGTGTCCTGAAGAATGGTT	TTAATTAGGACATGGTATTAGTTAAAACATAG	CCAGTCTTAACCATGGCTTTCTAGACCTTGG	-
Panel 1	rs7914558	*CNNM2*	104775908	Intron 2	tagSNP	CCCCATCTAGATTCCCACG	TAATACTATTCTACCACAAAAGCAACTG	CACGCCCCCAAGCCCCCCGCTTCCCCCAAG	act(gact)2
Panel 1	rs2073776	*DGCR2*	19024651	3'-UTR	eQTL & reported	TCCCCTTGGAGAAGGGCT	ACGGCATGCTTCCTGCTT	CTGCTACTTAACAAGGAAGTCCTGGCCACA	(gact)4
Panel 1	rs2277831	*DGCR2*	18299197	3' region	tagSNP	ATCACCTATGAAGAAAGACAATGC	ACAGCAAGGGCTTCTGATT	GCTGCCTCCTCCAGAGAAGGGGAGTGTGAC	(gact)9
Panel 1	rs2072123	*DGCR2*	19026613	Exon 10	tagSNP & reported	ACACCCCTCTGTCTCTTCC	TAATGCACCTTCACTCCCA	TTCCTGCAGACGATGATGCTTTTGAGCCTG	t(gact)10
Panel 1	rs1630675	*OPCML*	133096325	Intron 1	tagSNP & reported	CCCCCTTCTTTTTACCCAG	TTTAGGAAGTGTCTAGTTAATTAAGGGA	CCAGTCAAATGCTGATGGTTGTGATCCAAA	t(gact)11
Panel 1	rs4379857	*OPCML*	133183604	Intron 1	tagSNP	ATATGAGAGAAGAGGCCAGGA	TAGGATGGTGTGTGTTTCTGG	CAGGTGGCCTTTAGAAGCTGAAAAGACAAG	t(gact)12
Panel 2	rs1911255	*OPCML*	132791152	Intron 2	Reported	ATTGCATTCACTGCGTCG	GCTGCAGACATCATGGGT	CCTTACCGTGATCAAAAGCCAAACA	-
Panel 2	rs11191688	*NT5C2*	105192570	5' region	Reported	ATTCCTCTTCGATGAGCATCT	AAGGTCAGATTATGAAGAGCCTTAT	TCAGTTACCTATTGTCTATAGACATTCATAC	-
Panel 2	rs10883801	*CNNM2*	104677887	5' region	tagSNP	CACATACACAAGAGTCACGC	AATGGCAAGGTAAGTGGGT	ACGCGCCTTTCAGTTGGAAACTTGGCGTGC	act(gact)2
Panel 2	rs7831	*NT5C2*	105205302	5' region	tagSNP	CCAAGGCCCTGGAGTATG	TGGCCCATTGTTGACAGT	ATGTGGAGGCCAAGAGCTCAGTGCTAGAGG	t(gact)5
Panel 2	rs2360969	*CREB1*	208372996	5' region	tagSNP	AAAACTGGGGATGAGGGC	TTGCTATACTGGTTCTCACGC	GCCAGTCATCTGTATTTCAACAAGTCTTGC	ct(gact)6
Panel 2	rs807759	*DGCR2*	19113686	Intron 1	tagSNP	GTGCTGGGGTCGGGAAAG	AACTGCAGATAAAGCCTGC	AAGGGCAGGGAGAGGACAGGACCAAGAAGC	act(gact)7
Panel 2	rs1939498	*OPCML*	132558337	Intron 2	tagSNP	TAGCTTGAGACATTGAGTGTGAC	ATTACAGCCTTTAATTGGAGAGG	TGACCTTGGGAAATCAAAGAGCTTTCCTGG	t(gact)10
Panel 2	rs11223225	*OPCML*	132693210	Intron 2	tagSNP	ATCCCTTCCTTCCCAGTAGA	ACAGGTGCGGATAGTTTCC	GCTCTTGAGAACATGCCTTCGCATGGGGGA	t(gact)11
Panel 3	rs11598702	*NT5C2*	104897985	Intron 4	tagSNP	TGCAGCTGACCATTGGGA	CCCAAAAAACTCAGAGATGTACTG	ACTGGCTAAATTCCTGCATATACCC	-
Panel 3	rs11191686	*NT5C2*	105187746	5' region	tagSNP	TGGCTCAGTCCTGGGATC	TGCCTGGCTCTGGAAGTT	TCCAGCCAGGTTGTAAATCTTAATGCACCTA	-
Panel 3	rs3740387	*NT5C2*	104849468	Exon (D549E)	tagSNP	CAACCTCTTCCCACTGGC	CTCCTTATTCTTCCTCCTCCTC	GGCCCCCCAGGAAATTACACACTGCCATGA	ct(gact)
Panel 3	rs2216374	*CREB1*	208089904	5' region	tagSNP	GGAAAGTTTGAATTATTCATTTTCTG	TTTAAAAGATGGCTGAGACACAT	CTGACCCTGTACCAGGTACTTTTTACTGCT	act(gact)2
Panel 3	rs11191580	*NT5C2*	104906211	Intron 3	tagSNP	TATAAAATCTTCTTGGGCATCTGT	AAAAGCACTTGTAACTTTTTCTTTTC	GTTTTCCTTATGGGCTTGCATAATTACACA	(gact)4
Panel 3	rs11191612	*NT5C2*	104969578	5' region	tagSNP	TAAAAACTGTCCCACACTAGCTC	TAAAACAGTAAGGTTACATTGTTTGG	ACACACAGCCAAACTACAGGGCTGTGCGGT	act(gact)7
Panel 3	rs16839849	*CREB1*	208378981	5' region	tagSNP	TGGGTAATTGTACTGGGGG	TTAATCAAAGTCCACAATCTCCTT	GGGGAAGAGAGGGACAAAAAGAATGTAGTT	(gact)9
Panel 3	rs2238751	*DGCR2*	19057790	Intron 2	tagSNP	TGGAGAAGGGAAGAGCCA	AAAGCTGCTCCATCTGCC	CCACCTGAGCTTCTGGGCCTGAGACTGGAA	t(gact)10

Panel 3	rs11223064	*OPCML*	132289393	3'-UTR	tagSNP	ACGTTTATGAGTTGAACTTCTCTTTT	TTTGGGGCTATGCATGAA	TATCTTCAAAATGCCTCCCCCTTTTTTGGT	t(gact)11
Panel 3	rs11223249	*OPCML*	132724717	Intron 2	tagSNP	TTCCACTTGGGAGCTATTATGA	TAGAAGAAAGATGGAGTCTGTTATGC	AACAATTGCATGTAAGTCTTTGTGGGACAT	t(gact)12
Panel 4	rs2042484	*CREB1*	208363062	5' region	tagSNP	TATAAATAACACACGTGCCATTC	TAATATTCTGGTTAGCTGATGTCACT	TTCATGCCTTGTGTAGTTTGTTTTATTATTA	-
Panel 4	rs2281861	*NT5C2*	105201565	5' region	tagSNP	GCGCTTCTGCTGGCTTGG	ACGGGGGAACAGGGAGCA	TGGTTTGCTTCTCCTGGAGGAACACAGAGG	ct(gact)
Panel 4	rs11904814	*CREB1*	208426798	Intron 4	tagSNP	AAATCATCAGTTGGATCAATTTC	AATGGAATAAAAACTAAGGCAAAG	AAGATCTAAGAACTTGATATTTCTATGAAA	act(gact)2
Panel 4	rs4675690	*CREB1*	208507807	3' region	tagSNP	ATATTTACTGGGTGTCTTTTATCAAGC	AAATGATGGGTATGTTTCCCA	TGTGTTAGATGCTGAAACACAAAAAAAGAA	(gact)4
Panel 4	rs5992854	*DGCR2*	18300240	3' region	tagSNP	AGAAGACGGACTTCCACAGG	AAGAAGTCCAAAGGCGAGG	CAGGGACTTGGGTTTGGCGGCGGCTTCTTC	ct(gact)6
Panel 4	rs8399	*DGCR2*	19025330	3'-UTR	eQTL	TACTGCAGCTCAGTGCCC	AAAAGGGTGCCAGGAAGG	CCAGCGGCCGCATGCACACCTCCCGGCCCC	act(gact)7
Panel 4	rs1784519	*OPCML*	132526865	Intron 2	dbSNP	TTCTACCTGTGTGCGAAGAC	TTAGTTAAGTACCCATTCGATGG	AGACACAAAAGTCATCTTCCTCTACTGAAT	(gact)9
Panel 4	rs1894193	*OPCML*	132527512	Intron 2	tagSNP & GWAS	AGACATATTATGTATCTTTGCCATCC	TGGGCCACTTCTTCCCCA	ATTTTTTTCCCCATAGGCTTATATGAAAAT	t(gact)10
Panel 4	rs3016384	*OPCML*	132573390	Intron 2	tagSNP	AGACTGAGTGGGAAATGACTTC	TTCCCTGATATTTGCAATCC	CAAAGACAGTCTATACAGACCGGAAGAGAT	t(gact)11
Panel 5	rs12220375	*NT5C2*	104901491	Intron 3	tagSNP & reported	ATTACTTTTGGGATTTTATAACTCATTC	AGTATCTGGGTGTAATAGACAGTAGCTAT	TTAAAGCCCACTGAAGATAAAAATA	-
Panel 5	rs11191602	*NT5C2*	104954219	5' region	tagSNP & reported	ATACTGTTGACTCATATTGGGCTT	TCTATCCTTAGCTGACCTGAAGC	TTAATAAGTGCTCCGCCAGAGCACTTCTAC	ct(gact)
Panel 5	rs11191609	*NT5C2*	104963721	5' region	tagSNP & GWAS	TCCTTCACAGTGATATTGAGCTG	TTTTCAGGTGTGAGCCACA	AATTTACTTTGGGAGTCTCAGAGCACATTT	(gact)4
Panel 5	rs2253206	*CREB1*	208391978	5' region	dbSNP	AATTACATGGACACTAAACAAATAATGA	TTTTTATTCCCTACTGTAGGCTCTC	ATTGGAAAAGAAAGTGATAAGTTACAGTTA	ct(gact)6
Panel 5	rs2551941	*CREB1*	208492143	3' region	tagSNP	ATAAAGGGAAAGCTGAGAGGC	TTCGAGGGTAACCTCAGG	CAAAAAAGAAAACGAAAAAGGAGGTAAAAG	act(gact)7
Panel 5	rs4234080	*CREB1*	208489101	3' region	Reported	AATTTCTGCAACCCAAATTC	ATTCTTTTGCACGGCAGG	GTGGTCTCCTCATAGGGCACGAGGGCCATT	(gact)9
Panel 5	rs1034727	*DGCR2*	19097700	Intron 1	eQTL	AACATTCAAAGGAAACATTTGC	AGCACTTGGCATGTAATAGCTC	TTTAACACACGAAAAACTCTATCCAAAAAA	t(gact)12
Panel 6	rs35753505	*NRG1*	31474141	5' region	eQTL	CATCAGTTTTCAATAGCTTTTTTATG	AAGACAGATGTCTCAAGAGACTGG	ACTAAAAAAGAGATATATGATATTTGG	-
Panel 6	rs4268090	*NRG1*	31480888	5' region	eQTL	TTGCAAAAGGCTCAGGAA	CATAGGATCATAGAGCAGGTGAG	ATAAGAACTGTTTGCAATCTGATTTAGAAG	ct(gact)2
Panel 6	rs10109424	*NRG1*	31486788	5' region	eQTL	TAAGTTCAGACTGTTCAGTCAATTG	CAAACATTGATTTTATTCACAGAAGA	GAGATTATGCAGTAAAGTTATTTAGCTGAC	act(gact)3
Panel 6	rs7831093	*NRG1*	31488205	5' region	tagSNP	AGAAATGAGAGAGCCAGTACC	TTTCATATGTGATAAAAAGGTTCC	TACCTGAGTCCTAAACCTCCTTCTGACAAC	(gact)5
Panel 6	rs4457296	*NRG1*	31488638	5' region	Reported	AAGCCTTCACTGAAAACTTCG	ACATCTGAATAGTAGGATGCTTTGA	AATAGAAGTTGGCATAAAGTATAGTAATTA	(gact)6
Panel 6	rs4457297	*NRG1*	31488655	5' region	tagSNP & reported	TTCACTGAAAACTTCGGAATAGA	TCCACACATCTGAATAGTAGGATG	AGTATAGTAATTATGACTTTCGTGGACTGA	(gact)7
Panel 6	rs4281084	*NRG1*	31495374	5' region	Reported	AAATTGCCAACTTGCAGAAT	ACTTTTTGGGAGGTACAACTCTG	ATGAACCAACAGGTCACCAAATGTTGAAGT	(gact)8
Panel 6	rs6994992	*NRG1*	31495581	5' region	tagSNP & reported	AAATTAGTAGGATTGGATGTTTGAAC	ACCATTTAGCAGCATAGTTGG	GCTAGAAGCACCATGCAGGGTTCAAGTGAA	(gact)9

Panel 6	rs7014762	*NRG1*	31495668	5' region	tagSNP & reported	AACTATGCTGCTAAATGGTGATCT	TTTGGAGGGACAGGGTCA	AGCGCTCCATCAGGGTATGAGTAACAGGGA	(gact)10
Panel 6	rs1081062	*NRG1*	31500264	Intron 2	tagSNP & reported	AGCAATATCCTACTTTAAATCATATAACAC	TAATCGAGTATGTTAAGAAATCACAAA	ACAAAAATCGCCAAGCACACACAATAGCTA	(gact)11
Panel 6	rs776389	*NRG1*	31784340	Intron 2	tagSNP	AAGACCCAGAAAAGAGAGGATG	AAAATGGCCAAGAAGAAGACTC	CTGAAAGCTAACAAGAAGTACAGTGATGAA	(gact)12
Panel 7	rs6468119	*NRG1*	32401561	Intron 2	Reported	ACCTCAATTTCCTCCCTCAT	TATAAATAGGAGACATTAACAAGCAATC	TAGGATGGTCTTTGCTTATCTCAGTTT	-
Panel 7	rs10808318	*NRG1*	31919211	Intron 2	Reported	TCCTATGCTGTTTCCTAATGTACA	TTCACTAGCCTTGTTGTGGG	TTTTTCTCTTGTCACAACTAGGAAAGTCAC	act
Panel 7	rs16879552	*NRG1*	32411216	Intron 3	eQTL	AGAAGATTTTGAAACATGACTCTTTT	AGGTTGGTGCACACTTTTGT	ATAGAATACATAATCTTCTCTTTTGGGACT	ct(gact)2
Panel 7	rs6988339	*NRG1*	32545916	Intron 8	eQTL	CATTAATTGACATTATAACCATCTTTCA	CACATTCGAAGATAACCATTTTAAA	GGAGCAGATACGGCCCTTTTTCCCAGCTAC	act(gact)3
Panel 7	rs3924999	*NRG1*	32453358	Exon 7 (R38Q)	tagSNP & reported	TTTCACACCGAAGGACTAGTTT	TAACTCCAGATTAAATGTTTCTCTTTG	GCCGATTCCTGGCTTTTCATCTCTTTCAAT	(gact)5
Panel 7	rs4733376	*NRG1*	32613829	Intron 14	dbSNP	ATACAGTGGACATACTCCCTTCC	TTTGATTATTTTCATTTCATAGTAGAGC	CTCTTCCTCATCACCAGTCTACTGCCTTGA	(gact)6
Panel 7	rs4531002	*NRG1*	32501987	Intron 8	tagSNP & reported	TTCCTTTCTTACTCTTTTATTGCTATAGA	AAATTGGAAAAAGAATGTAAATGG	AGGCAGAAAGGCAACTTCTGGGTCCTAGTC	(gact)9
Panel 7	rs2954041	*NRG1*	32522626	Intron 8	tagSNP	TTAAATAATATTAATTGTTCTTCCCCTG	ATAAATTACCAGTGCCTATTAAAGGG	GACATTATTCATTGTTTGTTGCTGATGAAT	(gact)10
Panel 7	rs7005288	*NRG1*	32620467	Intron 17	tagSNP	ATTATCAGACATAACATGATAGAAGGTAAC	TTCCAAAACCCACAGCTATATC	TAACAGGGTACAGCCAACCCTTGAGGCATC	(gact)11
Panel 7	rs6992642	*NRG1*	32624387	3' region	tagSNP & reported	ATCCTTTGCACCCTTCATAAA	AATTTATTAACAGAGGTGGTCAAGAG	GGCCACTAGATTTGTGTTTACAGGGGTATC	(gact)12

[a] SNPs were selected from GWAS, reported, tagSNP, eQTL and dbSNP database.

All SNPs were assigned into seven panels for multiple amplifications.

Supplementary Table 2 The SNPs of *NRG1* with extreme iHS value in 14 populations of 1000 Genomes

SNP	P_{SCZ}	P_{BIP}	Asian			European					African			American		
			CHB	CHS	JPT	CEU	FIN	GBR	TSI	IBS	YRI	LWK	ASW	CLM	MXL	PUR
rs10503888	0.275	0.216	-3.449	-3.443	-3.395	-3.383	-3.395	-3.395	-3.451	NA	-3.385	-3.449	-3.498	-3.497	-3.534	-3.532
rs2008626	0.884	0.983	-3.020	-3.014	-3.089	-3.058	NA	-3.089	-3.016	NA	-3.065	-3.020	-3.114	-3.111	-3.126	-3.105
rs16879134	0.863	0.918	-3.020	-3.014	-3.089	-3.058	NA	-3.089	-3.016	NA	-3.065	-3.020	-3.114	-3.111	-3.126	-3.104
rs16879136	0.885	0.996	-3.020	-3.014	-3.089	-3.058	NA	-3.089	-3.016	NA	-3.065	-3.020	-3.114	-3.111	-3.126	-3.104
rs16879137	0.884	0.977	-3.020	-3.014	-3.089	-3.058	NA	-3.089	-3.016	NA	-3.065	-3.020	-3.114	-3.111	-3.126	-3.105
rs1487154	0.299	0.434	-3.043	-3.063	-2.959	-3.008	-2.963	-2.959	-3.067	-3.356	-3.045	-3.043	-3.168	-3.166	-3.009	-3.130
rs16879147	0.903	0.910	-3.026	-3.029	-3.095	-3.054	NA	-3.095	-3.028	NA	-3.072	-3.026	-3.100	-3.102	-3.108	-3.072
rs1685113	0.401	0.061	-3.365	-3.316	-3.401	-3.252	-3.371	-3.401	-3.367	-2.008	-3.272	-3.365	-3.670	-3.577	-3.528	-3.539
rs17705398	0.013	0.035	3.125	3.156	3.001	2.913	3.080	3.001	3.118	1.547	2.974	3.125	2.847	2.864	2.857	3.040
rs1685115	0.406	0.049	-3.460	-3.411	-3.479	-3.328	-3.452	-3.479	-3.460	-2.016	-3.349	-3.460	-3.738	-3.648	-3.598	-3.606
rs1685116	0.465	0.056	-3.461	-3.411	-3.479	-3.329	-3.452	-3.479	-3.460	-2.016	-3.349	-3.460	-3.738	-3.648	-3.598	-3.606
rs2347070	0.297	0.427	-3.297	-3.323	-3.203	-3.235	-3.212	-3.203	-3.325	-3.396	-3.288	-3.297	-3.388	-3.394	-3.221	-3.318
rs1685117	0.381	0.040	-3.462	-3.413	-3.481	-3.330	-3.454	-3.481	-3.462	-2.017	-3.351	-3.462	-3.740	-3.650	-3.600	-3.608
rs1685101	0.474	0.038	-3.452	-3.403	-3.470	-3.320	-3.445	-3.470	-3.452	-2.012	-3.341	-3.452	-3.729	-3.639	-3.588	-3.596
rs1685103	0.423	0.055	-3.214	-3.153	-3.236	-3.118	-3.201	-3.236	-3.206	-1.819	-3.100	-3.214	-3.490	-3.400	-3.368	-3.362
rs17706609	0.696	0.620	-3.379	-3.385	-3.318	-3.316	-3.325	-3.318	-3.378	-3.159	-3.290	-3.379	-3.231	-3.238	-3.248	-3.237
rs16879268	0.392	0.083	-3.152	-3.077	-3.156	-3.073	-3.139	-3.156	-3.143	-1.841	-3.034	-3.152	-3.287	-3.199	-3.166	-3.197
rs17634051	0.693	0.705	-3.404	-3.414	-3.342	-3.337	-3.348	-3.342	-3.404	-3.178	-3.309	-3.404	-3.290	-3.295	-3.312	-3.293
rs17634082	0.693	0.664	-3.404	-3.414	-3.342	-3.337	-3.348	-3.342	-3.404	-3.179	-3.309	-3.404	-3.290	-3.295	-3.312	-3.293
rs17634100	0.683	0.653	-3.404	-3.414	-3.342	-3.337	-3.348	-3.342	-3.404	-3.179	-3.309	-3.404	-3.290	-3.295	-3.312	-3.293
rs7840327	0.657	0.773	3.202	3.195	3.146	3.162	3.157	3.146	3.204	3.146	3.138	3.202	3.104	3.105	3.130	3.095
rs1487149	0.696	NA	3.202	3.195	3.146	3.162	3.157	3.146	3.204	3.146	3.138	3.202	3.104	3.105	3.130	3.095

NA, not available; P_{SCZ} and P_{BIP}, PGC GWAS results of SNPs spanning *NRG1* region with schizophrenia (Ripke et al. 2013) and bipolar disorders (Sklar 2011).

CHB, Han Chinese in Beijing, China; CHS, Han Chinese South; JPT, Japanese in Toyko, Japan; CEU, Utah residents (CEPH) with Northern and Western European ancestry; FIN, Finnish from Finland; GBR, British from England and Scotland; IBS, Iberian populations in Spain; TSI, Toscani in Italia; ASW, African Ancestry in Southwest US; YRI, Yoruba in Ibadan, Nigeria; LWK, Luhya in Webuye, Kenya; CLM, Colombian in Medellin, Colombia; MXL, Mexican Ancestry in Los Angeles, CA; PUR, Puerto Rican in Puerto Rico. The imputation data of 1000 Genomes were obtained from web (http://mathgen.stats.ox.ac.uk/impute/).

Supplementary Table 3 The distribution of allele frequencies of the *CREB1* pathway SNPs in 976 Han Chinese with schizophrenia and 1043 matched controls

N	Chr.	SNP	Position [a]	Allele [b]	Minor allele frequency		*P*-value [c]	OR (95% CI)	*P*-value [d]	*P*-value [e]	Frequency [f]	
					Case	Control					CHB	CEU
1	2	rs2216374	208089904	A/G	0.225	0.207	0.204	1.115 (0.943,1.319)	0.291	0.999	0.201	0.305
2	2	rs2042484	208363062	G/A	0.246	0.263	0.262	0.914 (0.781,1.069)	3.83×10^{-4}	3.74×10^{-3}	0.322	0.593
3	2	rs2360969	208372996	T/C	0.123	0.134	0.375	0.913 (0.747,1.116)	0.198	0.842	0.127	0.463
4	2	rs16839849	208378981	T/C	0.247	0.241	0.706	1.030 (0.882,1.203)	0.001	0.284	0.223	0.018
5	2	rs2253206	208391978	A/G	0.376	0.378	0.937	0.994 (0.865,1.143)	0.776	0.704	0.354	0.478
6	2	rs11904814	208426798	G/T	0.378	0.379	0.926	0.993 (0.862,1.144)	0.587	0.326	0.354	0.327
7	2	rs4234080	208489101	A/C	0.163	0.180	0.178	0.883 (0.737,1.058)	0.548	0.098	0.193	0.155
8	2	rs2551941	208492143	A/T	0.412	0.409	0.856	1.013 (0.883,1.162)	0.933	0.556	0.356	0.417
9	2	rs4675690	208507807	C/T	0.417	0.404	0.452	1.055 (0.918,1.213)	0.002	7.92×10^{-4}	0.367	0.600
10	8	rs35753505	31474141	T/C	0.430	0.413	0.301	1.075 (0.937,1.233)	0.145	NA	0.383	0.658
11	8	rs4268090	31480888	C/T	0.434	0.420	0.400	1.060 (0.925,1.215)	0.223	0.105	0.471	0.642
12	8	rs10109424	31486788	T/G	0.415	0.390	0.142	1.111 (0.965,1.278)	0.205	0.062	0.427	0.571
13	8	rs7831093	31488205	A/C	0.411	0.394	0.293	1.076 (0.938,1.235)	0.380	0.052	0.427	0.566
14	8	rs4457296	31488638	T/C	0.380	0.372	0.622	1.036 (0.901,1.190)	0.592	0.093	0.376	0.504
15	8	rs4457297	31488655	T/C	0.381	0.372	0.599	1.038 (0.903,1.193)	0.583	0.093	0.376	0.504
16	8	rs4281084	31495374	A/G	0.319	0.310	0.596	1.040 (0.900,1.202)	0.656	0.106	0.256	0.228
17	8	rs6994992	31495581	C/T	0.400	0.397	0.873	1.011 (0.881,1.161)	0.132	NA	0.308	0.583
18	8	rs7014762	31495668	A/T	0.336	0.333	0.874	1.012 (0.877,1.167)	0.899	NA	0.283	0.208
19	8	rs1081062	31500264	C/T	0.094	0.111	0.097	0.829 (0.664,1.035)	0.443	0.906	0.158	0.283
20	8	rs776389	31784340	T/G	0.370	0.385	0.362	0.938 (0.816,1.077)	0.600	0.752	0.405	0.867
21	8	rs10808318	31919211	A/G	0.316	0.298	0.239	1.092 (0.944,1.263)	0.936	0.562	0.314	0.681
22	8	rs6468119	32401561	T/C	0.196	0.201	0.723	0.970 (0.820,1.148)	0.117	NA	0.190	0.442
23	8	rs16879552	32411216	C/T	0.403	0.421	0.283	0.928 (0.809,1.064)	0.803	0.242	0.383	0.973

24	8	rs3924999	32453358	C/T	0.234	0.223	0.476	1.060 (0.903,1.245)	0.069	0.424	0.235	0.571
25	8	rs4531002	32501987	T/C	0.124	0.115	0.419	1.090 (0.885,1.341)	0.028	0.863	0.139	0.341
26	8	rs2954041	32522626	T/G	0.363	0.368	0.748	0.977 (0.850,1.124)	0.567	0.453	0.376	0.018
27	8	rs6988339	32545916	G/A	0.463	0.475	0.464	0.951 (0.831,1.088)	0.998	0.772	0.482	0.429
28	8	rs4733376	32613829	G/A	0.339	0.362	0.171	0.906 (0.787,1.043)	0.119	0.953	0.382	0.088
29	8	rs7005288	32620467	A/G	0.441	0.436	0.750	1.022 (0.893,1.171)	0.237	0.995	0.522	0.832
30	8	rs6992642	32624387	C/T	0.201	0.189	0.376	1.080 (0.911,1.280)	0.324	0.321	0.132	0.447
31	10	rs10786719	104637992	A/G	0.437	0.425	0.460	1.053 (0.918,1.208)	3.61×10^{-10}	0.362	0.434	0.619
32	10	rs10883801	104677887	C/A	0.283	0.305	0.169	0.900 (0.774,1.046)	0.002	NA	NA	NA
33	10	rs7914558	104775908	G/A	0.446	0.434	0.489	1.049 (0.916,1.202)	2.56×10^{-10}	0.278	0.427	0.615
34	10	rs3740387	104849468	T/C	0.463	0.472	0.606	0.965 (0.843,1.105)	2.26×10^{-10}	0.235	0.493	0.385
35	10	rs11598702	104897985	C/T	0.254	0.244	0.505	1.055 (0.902,1.233)	1.04×10^{-7}	0.837	0.226	0.387
36	10	rs12220375	104901491	C/T	0.253	0.264	0.459	0.944 (0.810,1.100)	2.01×10^{-10}	7.14×10^{-4}	0.293	0.075
37	10	rs11191580	104906211	C/T	0.253	0.265	0.432	0.941 (0.807,1.096)	2.09×10^{-10}	6.09×10^{-4}	0.294	0.075
38	10	rs11191602	104954219	T/C	0.222	0.222	0.976	0.998 (0.848,1.173)	0.006	NA	0.158	0.317
39	10	rs11191609	104963721	G/T	0.167	0.163	0.785	1.026 (0.854,1.233)	0.342	0.675	0.136	0.119
40	10	rs11191612	104969578	G/A	0.296	0.296	0.966	1.003 (0.864,1.165)	1.2×10^{-5}	0.912	0.311	0.367
41	10	rs11191686	105187746	A/G	0.309	0.307	0.895	1.010 (0.874,1.167)	0.016	0.119	0.299	0.376
42	10	rs11191688	105192570	G/A	0.323	0.321	0.895	1.010 (0.875,1.165)	0.015	0.148	0.314	0.389
43	10	rs2281861	105201565	T/C	0.327	0.331	0.792	0.981 (0.848,1.134)	0.016	0.183	0.314	0.397
44	10	rs7831	105205302	C/A	0.289	0.288	0.963	1.004 (0.865,1.164)	0.019	0.172	0.296	0.389
45	11	rs11223064	132289393	C/T	0.388	0.402	0.394	0.940 (0.814,1.084)	0.352	0.864	0.482	0.345
46	11	rs1784519	132526865	T/C	0.457	0.445	0.477	1.051 (0.917,1.204)	0.045	0.731	0.449	0.190
47	11	rs1894193	132527512	T/C	0.403	0.421	0.289	0.929 (0.810,1.065)	0.006	0.758	0.438	0.566
48	11	rs1939498	132558337	A/G	0.452	0.476	0.164	0.908 (0.792,1.040)	0.001	0.243	0.533	0.576
49	11	rs3016384	132573390	A/G	0.431	0.445	0.393	0.941 (0.819,1.082)	0.001	0.290	0.438	0.535

50	11	rs11223225	132693210	C/T	0.413	0.409	0.844	1.014 (0.881,1.167)	0.560	0.025	0.420	0.903
51	11	rs11223249	132724717	A/G	0.245	0.240	0.733	1.028 (0.878,1.203)	0.365	0.146	0.281	0.190
52	11	rs1911255	132791152	C/G	0.306	0.327	0.189	0.907 (0.784,1.049)	0.032	0.742	0.395	0.513
53	11	rs1630675	133096325	C/G	0.221	0.233	0.392	0.933 (0.797,1.093)	0.681	0.094	0.255	0.650
54	11	rs4379857	133183604	A/G	0.444	0.488	0.009	0.833 (0.725,0.956)	0.096	0.046	0.478	0.066
55	22	rs2277831	18299197	G/A	0.118	0.100	0.092	1.208 (0.970,1.504)	0.542	0.699	0.109	0.254
56	22	rs5992854	18300240	C/T	0.223	0.208	0.262	1.100 (0.931,1.300)	0.388	0.651	0.216	0.323
57	22	rs2073776	19024651	A/G	0.321	0.340	0.226	0.914 (0.791,1.057)	0.620	0.446	0.363	0.341
58	22	rs8399	19025330	G/T	0.121	0.144	0.053	0.821 (0.672,1.003)	0.141	NA	0.117	0.158
59	22	rs2072123	19026613	C/T	0.360	0.371	0.480	0.950 (0.825,1.095)	0.773	0.724	0.369	0.394
60	22	rs2238751	19057790	A/G	0.125	0.152	0.023	0.798 (0.657,0.969)	0.130	0.966	0.169	0.217
61	22	rs1034727	19097700	T/C	0.274	0.292	0.244	0.913 (0.783,1.064)	0.484	0.336	0.299	0.265
62	22	rs807759	19113686	G/A	0.408	0.426	0.275	0.926 (0.805,1.063)	0.509	0.741	0.412	0.469

Six SNPs were excluded from the analyses and were not listed in this table: rs4936180, rs7095304, rs2919370 and rs3808368 had a successful genotyping rate less than 95%, rs7569963 was deviated from Hardy-Weinberg equilibrium (HWE) in Hunan cases and Hunan controls, rs4513929 was failed to be genotyped. Chr., chromosome; OR, odds ratio; CI, confidence interval; *P*-value < 0.001 was highlighted in red.

[a] SNP position is defining according to the NCBI, GRCh37.p5 version.

[b] Minor allele/major allele.

[c] *P*-value were calculated by logistic model in which gender as covariate.

[d] Data come from PGC schizophrenia GWAS results in SWE+PGC1 samples (Ripke et al. 2013).

[e] Data come from PGC schizophrenia GWAS results in SWE+BIP samples (Sklar 2011).

[f] Data come from HapMap3 release 3 or dbSNP database, and data corresponding to minor allele of Hunan population.

Supplementary Table 4 The distribution of genotype frequencies of CREB1 pathway SNPs in 976 Han Chinese with schizophrenia and 1043 matched controls

Chr.	SNP	Gene	Location	Source [a]	Genotype	Number of sample		*P*-value [b]	HWE *P*-value
						Case	Control		
2	rs2216374	*CREB1*	5' region	eQTL	AA/AG/GG	40/330/543	29/268/498	0.346	0.300/0.384
2	rs2042484	*CREB1*	5' region	tagSNP	GG/GA/AA	56/339/522	46/332/420	0.950	0.929/0.069
2	rs2360969	*CREB1*	5' region	eQTL & reported	TT/TC/CC	14/197/701	16/180/607	0.406	1/0.537
2	rs16839849	*CREB1*	5' region	tagSNP	TT/TC/CC	65/321/527	46/294/456	0.320	0.109/0.923
2	rs2253206	*CREB1*	5' region	tagSNP & reported	AA/AG/GG	136/408/360	111/368/305	0.805	0.257/1
2	rs11904814	*CREB1*	Intron 4	tagSNP & reported	GG/GT/TT	135/412/356	106/364/290	0.776	0.396/0.700
2	rs4234080	*CREB1*	3' region	tagSNP	AA/AC/CC	20/257/636	26/240/530	0.167	0.394/0.906
2	rs2551941	*CREB1*	3' region	Reported	AA/AT/TT	161/426/321	126/394/274	0.613	0.338/0.462
2	rs4675690	*CREB1*	3' region	Reported	CC/CT/TT	162/440/314	117/407/271	0.250	0.734/0.077
8	rs35753505	*NRG1*	5' region	tagSNP	TT/TC/CC	174/434/301	135/371/271	0.577	0.457/0.711
8	rs4268090	*NRG1*	5' region	tagSNP	CC/CT/TT	180/436/301	140/384/267	0.594	0.347/0.942
8	rs10109424	*NRG1*	5' region	tagSNP	TT/TG/GG	155/419/305	119/354/287	0.348	0.627/0.593
8	rs7831093	*NRG1*	5' region	tagSNP	AA/AC/CC	157/438/319	127/368/295	0.557	0.785/0.503
8	rs4457296	*NRG1*	5' region	tagSNP	TT/TC/CC	137/423/357	114/361/317	0.886	0.529/0.494
8	rs4457297	*NRG1*	5' region	tagSNP	TT/TC/CC	136/426/355	114/361/317	0.856	0.674/0.494
8	rs4281084	*NRG1*	5' region	tagSNP	AA/AG/GG	94/395/426	80/330/380	0.812	0.879/0.506
8	rs6994992	*NRG1*	5' region	tagSNP	CC/CT/TT	156/418/339	127/371/289	0.812	0.168/0.656
8	rs7014762	*NRG1*	5' region	tagSNP	AA/AT/TT	106/403/407	92/343/356	0.965	0.711/0.522
8	rs1081062	*NRG1*	Intron 2	tagSNP	CC/CT/TT	9/154/753	9/158/625	0.228	0.697/1
8	rs776389	*NRG1*	Intron 2	tagSNP	TT/TG/GG	123/432/361	130/351/312	0.192	0.777/0.072
8	rs10808318	*NRG1*	Intron 2	tagSNP	AA/AG/GG	80/420/417	72/328/393	0.176	0.079/0.799
8	rs6468119	*NRG1*	Intron 2	tagSNP	TT/TC/CC	29/302/586	35/249/509	0.351	0.209/0.508
8	rs16879552	*NRG1*	Intron 3	tagSNP	CC/CT/TT	140/455/318	146/371/271	0.197	0.302/0.342
8	rs3924999	*NRG1*	Exon 7 (R38Q)	tagSNP	CC/CT/TT	50/327/537	37/277/472	0.728	1/0.758
8	rs4531002	*NRG1*	Intron 8	tagSNP	TT/TC/CC	7/213/696	11/160/621	0.157	0.032/0.861
8	rs2954041	*NRG1*	Intron 8	tagSNP	TT/TG/GG	123/419/374	108/363/315	0.948	0.775/0.819
8	rs6988339	*NRG1*	Intron 8	tagSNP	GG/GA/AA	199/450/267	182/390/221	0.767	0.740/0.722
8	rs4733376	*NRG1*	Intron 14	tagSNP	GG/GA/AA	99/424/394	103/366/322	0.317	0.377/1
8	rs7005288	*NRG1*	Intron 17	tagSNP	AA/AG/GG	175/459/283	141/409/243	0.744	0.688/0.194
8	rs6992642	*NRG1*	3' region	tagSNP	CC/CT/TT	47/275/595	28/244/521	0.273	0.050/1
10	rs10786719	*CNNM2*	5' region	eQTL	AA/AG/GG	171/456/286	147/394/261	0.497	0.687/1
10	rs10883801	*CNNM2*	5' region	dbSNP	CC/CA/AA	69/360/450	72/337/389	0.243	0.868/1
10	rs7914558	*CNNM2*	Intron 2	tagSNP & GWAS	GG/GA/AA	185/441/284	150/399/249	0.421	0.591/0.719
10	rs3740387	*NT5C2*	Exon (D549E)	tagSNP	TT/TC/CC	199/448/267	173/400/223	0.640	0.690/0.831
10	rs11598702	*NT5C2*	Intron 4	tagSNP & reported	CC/CT/TT	58/348/508	50/297/450	0.706	0.931/0.925
10	rs12220375	*NT5C2*	Intron 3	tagSNP & reported	CC/CT/TT	62/328/505	60/297/435	0.561	0.377/0.357
10	rs11191580	*NT5C2*	Intron 3	tagSNP & GWAS	CC/CT/TT	61/339/513	61/299/436	0.425	0.600/0.318
10	rs11191602	*NT5C2*	5' region	dbSNP	TT/TC/CC	54/296/562	34/283/479	0.190	0.083/0.409
10	rs11191609	*NT5C2*	5' region	tagSNP	GG/GT/TT	23/256/628	21/217/556	0.976	0.719/1
10	rs11191612	*NT5C2*	5' region	Reported	GG/GA/AA	79/383/451	63/344/384	0.720	0.937/0.269

10	rs11191686	*NT5C2*	5' region	eQTL	AA/AG/GG	83/399/432	90/310/396	0.381	0.536/0.016
10	rs11191688	*NT5C2*	5' region	eQTL	GG/GA/AA	89/410/411	98/321/381	0.356	0.404/0.019
10	rs2281861	*NT5C2*	5' region	eQTL	TT/TC/CC	89/422/406	92/346/360	0.383	0.178/0.524
10	rs7831	*NT5C2*	5' region	eQTL	CC/CA/AA	69/387/454	81/302/420	0.232	0.295/0.020
11	rs11223064	*OPCML*	3'-UTR	tagSNP	CC/CT/TT	122/458/324	121/392/274	0.323	0.050/0.336
11	rs1784519	*OPCML*	Intron 2	Reported	TT/TC/CC	200/438/278	146/413/238	0.340	0.287/0.173
11	rs1894193	*OPCML*	Intron 2	tagSNP & reported	TT/TC/CC	165/405/343	132/405/254	0.665	0.019/0.190
11	rs1939498	*OPCML*	Intron 2	Reported	AA/AG/GG	195/436/283	170/425/206	0.247	0.257/0.077
11	rs3016384	*OPCML*	Intron 2	tagSNP & reported	AA/AG/GG	172/445/299	138/432/224	0.657	0.788/0.005
11	rs11223225	*OPCML*	Intron 2	tagSNP & reported	CC/CT/TT	147/460/307	124/406/273	0.811	0.275/0.189
11	rs11223249	*OPCML*	Intron 2	tagSNP & reported	AA/AG/GG	55/333/516	53/276/457	0.832	0.928/0.208
11	rs1911255	*OPCML*	Intron 2	tagSNP	CC/CG/GG	80/400/434	92/344/367	0.091	0.393/0.424
11	rs1630675	*OPCML*	Intron 1	Reported	CC/CG/GG	49/306/560	51/274/476	0.380	0.388/0.170
11	rs4379857	*OPCML*	Intron 1	Reported	AA/AG/GG	176/461/279	182/421/200	0.011	0.593/0.180
22	rs2277831	*DGCR2*	3' region	eQTL	GG/GA/AA	10/197/709	9/148/646	0.818	0.431/0.848
22	rs5992854	*DGCR2*	3' region	eQTL	CC/CT/TT	45/319/552	31/274/492	0.173	1/0.395
22	rs2073776	*DGCR2*	3'-UTR	tagSNP & reported	AA/AG/GG	97/393/425	81/387/335	0.763	0.650/0.050
22	rs8399	*DGCR2*	3'-UTR	dbSNP	GG/GT/TT	13/195/706	16/190/584	0.254	1/0.883
22	rs2072123	*DGCR2*	Exon 10	tagSNP & reported	CC/CT/TT	118/423/375	105/388/309	0.729	1/0.365
22	rs2238751	*DGCR2*	Intron 2	tagSNP	AA/AG/GG	17/194/702	19/200/575	0.314	0.367/0.780
22	rs1034727	*DGCR2*	Intron 1	tagSNP	TT/TC/CC	62/376/475	59/344/395	0.486	0.318/0.197
22	rs807759	*DGCR2*	Intron 1	tagSNP & reported	GG/GA/AA	151/443/320	137/411/254	0.425	0.945/0.194

Chr., chromosome; HWE, Hardy-Weinberg equilibrium; UTR, untranslated region

[a] SNPs were selected from GWAS, reported, tagSNP, eQTL and dbSNP database.

[b] *P*-value were calculated by logistic model in which gender as covariate.

Supplementary references

Holland MM, Parson W (2011) GeneMarker® HID: A reliable software tool for the analysis of forensic STR data. J Forensic Sci 56: 29-35

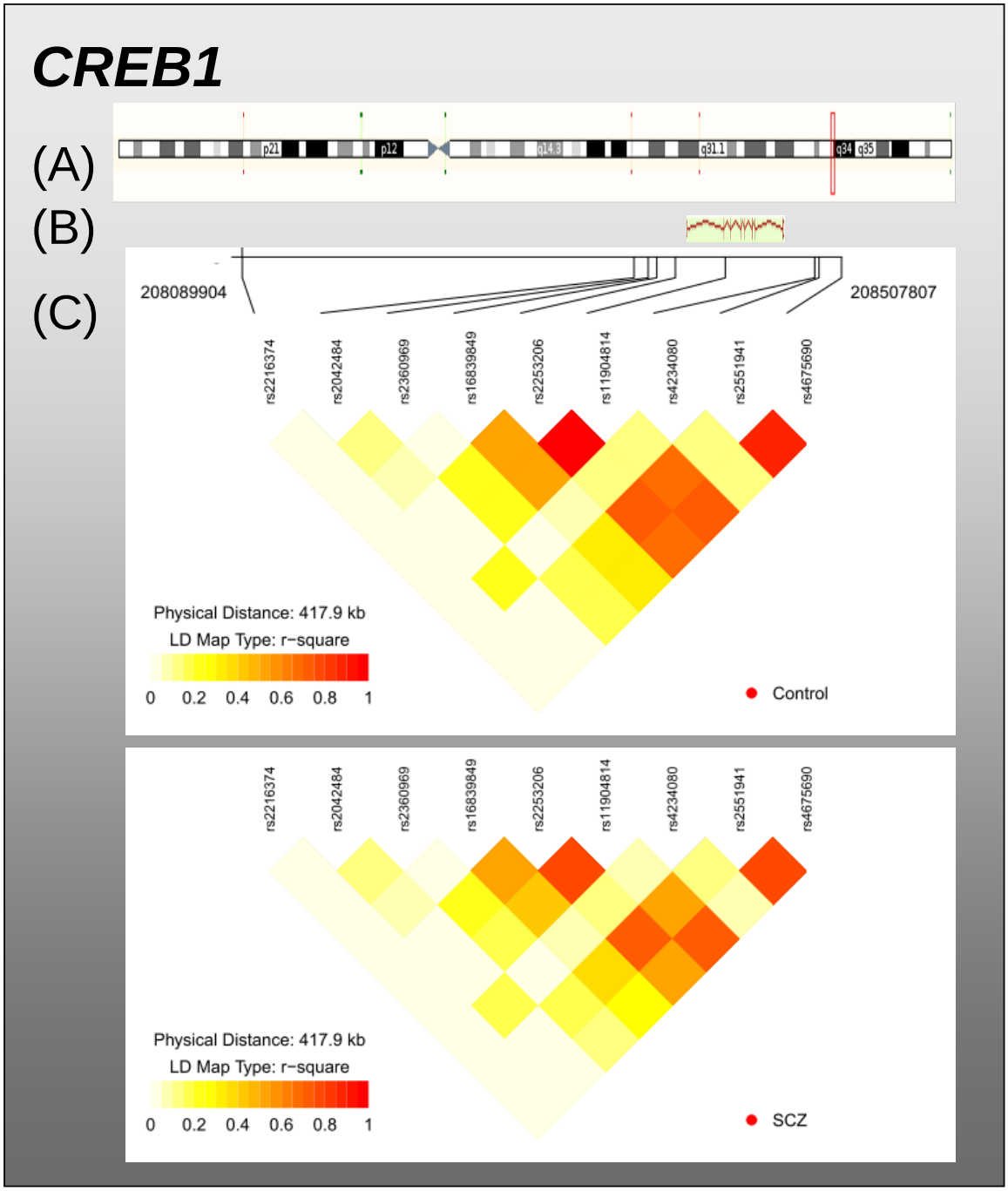
Ripke S, O'Dushlaine C, Chambert K, Moran JL, Kahler AK, Akterin S, Bergen SE, Collins AL, Crowley JJ, Fromer M, Kim Y, Lee SH, Magnusson PK, Sanchez N, Stahl EA, Williams S, Wray NR, Xia K, Bettella F, Borglum AD, Bulik-Sullivan BK, Cormican P, Craddock N, de Leeuw C, Durmishi N, Gill M, Golimbet V, Hamshere ML, Holmans P, Hougaard DM, Kendler KS, Lin K, Morris DW, Mors O, Mortensen PB, Neale BM, O'Neill FA, Owen MJ, Milovancevic MP, Posthuma D, Powell J, Richards AL, Riley BP, Ruderfer D, Rujescu D, Sigurdsson E, Silagadze T, Smit AB, Stefansson H, Steinberg S, Suvisaari J, Tosato S, Verhage M, Walters JT, Levinson DF, Gejman PV, Laurent C, Mowry BJ, O'Donovan MC, Pulver AE, Schwab SG, Wildenauer DB, Dudbridge F, Shi J, Albus M, Alexander M, Campion D, Cohen D, Dikeos D, Duan J, Eichhammer P, Godard S, Hansen M, Lerer FB, Liang KY, Maier W, Mallet J, Nertney DA, Nestadt G, Norton N, Papadimitriou GN, Ribble R, Sanders AR, Silverman JM, Walsh D, Williams NM, Wormley B, Arranz MJ, Bakker S, Bender S, Bramon E, Collier D, Crespo-Facorro B, Hall J, Iyegbe C, Jablensky A, Kahn RS, Kalaydjieva L, Lawrie S, Lewis CM, et al. (2013) Genome-wide association analysis identifies 13 new risk loci for schizophrenia. Nat Genet 45: 1150-9

Sklar P (2011) Large-scale genome-wide association analysis of bipolar disorder identifies a new susceptibility locus near ODZ4. Nat Genet 43: 977-83

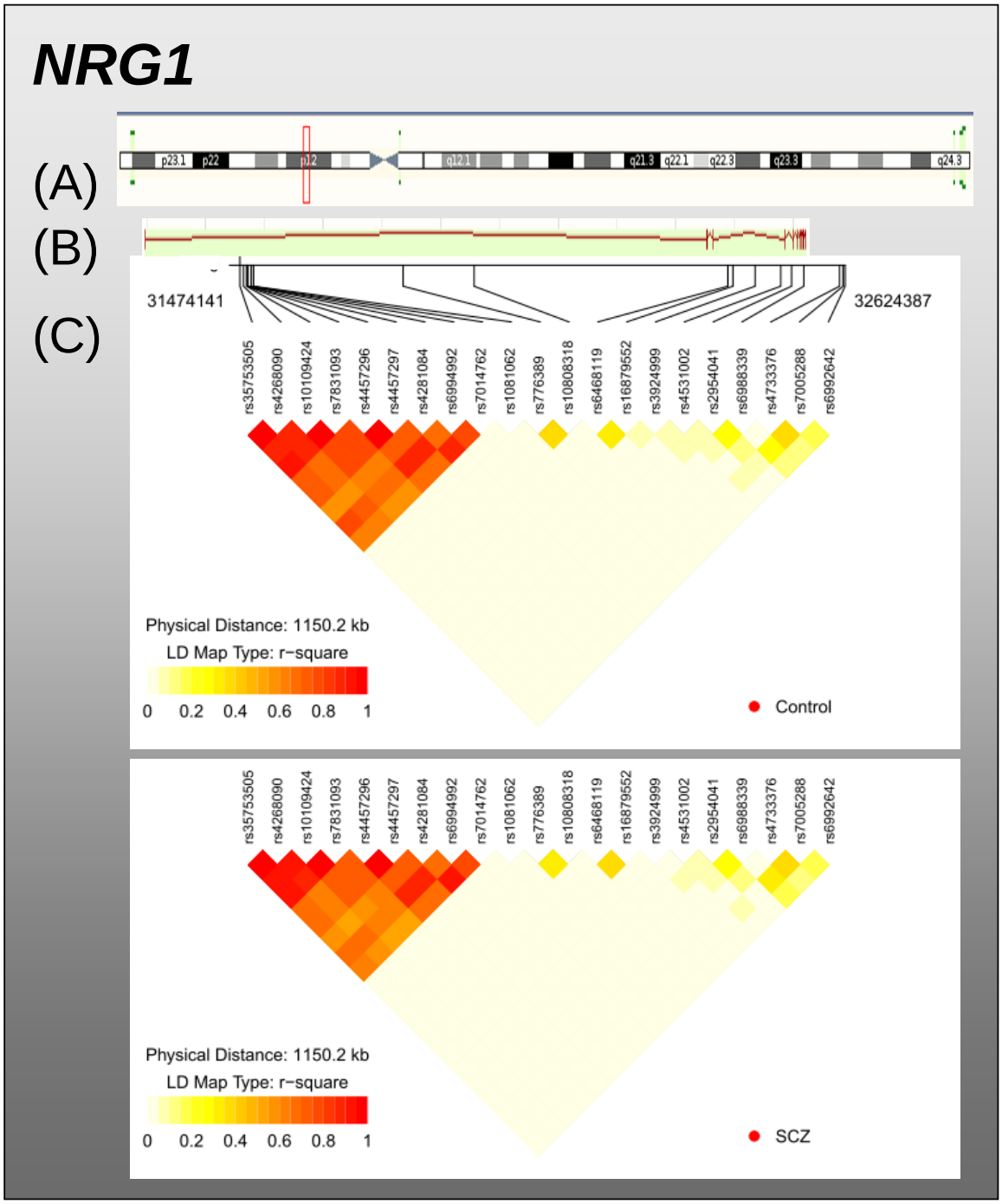
Figure legend

Supplementary Figure 1 Location of SNPs in the analyzed gene region (1. *CREB1*; 2. *NRG1*; 3. *CNNM2-NT5C2*; 4. *OPCML*; 5. *DGCR2*) and linkage disequilibrium pattern in the case and control samples from Hunan Province, China. Shown are the chromosome genetic loci (A), genetic structure (B), and linkage disequilibrium pattern (C) of each analyzed gene.

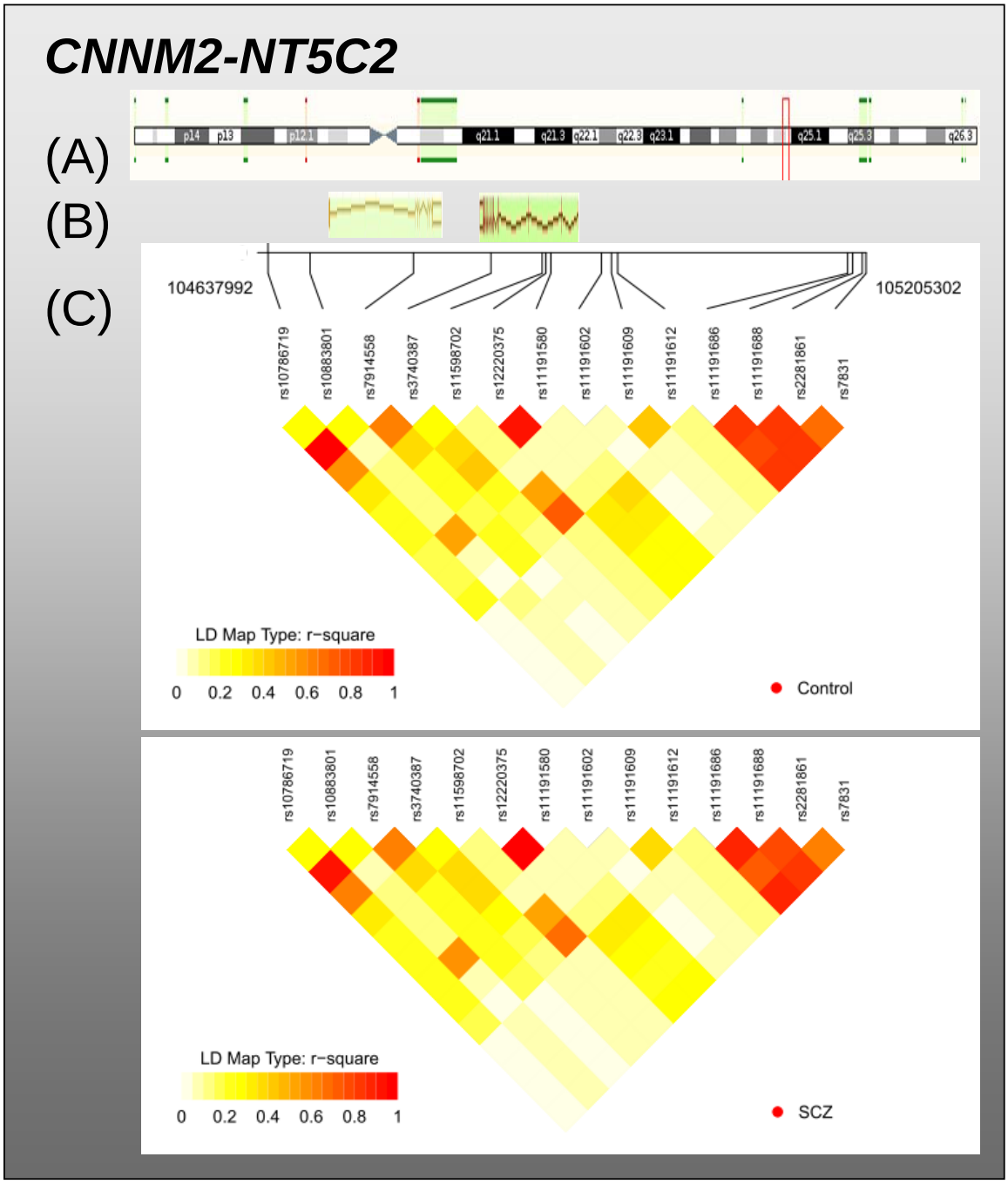
Supplementary Figure 2 Haploview plot of linkage disequilibrium pattern (r^2) of *CREB1* rare haplotype SNPs based on data from (A) schizophrenia and (B) control of the current study and data of the corresponding SNPs from (C) HapMap CHB (Data from HapMap 3 Release 2).



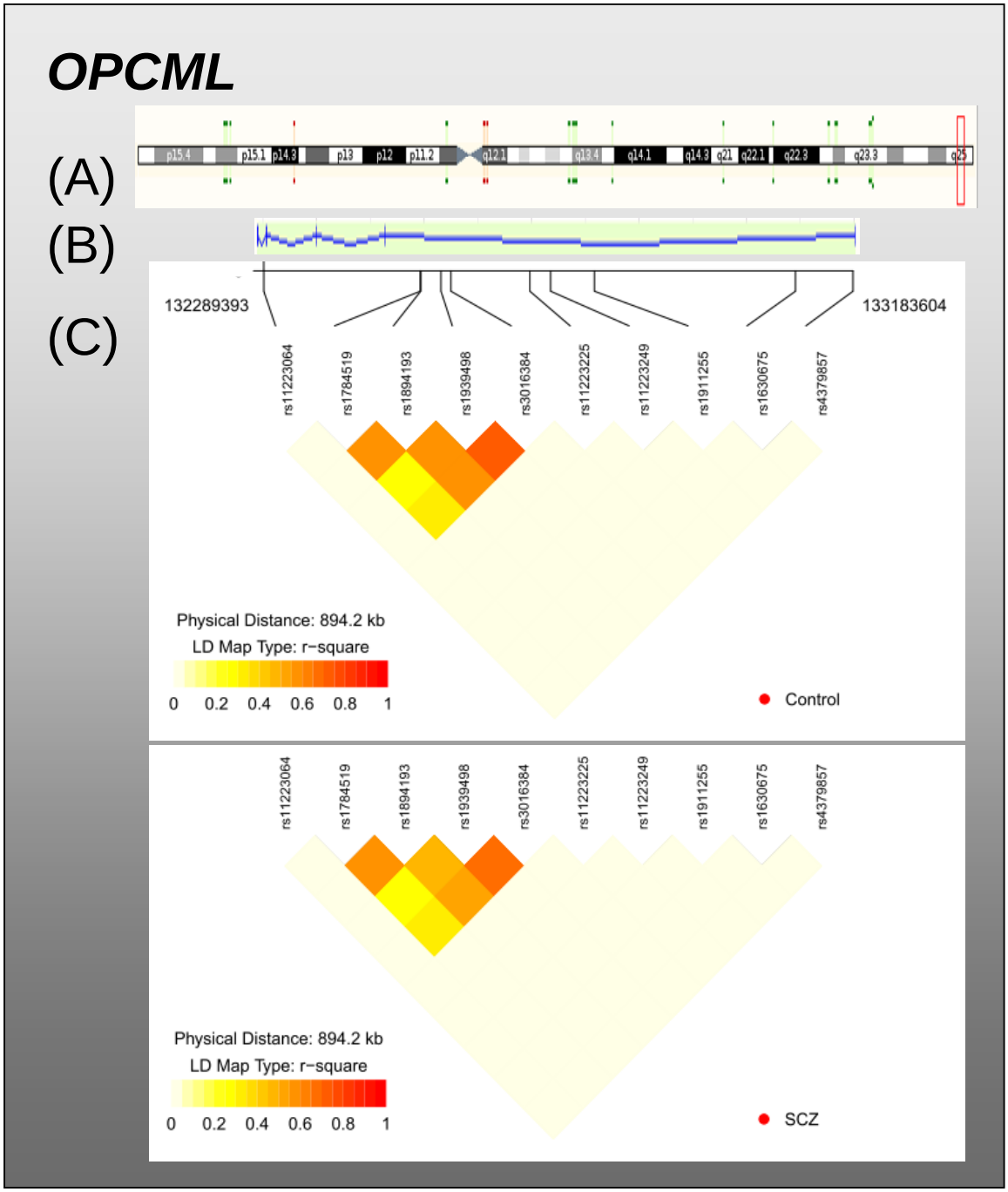
Supplementary Figure 1-1



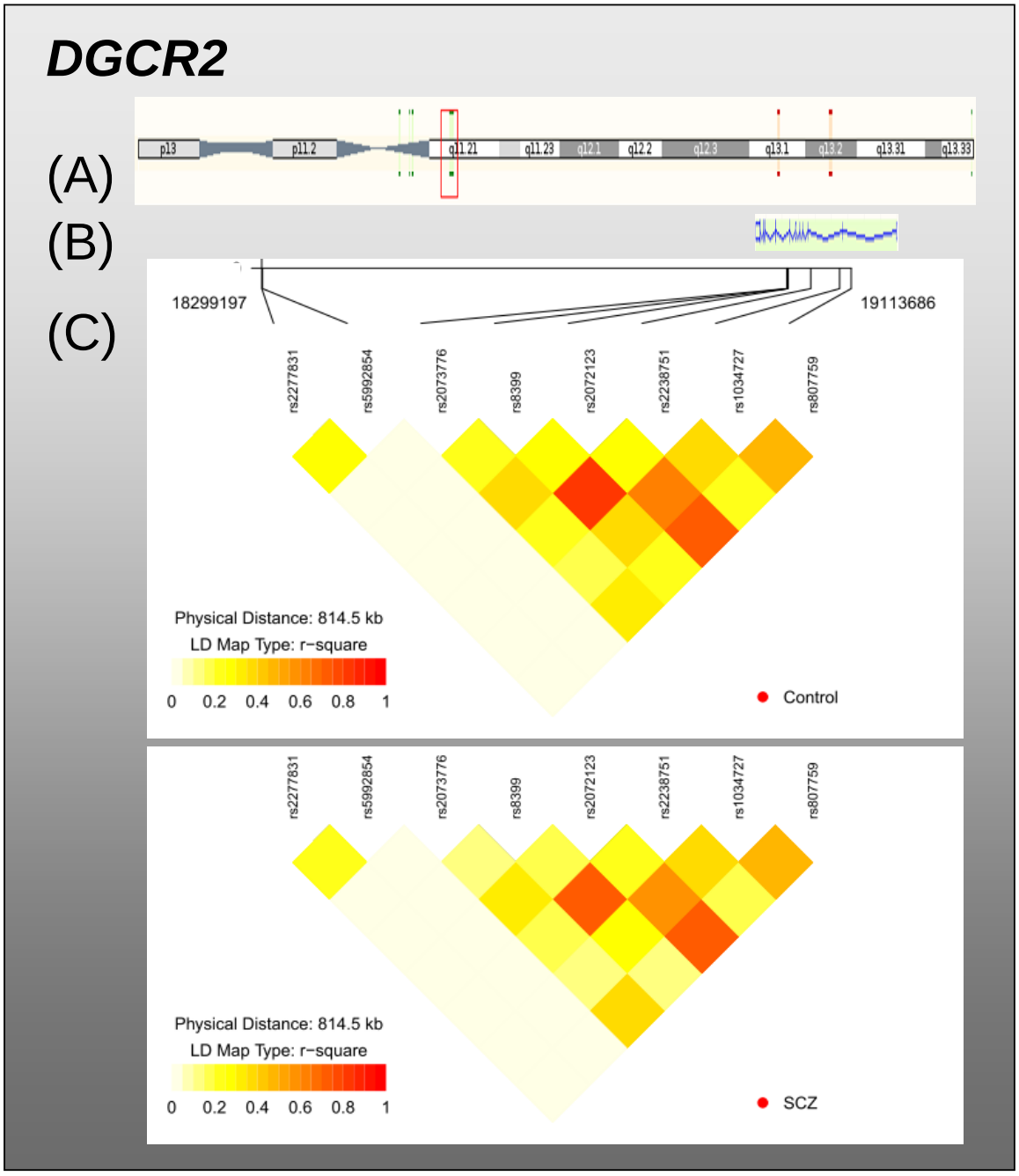
Supplementary Figure 1-2



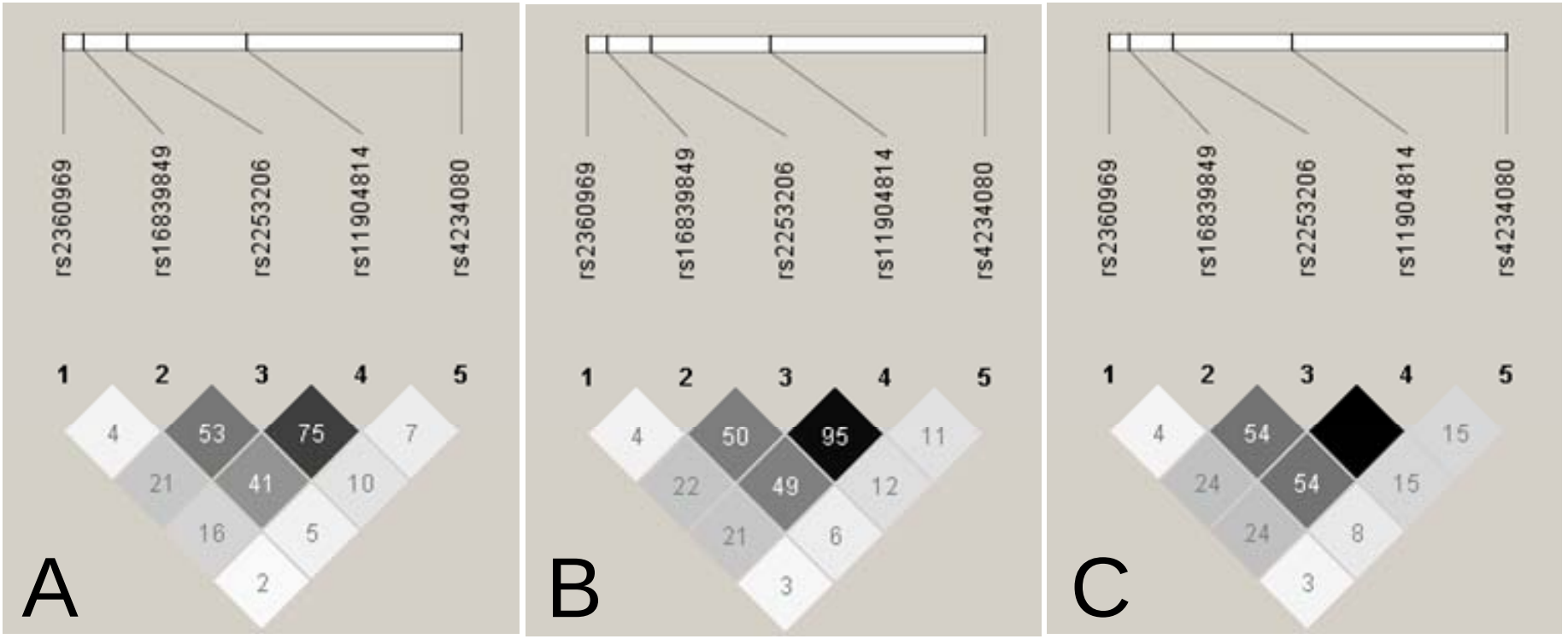
Supplementary Figure 1-3



Supplementary Figure 1-4



Supplementary Figure 1-5



Supplementary Figure 2