Sample description

We recruited 278 unrelated Han Chinese healthy subjects (174 females and 104 males, mean age 36.21 ±12.58 years). All volunteers were free from mental disorders, drug abuse, alcohol dependence, or brain injury. All individuals provided written informed consents for participation. The research protocol was approved by the internal review board of Kunming Institute of Zoology, Chinese Academy of Sciences. This imaging sample has been used previously and shown to be effective for the detection of genetic effects on the features extracted from magnetic resonance imaging (MRI) data, e.g., we have previously reported that a schizophrenia risk SNP rs2312147 in VRK2 was significantly associated with total white matter volume and total brain volume in this sample (1).

MRI acquisition and image preprocessing

Structural MRI data were acquired using a Philips MRI scanner (Achieva Release 3.2.1.0) operating at 3 Tesla. High-resolution whole-brain T1-weighted images were acquired sagittally with an inversion-recovery prepared 3-D spoiled gradient echo (SPGR) pulse sequence (Repetition Time=7.38ms, Echo Time=3.42ms, flip angle=8, voxel dimensions 1.04×1.04×1.80 mm^3, slice thickness=1.2mm).

All preprocessing was performed using SPM8 (Wellcome Trust Centre for Neuroimaging, London, UK. http://www.fil.ion.ucl.ac.uk/spm/software/spm8/). Specifically, high resolution T1 images were first segmented into gray matter, white matter and cerebrospinal fluid in the native space using SPM's new segmentation. The gray matter images were then iteratively aligned to an increasingly crisp average template by DARTEL (2). Finally, all images were normalized to the standard Montreal Neurological Institute (MNI) template, modulated to account for volume changes in the warping, and re-sampled to 2×2×2 mm^3. Modulated gray matter images were smoothed with 8 mm Gaussian kernels. A gray matter analysis mask was constructed by thresholding the averaged gray matter image that maximized the sample Pearson correlation between the binarized mask and the original average image (3).

Genotyping

Venous blood was collected from all the participants, and genomic DNA was extracted from the blood sample using the phenol-chloroform method. SNaPShot

Data Supplement for Li et al. (doi: 10.1176/appi.ajp.2013.12111458)

method was used to genotype SNP rs1545843 in the sample, as described in our previous studies (4,5). In brief, the genomic fragments which contain the selected SNP were amplified by polymerase chain reaction (PCR) with a total volume of 25 μl (including 10 ng of genomic DNA) in 96-well plates. The amplified fragments were purified and specific genotyping primers were used to amplify the target site. After one base extension, the reaction was terminated and the products were loaded on an ABI 3130 automatic sequencer (Applied Biosystems). Details of the primer and assay condition are available on request. The SNP genotype callings were automatically performed using ABI GeneMapper 4.0 and verified manually. The genotyping success rate is >95%. SNP rs1545843 is in Hardy-Weinberg equilibrium in the sample analyzed (AA=35, GA=133, GG=105; P_{HWE}=0.48).

Test for population stratification

We used PLINK v1.07 for the pair-wise identities by descent (IBD) analysis and calculated the multidimensional scaling of IBD distances in our imaging sample using 76 selected SNPs (**Table S1**) (6). We also obtained data on these 76 SNPs from 273 samples from the 1000-Human Genome Project (85 subjects for CEU [Utah residents with ancestry from northern and western Europe in the United States], 88 subjects for YRI [Yoruba in Ibadan, Nigeria], and 100 subjects for CHS [Southern Han Chinese] and found that these SNPs worked well to differentiate the samples into three clusters by ethnicity (**Figure S1A**). When merging the data with our samples, multidimensional scaling plot indicated that our sample clustered with the CHS sample and can be clearly differentiated from the CEU and YRI samples (**Figure S1B**).

We then calculated genomic inflation factor (λ) in our sample, using PLINK v1.07 (6) (the "--linear --adjust" command line option), with the same 76 SNPs to test potential population stratification. It turned out that λ was low (=1.00) when regional brain volumes (e.g., hippocampus), the total brain volume, and the total intracranial volume were used as phenotypes. We also studied Q-Q plots to ensure that the p-value distributions did not deviate from expectations (**Figure S2**). These results indicated that there is a negligible population stratification in our sample, and the findings are unlikely to be a false-positive association simply caused by population stratification.

Data Supplement for Li et al. (doi: 10.1176/appi.ajp.2013.12111458)

Statistical analysis

In this study, the association between rs1545843 and the brain structures was conducted using a linear regression model at each voxel, adjusting for age, gender and the total intracranial volume (the total volume of gray matter, white matter and cerebrospinal fluid). The regressor for the SNP was coded in a trinary fashion, with 0, 1, 2, corresponding to homozygotes for the major allele (GG), heterozygotes (AG) and homozygotes for the minor allele (AA), respectively. This corresponds to a commonly used additive model of allelic effect.

Voxel-wise p-values were then corrected over the brain to control family-wise error (FWE) rate. In particular, random field theory (7) was used to detect effect or activation at an unknown spatial location by assessing whether the test statistic value or the spatial extent of a set of contiguous voxels exceeding some predefined threshold is unusually large by chance alone, known as voxel-wise inferences and cluster size inferences respectively. In this study, a non-parametric permutation method was used to accurately estimate the probability of an observed peak or cluster (8). Specifically, the SNP regressor was randomly permuted and the maximal statistic (voxel or cluster) over the brain was saved for each permutation. The p-value for the observed statistic *T* was then computed as the proportion of the permutation distribution as or more extreme than *T*. 10,000 permutations were performed.

Data Supplement for Li et al. (doi: 10.1176/appi.ajp.2013.12111458)

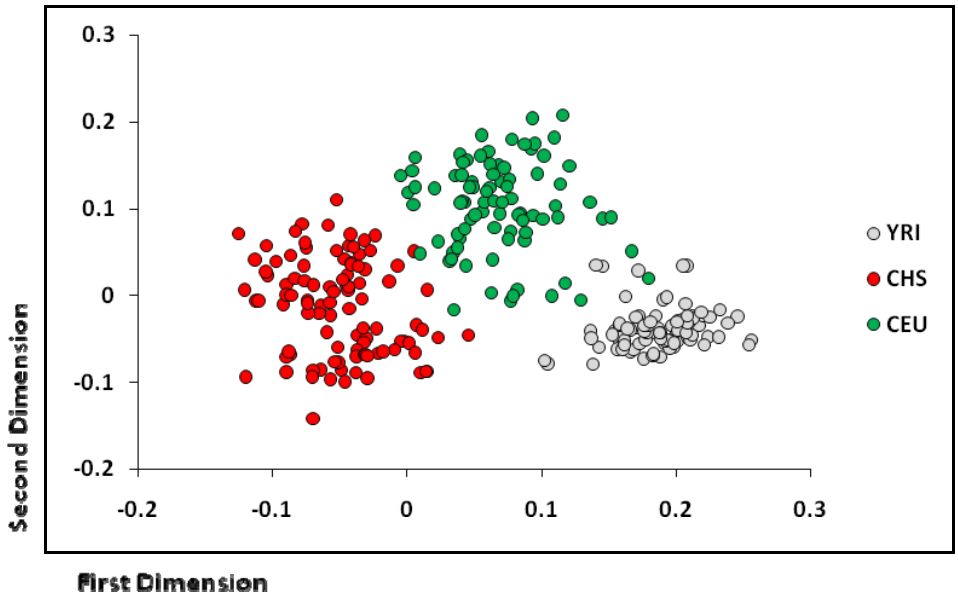
Table S1. A list of the 76 SNPs across the genome included in the population stratification analysis.

Chromosome	SNP	Chromosome	SNP
1	rs12138072	1	rs1625579
1	rs10489202	1	rs1127661
1	rs10494373	1	rs3818361
2	rs6741949	2	rs2678907
2	rs2312147	2	rs744373
2	rs359895	2	rs12693384
2	rs7557843	2	rs1344706
3	rs2251219	3	rs2239547
3	rs6782799	3	rs6438552
3	rs7431209	3	rs334558
3	rs3755557	5	rs3916441
5	rs743564	5	rs31480
6	rs6913660	6	rs1778477
6	rs13219354	6	rs13211507
6	rs2142731	6	rs4273712
8	rs2920603	8	rs2916656
8	rs7010345	8	rs4281084
8	rs2243755	8	rs1488935
8	rs16887244	9	rs7852872
9	rs7861226	10	rs10994336
10	rs10883854	10	rs7914558
10	rs11191580	11	rs835784
11	rs12576775	11	rs3851179
11	rs12807809	11	rs6265
12	rs6581612	12	rs17178006
12	rs7294919	12	rs1006737
12	rs1545843	12	rs10850828
13	rs3916967	13	rs9553470
13	rs9578796	13	rs9553471
13	rs9578798	13	rs1012053
14	rs2494731	14	rs1130233
14	rs10149779	14	rs3803300
15	rs2899424	15	rs7169052
17	rs9303525	18	rs9960767
19	rs1064395	19	rs7250448
19	rs3810448	22	rs737865
22	rs4680	22	rs165599

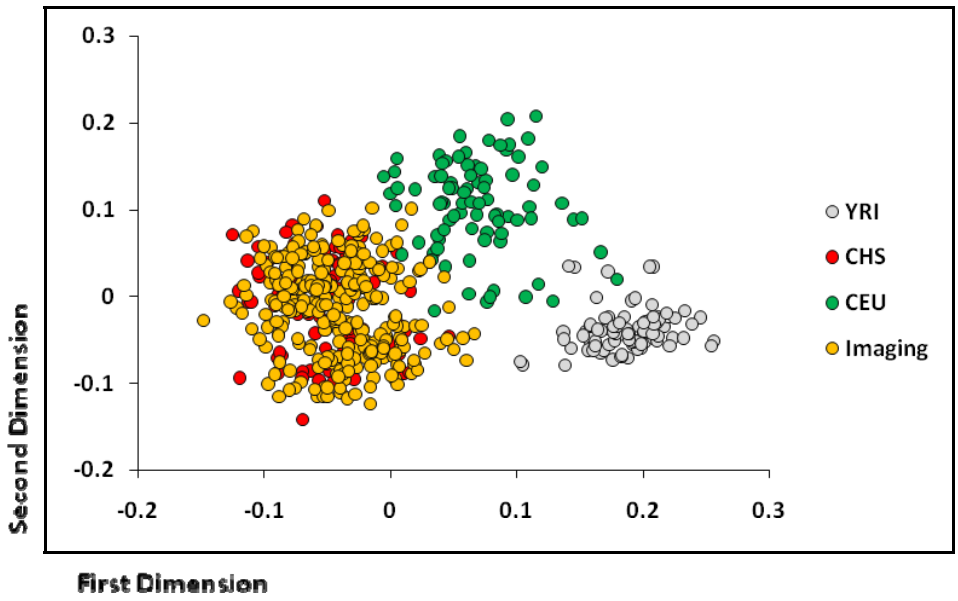
Data Supplement for Li et al. (doi: 10.1176/appi.ajp.2013.12111458)

Figure S1. Multidimensional scaling plots for the samples with three different ancestries from the 1000-Human Genome Project (A) and the same with our sample in this analysis overlaid (B).

A

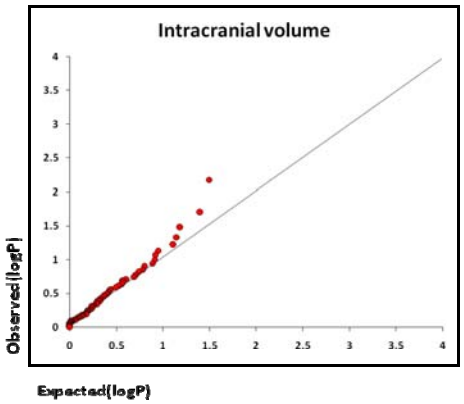


B

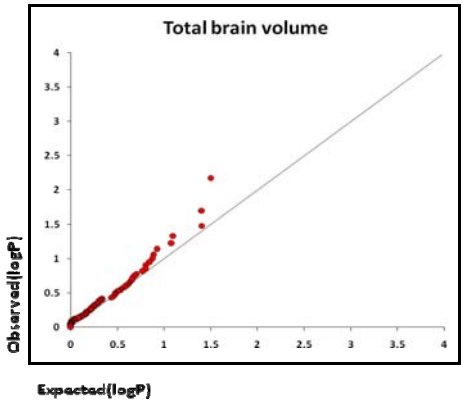


Data Supplement for Li et al. (doi: 10.1176/appi.ajp.2013.12111458)

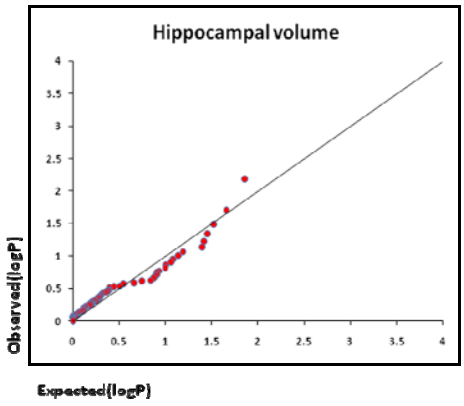
Figure S2. Q-Q plots of the p-values when testing the association between the 76 SNPs and three different phenotypes, the total intracranial volume, the total brain volume and the hippocampal volume, showing that the test statistics are not inflated.



λ=1.00



λ=1.00



λ=1.00

Data Supplement for Li et al. (doi: 10.1176/appi.ajp.2013.12111458)

References

1. Li M, Wang Y, Zheng XB, Ikeda M, Iwata N, Luo XJ, Chong SA, Lee J, Rietschel M, Zhang F, Muller-Myhsok B, Cichon S, Weinberger DR, Mattheisen M, Schulze TG, Martin NG, Mitchell PB, Schofield PR, Liu JJ, Su B: Meta-analysis and brain imaging data support the involvement of VRK2 (rs2312147) in schizophrenia susceptibility. Schizophr Res 2012; 142:200-205.
2. Ashburner J: A fast diffeomorphic image registration algorithm. Neuroimage 2007; 38:95-113.
3. Ridgway GR, Omar R, Ourselin S, Hill DL, Warren JD, Fox NC: Issues with threshold masking in voxel-based morphometry of atrophied brains. Neuroimage 2009; 44:99-111.
4. Li M, Mo Y, Luo XJ, Xiao X, Shi L, Peng YM, Qi XB, Liu XY, Yin LD, Diao HB, Su B: Genetic association and identification of a functional SNP at GSK3beta for schizophrenia susceptibility. Schizophr Res 2011; 133:165-171.
5. Li M, Luo XJ, Xiao X, Shi L, Liu XY, Yin LD, Ma XY, Yang SY, Pu XF, Yu J, Diao HB, Shi H, Su B: Analysis of common genetic variants identifies RELN as a risk gene for schizophrenia in Chinese population. World J Biol Psychiatry 2013; 14:91-99.
6. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de Bakker PI, Daly MJ, Sham PC: PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet 2007; 81:559-575.
7. Worsley KJ, Marrett S, Neelin P, Vandal AC, Friston KJ, Evans AC: A unified statistical approach for determining significant signals in images of cerebral activation. Human Brain Mapping 1996; 40:58-73.
8. Ge T, Feng J, Hibar DP, Thompson PM, Nichols TE: Increasing power for voxel-wise genome-wide association studies: the random field theory, least square kernel machines and fast permutation procedures. NeuroImage 2012; 63:858-873.

Data Supplement for Li et al. (doi: 10.1176/appi.ajp.2013.12111458)