nature research

Corresponding author(s): Hailiang Huang

Last updated by author(s): Jul 24, 2019

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☐	☒	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	No other direct software used for data collection. Genome Studio v1.9.4 used for extracting Plink binary genotype files.
Data analysis	Computer code used in this manuscript: RICOPILI (quality control, principal component analysis, pre-phasing, imputation, association test and meta-analysis) https://github.com/Nealelab/ricopili/wiki; embedded within RICOPILI (Eigenstrat https://github.com/DReichLab/EIG/tree/master/EIGENSTRAT; SHAPEIT https://mathgen.stats.ox.ac.uk/genetics_software/shapeit/shapeit.html; EAGLE https://github.com/poruloh/Eagle; IMPUTE https://mathgen.stats.ox.ac.uk/impute/impute_v2.html; Minimac https://genome.sph.umich.edu/wiki/Minimac); POPCORN (trans-ancestry genetic correlation): https://github.com/brielin/Popcorn; LDSC (heritability, partitioned heritability and within-ancestry genetic correlation): https://github.com/bulik/ldsc; MAGMA (pathway analysis): https://ctg.cncr.nl/software/magma; Fine-mapping (Fine-mapping and PAINTOR): https://github.com/hailianghuang/FM-summary, https://github.com/gkichaev/PAINTOR_V3.0; REHH (selection): https://cran.r-project.org/web/packages/rehh/index.html; B score (background selection): http://www.phrap.org/othersoftware.html; PRS analyses: https://github.com/armartin/pgc_scz_asia

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

Genome-wide summary statistics from EAS samples, EUR samples ("49 EUR samples") and all samples (EAS and EUR combined) in this study can be downloaded from https://www.med.unc.edu/pgc/results-and-downloads/. Individual-level genotype data for EAS samples are available upon request to contact authors

(Supplementary Information). Alternately, requests can be made to the Psychiatric Genomics Consortium (PGC). In this case, access to individual-level genotypes from samples recruited outside of mainland China will go through the PGC "fast-track" approval. Access to individual-level genotypes from samples recruited within mainland China has to be approved by the individual Chinese contact authors (Supplementary Information), and are subject to the policies and approvals from the Human Genetic Resource Administration, Ministry of Science and Technology of the People's Republic of China. Individual-level genotypes from samples recruited within mainland China have been stored and kept in a server physically located in mainland China. Analyses were performed on these samples using the same computer codes as those used for other EAS and EUR samples, which are available in the Code availability section.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Previous studies of European individuals suggest that a sample size of over 100,000 individuals may be needed to saturate the power of schizophrenia genetic loci discovery. We have not reached this sample size but have included all major East Asian samples that are available date in the current study.
Data exclusions	We used standard quality control procedures for genome-wide association studies. Quality control procedures were carried out as part of the RICOPILI pipeline (https://sites.google.com/a/broadinstitute.org/ricopili/home; Lam et al., 2019) with the following steps and parameters: 1) Excluding variants with call rate below 95%; 2) Excluding subjects with call rate below 98%; 3) Excluding monomorphic variants; 4) Excluding subjects with inbred coefficient above 0.2 and below -0.2; 5) Excluding subjects with mismatch in reported gender and chromosome X computed gender; 6) Excluding variants with missing rate differences greater than 2% between cases and controls; 7) Subsequent to step 6, exclude variants with call rate below 98%; and 8) Exclude variants in violation of Hardy-Weinberg equilibrium (P < 10-6 for controls or P < 10-10 for cases). Numbers of variants or subjects removed in each step were reported in Supplementary Table 2. Additionally, We used Eigenstrat66 to calculate the principal components for all the samples using the best guess genotypes from imputation (Supplementary Figure 9b). We computed the identity-by-descent matrix to identify intra- and inter- dataset sample overlaps. Samples with pi-hat > 0.2 were extracted, followed by Fisher-Yates shuffle on all samples. The number of times with which each sample was related to another sample was tracked and samples that were related to more than 25 samples were removed. When deciding which samples to retain, trio were preferred, followed by cases, and thereafter a random sample for each related pair was removed, 704 individuals were removed. To identify population outliers, k-means clustering was conducted using the first 20 PCs from PCA and covariates representing each of the 13 Stage 1 samples. Guided by results of k-means clustering and visual inspection of PCA plots, 46 individuals were identified as outliers and were excluded. These steps are considered standard procedures to retain high quality samples fro the data-analysis.
Replication	The current study employs a two-stage study. Stage 2 serves the purpose replication.
Randomization	No randomization was conducted. The current study is a large-scale genetics study. Randomness is achieved from the nature of how alleles are relatively distributed in the population. Alleles are randomly passed from parent to offspring during meiosis, hence the nature of these types of studies tend to be shielded from extraneous confounds of standard epidemiological studies.
Blinding	No blinding was carried out. The nature of the study is in of itself blind to study recruiters. No one would beforehand know the genotype make up of the cases and controls recruited in the current study.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology
☒	☐ Animals and other organisms
☐	☒ Human research participants
☒	☐ Clinical data

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Human research participants

Policy information about studies involving human research participants

Population characteristics	The following samples were used in this study: EAS samples, full-genome: genome-wide genotype data was obtained from 16 samples from East Asia (Supplementary Table 1). Two of these samples (TAI-1 and TAI-2) had parents off-spring trios and were processed as case/pseudo-controls. DSM-IV was used for diagnosing all schizophrenia cases in these samples except for the trios (TAI-1 and TAI-2), for which DIGS was used. EAS samples, selected variants: summary statistics was obtained for a set of variants from four EAS samples (BJM-2, BJM-3, BJM-4, BIX-5) which had been analyzed in published studies7,8. The summary statistics included odds ratio, standard error, reference and tested alleles for variants that have P<10-5 in either Stage 1 or the meta-analysis combining Stage 1 and EUR samples. Between 22,156 and 31,626 variants were available after the exclusion of strand ambiguous60 variants (Supplementary Table 2). EUR samples: Genotypes for EUR schizophrenia patients and controls were obtained from the Psychiatric Genomics Consortium as reported in ref 2. All samples of EUR ancestry were included in this study except for the deCODE samples (1,513 cases and 66,236 controls). We also like to note that three samples of EAS ancestry reported in this publication were not included in the EUR samples in our analysis but were included in the EAS samples (IMH-1, HNK-1 and JPN-1). The same procedures used in processing EAS samples were applied to the EUR samples.
Recruitment	East Asian Ancestry Cohorts: Stage 1 Discovery and Stage 2 Replication Sample: IMH-1, HNK-1 and JPN-1 Type: Case-Control Contact: IMH-1: Liu, Jianjun (liuj3@gis.a-star.edu.sg) and Lee, Jimmy (jimmy_lee@imh.com.sg); HNK-1: Sham, Pak (pcsham@hku.hk); JPN-1: Iwata, Nakao (nakao@fujita-hu.ac.jp) Description: These samples were included in a previous publication1. Description of these samples can be found on Page 11 of the Supplementary Information in the publication ("East Asian ancestry, case-control design" subsection). Mapping between the study identifiers: IMH-1=scz_tcr1_asn; HNK-1=scz_hok2_asn; JPN-1=scz_jpn1_asn. Sample: IMH-2 Type: Case-Control Contact: Liu, Jianjun (liuj3@gis.a-star.edu.sg) and Lee, Jimmy (jimmy_lee@imh.com.sg) Description: Subjects were part of the Singapore Translational and Clinical Research in Psychosis (STCRP) Study, and comprised of cases and healthy controls of Chinese ancestry. Schizophrenia cases were recruited from the Institute of Mental Health, Singapore. Diagnosis of schizophrenia was made using the Structured Clinical Interview for DSM-IV, Research Version, Patient Edition (SCID-I/P) by trained raters. And healthy controls were screened using the DSM-IV, Research Version, Non-Patient Edition (SCID-N/P). All participants gave consent according to IRB/DSRB guidelines prior to participating. Sample: BIX-1, BIX-2, BIX-3 and BIX-5 Type: Case-Control Contact: Shi, Yongyong (shiyongyong@gmail.com) Description: BIX1, BIX2 and BIX3: Participants in the schizophrenia group were in-patients or out-patients who were recruited from various mental health centres. The patients were interviewed by two independent psychiatrists and were diagnosed according to the DSM-IV criteria, and had two-year histories of the disorder. All met the following two criteria: preoccupation with one or more delusions and frequent auditory hallucinations. However, none of the following symptoms were prominent: disorganised speech, disorganised or catatonic behaviour, or flat or inappropriate affect. All healthy controls were randomly selected from Chinese Han volunteers (from hospitals and a community survey) who were asked to reply to a written invitation to evaluate their medical histories. Potential lists of controls were screened for suitable volunteers by excluding individuals with major mental illnesses. The study had been previously reported2. BIX5: Case were schizophrenia patients according to DSM-IV criteria, and controls were volunteers with no psychiatric history. All the cases and controls were of Han Chinese ancestry, and recruited in the mainland of China. Further details for ascertainment, inclusion/exclusion criteria, and other information have been previously described3,4. Approval was obtained from the Ethics Committee of Human Genetic Resources at the Bio-X Institutes of Shanghai Jiao Tong University, in accordance with the tenets of the Declaration of Helsinki, and all participants provided written, informed consent. The genotyping was conducted at Bio-X Institutes at Shanghai Jiao Tong University, and details were provided in previous publication3. Sample: XJU-1 and BIX-4 Type: Case-Control Contact: Qin, Shengying (chinsir@sjtu.edu.cn) Description: XJU-1: Schizophrenia patients were recruited from the inpatients in nine clinical centers in Shaan Xi Province, China, including 1) Department of Psychiatry of the First Affiliated Hospital of Xi'an Jiaotong University (TFAHXJTU) and Mental Health Centers of 2) Baoji City, 3) Xianyang City, 4) Weinan City, 5) Hanzhong City, 6) Huayin City, 7)Yulin City, 8)Yanan City. Healthy controls were recruited from native residents of communities and villages in the city that each clinical center located in. Two professional psychiatrists interviewed all participants and evaluated their mental status according to DSM-IV criteria using SCID and/or Mini-International Neuropsychiatric Interview. The present study was approved by the Medical Ethics Committee of TFAHXJTU. All subjects participated voluntarily and signed written informed consent prior to the enrollment. BIX-4: Subjects were recruited from the Shanghai Mental Health Center, Shanghai Jiao Tong University School of Medicine between August 2015 and August 2017. Patients were diagnosed by at least two experienced psychiatrists according to the DSM-IV based on structured Clinical Interview for DSM-IV Axis 1 Disorders (SCID-1). Healthy controls were selected from the general public of Chinese population in Shanghai. The control subjects were also excluded with major mental illness and matched with patients in terms of age and gender. All the subjects were Han Chinese in origin. The study was approved by the Ethical Committee of Human Genetic Resources in Shanghai. All subjects or their guardians understood the procedure and gave written informed

consent, according to the Declaration of Helsinki.

Sample: UMC-1 and SIX-1
Type: Case-Control
Contact: Kahn, Rene (rene.kahn@mssm.edu) and Yu, Xin (yuxin@bjmu.edu.cn)
Description: Funded by the Division of Neuroscience of the University Medical Center in Utrecht (UMC, Utrecht), a collaborative genetic study was developed between UMC Utrecht and Peking University Institute of Mental Health (PKUIMH). The goal of the study was to identify susceptibility genes that contribute to the development of schizophrenia and related disorders. Subjects were recruited from thirteen major cities across mainland China. All patients were interviewed with Structured Clinical Interview for DSM-IV axis I Disorder (SCID–I) by extensively trained Chinese psychiatrists and researchers involved in the study. All included patients fulfilled DSM-IV criteria for schizophrenia. The SCID interviews were tape-recorded and randomly checked by senior psychiatrists not involved in the recruitment of interviewing. Healthy controls were matched with patients in terms of age, gender, and geographic information, and educational level, ethnic and economic background and were free of a history of psychiatric illness based on GHQ-20 and SCL-90. Regular inter-rater reliability sessions and onsite check-ups were conducted to secure the long-term research quality. This study was approved by the institutional review boards of the participating hospitals, including that at UMC Utrecht. Written informed consent was obtained from all subjects after complete description of the study and its procedures.

Sample: UWA-1
Type: Case-Control
Contact: Schwab, Sibylle (schwab@uow.edu.au)
Description: Patients with schizophrenia admitted consecutively to psychiatric hospitals in the greater Jakarta area were informed about the study and were asked to sign the Informed Consent document for participation in interviews, blood withdrawal and subsequent genetic studies. The study was approved by the institutional review board of the University of Indonesia and by local Ethics Committees. Clinical consensus diagnosis of schizophrenia was made by psychiatrists according to the DSM-IV criteria. Non-psychiatric controls were recruited from students and staff of the University of Indonesia, Jakarta, and from the participating hospitals. Details of the study were previously reported5.

Sample: BJM-1, BJM-2, BJM-3 and BJM-4
Type: Case-Control
Contact: Yue, Weihua (puh60380@bjmu.edu.cn)
Description: All samples in BJM1 and BJM3 are of Northern Han Chinese origin, all participants in BJM2 are from northern China; and all participants in BJM4 are from a southern Han Chinese population. Consensus diagnoses were made by at least two experience psychiatrists according to the criteria for schizophrenia from DSM-IV. None of the subjects had sever medical complications or other psychiatric disorders. All healthy control individuals were clinically determined to be free of psychiatric disorders or family history of such disorders. The study was approved by the institutional review board of each hospital and written consent was obtained from all participants. Further information can be found in a previous publication6, with the following mapping between cohorts: BJM1 = GWAS3, BJM2 = GWAS1, BJM3 = GWAS2, and BMJ4 = GWAS4.

Sample: TAI-1 and TAI-2
Type: Trio
Contact: Tsuang, Ming (mtsuang@ucsd.edu) and Hwu, Hai-Gwo (haigohwu@ntu.edu.tw)
Description: Schizophrenia probands and their first-degree relatives were recruited from Schizophrenia Trio Genomic Research in Taiwan (S-TOGET) funded by National Institute of Mental Health. Schizophrenia patients were interviewed with the Diagnostic Interview for Genetic Studies (DIGS), which was designed specifically for family-genetic studies of schizophrenia and bipolar disorder. Research diagnostic assessment was made independently by two board-certified psychiatrists based on integrated clinical information of DIGS interview data and summary note of clinical course, symptom manifestations and social functioning derived from the records of the medical charts according to DSM-IV for schizophrenia. This study was approved by the institutional review boards of the participating hospitals. Written informed consent was obtained from all subjects after complete description. Study details have been reported elsewhere7.

Sample: KOR-1
Type: Case-Control
Contact: Hong, Kyung Sue (hongks@skku.edu)
Description: Patients who met the Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition (DSM-IV) criteria were recruited from the inpatient unit and outpatient clinic of Samsung Medical Center and Chonnam National University Hospital. South Korea. The healthy subjects consisted of volunteers from the community who were free of any history of clinically significant psychiatric symptoms. Written informed consent was obtained from all subjects after a complete explanation of the study. This study was approved by the institutional review boards of Samsung Medical Center and Chonnam National University Hospital. Details of the study have been previously reported8.

** Note: We have carried out large scale recruitment of case-control samples across these sites. As the majority of these samples are of East Asian descent, we do not foresee potential bias pertaining to our conclusions of the current study - at the genomic level.

Reference
1. Schizophrenia Working Group of the Psychiatric Genomics Consortium. Biological insights from 108 schizophrenia-associated genetic loci. Nature 511, 421–427 (2014).
2. Li, Z. et al. Loci with genome-wide associations with schizophrenia in the Han Chinese population. Br. J. Psychiatry 207, 490–494 (2015).
3. Li, Z. et al. Genome-wide association analysis identifies 30 new susceptibility loci for schizophrenia. Nat. Genet. 49, 1576–1583 (2017).
4. Shi, Y. et al. Common variants on 8p12 and 1q24.2 confer risk of schizophrenia. Nat. Genet. 43, 1224–1227 (2011).
5. Schwab, S. G. et al. Association of rs1344706 in the ZNF804A gene with schizophrenia in a case/control sample from Indonesia. Schizophr. Res. 147, 46–52 (2013).

6. Yue, W.-H. et al. Genome-wide association study identifies a susceptibility locus for schizophrenia in Han Chinese at 11p11.2. Nat. Genet. 43, 1228–1231 (2011).
7. Wang, S.-H. et al. Polygenic risk for schizophrenia and neurocognitive performance in patients with schizophrenia. Genes Brain Behav. (2017). doi:10.1111/gbb.12401
8. Yang, S. Y. et al. Association between ST8SIA2 and the Risk of Schizophrenia and Bipolar I Disorder across Diagnostic Boundaries. PLoS One 10, e0139413 (2015).

Ethics oversight

Ethics Oversight
The following institutions provided ethics oversight for data reported in the current manuscript:
Domain Specific Review Board (Singapore); University of Hong Kong; Fujita Health University; RIKEN Center for Integrative Medical Sciences; Nagoya University; Osaka University; Niigata University; Bio-X Institutes of Shanghai Jiao Tong University; University Medical Center Utrecht; The University of Western Australia; The University of Indonesia; Peking University Sixth Hospital; National Taiwan University; Samsung Medical Center; Chonnam National University Hospital.

Note that full information on the approval of the study protocol must also be provided in the manuscript.