nature research

Corresponding author(s): C.A. Rietveld

Last updated by author(s): Jul 26, 2021

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 our Editorial Policies 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	Participants of this study were sourced from UK Biobank, which has received ethical approval from the National Health Service North West Centre for Research Ethics Committee. UK Biobank is a prospective cohort study in the UK that collects physical, health and cognitive measures, and biological samples (including genotype data) in about 500,000 individuals.
Data analysis	MGREML is available on https://github.com/devlaming/mgreml as a ready-to-use command-line tool. The GitHub page comes with a full tutorial on the usage of this tool.

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 and 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

Individual-level genotype and phenotype data are available by application from the UKB Biobank (https://www.ukbiobank.ac.uk/). Main empirical results are available in the Supplementary Tables.

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	*Describe how sample size was determined, detailing any statistical methods used to predetermine sample size OR if no sample-size calculation was performed, describe how sample sizes were chosen and provide a rationale for why these sample sizes are sufficient.*
Data exclusions	*Describe any data exclusions. If no data were excluded from the analyses, state so OR if data were excluded, describe the exclusions and the rationale behind them, indicating whether exclusion criteria were pre-established.*
Replication	*Describe the measures taken to verify the reproducibility of the experimental findings. If all attempts at replication were successful, confirm this OR if there are any findings that were not replicated or cannot be reproduced, note this and describe why.*
Randomization	*Describe how samples/organisms/participants were allocated into experimental groups. If allocation was not random, describe how covariates were controlled OR if this is not relevant to your study, explain why.*
Blinding	*Describe whether the investigators were blinded to group allocation during data collection and/or analysis. If blinding was not possible, describe why OR explain why blinding was not relevant to your study.*

Behavioural & social sciences study design

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

Study description	UK biobank (UKB) is a prospective cohort study.
Research sample	UK Biobank is a prospective cohort study in the UK that collects physical, health and cognitive measures, and biological samples (including genotype data) in about 500,000 individuals.
Sampling strategy	We used all participants with complete data.
Data collection	Participants were invited to visit and complete a touchscreen questionnaire on several lifestyle questions and to undergo some medical examinations. Their records are linked to NHS hospital inpatient records. The MRI measures of brain structure and function, volumes of grey matter and the mapping of major brain connections are collected in a separate assessment.
Timing	Initial enrollment for UKB started in 2006 and took 4 years. The Imaging study started in 2014 and is still expanding.
Data exclusions	We selected 43,691 (all available at that time) individuals with available genotype data from the UK Biobank brain imaging study who self-identified as 'white British' and with similar genetic ancestry based on a principal component analysis. We dropped individuals that were too related (37,392 individuals after removing too closely related individuals). From these unrelated individuals, we retained the 20,190 individuals with complete information on all 86 traits (and all covariates) in our analyses.
Non-participation	Not applicable, since at the time of the study the recruitment is still going on for the imaging study.
Randomization	Not applicable.

Ecological, evolutionary & environmental sciences study design

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

Study description	*Briefly describe the study. For quantitative data include treatment factors and interactions, design structure (e.g. factorial, nested, hierarchical), nature and number of experimental units and replicates.*
Research sample	*Describe the research sample (e.g. a group of tagged Passer domesticus, all Stenocereus thurberi within Organ Pipe Cactus National Monument), and provide a rationale for the sample choice. When relevant, describe the organism taxa, source, sex, age range and any manipulations. State what population the sample is meant to represent when applicable. For studies involving existing datasets, describe the data and its source.*

Sampling strategy	*Note the sampling procedure. Describe the statistical methods that were used to predetermine sample size OR if no sample-size calculation was performed, describe how sample sizes were chosen and provide a rationale for why these sample sizes are sufficient.*
Data collection	*Describe the data collection procedure, including who recorded the data and how.*
Timing and spatial scale	*Indicate the start and stop dates of data collection, noting the frequency and periodicity of sampling and providing a rationale for these choices. If there is a gap between collection periods, state the dates for each sample cohort. Specify the spatial scale from which the data are taken*
Data exclusions	*If no data were excluded from the analyses, state so OR if data were excluded, describe the exclusions and the rationale behind them, indicating whether exclusion criteria were pre-established.*
Reproducibility	*Describe the measures taken to verify the reproducibility of experimental findings. For each experiment, note whether any attempts to repeat the experiment failed OR state that all attempts to repeat the experiment were successful.*
Randomization	*Describe how samples/organisms/participants were allocated into groups. If allocation was not random, describe how covariates were controlled. If this is not relevant to your study, explain why.*
Blinding	*Describe the extent of blinding used during data acquisition and analysis. If blinding was not possible, describe why OR explain why blinding was not relevant to your study.*

Did the study involve field work? ☐ Yes ☐ No

Field work, collection and transport

Field conditions	*Describe the study conditions for field work, providing relevant parameters (e.g. temperature, rainfall).*
Location	*State the location of the sampling or experiment, providing relevant parameters (e.g. latitude and longitude, elevation, water depth).*
Access & import/export	*Describe the efforts you have made to access habitats and to collect and import/export your samples in a responsible manner and in compliance with local, national and international laws, noting any permits that were obtained (give the name of the issuing authority, the date of issue, and any identifying information).*
Disturbance	*Describe any disturbance caused by the study and how it was minimized.*

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 and archaeology
☒	☐ Animals and other organisms
☐	☒ Human research participants
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

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

Antibodies

Antibodies used	*Describe all antibodies used in the study; as applicable, provide supplier name, catalog number, clone name, and lot number.*
Validation	*Describe the validation of each primary antibody for the species and application, noting any validation statements on the manufacturer's website, relevant citations, antibody profiles in online databases, or data provided in the manuscript.*

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)	*State the source of each cell line used.*
Authentication	*Describe the authentication procedures for each cell line used OR declare that none of the cell lines used were authenticated.*

Mycoplasma contamination	*Confirm that all cell lines tested negative for mycoplasma contamination OR describe the results of the testing for mycoplasma contamination OR declare that the cell lines were not tested for mycoplasma contamination.*
Commonly misidentified lines (See ICLAC register)	*Name any commonly misidentified cell lines used in the study and provide a rationale for their use.*

Palaeontology and Archaeology

Specimen provenance	*Provide provenance information for specimens and describe permits that were obtained for the work (including the name of the issuing authority, the date of issue, and any identifying information).*
Specimen deposition	*Indicate where the specimens have been deposited to permit free access by other researchers.*
Dating methods	*If new dates are provided, describe how they were obtained (e.g. collection, storage, sample pretreatment and measurement), where they were obtained (i.e. lab name), the calibration program and the protocol for quality assurance OR state that no new dates are provided.*

☐ Tick this box to confirm that the raw and calibrated dates are available in the paper or in Supplementary Information.

Ethics oversight	*Identify the organization(s) that approved or provided guidance on the study protocol, OR state that no ethical approval or guidance was required and explain why not.*

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

Animals and other organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

Laboratory animals	*For laboratory animals, report species, strain, sex and age OR state that the study did not involve laboratory animals.*
Wild animals	*Provide details on animals observed in or captured in the field; report species, sex and age where possible. Describe how animals were caught and transported and what happened to captive animals after the study (if killed, explain why and describe method; if released, say where and when) OR state that the study did not involve wild animals.*
Field-collected samples	*For laboratory work with field-collected samples, describe all relevant parameters such as housing, maintenance, temperature, photoperiod and end-of-experiment protocol OR state that the study did not involve samples collected from the field.*
Ethics oversight	*Identify the organization(s) that approved or provided guidance on the study protocol, OR state that no ethical approval or guidance was required and explain why not.*

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

Human research participants

Policy information about studies involving human research participants

Population characteristics	UK Biobank.
Recruitment	Prospective participants were invited to visit an assessment centre, at which they completed an automated questionnaire and were interviewed about lifestyle, medical history and nutritional habits; basic variables such weight, height, blood pressure etc. were also measured; and blood and urine samples were taken. These samples were preserved so that it was possible to later extract DNA and measure other biologically important substances. During the whole duration of the study it was intended that all disease events, drug prescriptions and deaths of the participants are recorded in a database, taking advantage of the centralized UK National Health Service.
Ethics oversight	UK Biobank Ethics and Governance Council

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

Clinical data

Policy information about clinical studies

All manuscripts should comply with the ICMJEguidelines for publication of clinical research and a completedCONSORT checklist must be included with all submissions.

Clinical trial registration	*Provide the trial registration number from ClinicalTrials.gov or an equivalent agency.*
Study protocol	*Note where the full trial protocol can be accessed OR if not available, explain why.*
Data collection	*Describe the settings and locales of data collection, noting the time periods of recruitment and data collection.*
Outcomes	*Describe how you pre-defined primary and secondary outcome measures and how you assessed these measures.*

Dual use research of concern

Policy information about dual use research of concern

Hazards

Could the accidental, deliberate or reckless misuse of agents or technologies generated in the work, or the application of information presented in the manuscript, pose a threat to:

No	Yes	
☐	☐	Public health
☐	☐	National security
☐	☐	Crops and/or livestock
☐	☐	Ecosystems
☐	☐	Any other significant area

Experiments of concern

Does the work involve any of these experiments of concern:

No	Yes	
☐	☐	Demonstrate how to render a vaccine ineffective
☐	☐	Confer resistance to therapeutically useful antibiotics or antiviral agents
☐	☐	Enhance the virulence of a pathogen or render a nonpathogen virulent
☐	☐	Increase transmissibility of a pathogen
☐	☐	Alter the host range of a pathogen
☐	☐	Enable evasion of diagnostic/detection modalities
☐	☐	Enable the weaponization of a biological agent or toxin
☐	☐	Any other potentially harmful combination of experiments and agents

ChIP-seq

Data deposition

☐ Confirm that both raw and final processed data have been deposited in a public database such as GEO.

☐ Confirm that you have deposited or provided access to graph files (e.g. BED files) for the called peaks.

Data access links *May remain private before publication.*	*For "Initial submission" or "Revised version" documents, provide reviewer access links. For your "Final submission" document, provide a link to the deposited data.*
Files in database submission	*Provide a list of all files available in the database submission.*
Genome browser session (e.g. UCSC)	*Provide a link to an anonymized genome browser session for "Initial submission" and "Revised version" documents only, to enable peer review. Write "no longer applicable" for "Final submission" documents.*

Methodology

Replicates	*Describe the experimental replicates, specifying number, type and replicate agreement.*
Sequencing depth	*Describe the sequencing depth for each experiment, providing the total number of reads, uniquely mapped reads, length of reads and whether they were paired- or single-end.*
Antibodies	*Describe the antibodies used for the ChIP-seq experiments; as applicable, provide supplier name, catalog number, clone name, and lot number.*
Peak calling parameters	*Specify the command line program and parameters used for read mapping and peak calling, including the ChIP, control and index files used.*
Data quality	*Describe the methods used to ensure data quality in full detail, including how many peaks are at FDR 5% and above 5-fold enrichment.*
Software	*Describe the software used to collect and analyze the ChIP-seq data. For custom code that has been deposited into a community repository, provide accession details.*

Flow Cytometry

Plots

Confirm that:

☐ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).

☐ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).

☐ All plots are contour plots with outliers or pseudocolor plots.

☐ A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation	*Describe the sample preparation, detailing the biological source of the cells and any tissue processing steps used.*
Instrument	*Identify the instrument used for data collection, specifying make and model number.*
Software	*Describe the software used to collect and analyze the flow cytometry data. For custom code that has been deposited into a community repository, provide accession details.*
Cell population abundance	*Describe the abundance of the relevant cell populations within post-sort fractions, providing details on the purity of the samples and how it was determined.*
Gating strategy	*Describe the gating strategy used for all relevant experiments, specifying the preliminary FSC/SSC gates of the starting cell population, indicating where boundaries between "positive" and "negative" staining cell populations are defined.*

☐ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.

Magnetic resonance imaging

Experimental design

Design type	The protocol includes three structural MRI scans; T1, T2 fluid attenuation inversion recovery (FLAIR) and susceptibility-weighted MRI (swMRI), as well as diffusion MRI (dMRI) and resting and task functional MRI (fMRI).
Design specifications	The brain MRI protocol is performed using a 3 Tesla Siemens Skyra scanner (Siemens Healthineers, Erlangen, Germany) with VD13 software and a 32-channel head coil. The full examination lasts approximately 35 min.
Behavioral performance measures	Most behavioral measures were direclty available in the UK Biobank data. To measure intelligence, however, we combined several behavioral measures. In the inital UK Biobank assesment, a visual memory test was administered, labelled 'pairs-matching' (http://biobank.ctsu.ox.ac.uk/crystal/label.cgi?id=100030), where participants were asked to memorize the positions of six card pairs, and then match them from memory while making as few errors as possible. Scores on the pairs-matching test are for the number of errors that each participant made; therefore, higher scores reflect poorer cognitive function. Participants completed a timed test of symbol matching, similar to the common card game 'Snap' hereafter referred to as reaction time (RT) (http://biobank.ctsu.ox.ac.uk/crystal/field.cgi?id=20023). The score on this task was the mean response time in milliseconds across trials which contained matching pairs. The fluid intelligence score (UK Biobank field ID: 20016, 20191) http://biobank.ctsu.ox.ac.uk/crystal/refer.cgi?id=100231) is used to measure intelligence. Since there is a learning effect, we used a weighted average that corrects for the number of waves individuals participated in.

Acquisition

Imaging type(s)	Structural, diffusion, resting and task.
Field strength	3
Sequence & imaging parameters	Modality Duration (mins) Resolution (mm3) Matrix Other parameters T1 MPRAGE 4:54 1.0x1.0x1.0 256x256x208 TI/TR=880/2000 ms, R=2 Resting fMRI 6:10 2.4x2.4x2.4 88x88x64 TE/TR=39/735 ms, α=51°, MB=8 T2 FLAIR 5:52 1.0x1.0x1.05 256x256x192 TI/TR=1800/5000 ms, R=2 Diffusion MRI1 7:08 2.0x2.0x2.0 104x104x72 TR=3600 ms, 50 directions/shell, b=0, 1000, 2000 s/mm2, α=51°, MB=3 Susceptibility-weighted 2:34 0.8x0.8x3.0 288x256x48 TE1/TE2/TR=9.4/20/27 ms, R=2 Task fMRI 4:13 2.4x2.4x2.4 88x88x64 TE/TR=39/735 ms, α=51°, MB=8 For more information, see https://www.nature.com/articles/s41467-020-15948-9#Tab2
Area of acquisition	Whole brain

Diffusion MRI ☒ Used ☐ Not used

Preprocessing

Preprocessing software: An automated processing pipeline for brain image analysis and quality control was established for UKB at the University of Oxford's Wellcome Centre for Integrative Neuroimaging (WIN/FMRIB). This pipeline is primarily based around FSL (FMRIB's Software Library), and other packages such as FreeSurfer. When acquired at the imaging centres, the images are reconstructed from k-space on the scanner computer and saved initially as DICOM files. The processing pipeline then converts these files to the NIFTI format and undertakes pre-processing (e.g., correcting for head motion and other artefacts) as well as automated quality control that identifies issue with the equipment (e.g., coil failure) and artefacts specific to the participant or scanning session (e.g., excessive head movement). SeeMiller, K. L. et al. Multimodal population brain imaging in the UK Biobank prospective epidemiological study. Nat. Neurosci. 19, 1523–1536 (2016). and Alfaro-Almagro, F. et al. Image processing and quality control for the first 10,000 brain imaging datasets from UK Biobank. Neuroimage 166, 400–424 (2017).

Normalization: See above

Normalization template: See above

Noise and artifact removal: See above

Volume censoring: See above

Statistical modeling & inference

Model type and settings: We estimate a (multivariate) linear mixed model using restricted maximum likelihood estimation. See Methods/ Supplementary Information (SI) for details.

Effect(s) tested: We estimate heritability and genetic correlations between brain regions of interest (ROIs) and behavioural outcomes.

Specify type of analysis: ☐ Whole brain ☐ ROI-based ☒ Both

Statistic type for inference (See Eklund et al. 2016): We estimate a (multivariate) linear mixed model using restricted maximum likelihood estimation. See online methods/SI for details.

Correction: We use a significance level of 5%, as the results presented are descriptive (we don't test particular hypotheses). Using the empirical as results presented in the Supplementary Tables, one could easily perform Bonferonni corrections.

Models & analysis

n/a	Involved in the study
☒	☐ Functional and/or effective connectivity
☒	☐ Graph analysis
☒	☐ Multivariate modeling or predictive analysis

Functional and/or effective connectivity: *Report the measures of dependence used and the model details (e.g. Pearson correlation, partial correlation, mutual information).*

Graph analysis: *Report the dependent variable and connectivity measure, specifying weighted graph or binarized graph, subject- or group-level, and the global and/or node summaries used (e.g. clustering coefficient, efficiency, etc.).*

Multivariate modeling and predictive analysis: *Specify independent variables, features extraction and dimension reduction, model, training and evaluation metrics.*