nature research

Corresponding author(s): Dajiang Liu
Scott Vrieze

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.

Statistical parameters

When statistical analyses are reported, confirm that the following items are present in the relevant location (e.g. 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
☐	☒	An indication of 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 statistics 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
☐	☒	Clearly defined error bars *State explicitly what error bars represent (e.g. SD, SE, CI)*

Our web collection on statistics for biologists may be useful.

Software and code

Policy information about availability of computer code

Data collection	No software was used.
Data analysis	All studies used either ShapeIt2, EAGLE or Finch to phase genotypes and used either Minimac3 or IMPUTE2V3 for imputation. Summary statistics were generated using RVTESTS release v1.9.7 or v1.9.9 or BOLT-LMM v2.2. Meta-analysis and conditional analysis was performed using rareGWAMA_0.4 in R. LD Score Regression v1.0.0 was used to measure heritability, test for population stratification and cryptic relatedness, estimate genetic correlations and enrichment analyses. RiVIERA-ridge was also used for enrichment analyses. LDpred v0.9.09 was used to construct the polygenic scores. PASCAL was used for gene based and pathway analysis and DEPICT was used to identify enrichment within tissues/cell types and reconstituted gene sets. Locuszoom plots were made using LocusZoom standalone software v1.3. GenomicSEM was used for the Genomic SEM analyses. MTAG software was used for the MTAG analysis.

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 upon request. 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

Upon acceptance, results excluding the 23andMe substudy will be available from the GSCAN Wiki page (https://genome.psych.umn.edu/), and posted on dbGaP. The 23andMe substudy itself is available upon request to 23andMe.

Field-specific reporting

Please select 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/authors/policies/ReportingSummary-flat.pdf

Life sciences study design

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

Sample size	No sample size calculation was done but we tried to increase our sample size as much as possible. We contacted as many studies (with our phenotypes of interest) as possible and applied for relevant studies available in public repositories. Our meta-analysis includes the largest sample size of similar phenotypes to date and therefore, we believe our results are sufficiently powered.
Data exclusions	We excluded any non-European sample as population differences may lead to spurious results. We also excluded results for some phenotypes from smaller studies when those results were severely inflated or deflated per the genomic control, and there was no alternative explanation (e.g., inflation was due to polygenic signal). We applied filters to the genomic data post meta-analysis (minor allele frequency > .1%, effective sample size of at least 10% per phenotype and at least 3 studies must be included for each variant) in order to only report variants on which we had robust results.
Replication	Our results have replicated 26/27 previous known loci as detailed in the manuscript. In order to maximize power to detect the variants, we did not separate our sample into a separate discovery and replication set.
Randomization	N/A
Blinding	N/A

Reporting for specific materials, systems and methods

Materials & experimental systems

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

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	European ancestry with 52.2% female.
Recruitment	We did not do any recruitment. Analysis was of existing de-identified data.