nature research

Corresponding author(s): Danielle Posthuma

☐ Initial submission ☐ Revised version ☒ Final submission

Life Sciences Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form is intended for publication with all accepted life science papers and provides structure for consistency and transparency in reporting. Every life science submission will use this form; some list items might not apply to an individual manuscript, but all fields must be completed for clarity.

For further information on the points included in this form, see Reporting Life Sciences Research. For further information on Nature Research policies, including our data availability policy, see Authors & Referees and the Editorial Policy Checklist.

Experimental design

1. Sample size

Describe how sample size was determined.

Data used in this study was collected by external sources and used for secondary analysis. Sample size was not pre-determined and was chosen based on all known available cohorts with relevant data collected to date, after quality control steps were performed in each cohort (described in detail in Online Methods section "Study Cohorts" and Supplementary Information section 1.1). Power calculations using the Genetic Power Calculator indicated that we had virtually 100% power to detect SNPs accounting for >0.05% of trait variance in our sample size of N=269867.

2. Data exclusions

Describe any data exclusions.

Each cohort in the meta-analysis applied their own quality control criteria (described in Supplementary Information section 1.1), with samples being excluded for study design purposes (i.e. based on age), criteria that would affect normal intelligence scores such as the presence of dementia, poor quality DNA sample, or if consent was withdrawn after enrollment. Genetic variants were excluded using standard quality control metrics. At the meta-analysis level, we applied a set of pre-determined quality control filters for the removal of SNPs based on minor allele frequency, imputation quality, mismatch from a known reference panel, etc. as described fully in the Online Methods (section "Meta-analysis").

3. Replication

Describe whether the experimental findings were reliably reproduced.

The meta-analysis strategy includes replication by default, as it weights the reported test statistics by the evidence of association across multiple samples. Further, SNP-based replication was carried out using a proxy phenotype of educational attainment in an independent sample. Aggregate genomic associations were replicated using polygenic score validation in four holdout samples.

4. Randomization

Describe how samples/organisms/participants were allocated into experimental groups.

Not relevant; no experimental procedures were performed as this was a study of genetic association between genotypes and non-manipulated phenotypes.

5. Blinding

Describe whether the investigators were blinded to group allocation during data collection and/or analysis.

Not relevant; there were no experimental groups.

Note: all studies involving animals and/or human research participants must disclose whether blinding and randomization were used.

6. Statistical parameters

For all figures and tables that use statistical methods, confirm that the following items are present in relevant figure legends (or in the Methods section if additional space is needed).

n/a	Confirmed
☐	☒ The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement (animals, litters, cultures, etc.)
☐	☒ A description of how samples were collected, noting whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☒	☐ A statement indicating how many times each experiment was replicated
☐	☒ The statistical test(s) used and whether they are one- or two-sided (note: only common tests should be described solely by name; more complex techniques should be described in the Methods section)
☐	☒ A description of any assumptions or corrections, such as an adjustment for multiple comparisons
☐	☒ The test results (e.g. *P* values) given as exact values whenever possible and with confidence intervals noted
☐	☒ A clear description of statistics including central tendency (e.g. median, mean) and variation (e.g. standard deviation, interquartile range)
☐	☒ Clearly defined error bars

See the web collection on statistics for biologists for further resources and guidance.

Software

Policy information about availability of computer code

7. Software

Describe the software used to analyze the data in this study.

We used standard, publicly available statistical genetics software packages, which are described and linked to in the Online Methods. The packages we used included:

R - Data management and statistical analyses (R Core Team, 2016)
SHAPEIT2/IMPUTE2 - Genotype imputation (Howie et al., 2009)
PLINK - Genetic association testing (Purcell et al., 2007; Chang et al., 2015)
SNPTEST - Genetic association testing (Marchini et al., 2007)
RAREMETALWORKER - Genetic association testing in related individuals (Liu et al., 2014)
METAL - GWAS meta-analysis (Willer et al., 2010)
FUMA - Online platform for functional annotation of GWAS results (Watanabe et al., 2017)
MAGMA - Gene-based association testing (de Leeuw et al., 2015)
LD score regression - SNP-based heritability and genetic correlations from GWAS summary statistics (Bulik-Sullivan et al., 2015)
PRSice - Polygenic score analysis (Eusden et al., 2015)
LDpred - Polygenic score analysis (Vilhjalmsson et al., 2015)

For manuscripts utilizing custom algorithms or software that are central to the paper but not yet described in the published literature, software must be made available to editors and reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). *Nature Methods* guidance for providing algorithms and software for publication provides further information on this topic.

Materials and reagents

Policy information about availability of materials

8. Materials availability

Indicate whether there are restrictions on availability of unique materials or if these materials are only available for distribution by a for-profit company.

Summary statistics from the GWAS meta-analysis will be made freely available for download at https://ctg.cncr.nl/.

9. Antibodies

Describe the antibodies used and how they were validated for use in the system under study (i.e. assay and species).

No antibodies were used in this study.

10. Eukaryotic cell lines

a. State the source of each eukaryotic cell line used.	No eukaryotic cell lines were used in this study.
b. Describe the method of cell line authentication used.	No eukaryotic cell lines were used in this study.
c. Report whether the cell lines were tested for mycoplasma contamination.	No eukaryotic cell lines were used in this study.
d. If any of the cell lines used are listed in the database of commonly misidentified cell lines maintained by ICLAC, provide a scientific rationale for their use.	No eukaryotic cell lines were used in this study.

Animals and human research participants

Policy information about studies involving animals; when reporting animal research, follow the ARRIVE guidelines

11. Description of research animals

Provide details on animals and/or animal-derived materials used in the study.	No animals were used in this study.

Policy information about studies involving human research participants

12. Description of human research participants

Describe the covariate-relevant population characteristics of the human research participants.	Human research participants of all ages (approximately 6 too 100 years old) were included, with approximately equal proportions of males and females. Participants came from 14 independent cohorts previously collected by external sources, which were defined by different criteria based on the original study goals but were generally population-based cohorts of healthy individuals. Participants were not included/excluded for specific criteria except the presence of Alzheimer's disease, dementia, or other cognitive impairments. Population characteristics of the sample are described extensively in the Supplementary Information section 1.1 and summarized in Supplementary Table 1.