Corresponding Author: Daniel J. Benjamin

Date: March 14, 2018

Life Sciences Reporting Summary

Nature Research wishes to improve the reproducibility of the work we publish. This form is published with all life science papers and is intended to promote consistency and transparency in reporting. All life sciences submissions use this form; while some list items might not apply to an individual manuscript, 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.

   For each of the phenotypes, we combined all publicly available summary statistics with summary statistics from new association analyses. For the GWAS of EA, we had a combined sample size of 1,131,881 individuals. For the supplementary GWASs of CP, Math Ability, and Highest Math, we had sample sizes of 257,841, 430,445, and 564,698, respectively. For EA and CP, the sample sizes are larger than previously published papers that have identified replicable loci. For the two math-related phenotypes, they each have a heritability greater than that of EA (Math Ability: 0.156, Hardest Math: 0.165, EA: 0.122) and have sample sizes greater than previous GWASs of EA. This suggests that we will be well powered to find replicable loci for each of the four phenotypes considered.

2. Data exclusions

   Describe any data exclusions.

   No data were excluded from the analysis, except for standard quality-control filters applied to the SNP data. The main filtering steps involved dropping SNPs that: (i) are known to have strand issues in some imputation programs, (ii) have missing or incorrect numerical values supplied for some variables (e.g., a P value of association outside the range 0 to 1), (iii) have a minor allele count below 25, (iv) have poor imputation accuracy, (v) are indels or not located on the autosomes, or (vi) have invalid or duplicated chromosomal coordinates or whose alleles do not match those in the reference file. In association results from analyses of the full release of the UK Biobank data, we further filter out all SNPs that are not in the Haplotype Reference Consortium's reference panel.

3. Replication

   Describe whether the experimental findings were reliably reproduced.

   We test the lead SNPs from a previous GWAS of educational attainment, Okbay et al. (2016) (Supplementary Note section 1.10). We also replicate the lead SNPs for the MTAG of cognitive performance (Supplementary Note section 1.14). Finally, we test whether polygenic predictors based on both the GWAS and MTAG results have predictive power in a out-of-sample prediction cohort consistent with replication (Supplementary Note section 6.4, 6.5, and 6.6). In all cases, the replication record is strong.

4. Randomization

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

   Not relevant because the study is not experimental.

5. Blinding

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

   Not relevant because the study is not experimental.

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 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 summary of the descriptive 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.

The software used for imputation by each cohort included Minimac2, BEAGLE v2.1.2, IMPUTE2 v2.3.1, PBWT, and IMPUTE4 ShapeIT v2.r790. Associations analyses software used by each cohort included BOLT-LMM, SNPTEST v2.4.1, and REGSCAN v0.2.0. Meta-analyses were performed with Metal, release 2011-03-25. QC was run with EasyQC v9.0. LD score regressions were done using ldsc v1.0.0. Clumping was perfomed with Plink, 1.90b3p. Polygenic score weights were generated using LDpred v0.9.09, and the prediction analyses were executed in Stata v14.2. Biological annotation was completed using DEPICT (downloaded Feb 2015), MAGMA v1.06b, PANTHER release 20170403, and CAVIARBF v0.2.1. MTAG analyses were conducted using the MTAG software v1.0.1.

For all studies, we encourage code deposition in a community repository (e.g. GitHub). Authors must make computer code available to editors and reviewers upon request. The *Nature Methods* guidance for providing algorithms and software for publication may be useful for any submission.

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.

No unique materials were used.

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.

10. Eukaryotic cell lines

a. State the source of each eukaryotic cell line used.

No eukaryotic cell lines were used.

b. Describe the method of cell line authentication used.

No eukaryotic cell lines were used.

c. Report whether the cell lines were tested for mycoplasma contamination.

No eukaryotic cell lines were used.

d. If any of the cell lines used in the paper are listed in the database of commonly misidentified cell lines maintained by ICLAC, provide a scientific rationale for their use.

No cell lines were used.

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.

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.	Analyses were conducted on GWAS summary statistics. Of the individuals included in the current meta-analysis and not already included in Okbay et al.'s (2016) GWAS of EA, 54% are female. The mean birth year is 1955 with a range from 1901 to 1989.