nature research

Corresponding Author: Daniel J. Benjamin

Date: Oct 23, 2017

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 three phenotypes, we combined all publicly available summary statistics with summary statistics from new association analyses. Details are reported in Section 3.2 of the Supplementary Note.

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, described in Supplementary Note sections 3.2 and 3.3 and Supplementary Table 10).

3. Replication

   Describe whether the experimental findings were reliably reproduced.

   We test the MTAG-identified lead SNPs jointly for replication. Their replication record is strong; see Figure 5 and Supplementary Note section 5.

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 GWASs in the UKB and replication cohorts were done with SNPtest v2.5.2. Meta-analyses were performed with Metal, release 2011-03-25. QC was run with EasyQC v9.0. Simulated results were generated and replication analyses were conducted using Python v2.7. 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). The comparative analyses estimates for Shom and Shet were calculated using the R package CPASSOC v1.01. MTAG analyses were conducted in Python v2.7.

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. References to the studies that report covariate-relevant population characteristics are in Supplementary Table 2 and Supplementary Note section 3.2.