nature research

Corresponding author(s): Abraham A Palmer

☒ 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.

Sample size was set at 25,000 because this was larger than comparable well-powered studies of similar phenotypes. Using Quanto (http://biostats.usc.edy/Quanto.html) we estimated that with a sample size of 25,000, allele frequency of 0.05, mean trait value of -2.25 (SD= 0.73) and effect size (beta) of -0.1 to 0.1, we had 89% power to detect significant (5 x 10-8) associations.

2. Data exclusions

Describe any data exclusions.

Exclusion criteria were pre established based on genotype data quality, and phenotypic accuracy, as described in the Online Methods. In addition, we added three items to the 27-items that make up the MCQ that were intended to detect careless responding (items 8, 17 and 24 in Supplementary Table 3). Research participants who answered any of these items inappropriately were excluded. If subjects responded randomly, 87.5% would have answered one of these three items inappropriately, whereas the actual rate of inappropriate responses was 2.10%. A total of 550 research participants were excluded from our sample due to low consistency (< 80%) or inappropriate responses. Only unrelated participants of European ancestry were included in the study.

3. Replication

Describe whether the experimental findings were reliably reproduced.

The top GWAS result was replicated in an independent sample of 928 individuals. The replication showed that the direction of effect was the same as in our original study. A meta-analysis using the independent dataset was more significant than our original result (aww Online Methods).

4. Randomization

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

Participants were not allocated into groups. We controlled for sex, age, genotype and batch effects.

5. Blinding

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

Blinding was not relevant to our study; all participants were exposed to the same experimental battery. Because of the amount of data and the automatic data collection and analysis pipeline, blinding is not an meaningful consideration.

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.	A custom analysis pipeline that was developed by 23andMe was used for the GWAS analysis, and EPACTS was used for the Genes for Good dataset; phasing was performed using Finch; imputation was performed using Minimac2; imputed HLA allele dosages were performed using HIBAG; python packages were used for LDSC genetic correlations and MetaXcan analyses; meta-analysis was conducted using METAL. Full description of the materials, and citations including software version, are described in the Online Methods.

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.	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 are listed in the database of commonly misidentified cell lines maintained by ICLAC, provide a scientific rationale for their use.	No commonly misidentified 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.

All participants were drawn from the customer base of 23andMe, Inc., a consumer genetics company. Participants provided informed consent and participated in the research online, under a protocol approved by the external AAHRPP-accredited IRB, Ethical & Independent Review Services (www.eandireview.com). Over the course of approximately four months in 2015, more than 25,000 23andMe research participants responded to survey questions that were part of a study of the genetics of decision-making designed by Drs. James MacKillop and Abraham A Palmer. In this paper we examine only the 23,127 participants who were of European ancestry (>97% as determined through an analysis of local ancestry), for whom DD data were available, and who were not excluded for any of the reasons described elsewhere in this paper. Participants mean age was 53.79 (SD=16.08), 55.3% were females; we included age and sex as covariates in our GWAS analysis. Socio-demographic details of this cohort are described in the Supplementary Table 2.