Corresponding author(s): Jonathan P Beauchamp
jonathan.pierre.beauchamp@gmail.com

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.

Please do not complete any field with "not applicable" or n/a. Refer to the help text for what text to use if an item is not relevant to your study.
For final submission: please carefully check your responses for accuracy; you will not be able to make changes later.

Experimental design

1. Sample size

Describe how sample size was determined.

The appropriate sample size was determined based on power calculations reported in the preregistered analysis plan on Open Science Framework (https://osf.io/cjx9m/). That information is also reported in Supplementary Note section 2:
"The analysis plan included power calculations assuming that 100,000 individuals in the UKB answered “Yes” (“cases”) to the general-risk-tolerance question, and 270,000 individuals answered “No” (“controls”). Under this assumption, our study would have 73% power to detect single nucleotide polymorphisms (SNPs) with a minor allele frequency (MAF) of 0.3 and an odds ratio of 1.05 with a genome-wide significance threshold of P = 5×10–8."
Ultimately, for the general risk tolerance phenotype, we combined data from the UKB, from a sample of research participants from 23andMe, and from 10 smaller replication cohorts.

2. Data exclusions

Describe any data exclusions.

Exclusion criteria were specified in the preregistered analysis plan on Open Science Framework. That information is also reported in Supplementary Note section 2:
"The analysis plan instructed all cohorts to limit the analysis to individuals of European ancestry, to exclude individuals with missing covariates, to remove samples that displayed a SNP call rate of less than 95%, and to apply cohort-specific standard quality control filters before imputation."

3. Replication

Describe the measures taken to verify the reproducibility of the experimental findings.

Supplementary Information section 5 reports holistic, out-of-sample replication in ~35,000 individuals. Also, we test whether polygenic scores have predictive power in multiple independent prediction cohorts (Supplementary Information section 10). 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 in vivo studies must report how sample size was determined and 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 *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒ A description of any assumptions or corrections, such as an adjustment for multiple comparisons
☐	☒ Test values indicating whether an effect is present *Provide confidence intervals or give results of significance tests (e.g. P values) as exact values whenever appropriate and with effect sizes 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 in all relevant figure captions (with explicit mention of central tendency and variation)

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 URLs section in the main text contains links to software used in this study, and the Supplementary Note describes the software that were used for the various analyses. In summary:
Cohort-level imputation was performed with MaCH/Minimac, BEAGLE, IMPUTE2/4, PBWT, and ShapeIT.
Genetic data was managed with QCtool and BCFtools.
GWAS association was performed with PLINK, SNPtest, Mach2QTL, GCTA, BOLT-LMM, regscan, and R.
Meta-analyses were performed with METAL.
QC of GWAS summary statistics was performed with EasyQC and R.
LD score regressions were done using ldsc v1.0.0 and Python.
Clumping was performed with Plink, 1.90b3p.
Conditional association analysis was performed with GCTA.
Polygenic score weights were generated using LDpred v0.9.09.
Polygenic scores were calculated with PLINK.
SNP heritability was estimated with ldsc v1.0.0, HESS, and GCTA.
Biological annotation and gene-based analyses were conducted using SMR, DEPICT (downloaded Feb 2015), MAGMA v1.06b.
MTAG analyses were conducted using the MTAG software v1.0.1.
Auxiliary statistical analyses were performed with Python, STATA, and R.

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 third party.

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 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 all relevant details on animals and/or animal-derived materials used in the study. — No research 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. — Analyses were conducted on GWAS summary statistics. References to the studies that report covariate-relevant population characteristics are in Supplementary Tables 4-5.