

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our [Editorial Policies](#) and the [Editorial Policy Checklist](#).

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a Confirmed

- ☐ ☒ The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
- ☐ ☒ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- ☐ ☒ 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 all covariates tested
- ☐ ☒ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- ☐ ☒ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- ☐ ☒ For null hypothesis testing, the test statistic (e.g. F , t , r) with confidence intervals, effect sizes, degrees of freedom and P value noted
Give P values as exact values whenever suitable.
- ☐ ☒ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- ☐ ☒ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- ☐ ☒ Estimates of effect sizes (e.g. Cohen's d , Pearson's r), indicating how they were calculated

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

Software and code

Policy information about [availability of computer code](#)

Data collection

No software was used for data collection.

Data analysis

Analyses were conducted using the Howe and colleagues (2022) within-sibship analytic pipeline, the programs fastGWA (Jiang et al., 2019), PLINK version 1.9 (Purcell et al., 2007), PLINK2 (Chang et al., 2015), GCTA version 1.94 (Yang et al., 2011), BOLT-LMM version 2.4.1 (Loh et al., 2015), FUMA (Watanabe et al., 2017), POPCORN (Brown et al., 2016), MAGMA (de Leeuw et al., 2015), PRS-CSx (Ruan et al., 2022), SBayesR (Lloyd-Jones et al., 2019), SNIIPAR (Young et al., 2022), SLD4M (O'Connor et al., 2019), and R version 4.3.1 (R Core Team, 2023) using R data-management packages dplyr (Wickham et al., 2019), data.table (Dowle et al., 2019), and psych (Revelle, 2017), R visualization packages corplot (Wei et al., 2017), ggplot2 (Wickham et al., 2019), and qqman (Turner, 2014), and R analysis packages lavaan (Rosseel, 2012), lme4 (Bates, 2010), metafor (Viechtbauer, 2010), EasyQC (Winkler et al., 2014), GenomicSEM (Grotzinger et al., 2019), and ieugwasr and TwoSampleMR (Hemani et al., 2018).

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio [guidelines for submitting code & software](#) for further information.

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our [policy](#)

Summary statistics for GWAS of each Big Five trait, excluding 23andMe participants, are available at <https://doi.org/10.17605/OSF.IO/HGNSM>. This includes EUR, AFR, and trans-ancestry meta-analytic population-level summary statistics, polygenic index weights, and EUR within-family summary statistics, neuroticism summary statistics both including and excluding UK Biobank participants, and country-, age-, perspective- and questionnaire-stratified summary statistics. Other data used in this study cannot be made public and are available, within the constraints of the relevant regulations and data use agreements, upon request.

Code Availability Code for all analyses performed this study is available at <https://doi.org/10.17605/OSF.IO/HGNSM>

Research involving human participants, their data, or biological material

Policy information about studies with [human participants or human data](#). See also policy information about [sex, gender \(identity/presentation\), and sexual orientation](#) and [race, ethnicity and racism](#).

Reporting on sex and gender	In this study, chromosomal sex was coded for each participant and used as a control variable in polygenic index analyses, following field-standard procedures, as well as a stratifying variable in some X-chromosomal analyses, when noted.
Reporting on race, ethnicity, or other socially relevant groupings	Some analyses require the stratification of participants by recent genetic ancestry. To refer to ancestry groups while minimizing potential stigma, we follow guidance from National Academy of Sciences (https://nap.nationalacademies.org/read/26902/chapter/9) and use broad, overarching terms like EUR and AFR, or European-like and African-like genomes. Specifically, the term EUR/AFR refers to participants whose genomes are most similar to the participants in the 1000 Genomes European/African superpopulation sample (https://www.internationalgenome.org/).
Population characteristics	Demographic characteristics of each of the 46 cohorts, including sample size, age, and site of geographical ascertainment are described in full in the supplementary text and supplementary tables 1 and 2.
Recruitment	Recruitment was handled per-cohort, each using different procedures. Full information is contained in the supplementary text.
Ethics oversight	As this meta-analysis used de-identified summary-level data it was deemed exempt from ethics board approval by the University of Texas at Austin Institutional Review Board (STUDY00001941) ethics approval for individual cohorts that collected data and contributed to the meta-analysis is presented in the Supplementary Text. All participants consented to the research, and this research was performed in accordance with all relevant guidelines and regulations.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☐ Life sciences ☒ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Behavioural & social sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	This is a genome-wide association study using quantitative genomic and personality trait data
Research sample	The sample for this study is a meta-analysis of 46 different study cohorts who are members of the Revived Genomics of Personality Consortium. Sample size, demographic information, and sampling strategy differs by cohort. See supplementary text for complete information.
Sampling strategy	Sampling strategy varied by cohort; full information on sampling is provided in the cohort information at the end of the supplementary note. Most cohorts were not representative of their respective populations and were convenience samples or samples designed to collect data on twin sibling dyads. No pre-determined sample size was calculated; analyses make use of all available data to maximize power.
Data collection	Data collection varies by cohort; full information on data collection is provided in the cohort information at the end of the supplementary note. Each cohort collected a genomic sample and measured the Big Five personality traits using a valid questionnaire. There was no blinding nor randomization, as participants were not stratified into conditions.
Timing	No primary data was collected for this study. Start and stop of data collection varies by cohort. See supplementary text for complete information for each cohort
Data exclusions	Genotypes were aligned to build 37/hg19 on the forward strand, excluding SNPs with Minor Allele Frequency (MAF) < 1%, Call rate < 95%, those that deviated from Hardy-Weinberg Equilibrium ($p < 1 \times 10^{-6}$), and those with poor clustering on visual inspection of intensity plots. SNPs were excluded per analysis if they were below a minimum effective sample size threshold (e.g., 5,000). Participants were excluded prior to cohort-level analysis if they had low overall call rates (< 95%), excess autosomal heterozygosity or homozygosity, were duplicate samples, had sex inconsistent with gender, or had chromosomal abnormalities. As such exclusion rate cannot be calculated at the meta-analytic level.

Non-participation

Non-participation varies by cohort and was handled on a per-cohort basis.

Randomization

There was no randomization to condition in this study. Genome-wide analyses included controls for age, age2, birth cohort, sex, and genetic principal components, to address potential lifespan and birth cohort differences in personality traits. Within-family analyses included controls for genomes of family members to reduce confounding due to shared environment.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involvement in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Clinical data
☒	☐ Dual use research of concern
☒	☐ Plants

Methods

n/a	Involvement in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Plants

Seed stocks

Report on the source of all seed stocks or other plant material used. If applicable, state the seed stock centre and catalogue number. If plant specimens were collected from the field, describe the collection location, date and sampling procedures.

Novel plant genotypes

Describe the methods by which all novel plant genotypes were produced. This includes those generated by transgenic approaches, gene editing, chemical/radiation-based mutagenesis and hybridization. For transgenic lines, describe the transformation method, the number of independent lines analyzed and the generation upon which experiments were performed. For gene-edited lines, describe the editor used, the endogenous sequence targeted for editing, the targeting guide RNA sequence (if applicable) and how the editor was applied.

Authentication

Describe any authentication procedures for each seed stock used or novel genotype generated. Describe any experiments used to assess the effect of a mutation and, where applicable, how potential secondary effects (e.g. second site T-DNA insertions, mosaicism, off-target gene editing) were examined.