

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
<input type="checkbox"/>	<input checked="" type="checkbox"/> The exact sample size (<i>n</i>) for each experimental group/condition, given as a discrete number and unit of measurement
<input checked="" type="checkbox"/>	<input type="checkbox"/> A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
<input type="checkbox"/>	<input checked="" type="checkbox"/> The statistical test(s) used AND whether they are one- or two-sided <i>Only common tests should be described solely by name; describe more complex techniques in the Methods section.</i>
<input type="checkbox"/>	<input checked="" type="checkbox"/> A description of all covariates tested
<input type="checkbox"/>	<input checked="" type="checkbox"/> A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
<input type="checkbox"/>	<input checked="" type="checkbox"/> 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)
<input type="checkbox"/>	<input checked="" type="checkbox"/> For null hypothesis testing, the test statistic (e.g. <i>F</i> , <i>t</i> , <i>r</i>) with confidence intervals, effect sizes, degrees of freedom and <i>P</i> value noted <i>Give P values as exact values whenever suitable.</i>
<input checked="" type="checkbox"/>	<input type="checkbox"/> For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
<input checked="" type="checkbox"/>	<input type="checkbox"/> For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
<input type="checkbox"/>	<input checked="" type="checkbox"/> Estimates of effect sizes (e.g. Cohen's <i>d</i> , Pearson's <i>r</i>), 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 were used for data collection.
Data analysis	Quality control of genetic data and calculation of linkage disequilibrium (LD) score were performed using PLINK (v1.90b6.20 and v2.00a6LM). Genome-wide association analyses were performed REGENIE (v 4.0). Estimation of variance components associated with genetic relationship matrices were performed using the software MPH (v0.54.0). Estimation of heritability from close relatives was done using a custom R function available on GitHub (https://github.com/loic-yengo/REML-with-sparse-relationship-matrices) for quantitative traits and using the TetraHer module in LDAK 5.2. Firth's penalized logistic regression was implemented using the R package logistf (v1.26.1). Other statistical analyses and figure generation were performed with R (v4.1.0 and v4.2.1). LD scores calculation were mainly done using PLINK 2.0 except for those used in Supplementary Note 6 to estimate heritability for variants in the Telomere-to-telomere genome build. For the latter, we used a custom C++ code available on Zenodo at https://doi.org/10.5281/zenodo.16550864 . Fine-mapping analyses were performed using the SuSiE (https://stephenslab.github.io/susieR/index.html) implemented into the susieR package v0.14.2.

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](#)

Individual-level data of UK Biobank participants can be accessed upon application to the UK Biobank (<http://www.ukbiobank.ac.uk>). Results of fine-mapping analyses of GWAS loci identified in this study are available in Supplementary Data. Data and (R scripts) used to generate all the main text and extended data figures in the manuscript are available on Zenodo (<https://zenodo.org/records/17255323>). Due to the nature of the AGD dataset and commercial limitations, individual-level raw data are not available. Genotypes of participants in the 1000 Genomes Project were downloaded under the hg38 genome build (<https://www.cog-genomics.org/plink/2.0/resources>) and the T2T genome build (chm13v2.0) (https://s3-us-west-2.amazonaws.com/human-pangenomics/index.html?prefix=T2T/CHM13/assemblies/variants/1000_Genomes_Project/chm13v2.0/).

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	We analyzed data from all participants regardless of their sex or gender. Analyses of autosomal variants were adjusted for sex by using it as a covariate in regression models or by standardizing phenotypes within each sex-group prior to analyses.
Reporting on race, ethnicity, or other socially relevant groupings	Study participants were assigned European ancestry using principal components analyses of their genotypes. Principal components were used to determine genetic proximity to a reference sample of 504 individuals in the 1000 Genomes Project, whose ancestry group was broadly defined as European.
Population characteristics	We included as base covariates sex, year of birth, assessment centres, fasting time at blood sample collection, month of assessment and prescription drug usage. For the drug usage information, we extracted the field 20003 of the UK Biobank, mapped it to ATC codes and grouped in large categories (statins, diuretics, anti-hypertensive, beta-blockers, calcium blockers, angiotensin). Additionally, we also grouped individuals based on their north and east birth coordinates (UK Biobank fields 129 and 130) with a k-means clustering, with different number of clusters (10, 20, 50, 100). Individuals with missing birth location (typically, those born outside of the UK) were grouped into a separate cluster. All fasting times > 24h were merged into a single group. Similarly, missing data for assessment centres and month of assessment were grouped into distinct groups. We binarized each of these sets of covariates including each possible year of birth, dropped unused levels for each phenotype, and standardized each covariate to have a mean of 0 and a variance of 1. To reduce data dimensionality (and reduce collinearity), we applied a singular-value decomposition (SVD) on the covariate matrix from which we selected the top singular vectors associated with eigenvalues explaining in total >99% of the total variance.
Recruitment	UK Biobank investigators sent postal invitations to 9,238,453 individuals registered with the UK's National Health Service who were aged 40–69 years and lived within approximately 25 miles (40 km) of one of 22 assessment centers located throughout England, Wales, and Scotland. Overall, 503,317 participants consented to join the study cohort and visited an assessment center between 2006 and 2010, resulting in a participation rate of 5.45%. (Fry et al. Am J Epidemiol. 2017 Nov 1;186(9):1026-1034).
Ethics oversight	This research used data from participants in the UK Biobank study for discovery and from the Vanderbilt University's biorepository of DNA (BioVU) linked to de-identified medical records for replication of specific results. Written informed consent was obtained from every participant in UK Biobank study. The BioVU study was designed as an opt-out biobank. The UK Biobank study received ethics approval from the North West Centre for Research Ethics Committee (no. 11/NW/0382) and the BioVU study from the Vanderbilt Institutional Review Board.

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](https://www.nature.com/documents/nr-reporting-summary-flat.pdf)

Life sciences study design

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

Sample size	We analyze data from all European ancestry participants (N=452,618) available in the UK Biobank.
Data exclusions	We focused on European ancestry participants (to ensure the largest sample size) with available measures across 43 phenotypes. For

Data exclusions	quantitative traits, phenotypic values larger than 6 standard deviation away from the mean (in the sample) were discarded.
Replication	Replication of GWAS results for LDL was conducted in the Alliance for Genomic Discovery (AGD) dataset consisting of 191,454 European ancestry samples and 28,232 African ancestry samples.
Randomization	N/A - Rationale: No intervention was implemented on study participants. We used all available data whenever available.
Blinding	N/A - Rationale: No intervention was implemented on study participants.

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	Involved in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> Antibodies
<input checked="" type="checkbox"/>	<input type="checkbox"/> Eukaryotic cell lines
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology and archaeology
<input checked="" type="checkbox"/>	<input type="checkbox"/> Animals and other organisms
<input checked="" type="checkbox"/>	<input type="checkbox"/> Clinical data
<input checked="" type="checkbox"/>	<input type="checkbox"/> Dual use research of concern
<input checked="" type="checkbox"/>	<input type="checkbox"/> Plants

Methods

n/a	Involved in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> ChIP-seq
<input checked="" type="checkbox"/>	<input type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging

Plants

Seed stocks	N/A
Novel plant genotypes	N/A
Authentication	N/A