

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 (<i>n</i>) 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 <i>Only common tests should be described solely by name; describe more complex techniques in the Methods section.</i>
☐	☒ 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. <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>
☒	☐ 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 <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	Data collection was not performed by any software. All data underwent standard quality check procedures.
Data analysis	The software and resources used in this study are all publicly available: • MLN167 (v0.1.4): https://github.com/anbai106/mln1, DNEs for ASD1-3, LLD1-2, SCZ1-2 • Ukbnmf (v3.2.0): https://github.com/sritchie73/ukbnmf, Quality checks for the raw metabolites; • FUMA (v1.5.0): https://fuma.ctglab.nl/, Gene mapping, genomic locus annotation; • GCTA (v1.94.1): https://yanglab.westlake.edu.cn/software/gcta/#Overview, fastGWA; • LDSC (git version aa33296): https://github.com/bulik/ldsc, genetic correlation • TwoSampleMR (v0.5.6): https://mrcieu.github.io/TwoSampleMR/index.html, Mendelian randomization; • PRSs (release date: Aug 10, 2023): https://github.com/getian107/PRSs, PRS calculation; • Lifelines (v0.27.8): https://lifelines.readthedocs.io/en/latest/, Survival analysis; • GCTB (v2.0.5): https://cnsgenomics.com/software/gctb/; SBayesS.

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](#)

The GWAS summary statistics and pre-trained AI models from this study are publicly accessible via the MEDICINE Knowledge Portal (<https://labs-laboratory.com/medicine/>) and Synapse (<https://www.synapse.org/Synapse:syn64923248/wiki/63099266>). Our study used data generated by the human protein atlas (HPA: <https://www.proteinatlas.org>). GWAS summary data for the DEs were downloaded from the official websites of FinnGen (R9: https://www.finnngen.fi/en/access_results) and PGC (<https://pgc.unc.edu/for-researchers/download-results/>). Individual data, including plasma proteomics and metabolomics, from UKBB can be requested with proper registration at <https://www.ukbiobank.ac.uk/>. Certain sensitive data (e.g., allele frequency information) supporting the findings are also available from the author upon request. All original data for the main figures, along with the large-sized tables for Supplementary Data 1-5, are available on Zenodo (<https://zenodo.org/records/1512818567>).

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

Throughout the manuscript, we consistently employed the variable of sex, incorporating both self-reported sex and genetic sex as covariates and quality-check criteria. Additionally, we conducted sex-stratified analyses to investigate potential biases (sensitivity check) and explore sex-specific biological insights in males and females.

Reporting on race, ethnicity, or other socially relevant groupings

For GWAS, the primary analysis was on European ancestries

Population characteristics

Study population is detailed in Method 1 in the main manuscript and Supplementary eTable in supplement

Recruitment

All individual-level data were initially recruited by respective studies, underwent ethnic scrutinization. In addition, we performed standard quality check for all imaging, genetics, and proteomics to ensure the data quality.

Ethics oversight

The protocols of this study were approved by the Columbia University 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

No statistical methods were used to determine the final sample size used. All data from all participants were subjected to a quality check procedure to exclude unqualified data.

Data exclusions

All genetic, and metabolomics were quality-checked to exclude data that did not qualify for this study.

Replication

We performed several sensitivity check analysis to replicate the findings in the main manuscript. This include within-study by employing a cross-validation procedure if applicable.

Randomization

No randomization was performed to allocate participants into different groups.

Blinding

The Investigators were not blinded to allocation during experiments and outcome assessment.

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
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Clinical data
☒	☐ Dual use research of concern
☒	☐ Plants

Methods

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

Plants

Seed stocks

NA

Novel plant genotypes

NA

Authentication

NA