nature portfolio

Corresponding author(s): Carlos Cruchaga

Last updated by author(s): Apr 4, 2024

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	We used the SomaScan7k platform from SomaLogic for profiling the proteomics data; We used the Metabolon HD4 platform from Metabolon for profiling the metabolomics data; Genotype data were first measured by multiple genotype array types (TableS1) and then imputed with TOPMed Imputation Server.
Data analysis	QC steps for proteomic and metabolomic datasets were performed with R (v4.3.0); The specific R packages include: SomaDataIO (v6.0.0); Biobase (v2.62.0); QC steps for genotype data were performed with PLINK2 (v20220814). Imputation steps for genotype data were uploaded to the TOPMed Imputation Server. QTL analyses were performed using PLINK2 (v20220814). Annotation of the QTL findings were performed using VEP (v107). Visualization figures were generated with R packages: ggplot2 (v3.5.0); hudson (v1.0.0); karyoploteR (v1.26.0); circlize (v0.4.15). MASH (mashr, v0.2.71) method was used for the alternative ancestry-specificity analysis. Power analyses were using R package powerEQTL (V0.3.4). The custom scripts for performing TWAS with both cis and trans regions are available at https://github.com/cyang-2014/Plasma2omic2pops/tree/master; The scripts were extended based on FUSION R package (released in 2016). Functionally-informed Z-score Imputation (FIZI) was used to imputed the missing variant from the summary statistics of T2D GWAS and the genotype data of the protein/metabolite QTL from this study. Colocalization analyses were using R package coloc (v5.1) for the function coloc.susie and coloc (v3.1) for the function coloc.abf. Mendelian randomization was performed with the R package TwoSampleMR (v0.5.7). Alternative integration of TWAS and colocalization was performed with R package INTACT (v1.0.1). Cell-type specificity analysis was performed with S-LDSC (v1.0.1).

Druggable target was queried with the drugbank database (v5.1.10).
The QTL summary statistics were also processed with the PheWeb (v1.3.16) for online browsing.

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 four datasets generated during this study (QTL summary statistics and TWAS weights per feature) are available in the BOX folders listed below:
EUR-proteomics: https://wustl.app.box.com/folder/246498133407;
AFR-proteomics: https://wustl.app.box.com/folder/246495587213;
EUR-metabolomics: https://wustl.app.box.com/folder/246497314974;
AFR- metabolomics: https://wustl.app.box.com/folder/246496814203;
The PheWeb browser for visualizing all plasma omics QTL results is online at https://ontime.wustl.edu/.

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 the Figure-S3E, we presented the sex distribution for each of the four sets analyzed. Overall, the female percentages for all datasets were from 54% to 71%.
Reporting on race, ethnicity, or other socially relevant groupings	In this study, we aim to uncover the ancestary-unique findings between participants of European and African ancestry (Figure-3A). We did not use the terms of "race" or "ethnicity", as they are not genetically defined.
Population characteristics	The age distribution of this cohort can be found in the Figure-S3F. The average ages across all four datasets were 74 to 75. In total, participants in this study included 1,254 AD patients, 1,720 healthy controls, 34 frontotemporal dementia patients, and 162 individuals with an unclassified neurodegenerative disease.
Recruitment	All participants were recruited at the Knight Alzheimer Disease Research Center (Knight ADRC) at Washington University School of Medicine in St. Louis.
Ethics oversight	This project was approved by the ethics committee of the Washington University School of Medicine in St. Louis.

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

Life sciences study design

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

Sample size	We did not use statistical methods to pre-determine sample sizes, but our sample sizes are similar to those reported. We also performed a post-hoc power analysis on the QTL given the input of MAF, the average standardized effect size per ancestry-specificity, the sample size, and using the genome-wide significance thresholds (FWER=0.05, nTests=1e6), we found all variant-feature pairs with a power of 0.8 or more.
Data exclusions	Samples failed QC of proteomics, metabolomics, genotyping were excluded per the Methods section. Table S3 listed the sample size after QC. Proteins and metabolites failed QC were also documented in the Methods section and Tables S4, s5, S6, S7 listed whether the molecular traits passed QC or not.
Replication	We used the published proteomic and metabolomic QTL studies to replicate our QTL findings. For the proteomics, we used EUR set from Ferkingstad et al 2021, AFR set from Surapaneni et al 2022, both EUR and AFR sets from Sun et al 2023. For the metabolomics, we used EUR set from Yin et al 2022, EUR set from Chen et al 2023, AFR set from Rhee et al 2022.

To determine the proportion of the ancestry-specific findings, we used the previous results by Zhang et al 2022 for proteomics and Rhee et al 2022 for metabolomics datasets).

Randomization

The samples were randomized when shipping to profile the proteomics and metabolomics.
For proteomics, 3170 samples were randomly distributed across 38 batches. The age, sex, diagnosis effects were tested to ensure not significant across batches after randomization.
For metabolomics, 3170 samples were randomly distributed across 34 batches. The age, sex, diagnosis effects were tested to ensure not significant across batches after randomization.

Blinding

To ensure the plate compositions were balanced, investigators were not blinded to the batch allocation of each specimen. Researchers had access to unique barcodes assigned to the donor or participant of each specimen as well as other metadata (e.g. age, sex, diagnosis)

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

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, mosiacism, off-target gene editing) were examined.