

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 involved in data collection (all data used is directly available from the UK Biobank under application number 94939).

Data analysis

- plink 1.9
- SNPEFF 5.1
- Bedtools 2.29.1
- R 4.1
- GCTA 1.93
- LDSC 1.0.1
- MatrixEQTL R package 2.3
- coloc R package 5.2.0
- TwoSampleMR R package 0.5.6
- MRPRESSO R package 1.0
- survival R package 3.3-1
- survminer R package 0.4.9
- rmda R package 1.6
- circlize R package 0.4.15
- clusterProfiler R package 4.2.2
- AnnotationDbi R package 1.56.2
- org.Hs.eg.db R package 3.14.0
- GO.db R package 3.14.0

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](#)

All results of pQTL associations, survival-pQTL associations, protein-disease associations, causal proteins for diseases, and drug target have been deposited through an interactive portal (<https://pqtl.whu.edu.cn/pqtl>) and are publicly available. Individual-level genetic and Olink proteomics data from the UKB are available at <https://biobank.ndph.ox.ac.uk/> by application. This research has been conducted using the UKB Resource under application number 94939. The NHGRI-EBI GWAS Catalog is accessible at <https://www.ebi.ac.uk/gwas/>. FinnGen disease GWAS is available at https://www.finnngen.fi/en/access_results. The GTEx v8 release data can be found at <https://www.gtexportal.org/home/datasets>. The PancanQTL database can be found at <http://bioinfo.life.hust.edu.cn/PancanQTL/>. The eQTLGen Consortium is accessible at <https://www.eqtlgen.org/cis-eqtls.html>. The ENCODE portal is accessible at <https://www.encodeproject.org>. The UniProt database can be found at <https://www.uniprot.org/uniprot/>. The drug-gene interactions can be found at <https://www.dgdb.org/>. Source data are provided with this paper

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 considered gender factor as a covariate in our study. Sex in the UK Biobank was determined based on self-reporting data via questionnaire (UK Biobank Field ID: 31), and all included participants gave written informed consent for sharing of individual-level data.

Reporting on race, ethnicity, or other socially relevant groupings

Ethnic background information of the UKB participants was collected based on self-reported data via questionnaires. The majority of UK Biobank participants are of white European descent, and in our study, samples were not excluded or categorized based on ethnicity.

Population characteristics

The UK Biobank cohort participants were largely European in ancestry. In our study, 35,690 participants were included and the average age of the cohort at the time of enrollment was 57.44 years, and 62.54% were women, more detailed information can be found in the Supplementary Table1.

Recruitment

UKB is a population-based prospective cohort consisting of more than 500,000 individuals aged 40 to 69 years at baseline. Participants were enlisted from 22 assessment centers across the UK between 2006 and 2010 and underwent extended monitoring. The database gathered a wide range of health information, including participants' socio-economic and lifestyle factors, physical measurements, blood and urine biomarkers and imputed genomic data.

Ethics oversight

Participants provided informed consent via electronic signatures. Ethics approval was granted by the North West Multi-Centre Research Ethics Committee.

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 restricted the pQTL analysis to individuals with confirmed disease diagnoses prior to enrolment (prevalent cases), rather than the entire PPP cohort. These disease cases were identified and categorized based on the ICD 10 codes. For the healthy control group, we included participants who showed no recorded incidence of the 42 specified disease types or related medication usage, both at baseline and throughout the six-month follow-up period. After excluding participants with a high genetic missingness rate ($\geq 10\%$), with over 20% missing protein data, with missing covariate data, or with official UK Biobank withdrawal, 35,690 participants were included in our pQTL study. Supplementary Data4 provides sample size for each disease. Our study includes disease cohorts with sample sizes in the dozens to ensure broad coverage. This scale is comparable to landmark QTL projects such as GTEx and TCGA, which have been widely applied.

Data exclusions	The final analytical cohort comprised 35,690 participants following exclusions for low-quality proteomic data, low-quality genetic data, missing covariate data, or official withdrawal from the UK Biobank. For full details, please refer to Supplementary Fig. 1 and the Methods section.
Replication	To validate our genetic findings, we compared our results with 23 previously published pQTL studies. Basic characteristics of these studies, including detection methods, sample sizes, and numbers of proteins and pQTLs, are summarized in Supplementary Data 5. More detailed information can be found in the Methods.
Randomization	Given the observational nature of this study, randomization was not applicable.
Blinding	Blinding is not applicable to this study as this study is observational.

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