nature portfolio

Corresponding author(s): Shinichi Namba
Yukinori Okada

Last updated by author(s): 2025/11/14

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	We conducted the analyses using the following publicly available tools: GEM v1.1 (https://github.com/large-scale-G×E-methods/GEM), OSCA v0.46 (https://cnsgenomics.com/software/osca), bgenie v1.3 (https://jmarchini.org/bgenie/), LocusZoom v1.4 (locuszoom.org), LDSC v1.0.0 (https://github.com/bulik/ldsc), PRScs version June 4th, 2021 (https://github.com/getian107/PRScs), MAGMA v1.09a (https://ctg.cncr.nl/software/magma), scDRS v1.0.2 (https://martinjzhang.github.io/scDRS/), enformer (https://github.com/google-deepmind/deepmind-research/tree/master/enformer), shapeit4 v4.1.2 (https://odelaneau.github.io/shapeit4/), minimac4 v1.0.1 (https://github.com/statgen/Minimac4), plink2 v2.00a6LM (https://www.cog-genomics.org/plink/2.0/), king v2.2.5 (https://www.kingrelatedness.com/), FINEMAP v1.4.2 (http://www.christianbenner.com/), GenomeStudio v2.0.5 (https://support.illumina.com/downloads/genomestudio-2-0.html), Applied Biosystems Analysis Power Tools v2.12.0 (https://www.thermofisher.com/jp/en/home/life-science/microarray-analysis/microarray-analysis-partners-programs/affymetrix-developers-network/affymetrix-power-tools.html), BCFtools/liftover (https://github.com/freeseek/score), Transanno v0.4.5 (https://github.com/informationsea/transanno), and R packages (logistf v1.24, mixmeta v1.2.0, GxEprs v1.2, survival v3.2-3, ukbnmr v2.2, glmnet v3.1-1, Rmpfr v0.9-4 [https://cran.r-project.org/web/packages], ConsensusClusterPlus v1.50.0 [https://bioconductor.org/packages/release/bioc/html/ConsensusClusterPlus.html], ACAT v0.91 [https://github.com/yaowuliu/ACAT], Seurat v4.3.0.1 [https://satijalab.org/seurat/], and susieR v0.12.35 [https://stephenslab.github.io/susieR/]).

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 genome-wide summary statistics of G×E interactions for both clinical phenotypes and the metabolome are publicly available at the NBDC Human Database (https://humandbs.dbcls.jp/en) with the accession ID hum0197.v26.374-traits.v1 and at the NHGRI-EBI GWAS Catalog (https://www.ebi.ac.uk/gwas) with the accession IDs GCST90681837–GCST90690020. The UK Biobank analysis was conducted under application number 47821 (https://www.ukbiobank.ac.uk/). The Biobank Japan data are available at the NBDC Human Database (https://humandbs.biosciencedbc.jp/en/) with the accession IDs JGAS000114 and JGAS000412 (genotype), JGAS000561 (metabolome), and JGAS000785 (proteome). Human reference genome GRCh38, http://ftp.1000genomes.ebi.ac.uk/vol1/ftp/technical/reference/GRCh38_reference_genome/; Japanese human reference genome JG2.1.0, https://jmorp.megabank.tohoku.ac.jp/downloads/tommo-jg2.1.0-20211208; the enformer model version 1, https://www.kaggle.com/models/deepmind/enformer/tensorFlow2/enformer/1; East Asian LD block data, https://github.com/jmacdon/LDblocks_GRCh38/blob/master/data/pyrho_EAS_LD_blocks.bed.

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 estimated genetic sex and used it for the analyses throughout the manuscript.
Reporting on race, ethnicity, or other socially relevant groupings	We used the term "population" to refer clusters of human genetic variations, avoiding to use the terms of race and ethnicity.
Population characteristics	BBJ is a prospective biobank that collaboratively collected DNA and serum samples from 12 medical institutions in Japan and recruited approximately 270,000 participants, mainly of Japanese ancestry. Mean age of participants at recruitment was 63.0 years old, and 46.3% were female. All study participants had been diagnosed with one or more of 51 target diseases by physicians at the cooperating hospitals. The UK Biobank project is a population-based prospective cohort that recruited approximately 500,000 people across the United Kingdom. Mean age of participants at recruitment was 56.8 years old, and 53.8% were female. We tested gene–environment interactions for age, sex, and lifestyle exposures. We controlled for age at baseline assessment and biological sex by including age2, age × sex, age2 × sex as covariates. For controlling for population structure, we used genetic principal components as covariates derived from the genotype information.
Recruitment	UK biobank is a population-based biobank of United Kingdom recruited between 2006 and 2010, aged 40–69 years. Biobank Japan is a hospital-based biobank recruited from 2003 to 2008 and from 2013 to 2017, and all participants were diagnosed with at least one of the target diseases of Biobank Japan by physicians at the cooperating hospitals.
Ethics oversight	This study was approved by the ethics committee of Osaka University (approval number: 734-18) and the ethics committee of the Graduate School of Medicine, the University of Tokyo (2023405G-(4)).

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 used UK Biobank and Biobank Japan for discovery and replication. We analyzed unrelated individuals from all genetically European participants in the former and all genetically Japanese participants in the latter. We also used unrelated individuals from all genotyped participants from four external cohorts for replication with total sample size of 436,272: Japan Multi-Institutional Collaborative Cohort (J-MICC) Study and Hospital-based Epidemiologic Research Program at Aichi Cancer Center (HERPACC), Japan Public Health Center-based Prospective (JPHC) Study, the National Institutes of Health's All of Us Research Program, and the Human Phenotype Project (HPP). We did not use any statistical methods to predetermine sample size, but included the maximum number of individuals in each cohort who passed the quality control. This strategy maximized the statistical power in each cohort and we also performed the cross-population meta-analysis to further increase the power.
Data exclusions	For the discovery cohort of Biobank Japan (BBJ1), we performed quality control and excluded sample relatedness as described in Kanai et al.

Data exclusions	Nature Genetics (2018). We analyzed 166,757 participants of the Japanese population as estimated by the visual inspection of principal component analysis (PCA). For the replication cohort of Biobank Japan (BBJ2), we excluded participants with a low call rate (<0.98) and outliers from the Japanese Hondo (i.e., main islands) cluster estimated based on PCA. We used King to exclude relatives within second degrees, resulting in 65,373 participants being analyzed. For the discovery cohort of UK Biobank (UKB1), we included British Europeans, defined as the intersection of the self-reported British participants and the genetically "Caucasian" participants, to strictly reduce the inflation of test statistics due to population stratification. We excluded one participant from every related pair within the third degree. For the replication cohort (UKB2), we included all other genetically "Caucasian" participants and excluded one participant from every related pair within the second degree. We also excluded the participants related to any participants in UKB1 within the second degree. For all cohorts, we excluded participants if their phenotype or environment data were unavailable.
Replication	In UK Biobank, we used British European participants as a discovery cohort (UKB1, Nmax = 273,453) and other European participants as a genetically independent replication cohort (UKB2, Nmax = 8,149). Biobank Japan was comprised of its first and second cohorts (BBJ1, Nmax = 166,757; BBJ2, Nmax = 65,373), and we used them as discovery and replication cohorts, respectively. We also used the external cohorts for replication. We observed replication rates as comparable to those of GWAS, which were summarized in Supplementary Notes and Supplementary Tables.
Randomization	We did not need to use randomization in this study because this is a genotype-phenotype association study with interaction terms with environmental exposures. All samples with available accessibility to genotype, phenotype, and environment data were included in the analysis. In UK Biobank, we used British European participants as a discovery cohort (UKB1, Nmax = 273,453) and other European participants as a genetically independent replication cohort (UKB2, Nmax = 8,149) to reduce population stratification particularly at the discovery stage. We did not apply any allocation to the other biobanks and cohorts. For controlling for population structure, we used genetic principal components as covariates derived from the genotype information.
Blinding	We did not apply blinding of the samples because this is a genotype-phenotype association study with interaction terms with environmental exposures and no intervention was conducted in our study.

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.*