

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
<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. F , t , r) with confidence intervals, effect sizes, degrees of freedom and P 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 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 The following publicly available software was used for data analysis: Illumina BCL Convert v4.2.7, Cell Ranger count v7.0.1, Cell Ranger ARC v2.0.2, Cumulus v2.1.1, Scanpy v1.10.1, anndata v0.10.7, Signac v1.14, MACS2 v2.2.9.1, SnapATAC2 v2.8.0, Eagle v2.3.5, Beagle v4.1, peerR v1.3, tensorQTL v1.0.10, SAIGE-QTL v0.3.0, PLINK2 v2.00a3.3LM, REGENIE v3.3, SuSiE v0.14.2 for molQTL fine-mapping and v0.11.92 for pQTL and complex-trait GWAS fine-mapping, LDSC v1.0.1, SMR v1.4.0.

The following analysis workflows and custom codes were used for data analysis:
<https://github.com/mkanai/finngen-multiome-flagship> (<https://doi.org/10.5281/zenodo.21980018>)
<https://github.com/mkanai/cascade> (<https://doi.org/10.5281/zenodo.21925417>)
<https://github.com/mkanai/fasthurdle> (<https://doi.org/10.5281/zenodo.21932523>)
<https://github.com/mkanai/ldcov> (<https://doi.org/10.5281/zenodo.21934450>)
<https://github.com/FINNGEN/singlecell-pipeline> (<https://doi.org/10.5281/zenodo.21979829>)
<https://github.com/FINNGEN/regenie-pipelines> (<https://doi.org/10.5281/zenodo.21934766>)
<https://github.com/FINNGEN/finemapping-pipeline> (<https://doi.org/10.5281/zenodo.21934769>)
<https://github.com/FINNGEN/coloc.susie.direct> (<https://doi.org/10.5281/zenodo.21933709>)
https://github.com/FINNGEN/kanta_lab_preprocessing (<https://doi.org/10.5281/zenodo.21938286>)

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 FinnGen single-nucleus caQTL, eQTL, peak-gene link summary statistics generated in this study are publicly available at https://www.finnngen.fi/en/access_results and interactively browsable at <https://cascade.finnngen.fi/>. The FinnGen GWAS summary statistics (R12) are available at <https://r12.finnngen.fi/>. The KANTA lab value GWAS summary statistics are available at <https://labvalues.finnngen.fi/>. The FinnGen (R12) + MVP + UKBB meta-analysis results are available at <https://mvp-ukbb.finnngen.fi/>.

Individual-level data, including single-nucleus count matrices, are available to approved researchers. FinnGen as a research project is granted use of national healthcare data and biospecimens according to national (Finland's Act on the Secondary Use of Health and Social Data) and European (the General Data Protection Regulation) regulations, which preclude the research project from distributing individual-level data. Those wishing to work with individual-level data can apply through three independent channels. Health register data are requested from Findata (<https://findata.fi/en/>). Individual-level genotype and other profiling data generated by the project from Finnish biobanks are requested through the Fingenious portal operated by the Finnish Biobank Cooperative (<https://site.fingenious.fi/en/>). The Blood Service Biobank is not a member of the Finnish Biobank Cooperative, and its materials, including the individual-level single-nucleus multiome data generated in this study, are requested directly at biopankki@bloodservice.fi. Academic users wishing instead to collaborate with the FinnGen project directly can fill in a collaboration request form at <https://www.finnngen.fi/en/how-we-collaborate>. Other enquiries should be addressed to finngen-info@helsinki.fi and receive a response within two weeks. Data use is governed by the permit or agreement granted by the relevant authority in each case. Analyses of individual-level data are performed within FinnGen's secure computing environment by a researcher affiliated with the project, and access to that environment is not granted to external collaborators.

The BBJ GWAS summary statistics are available at <https://pheweb.jp/> and at the National Bioscience Database Center (NBDC) Human Database (accession code: hum0197; <https://humandbs.dbcls.jp/en/research/hum0197>). The Japan COVID-19 Task Force whole-blood eQTL summary statistics are available at <https://japan-omics.jp/> and at the NBDC Human Database (accession code: hum0343; <https://humandbs.dbcls.jp/en/research/hum0343>). Additionally, the following datasets were obtained and re-analyzed in this study: eQTL Catalogue, <https://www.ebi.ac.uk/eqtl/>; TenK10K, <https://doi.org/10.5281/zenodo.17474112>; SingleBrain, <https://doi.org/10.5281/zenodo.14907598>; Chinese Immune Multi-Omics Atlas, <https://db.cngb.org/trueblood/cima/>; Alzheimer's disease GWAS meta-analysis from Bellenguez, et al. (2022), GWAS Catalog (accession code: GCST90027158; <https://www.ebi.ac.uk/gwas/studies/GCST90027158>); MPRA data from Siraj, et al. (2026), <https://doi.org/10.5281/zenodo.15297964>.

All analyses used the human genome assembly GRCh38 (GCA_000001405.15), with multiome sequencing data aligned to the 10x Genomics Cell Ranger ARC GRCh38 reference (refdata-cellranger-arc-GRCh38-2020-A, <https://www.10xgenomics.com/support/software/cell-ranger-arc/downloads>). Additional reference resources include: GENCODE v39, https://www.encodegenes.org/human/release_39.html; Ensembl Variant Effect Predictor v105, <https://ftp.ensembl.org/pub/release-105/>; gnomAD v3 Gnomon non-coding constraint and v4.1 gene constraint, <https://gnomad.broadinstitute.org/downloads>; ENCODE4 registry of candidate cis-regulatory elements, <https://screen.wenglab.org/downloads>.

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 used biological sex as a covariate for GWAS and molQTL analyses to correct for potential confounding

Reporting on race, ethnicity, or other socially relevant groupings

N/A

Population characteristics

FinnGen is a public-private partnership combining genotype data from Finnish biobanks with digital health records from Finnish health registries. Genotype and phenotype data in this study are based on Data Freeze 12 (October 2023), which contains 500,348 individuals of Finnish ancestry. The 1,171 participants profiled by single-nucleus multiome are healthy adult volunteer blood donors recruited through the Finnish Red Cross Blood Service, of whom 1,108 donors passed the QC criteria. Their median age at sample collection was 52.8 (min: 19, max: 71), and 52% were female. Participants were not specifically selected on any disease status for the multiome study; however, blood donors are screened for eligibility and are on average healthier than the general population.

Recruitment

FinnGen includes recruitment from population-based biobanks as well as hospital-based legacy cohorts. A subset of FinnGen participants who donated blood through the Finnish Red Cross Blood Service (n = 1,171) were included for single-nucleus multiome experiments.

Ethics oversight

Study subjects in FinnGen provided informed consent for biobank research, based on the Finnish Biobank Act. Alternatively, separate research cohorts, collected prior to the Finnish Biobank Act came into effect (in September 2013) and start of FinnGen (August 2017), were collected based on study-specific consents and later transferred to the Finnish biobanks after approval by Fimea (Finnish Medicines Agency), the National Supervisory Authority for Welfare and Health. Recruitment protocols followed the biobank protocols approved by Fimea. The Coordinating Ethics Committee of the Hospital District of Helsinki and Uusimaa (HUS) statement number for the FinnGen study is Nr HUS/990/2017. The FinnGen study is approved by Finnish Institute for Health and Welfare (permit numbers: THL/2031/6.02.00/2017, THL/1101/5.05.00/2017, THL/341/6.02.00/2018, THL/2222/6.02.00/2018, THL/283/6.02.00/2019, THL/1721/5.05.00/2019).

and THL/1524/5.05.00/2020), Digital and population data service agency (permit numbers: VRK43431/2017-3, VRK/6909/2018-3, VRK/4415/2019-3), the Social Insurance Institution (permit numbers: KELA 58/522/2017, KELA 131/522/2018, KELA 70/522/2019, KELA 98/522/2019, KELA 134/522/2019, KELA 138/522/2019, KELA 2/522/2020, KELA 16/522/2020), Findata permit numbers THL/2364/14.02/2020, THL/4055/14.06.00/2020, THL/3433/14.06.00/2020, THL/4432/14.06/2020, THL/5189/14.06/2020, THL/5894/14.06.00/2020, THL/6619/14.06.00/2020, THL/209/14.06.00/2021, THL/688/14.06.00/2021, THL/1284/14.06.00/2021, THL/1965/14.06.00/2021, THL/5546/14.02.00/2020, THL/2658/14.06.00/2021, THL/4235/14.06.00/2021, Statistics Finland (permit numbers: TK-53-1041-17 and TK/143/07.03.00/2020 (earlier TK-53-90-20) TK/1735/07.03.00/2021, TK/3112/07.03.00/2021) and Finnish Registry for Kidney Diseases permission/extract from the meeting minutes on 4th July 2019.

The Biobank Access Decisions for FinnGen samples and data utilized in FinnGen Data Freeze 12 include: THL Biobank BB2017_55, BB2017_111, BB2018_19, BB_2018_34, BB_2018_67, BB2018_71, BB2019_7, BB2019_8, BB2019_26, BB2020_1, BB2021_65, Finnish Red Cross Blood Service Biobank 7.12.2017, Helsinki Biobank HUS/359/2017, HUS/248/2020, HUS/430/2021 §28, §29, HUS/150/2022 §12, §13, §14, §15, §16, §17, §18, §23, §58, §59, HUS/128/2023 §18, Auria Biobank AB17-5154 and amendment #1 (August 17 2020) and amendments BB_2021-0140, BB_2021-0156 (August 26 2021, Feb 2 2022), BB_2021-0169, BB_2021-0179, BB_2021-0161, AB20-5926 and amendment #1 (April 23 2020) and it's modifications (Sep 22 2021), BB_2022-0262, BB_2022-0256, Biobank Borealis of Northern Finland_2017_1013, 2021_5010, 2021_5010 Amendment, 2021_5018, 2021_5018 Amendment, 2021_5015, 2021_5015 Amendment, 2021_5015 Amendment_2, 2021_5023, 2021_5023 Amendment, 2021_5023 Amendment_2, 2021_5017, 2021_5017 Amendment, 2022_6001, 2022_6001 Amendment, 2022_6006 Amendment, 2022_6006 Amendment, 2022_6006 Amendment_2, BB22-0067, 2022_0262, 2022_0262 Amendment, Biobank of Eastern Finland 1186/2018 and amendment 22§/2020, 53§/2021, 13§/2022, 14§/2022, 15§/2022, 27§/2022, 28§/2022, 29§/2022, 33§/2022, 35§/2022, 36§/2022, 37§/2022, 39§/2022, 7§/2023, 32§/2023, 33§/2023, 34§/2023, 35§/2023, 36§/2023, 37§/2023, 38§/2023, 39§/2023, 40§/2023, 41§/2023, Finnish Clinical Biobank Tampere MH0004 and amendments (21.02.2020 & 06.10.2020), BB2021-0140 8§/2021, 9§/2021, 9§/2022, §10/2022, §12/2022, 13§/2022, §20/2022, §21/2022, §22/2022, §23/2022, 28§/2022, 29§/2022, 30§/2022, 31§/2022, 32§/2022, 38§/2022, 40§/2022, 42§/2022, 1§/2023, Central Finland Biobank 1-2017, BB_2021-0161, BB_2021-0169, BB_2021-0179, BB_2021-0170, BB_2022-0256, BB_2022-0262, BB22-0067, Decision allowing to continue data processing until 31st Aug 2024 for projects: BB_2021-0179, BB22-0067, BB_2022-0262, BB_2021-0170, BB_2021-0164, BB_2021-0161, and BB_2021-0169, and Terveystalo Biobank STB 2018001 and amendment 25th Aug 2020, Finnish Hematological Registry and Clinical Biobank decision 18th June 2021, Arctic biobank P0844: ARC_2021_1001.

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	FinnGen enrolled 500,348 samples for genetic analyses at Data Freeze 12 (October 2023). No statistical method was used to predetermine sample size for the multiome study. Through the Finnish Red Cross Blood Service, we recruited 1,171 healthy Finnish donors enrolled in FinnGen for single-nucleus multiome sequencing. After the fixed QC thresholds described in the Methods, 1,108 donors contributed to the snRNA-seq atlas and 1,093 of them contributed paired snATAC-seq data. This sample size is deemed sufficient because statistical power was demonstrated empirically from the observed molQTL yield rather than assumed from a prespecified effect size.
Data exclusions	All QC exclusion criteria were pre-established and applied uniformly to every sample, as described in the Methods. At the nucleus level, we excluded nuclei with fewer than 400 unique UMIs, fewer than 200 detected genes, or more than 20% of unique UMIs mapping to mitochondrial genes, then Cumulus-predicted doublets, then nuclei whose SCTransform-normalized MALAT1 expression fell below a channel-specific adaptive threshold. For snATAC-seq, we retained only barcodes present in the snRNA-seq atlas and further excluded nuclei with fewer than 2,000 or more than 100,000 fragments in peaks, a ratio of reads in ENCODE blacklist regions of 0.01 or more, a nucleosome signal of 2 or more, or a transcription start site enrichment score of 4 or below. For multiplexed channels, only demuxlet-predicted singlets were retained. At the donor level, 1,171 donors were recruited, 1,108 yielded a QC-passing snRNA-seq library, and 1,093 of those also yielded paired snATAC-seq data. At the feature level, we tested genes expressed in more than 50% of samples and peaks meeting the same criterion within a cell type, requiring at least 1% nonzero expression or accessibility for candidate peak-gene pairs and at least 5,000 expressing cells for the hurdle count component. Variant-level thresholds were MAC above 20 for cis-eQTL and cis-caQTL mapping, above 8 for pQTL mapping, and at least 5 among cases and controls for the complex-trait GWAS.
Replication	No experiment was repeated, because no measurement in this study is a repeatable assay. Each nucleus was captured in exactly one channel, so there are no technical replicates. Reproducibility was established instead by external replication in independent cohorts. Fine-mapped cis-eQTLs were replicated in five independent single-cell PBMC datasets (TenK10K, OneK1K, CIMA, AIDA, and BBJ), and fine-mapped cis-caQTLs in two of them (CIMA and BBJ). Peak-gene links were replicated by rerunning the identical hurdle model in BBJ. GWAS-eQTL colocalizations were replicated against eQTL Catalogue and TenK10K. The constraint gradients were replicated in eQTL Catalogue, TenK10K, and SingleBrain.
Randomization	Not applicable to this study. This is an observational human genetics study with no experimental groups, interventions, or allocation of participants. The randomization that underpins genetic association analysis is meiotic, in that genotypes are assigned to individuals independently of later environment and disease status. Confounding was controlled through covariates rather than allocation. Molecular QTL models included age at sample collection, sex, the top four genetic PCs, and cell-type-specific PEER factors for expression or covariate-residualized accessibility PCs for chromatin. Complex-trait GWAS in FinnGen R12 included sex, age at the end of follow-up or death,

genotyping batch, and the top ten genotype PCs, with relatedness handled by LOCO predictions in regenie.

Blinding

Not applicable to this study. There were no treatment or experimental groups, and therefore no group allocation to be blinded to. Donors were not selected on disease status for single-nucleus multiome sequencing. All downstream analyses were applied uniformly across all donors, using QC criteria and significance thresholds described in the Methods.

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