nature portfolio

Corresponding author(s): Gudny E Thorlacius and Kari Stefansson

Last updated by author(s): Oct 3, 2025

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 used publicly available software together with software/methods developed at deCODE genetics as described in Methods. The publicly available software are: R, version 3.6.0 to analyze data and create plots; Graphtyper version 2, https://github.com/DecodeGenetics/graphtyper; Variant Effect Predictor (release 100), https://github.com/Ensembl/ensembl-vep; IMPUTE2 version 2.3.1, https://mathgen.stats.ox.ac.uk/impute/impute_v2.html; BOLT-LMM version 2.1, http://www.hsph.harvard.edu/alkes-price/software/ ADMIXTURE v1.23, https://dalexander.github.io/admixture/

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 are available at (https://download.decode.com/form/2025-Thorlacius-Ivarsdottir). This study used data from the All of Us Research Program's Controlled Tier Dataset V7, available to authorized users on the Researcher Workbench as well as summary statistics from the VA Million Veterans Program availabe through dbGAP (accession number phs002453). The UK Biobank data were downloaded under application no. 56270. Individual-level genomic and phenotypic data from the UK Biobank are available to researchers upon application (https://ukbiobank.ac.uk).

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 use the term sex to refer to biological sex, determined from genotype data, and the terms biologically female/male.
Reporting on race, ethnicity, or other socially relevant groupings	The study sample set included 27,871 study participants that were assessed as carrying African genomic ancestry by analyzing their genotypes with supervised ADMIXTURE v1.23 using the 1000G (1000 Genomes) populations CEU (Utah Caucasian), CHB (Beijing Han), ITU (Indian Telugu), PEL (Peruvian), and YRI (Nigerian Yoruba) as training samples. Individuals modeled by ADMIXTURE as either 1) carrying >90% YRI-like ancestry, or 2) carrying >30% YRI-like ancestry, >2% CEU-like ancestry, and YRI-like plus CEU-like ancestry proportions summing to >90% were included and referred to as Americans with African ancestry.
Population characteristics	All 27,871 study participants were whole-genome sequenced. Average YOB for the SLE cases was 1972, 90.2% were female. Average YOB of the controls was 1974, 59.7% were female.
Recruitment	Nashville Biosciences (NashBio) is a data and analytics provider wholly owned by Vanderbilt University Medical Center (VUMC, Tennessee, USA). NashBio harnesses VUMC's extensive genomic and bioinformatics resources, including its biobank collection BioVU®. BioVU® includes a collection of de-identified DNA samples that are linked to de-identified version of the VUMC's EHRs, referred to as Synthetic Derivative (SD) database. The Diverse Ancestry Cohort (DAC) included in this study is a subset of the BioVU® cohort, consisting of 35,080 WGS individuals of diverse ancestries, primarily African ancestry. We further restricted our study sample set to 27,871 study participants that we assessed as carrying African genomic ancestry, as decribed above. Therefore the DAC database includes data on patients of the VUMC who consented to their residual samples from routine clinical testing and data being contributed to BioVU®.
Ethics oversight	All participants consented to their residual samples from routine clinical testing and data being contributed to BioVU®. BioVU® extracts and banks germline DNA samples that are de-identified and only linked to the SD through a randomly assigned unique identifier, not back to the patient or their underlying medical record. The use of BioVU® is classified as non-human subject research by VUMC's Institutional Review Board (IRB), and the research conducted using the dataset generated by the Alliance for Genomic Discovery was approved by VUMC's IRB. The overall biobanking program is reviewed annually by the IRB to maintain this determination and make decisions about patient protections, privacy, and ethical issues. Each individual study seeking to use the SD database and BioVU® biobank is filed with VUMC's IRB to maintain this non-human subject determination and appropriate use of the data. NashBio takes steps to protect client confidentiality during this submission.

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	Sample sizes are reported in the article and correspond to the data that was available.
Data exclusions	No available data was excluded.
Replication	We replicated the novel variant in 2 additional datasets; the subset of the All of Us Research Program and the VA Million Veteran Program

Replication	that are defined as having African ancestry.
Randomization	Not applicable.
Blinding	Not applicable.

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.*