nature research

Corresponding author(s): Ayellet V. Segrè, Eric R. Gamazon

☐ Initial submission ☐ Revised version ☒ Final submission

# Life Sciences Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form is intended for publication with all accepted life science papers and provides structure for consistency and transparency in reporting. Every life science submission will use this form; some list items might not apply to an individual manuscript, but all fields must be completed for clarity.

For further information on the points included in this form, see Reporting Life Sciences Research. For further information on Nature Research policies, including our data availability policy, see Authors & Referees and the Editorial Policy Checklist.

## Experimental design

### 1. Sample size

| | |
|---|---|
| Describe how sample size was determined. | Our analyses are based on the Genotype Tissue Expression (GTEx) study, publicly available genome-wide association study (GWAS) results, and Biobank data. Sample sizes were hence determined by these studies. A description of the experimental design of GTEx is described in "GTEx Consortium, Genetic effects on gene expression across human tissues. Nature 550, 204-213 (2017)". References to all GWAS analyzed in the paper are provided in Supplementary Table 1, and references to the UK and BioVU Biobank resources are presented in the URLs and in our paper. |

### 2. Data exclusions

| | |
|---|---|
| Describe any data exclusions. | No data were excluded from the analysis. |

### 3. Replication

| | |
|---|---|
| Describe whether the experimental findings were reliably reproduced. | Experimental replication was not attempted. All replication analyses that we performed of variant and gene associations in separate biobank GWAS studies are reported in the text and supplementary tables. |

### 4. Randomization

| | |
|---|---|
| Describe how samples/organisms/participants were allocated into experimental groups. | The order of sample processing for library preparation and sequencing in GTEx was randomized to avoid batch effects as described in GTEx Consortium, Nature 2017. The genome-wide studies analyzed in this paper were designed and genotyped by other consortia. |

### 5. Blinding

| | |
|---|---|
| Describe whether the investigators were blinded to group allocation during data collection and/or analysis. | Blinding was not performed in the GTEx study or GWAS meta-analyses. |

Note: all studies involving animals and/or human research participants must disclose whether blinding and randomization were used.

### 6. Statistical parameters

For all figures and tables that use statistical methods, confirm that the following items are present in relevant figure legends (or in the Methods section if additional space is needed).

| n/a | Confirmed | |
|---|---|---|
| ☐ | ☒ | The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement (animals, litters, cultures, etc.) |
| ☐ | ☒ | A description of how samples were collected, noting whether measurements were taken from distinct samples or whether the same sample was measured repeatedly |
| ☒ | ☐ | A statement indicating how many times each experiment was replicated |
| ☐ | ☒ | The statistical test(s) used and whether they are one- or two-sided (note: only common tests should be described solely by name; more complex techniques should be described in the Methods section) |
| ☐ | ☒ | A description of any assumptions or corrections, such as an adjustment for multiple comparisons |
| ☐ | ☒ | The test results (e.g. *P* values) given as exact values whenever possible and with confidence intervals noted |
| ☐ | ☒ | A clear description of statistics including central tendency (e.g. median, mean) and variation (e.g. standard deviation, interquartile range) |
| ☐ | ☒ | Clearly defined error bars |

*See the web collection on statistics for biologists for further resources and guidance.*

## Software

Policy information about availability of computer code

### 7. Software

| | |
|---|---|
| Describe the software used to analyze the data in this study. | PLINK 1.90: https://www.cog-genomics.org/plink2<br>eCAVIAR: https://github.com/fhormoz/caviar<br>Regulatory Trait Concordance (RTC): https://qtltools.github.io/qtltools/<br>TORUS: https://github.com/xqwen/torus<br>PrediXcan: https://github.com/hakyim/PrediXcan<br>Storey's qvalue R package: https://github.com/StoreyLab/qvalue<br>LD score regression (LDSR): https://github.com/bulik/ldsc<br>GCTA: http://cnsgenomics.com/software/gcta/#Download<br>eGeneEnrich: https://segrelab.meei.harvard.edu/software/<br>eQTLEnrich: https://segrelab.meei.harvard.edu/software/<br>GTEx Portal: http://www.gtexportal.org/<br>Gene Ontology: http://geneontology.org/<br>UK Biobank: http://www.ukbiobank.ac.uk/<br>BioVU: https://victr.vanderbilt.edu/pub/biovu/?sid=194<br>NHGRI-EBI GWAS Catalog: http://www.ebi.ac.uk/gwas<br>Mouse Genome Informatics: http://www.informatics.jax.org/downloads/reports/index.html |

For manuscripts utilizing custom algorithms or software that are central to the paper but not yet described in the published literature, software must be made available to editors and reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). *Nature Methods* guidance for providing algorithms and software for publication provides further information on this topic.

## Materials and reagents

Policy information about availability of materials

### 8. Materials availability

| | |
|---|---|
| Indicate whether there are restrictions on availability of unique materials or if these materials are only available for distribution by a for-profit company. | Samples from the GTEx Biobank can be requested via a Sample Request Form: https://gtexportal.org/home/samplesPage. |

### 9. Antibodies

| | |
|---|---|
| Describe the antibodies used and how they were validated for use in the system under study (i.e. assay and species). | No antibodies were used in the study. |

### 10. Eukaryotic cell lines

| | |
|---|---|
| a. State the source of each eukaryotic cell line used. | Lymphoblastoid cell lines (LCLs) were extracted from GTEx donors. |
| b. Describe the method of cell line authentication used. | The LCLs were authenticated by confirming that they can grow and replicate, and by genotyping them to confirm that they belong to the right donor and match the other samples from that donor. |
| c. Report whether the cell lines were tested for mycoplasma contamination. | All lymphoblastoid cell lines (LCLs) extracted from GTEx donors tested negative for mycoplasma contamination. |
| d. If any of the cell lines used are listed in the database of commonly misidentified cell lines maintained by ICLAC, provide a scientific rationale for their use. | No commonly misidentified cell lines were used. |

## Animals and human research participants

Policy information about studies involving animals; when reporting animal research, follow the ARRIVE guidelines

### 11. Description of research animals

| | |
|---|---|
| Provide details on animals and/or animal-derived materials used in the study. | No animals were used in the study. |

Policy information about studies involving human research participants

### 12. Description of human research participants

| | |
|---|---|
| Describe the covariate-relevant population characteristics of the human research participants. | The Genotype Tissue Expression (GTEx) Project release v6p eQTLs are based on 449 postmorterm donors, with age ranging between 20-70, about one third females and two third males and about 84% Europeans, 15% African Americans and 1% Asian or other race. |