

Reporting Summary

Nature Research 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 Research policies, see [Authors & Referees](#) 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	HipSTR v0.5, bedtools v2.28.0, plink v1.90b3.44, Smartpca v13050, PEER (https://github.com/PMBio/peer), Python statsmodels package v0.8.0, CAVIAR v2.2, mashR v0.2.21, Python seaborn package v0.9.0, Python scipy package v1.2.1, Homer v4.10, Python sklearn package 0.20.3, Mfold v3.6, Custom code for additional analyses is available at https://github.com/gymreklab/gtex-estrs-paper

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/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research [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 list of figures that have associated raw data
- A description of any restrictions on data availability

All eSTR summary statistics are available for download on WebSTR <http://webstr.ucsd.edu/downloads>.

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	We used 652 samples with available whole genome sequencing from GTEx. This sample size was based on the sample size in the data available from GTEx (phs000424.v7.p2).
Data exclusions	No samples were excluded
Replication	We compared eSTR summary statistics to previously published results. Replication was measured using Pearson correlation between effect sizes previously reported and those reported in this study (computed for each tissue, mean Pearson $r=0.45$).
Randomization	Our study primarily relies on association testing. Analyses where randomization would typically be used (e.g. trials) were not performed here.
Blinding	Our study primarily relies on association testing. Analyses where blinding would typically be used (e.g. trials) were not performed here.

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

Methods

n/a	Involved in the study	n/a	Involved in the study
☒	☐ Antibodies	☒	☐ ChIP-seq
☐	☒ Eukaryotic cell lines	☒	☐ Flow cytometry
☒	☐ Palaeontology	☒	☐ MRI-based neuroimaging
☒	☐ Animals and other organisms		
☒	☐ Human research participants		
☒	☐ Clinical data		

Eukaryotic cell lines

Policy information about [cell lines](#)

Cell line source(s)	HEK293T cells from ATCC
Authentication	Cells were purchased from ATCC. We did no additional authentication.
Mycoplasma contamination	Mycoplasma was last tested using PCR in October 2018. Experiments were performed November 2018. All cell lines tested negative for mycoplasma contamination.
Commonly misidentified lines (See ICLAC register)	No cell lines used are listed in the ICLAC register. Experiments were performed in HEK293T cells (ATCC CRL-3216)