nature research

Corresponding Author: ALEXANDER GUSEV

Date: 3/20/2018

Life Sciences Reporting Summary

Nature Research wishes to improve the reproducibility of the work we publish. This form is published with all life science papers and is intended to promote consistency and transparency in reporting. All life sciences submissions use this form; while some list items might not apply to an individual manuscript, 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.

   No data was generated for this study, the largest sample sizes from publicly available data were used.

2. Data exclusions

   Describe any data exclusions.

   Pre-established quality control criteria were used for expression data as defined in previous publications from the GTEx Consortium.

3. Replication

   Describe whether the experimental findings were reliably reproduced.

   No individual variant replication was carried. Gene prediction accuracy was established by cross-validation and reported for all genes. Aggregate replication of significant TWAS Associations was performed using risk scores and shown to be significant. Individual and summary-based predictions were shown to be highly correlated.

4. Randomization

   Describe how samples/organisms/participants were allocated into experimental groups.

   Randomization was not necessary for the statistical analyses, and the functional analyses were performed with blinding.

5. Blinding

   Describe whether the investigators were blinded to group allocation during data collection and/or analysis.

   For functional interrogation of zebrafish, all experiments were repeated in duplicate and were scored by investigators blinded to injection cocktail.

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 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 summary of the descriptive 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.	TWAS/FUSION software was developed and used for analyses. The software + documentation is available on the web: http://gusevlab.org/projects/fusion/ and the source is available in a GitHub repository: https://github.com/gusevlab/fusion_twas

For all studies, we encourage code deposition in a community repository (e.g. GitHub). Authors must make computer code available to editors and reviewers upon request. The *Nature Methods* guidance for providing algorithms and software for publication may be useful for any submission.

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.	no restrictions

9. Antibodies

Describe the antibodies used and how they were validated for use in the system under study (i.e. assay and species).	a phosphor histone 3 (PH3) antibody (ser10)-R (sc-8656-R, rabbit, Santa Cruz; 1:500) was used and experiments were performed as previously described in reference 57: Jordan, D. M. et al. Identification of cis-suppression of human disease mutations by comparative genomics. Nature 524, 225–229 (2015).

10. Eukaryotic cell lines

a. State the source of each eukaryotic cell line used.	n/a
b. Describe the method of cell line authentication used.	n/a
c. Report whether the cell lines were tested for mycoplasma contamination.	n/a
d. If any of the cell lines used in the paper are listed in the database of commonly misidentified cell lines maintained by ICLAC, provide a scientific rationale for their use.	n/a

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.	Zebrafish embryos and adults were maintained and mated according to standard procedures and all experiments were carried out with the approval of the Institutional Animal Care and Use Committee (IACUC).

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.	Human data was previously collected and de-identified. Sex, age, genetic ancestry, and multiple measures of assay quality were always included as covariates.