

Clearly defined error bars

Corresponding Author:	Jeremy Schwartzentruber
Date:	23/10/2017

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.	QTL studies of fewer than 100 samples are common. The number of associations discovered is a function of sample size, so the sample size chosen is a function of cost and scope rather than power for specific effect sizes.
2.	2. Data exclusions	
	Describe any data exclusions.	Sensory neuron differentiation and characterisation, paragraph 1, and Supplementary Note.
3.	3. Replication	
	Describe whether the experimental findings were reliably reproduced.	QTLs were called using criteria for genome-wide significance at a chosen false discovery rate. Other analyses have not been replicated in independent samples.
4.	Randomization	
	Describe how samples/organisms/participants were allocated into experimental groups.	N/A
5.	Blinding	
	Describe whether the investigators were blinded to group allocation during data collection and/or analysis.	N/A
	Note: all studies involving animals and/or human research participants must di	sclose whether blinding and randomization were used.
6.	5. 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 <u>exact</u> sample size (<i>n</i>) 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	

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.

Scripts were used for processing raw sequencing reads to obtain expression / peak counts. Additional analyses were done with R. Code examples for these analyses are provided at https://github.com/js29/ insdsn.

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.

iPSC lines used for our differentiations are available at www.hipsci.org.

9. Antibodies

the system under study (i.e. assay and species).

Describe the antibodies used and how they were validated for use in | For all antibodies, as applicable, provide supplier name, catalog number, clone name, and lot number. Also describe the validation of each primary antibody for the species and application, noting any validation statements on the manufacturer's website, relevant citations, antibody profiles in online databases, or data provided in the manuscript OR state that no antibodies were used.

10. Eukaryotic cell lines

- a. State the source of each eukaryotic cell line used.
- b. Describe the method of cell line authentication used.
- c. Report whether the cell lines were tested for mycoplasma contamination.
- 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.

Cell lines used are described in supplementary tables 1-2, and further details are available at www.hipsci.org.

Genotypes from RNA-seq confirmed the identity of cell lines relative to genotypes reported by HIPSCI.

Confirm that all cell lines tested negative for mycoplasma contamination OR describe the results of the testing for mycoplasma contamination OR declare that the cell lines were not tested for mycoplasma contamination OR state that no eukaryotic cell lines were used.

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.

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.

IPSCs used in this study were derived from healthy individuals enrolled in the HIPSCI project.