nature research

Corresponding author(s): Rahul M. Kohli; Hao Wu

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.

Please do not complete any field with "not applicable" or n/a. Refer to the help text for what text to use if an item is not relevant to your study. For final submission: please carefully check your responses for accuracy; you will not be able to make changes later.

Experimental design

1. Sample size

Describe how sample size was determined.

No sample size pre-calculations were performed.

2. Data exclusions

Describe any data exclusions.

From next generation sequencing data, PCR duplicates were removed and the sequences further filtered to remove strands with three consecutive non-converted CpHs. The complete statistics for removal of filtered data are provided in SI Table 1. In ICR analysis, the Dlk1-Gtl2 IG and Igf2r loci were excluded due to aberrantly high 5mC levels and low coverage, respectively

3. Replication

Describe the measures taken to verify the reproducibility of the experimental findings.

The data was reliably reproduced. Comparison of different gDNA inputs is shown in SI Fig 6.

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 in vivo studies must report how sample size was determined and 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 *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of any assumptions or corrections, such as an adjustment for multiple comparisons
☐	☒	Test values indicating whether an effect is present *Provide confidence intervals or give results of significance tests (e.g. P values) as exact values whenever appropriate and with effect sizes 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 in all relevant figure captions (with explicit mention of central tendency and variation)

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.

The software used is described in the Methods section in detail. The programs used include:
1. Trim Galore (v0.4.1): https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/
2. FastQC (v0.11.5): https://www.bioinformatics.babraham.ac.uk/projects/fastqc/
3. Bismark (v0.14.3): https://www.bioinformatics.babraham.ac.uk/projects/bismark/
4. Picard (v2.5.0): https://broadinstitute.github.io/picard/
5. methpipe/MLML (v3.4.3): http://smithlabresearch.org/software/mlml/
6. R/ggplot2 (v2.2.1): http://ggplot2.tidyverse.org/
7. Integrative Genomics Viewer (v2.3.32): https://software.broadinstitute.org/software/igv/

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 third party.

All materials are available commercially, with the exception of APOBEC3A and genomic DNA derived from cell lines or mice. The APOBEC3A expression plasmid has been deposited in Addgene for release upon publication to academic users or can be requested from authors.

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.

10. Eukaryotic cell lines

a. State the source of each eukaryotic cell line used.

ESCs (J1) were originally purchased from ATCC and no additional authentication was performed. TET TKO cells were derived from two independent methods. One sample was obtained from the Zhang lab, with generation described in Lu et al, Genes Dev. (2014) 28: 2103-2119. The second TET TKO cell line was derived in house using the method from Wang et al, Cell (2013) 153:910-918. Both cells lines were confirmed as noted below.

b. Describe the method of cell line authentication used.

TET TKO ESCs were verified by Sanger sequencing at Tet1, Tet2 and Tet3. The lack of hmC was confirmed by mass spectrometry.

c. Report whether the cell lines were tested for mycoplasma contamination.

Not tested.

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.

None of the cell lines used were listed in ICLAC.

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 all relevant details on animals and/or animal-derived materials used in the study.

Purified mouse neuronal gDNA was provided by Z. Zhou, as per Johnson BS et al, Nat Med (2017) 23:1203-1214. The samples were taken from 6-week old C57BL/6 male mice.

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.

This study did not involve human participants.