nature research

Corresponding author(s): Thorsten Stafforst

☐ 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.

All experiments for evaluating editing yields of endogenous targets via Sanger sequencing were done in triplicate (independent experiments) to validate reproducibilty and to provide appropriate standard deviations
NGS analysis was performed with two independent replicates per sample; the required sequencing depth was determined in a pilot experiment and saturated with 50 Mio 100 bp paired-end reads at 25 000 detected transcripts. This sequencing depths was also similar to other very recent papers on global off-target effects of site-directed RNA editing (Cox et al. Science 2017, Rosenthal et al. RNA Biol. 2018)

2. Data exclusions

Describe any data exclusions.

There was no data exclusion; details for the NGS analysis pipeline and filter settings are given in the online methods in full detail

3. Replication

Describe whether the experimental findings were reliably reproduced.

all experiments could relialbly be reproduced

4. Randomization

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

no randomization was performed, samples were treated according to the same protocols side-by-side with the respective controls

5. Blinding

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

no blinding was performed, editing experiments were allocated to several experimentators

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.	As outlined in full detail in the online methods all software tools used for NGS are publically available: Mapping of RNA-seq and reads: BWA was used to align the reads to a combination of the reference genome sequences (hg19) and exonic sequences surrounding known splicing junctions from known gene models, obtained through the UCSC Genome Browser for Gencode, RefSeq, Ensembl, and UCSC Genes. Unique and non-duplicate reads were subjected to local realignment and base score recalibration using the IndelRealigner and TableRecalibration from the Genome Analysis Toolkit (GATK). Identification of editing sites from RNA-seq data: We used the UnifiedGenotyper from GATK27 to call variants from the mapped RNA-seq reads. In contrast to the usual practice of variant calling, we identified the variants with relatively loose criteria by using the UnifiedGenotyper tool. We first removed all known human SNPs present in dbSNP (except SNPs of molecular type "cDNA"; database version 135; http://www.ncbi.nlm.nih.gov/SNP/), the 1000 Genomes Project, and the University of Washington Exome Sequencing Project (http://evs.gs.washington.edu/EVS/). Finally, variants were annotated using ANNOVAR based on gene models from Gencode, RefSeq, Ensembl, and UCSC. The resulting sets of sites identified from RNA-seq data were compared with all sites available in the RADAR database and were subsequently referred to as 'known' sites if also found in RADAR, or 'novel' sites if not found. The manuscript and Supplementary Information were written with Microsoft Word 2016, Sanger editing data was analyzed (mean, SD) and plotted with GraphPad Prism 7.04 and Excel 2016. Figures were prepared with CorelDraw 2017.

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.	no unique material was used

9. Antibodies

Describe the antibodies used and how they were validated for use in the system under study (i.e. assay and species).

The protocol is provided in the manuscript: primary antibodies have been used against the SNAP-tag (1:1000, P9310S, New England Biolabs, USA) and ß-actin (1:40000, A5441, Sigma Aldrich, USA). Both antibodies are well established commercial antibodies. The SNAP-tag antibody was validated by the fact that total protein from parental empty cells (not expressing a SNAP-tagged protein) do not stain in the immunoblot. After integration of the SNAP-tagged protein, the total protein showed a clear doxycycline-inducible protein band of the expected size. The secondary antibodies were HRP-conjugated anti-rabbit (1:10000, 111-035-003, Jackson Immuno Research Laboratories, USA) and anti-mouse (1:10000, 115-035-003, Jackson Immuno Research Laboratories, USA). Both are well-known and validated commercial secondary antibodies.

10. Eukaryotic cell lines

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

We generated cell lines derived from the parental Flip-In T-REx cell line (Catalog no. R78007, Thermo Fisher scientific)

b. Describe the method of cell line authentication used.

Cell line authentication was confirmed by antibiotic selection before and after recombination

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

all cell lines were 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

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 study did not involve human research participants