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.

Statistical parameters

When statistical analyses are reported, confirm that the following items are present in the relevant location (e.g. 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 |
| ☑   | An indication of 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 |
| ☑   | Only common tests should be described solely by name; describe more complex techniques in the Methods section. |
| ☑   | 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 statistics 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. Give P values as exact values whenever suitable. |
| ☑   | 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 |
| ☑   | Clearly defined error bars |
|     | State explicitly what error bars represent (e.g. SD, SE, CI) |

Our web collection on statistics for biologists may be useful.

Software and code

Policy information about availability of computer code

Data collection

Illumina MiSeq was used to generate targeted amplicon sequencing data. Demultiplexed data were downloaded from Illumina Basespace.

Data analysis

CRISPRESSO2.0.31, R version 3.5.1

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 upon request. 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

Plasmids encoding CG9E1 [Addgene #140252] and miniCG9E1 [Addgene #140253], as well as other constructs used in this work are available on Addgene via https://www.addgene.org/Keith_Joung/
Targeted amplicon sequencing data (obtained from Illumina Basespace) have been deposited at the Sequence Read Archive [SRA]: https://www.ncbi.nlm.nih.gov/sra/PRJNA622835. All other relevant data are available from the corresponding authors upon request.

Field-specific reporting

Please select 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/authors/policies/reportingSummary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

- Sample size: Sample sizes are based on the work and experience of other groups in the field who generate reproducible results in similar experiments.
- Data exclusions: One experiment with a gRNA targeting a new genomic site (ABE site 22) was excluded because the positive control did not show editing.
- Replication: We used cells from different passages for independent replicates (n = 3 or 4, as indicated in the paper). All attempts at replication were successful.
- Randomization: Randomization was not performed.
- Blinding: Blinding was not performed.

Reporting for specific materials, systems and methods

| Materials & experimental systems | Methods |
|---------------------------------|---------|
| n/a                             | n/a     |
| - Involved in the study         | Involved in the study |
| - Unique biological materials   | - CHIP-seq |
| - Antibodies                    | - Flow cytometry |
| - Eukaryotic cell lines         | - MRI-based neuroimaging |
| - Palaeontology                 |         |
| - Animals and other organisms   |         |
| - Human research participants   |         |

Eukaryotic cell lines

Policy information about cell lines.

- Cell line source(s): HEK293T [ATCC#: CRL-3216], K562 [ATCC#: CCL-243], and U2OS [ATCC#: HTB-96] cells were obtained from ATCC. We believe that our HeLa cells were also obtained from ATCC but we do not have original documentation to confirm this - however, we note that STR profiling of our HeLa cells by ATCC showed an exact match with ATCC's reference profile for HeLa cells [ATCC#: CCL-2; see Authentication below].

- Authentication: We obtained STR profiling by ATCC for all 4 cell lines. HEK293T, K562, and HeLa cells showed an exact match with ATCC's reference profiles. Our U2OS cells were a similar match to ATCC# HTB-96. When compared to the reference profile, our cells showed a gain of the #8 allele at the DSS818 locus.

- Mycoplasma contamination: Supernatant was analyzed every 4 weeks using MycoAlert PLUS (Lonza). Cells continuously tested negative.

- Commonly misidentified lines (See ICCLAC register): HEK293T, U2OS, K562, and HeLa are not listed as misidentified cell lines in the ICCLAC register (version 10, March 2020).