

Efficient, continuous mutagenesis in human cells using a pseudo-random DNA editor

Haiqi Chen¹, Sophia Liu^{1,2,3,11}, Samuel Padula^{1,11}, Daniel Lesman¹, Kettner Griswold^{4,5,6,7}, Allen Lin^{8,9}, Tongtong Zhao^{1,10}, Jamie L. Marshall¹ and Fei Chen^{1*}

Here we describe TRACE (T7 polymerase-driven continuous editing), a method that enables continuous, targeted mutagenesis in human cells using a cytidine deaminase fused to T7 RNA polymerase. TRACE induces high rates of mutagenesis over multiple cell generations in genes under the control of a T7 promoter integrated in the genome. We used TRACE in a MEK1 inhibitor-resistance screen, and identified functionally correlated mutations.

Methods for studying the dynamics of eukaryotic cells, such as directed evolution, lineage tracing, and molecular recording, depend on developing tools for targeted, continuous mutagenesis¹. However, existing tools rely on non-physiological environments, rapidly saturate mutagenized sites, or have been adapted only in bacterial or yeast systems^{2–8}. Phage-assisted continuous evolution, for example, evolves in non-native environments, which may result in mutations that are less suited to physiological systems than to those in which they evolved³. Tools using non-homologous end joining-based DNA diversification from Cas9 cleavage saturate quickly after protospacer adjacent motif sequences are destroyed³, and make continuous, self-recurring mutagenesis impossible. Moreover, CRISPR–Cas9 systems such as CRISPR–X⁴ and EvolvR⁵ target narrow genomic windows near to the sgRNA-binding site, or require the design, synthesis, and cellular delivery of numerous sgRNAs that tile the regions of interest. Although a T7 RNA polymerase (T7 RNAP)-deaminase fusion approach was recently used for directed evolution in *Escherichia coli*⁶, and longer editing regions have been demonstrated^{6,7}, an editor system that is efficient in inducing continuous nucleotide diversification in eukaryotic cells, especially in human cells, has not been demonstrated.

Here we demonstrate TRACE, a system for continuous, targeted mutagenesis in eukaryotic cells that combines the DNA processivity of bacteriophage DNA-dependent RNAPs with the somatic hypermutation capability of cytidine deaminases. Bacteriophage RNAPs transcribe DNA sequences under the control of a specific promoter without auxiliary transcription factors⁹. In particular, the T7 RNAP-T7 promoter system can serve as an orthogonal gene expression system in mammalian cells^{10,11}. We reasoned that by combining T7 RNAP with a cytidine deaminase, TRACE could continuously diversify DNA nucleotides downstream of a T7 promoter (Fig. 1a). To test this hypothesis, we devised a dual-plasmid system (pTarget and pEditor), in which the pTarget contained an enhanced green fluorescent protein (*EGFP*) gene downstream of a T7 promoter and

the pEditor contained the T7 RNAP-cytidine deaminase fusion gene with a nuclear localization signal (Fig. 1b). Two variants of the cytidine deaminase, rat APOBEC1 (apolipoprotein B mRNA editing enzyme, catalytic polypeptide 1) and a hyperactive mutant of activation-induced cytidine deaminase, (AID*Δ)^{4,12}, were chosen for the pEditor. We also tested variants containing a uracil DNA glycosylase inhibitor (UGI), which has been shown to facilitate C:G>T:A mutations¹², fused to the 3' end of the T7 RNAP (Fig. 1b).

We first confirmed that T7 RNAP-deaminase fusions maintain the transcriptional activity of T7 RNAP (Supplementary Fig. 1). We then tested the ability of the T7 RNAP-deaminase fusion to induce mutations within a targeted region. HEK293T cells transfected with both pTarget and pEditor were collected 3 d after transfection. DNA from pTarget plasmids was extracted, and a window of ~2,000 base pairs (bp) downstream of the T7 promoter was amplified by PCR for sequencing (Fig. 1c and Methods). The most efficient variant, pAID-T7-UGI, showed an average C>T mutation rate of 1.47 mutations per 1,000 bp (kb⁻¹) and an average G>A mutation rate of 3.1 kb⁻¹ (Fig. 1d). Both C>T and G>A substitutions were observed in the data, which suggested that there was no obvious mutational strand bias. AID constructs (pAID-T7-UGI, and pAID-T7) exhibited higher enzymatic activity than did APOBEC, as shown by the average mutation rate (0.76 kb⁻¹ C>T and 0.7 kb⁻¹ G>A in pAPOEC-T7; 0.8 kb⁻¹ C>T and 0.74 kb⁻¹ G>A in pAPOBEC-T7-UGI; Fig. 1d). Importantly, cells transfected with only cytidine deaminase (pAPOBEC or pAID) showed C>T and G>A mutation rates similar to that of pT7 (Supplementary Fig. 2a). The rate of mutation to other bases from C and G was comparable between conditions (Supplementary Fig. 2b). We expect the base substitution rate measured in the absence of editor (pT7) to result from Illumina sequencing errors¹³. To examine the extent of off-target effects of TRACE in the genome beyond this measurement noise, we developed a barcoded, transposase-based, targeted sequencing approach which allows for unique molecular identifier (UMI)-based correction of sequencing errors (Supplementary Fig. 3a and Methods), with a detection lower limit of 10⁻⁶ mutations per bp. The measured genomic C>T and G>A substitution rate was in the order of 10⁻⁶ bp⁻¹ after 6 d of diversification for control groups (pT7 and pAID-UGI) as well as for pAID-T7-UGI, and no statistical significance was observed between the control and experimental conditions (Dunnett's test, pAID-T7-UGI versus

¹Broad Institute of MIT and Harvard, Cambridge, MA, USA. ²Biophysics Program, Harvard University, Boston, MA, USA. ³Harvard-MIT Division of Health Sciences and Technology, Massachusetts Institute of Technology, Cambridge, MA, USA. ⁴MIT Media Lab, Massachusetts Institute of Technology, Cambridge, MA, USA. ⁵Wyss Institute for Biologically Inspired Engineering, Harvard University, Boston, MA, USA. ⁶Department of Genetics, Harvard Medical School, Boston, MA, USA. ⁷Charles Stark Draper Laboratory, Cambridge, MA, USA. ⁸Department of Systems Biology, Harvard Medical School, Boston, MA, USA. ⁹Ragon Institute of MGH, MIT and Harvard, Cambridge, MA, USA. ¹⁰Department of Stem Cell and Regenerative Biology, Harvard University, Cambridge, MA, USA. ¹¹These authors contributed equally: Sophia Liu, Samuel Padula. *e-mail: chenf@broadinstitute.org

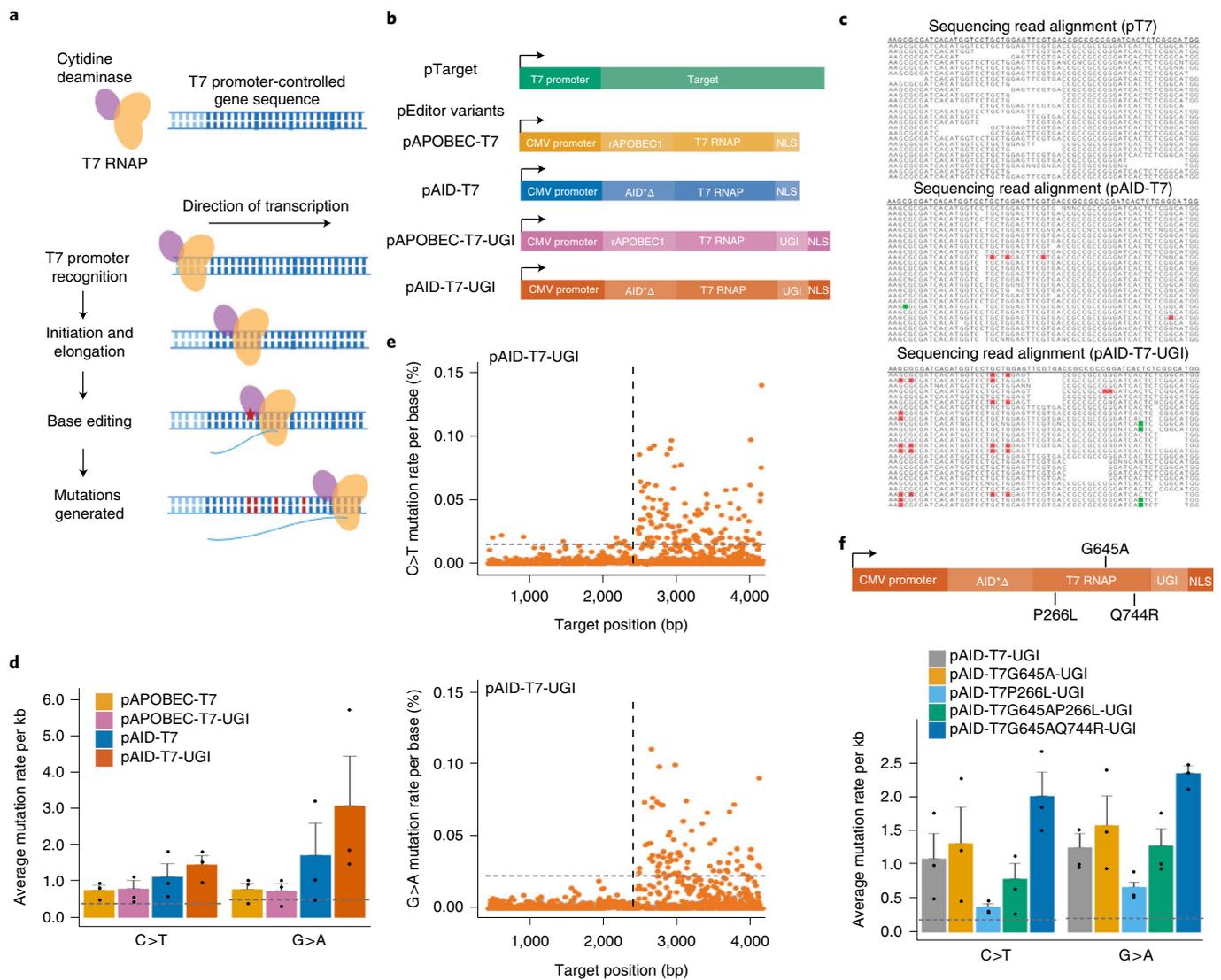


Fig. 1 | TRACE enables highly efficient targeted mutagenesis. **a**, Schematic of TRACE. The recombinant fusion of cytidine deaminase and T7 RNAP recognizes a T7 promoter inserted upstream of a target gene. As the T7 RNAP transcribes DNA, the deaminase introduces point mutations (red star). **b**, Constructs tested in **c–e**. **c**, Representative high-throughput sequencing reads after diversification. C>T (green) and G>A (red) mutations are highlighted. **d**, Mean C>T and G>A mutation rate of the target region. Dashed line, mean sequencing error rate. **e**, Mutation rate per base across an ~4-kb region. Vertical dashed line, end position of T7 promoter. Horizontal dashed line, mean sequencing error rate. **f**, Tested T7 RNAP mutations (top) and their effect on mutation rate (bottom). Dashed line, mean sequencing error rate. All bars in the figure (**d** and **f**) represent mean \pm s.e.m., $n=3$ independent experiments. NLS, nuclear localization signal.

pT7, $P=0.2125$ for C>T; $P=0.2790$ for G>A). By contrast, TRACE induced on-target (T7 promoter driven) C>T and G>A mutation rates in the order of 10^{-3} bp $^{-1}$ (Supplementary Fig. 3b).

Next, we developed TRACE to diversify targeted gene loci downstream of an integrated T7 promoter within the human genome. We observed a significant increase in mutation rates within a window of ~2,000bp downstream of the T7 promoter with pAID-T7-UGI (number of bases with above-noise mutation rate, 112 of 459 Cs and 83 of 406 Gs downstream; 7 of 407 Cs and 3 of 545 Gs upstream; Fig. 1e and Supplementary Fig. 4a). For pAID-T7-UGI, the average C>T and G>A mutation rate downstream of the T7 promoter was significantly higher than that of the upstream region (Supplementary Fig. 4a). Examination of the rate of mutation to T and to A in the T7 promoter sequence revealed an elevation of the C>T mutation rate at base -5 C in the pAID-T7-UGI condition, reaching on average approximately 0.5% of bases (Supplementary Fig. 4b and Supplementary Table 1). Although this substitution has

been previously shown to weaken the T7 promoter binding by T7 RNAP¹⁴, these data suggest that functional copies of the T7 promoter remain largely intact, with a potential loss at a rate of approximately 1 in 200 copies over 6 d.

We then reasoned that engineering of either the elongation rate or the processivity of T7 RNAP could tune the probability of a cytidine deaminase–DNA interaction, and thus tune the editing rate of TRACE. To this end, we tested three mutations (P266L, G645A and Q744R) within the wild type T7 RNAP based on previous studies^{15–17} (Fig. 1f). The G>A mutation rate in pAID-T7G645AQ744R-UGI was significantly higher than that of the wild type (two-sided t -test, $P=0.0051$), whereas the mutation rate of pAID-T7P266L-UGI was significantly lower than that of the wild type (two-sided t -test, $P=0.0424$; Fig. 1f). These results demonstrate the tunability of TRACE editing rates by T7 RNAP mutants.

Next, we asked whether TRACE could induce continuous nucleotide diversification over multiple cellular generations. A HEK293T

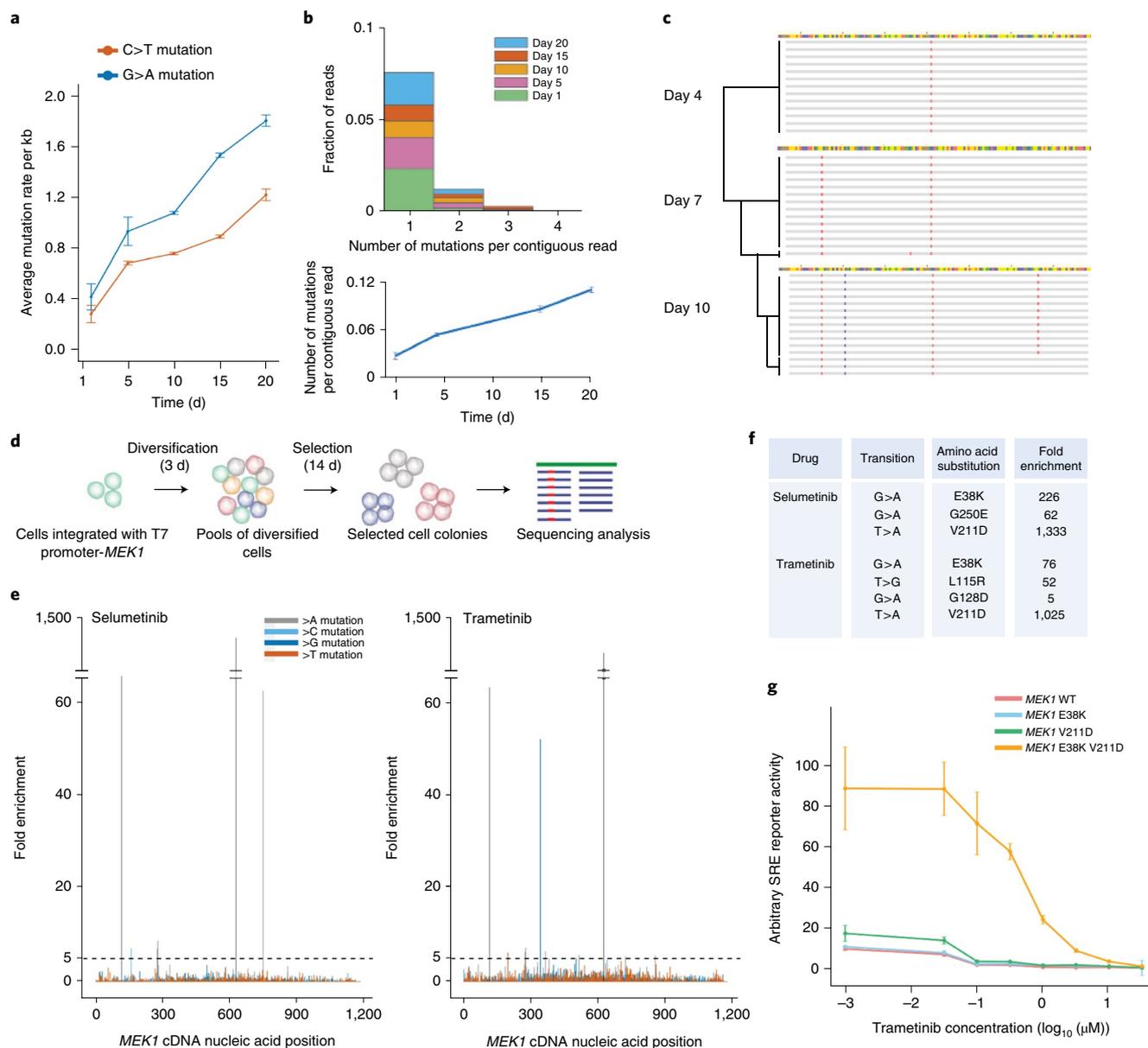


Fig. 2 | TRACE enables continuous somatic mutations in targeted gene loci and identification of correlated MEK1 inhibitor-resistance mutations. a, C>T and G>A mutations over 20 d with integrated T7 promoter-target (mean \pm s.e.m., $n=3$ independent experiments). **b**, Top, distribution of the number of mutations per contiguous read over 20 d. Bottom, mean mutations per contiguous read over time points (mean \pm s.e.m., $n=3$ independent experiments). **c**, Sequencing read alignment of a unique barcoded target detected over three time points. Dendrograms showing the hierarchical relationship of the sequencing reads across time points are on the left of each alignment. A, Red; T, Green; C, Blue; G, Yellow. **d**, Workflow of TRACE MEK1 inhibitor-resistance screen. **e**, Fold enrichment of mutations across MEK1 cDNA sequence in selumetinib-treated (left) and trametinib-treated (right) samples. Dashed line, fivefold enrichment. **f**, Summary of identified mutations in TRACE MEK1 inhibitor-resistance screen. **g**, MAPK-ERK signaling activity as measured by luciferase SRE reporter activity for E38K, V211D and E38KV211D mutants (mean \pm s.e.m., $n=3$ independent experiments).

clone with the T7 promoter-controlled target integrated into the genome was transfected with pAID-T7-UGI in a backbone containing the SV40 origin (Methods). The TRACE-induced mutation rate over 20 d exhibited a marked, monotonic, continuous increase in mutation rate (Fig. 2a and Supplementary Table 2). Examination of the distribution of mutations per read revealed that the average number of mutations per contiguous Illumina sequencing read increased monotonically over time points (Fig. 2b, $P<0.01$, one-way ANOVA). These data suggest that TRACE is able to induce continuous nucleotide diversification across 20 cell generations (~ 1 d per generation). In addition, using barcoded lentiviral TRACE templates, we demonstrated mutation accumulation over time

within the same molecular lineage (Fig. 2c and Supplementary Fig. 5). Reads that share a unique lentiviral barcode also share private clonal and hierarchical sub-clonal mutations, which accumulate over time; this suggests the possibility of future applications of TRACE for lineage tracing. Moreover, using tetracycline-inducible expression of pEditor, we demonstrated dynamic control of editing over 20 d (Supplementary Fig. 6a and Supplementary Table 3). Automated live or dead cell quantification showed no significant difference in cell viability between cells with induced expression of TRACE compared to uninduced cells not expressing TRACE (Supplementary Fig. 6b).

Finally, we applied TRACE to perform functional mutagenesis in mammalian systems in two independent contexts. First, TRACE

was used to shift the fluorescence spectra of blue fluorescent protein (BFP). A single H66Y amino acid substitution causes a shift in the fluorescence excitation and emission spectra of BFP to that of green fluorescent protein (GFP)¹⁸ (Supplementary Fig. 7a). We integrated the *BFP* gene under the control of a T7 promoter into the HEK293T cell genome, and transfected with pEditor variants. After 3 d, automated cell counting was used to assay the ratio of GFP-positive cells to BFP-positive cells (representative images in Supplementary Fig. 7a). Overall, a significant percentage of BFP sequences were converted to GFP in pAID-T7 and pAID-T7-UGI groups (Supplementary Fig. 7b and counts in Supplementary Fig. 7c). Second, we applied TRACE to screen for mutations within mitogen-activated protein kinase kinase 1 (MEK1 kinase, also known as MAP2K1) that promote resistance to pharmacologic inhibition in mammalian cells. We diversified a T7 promoter-targeted *MEK1* open reading frame integrated into the genome of A375 cells for 3 d and then selected with two MEK1 inhibitors, selumetinib and trametinib (Fig. 2d). We identified surviving *MEK1* mutants using high-throughput sequencing (Methods; Fig. 2e,f), summarized in Fig. 2f and described in Supplementary Table 4. Two mutations (E38K and V211D) conferred resistance to both selumetinib and trametinib, which suggests that a similar resistance mechanism or mechanisms may be involved. To investigate whether mutations are correlated beyond the read lengths obtained by Illumina sequencing, we examined the mutation profiles of individual *MEK1* molecules from trametinib-resistant cells via Sanger sequencing (Supplementary Fig. 8). We found that 57 of 152 single DNA clones sequenced had more than two mutations in the *MEK1* molecule. In particular, E38K mutants existed only as double mutants with V211D across several *MEK1* molecules, and not vice versa (Supplementary Fig. 8), which suggests the occurrence of a potential stepwise mutation fixation under selection. To examine the functional correlation between E38K and V211D, we performed a luciferase serum response element (SRE) reporter assay to test the effects of the two mutations, both individually and combined, on the activity of the MAPK–ERK signaling pathway. Notably, MEK1(E38K V211D) induced significantly higher reporter activity than did MEK1 wild type, MEK1(E38K) or MEK1(V211D) alone (Fig. 2g and Supplementary Table 5), which suggests a significant elevation in MAPK–ERK signaling pathway activity as a mechanism of trametinib-resistance for the MEK1(E38K V211D) mutant. To our knowledge, no functional correlation between MEK1(E38K) and MEK1(V211D) has been characterized. In contrast with mutation library screening methods, TRACE, when combined with long-read sequencing, can readily identify correlated mutations.

In summary, we demonstrate that TRACE generates nucleotide diversity within the human genome at average C>T and G>A mutation rates in a tunable range from ~0.5 kb⁻¹ to 4 kb⁻¹ within a week. The high editing rate and large editing window (at least 2,000 bp) exhibited by TRACE make it possible to target long genomic regions over multiple cellular generations, such that this method is on a par with, or exceeds, other technologies (Supplementary Table 6). Although insertion of the T7 promoter site is required, we anticipate that TRACE can diversify multiple genomic loci without disruption of reading frames, by avoiding the insertions and deletions that are observed with other DNA editors^{19,20}. TRACE can also be engineered for dynamic control, and although further testing of the long-term off-target effects when integrated into animal models is needed, TRACE neither elevates global mutation rates within 6 d (with a current detection limit at 10⁻⁶ bp⁻¹) nor observably affects cell health over 20 d.

In the future, the base-editing profile of the system can be expanded by using other base-editing enzymes²¹, and orthogonal bacteriophage polymerase systems (for example, SP6 RNAP) may permit differential editing on multiple loci. Furthermore, we anticipate that TRACE is well suited to serve as a long-term cellular

recorder owing to its continuity, its ability to be engineered, and its wide recording window. For these reasons, we envision TRACE as an engineerable and generalized platform for nucleotide diversification in mammalian systems.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41587-019-0331-8>.

Received: 10 December 2018; Accepted: 28 October 2019;

Published online: 16 December 2019

References

- Farzadfard, F. & Lu, T. K. Emerging applications for DNA writers and molecular recorders. *Science* **361**, 870–875 (2018).
- Esvelt, K. M., Carlson, J. C. & Liu, D. R. A system for the continuous directed evolution of biomolecules. *Nature* **472**, 499–503 (2011).
- Su, T. et al. A CRISPR-Cas9 assisted non-homologous end-joining strategy for one-step engineering of bacterial genome. *Genome Sci. Rep.* **6**, 37895 (2016).
- Hess, G. T. et al. Directed evolution using dCas9-targeted somatic hypermutation in mammalian cells. *Nat. Methods* **13**, 1036–1042 (2016).
- Halperin, S. O. et al. CRISPR-guided DNA polymerases enable diversification of all nucleotides in a tunable window. *Nature* **560**, 248–252 (2018).
- Moore, C. L., Papa, L. J. 3rd & Shoulders, M. D. A processive protein chimera introduces mutations across defined DNA regions in vivo. *J. Am. Chem. Soc.* **140**, 11560–11564 (2018).
- Alexander, D. L. et al. Random mutagenesis by error-prone pol plasmid replication in *Escherichia coli*. *Methods Mol. Biol.* **1179**, 31–44 (2014).
- Ravikumar, A., Arzumanyan, G. A., Obadi, M. K. A. & Liu, C. C. Scalable, continuous evolution of genes at mutation rates above genomic error thresholds. *Cell* **175**, 1–12 (2018).
- Chamberlin, M., Kingston, R., Gilman, M., Wiggs, J. & deVera, A. Isolation of bacterial and bacteriophage RNA polymerases and their use in synthesis of RNA in vitro. *Methods Enzymol.* **101**, 540–568 (1983).
- Lieber, A., Kiessling, U. & Strauss, M. High level gene expression in mammalian cells by a nuclear T7-phase RNA polymerase. *Nucleic Acids Res.* **17**, 8485–8493 (1989).
- Ghaderi, M. et al. Construction of an eGFP expression plasmid under control of T7 promoter and IRES Sequence for assay of T7 RNA polymerase activity in mammalian cell lines. *Iran. J. Cancer Prev.* **7**, 137–141 (2014).
- Komor, A. C., Kim, Y. B., Packer, M. S., Zuris, J. A. & Liu, D. R. Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. *Nature* **533**, 420–424 (2016).
- Schirmer, M. et al. Insight into biases and sequencing errors for amplicon sequencing with the Illumina MiSeq platform. *Nucleic Acids Res.* **43**, e37 (2015).
- Imburgio, D., Rong, M., Ma, K. & McAllister, W. T. Studies of promoter recognition and start site selection by T7 RNA polymerase using a comprehensive collection of promoter variants. *Biochemistry* **39**, 10419–10430 (2000).
- Guillerez, J., Lopez, P. J., Proux, F., Launay, H. & Dreyfus, M. A mutation in T7 RNA polymerase that facilitates promoter clearance. *Proc. Natl Acad. Sci. USA* **102**, 5958–5963 (2005).
- Bonner, G., Lafer, E. M. & Sousa, R. Characterization of a set of T7 RNA polymerase active site mutant. *J. Biol. Chem.* **269**, 25120–25128 (1994).
- Boulin, J. C. et al. Mutants with higher stability and specific activity from a single thermosensitive variant of T7 RNA polymerase. *Protein Eng. Des. Sel.* **26**, 725–734 (2013).
- Glaser, A., McColl, B. & Vadolas, J. GFP to BFP conversion: a versatile assay for the quantification of CRISPR/Cas9-mediated genome editing. *Mol. Ther. Nucleic Acids* **5**, e334 (2016).
- Jakociunas, T., Pedersen, L. E., Lis, A. V., Jensen, M. K. & Keasling, J. D. CasPER, a method for directed evolution in genomic contexts using mutagenesis and CRISPR/Cas9. *Metab. Eng.* **48**, 288–296 (2018).
- Spanjaard, B. et al. Simultaneous lineage tracing and cell-type identification using CRISPR-Cas9-induced genetic scars. *Nat. Biotechnol.* **36**, 469–473 (2018).
- Gaudelli, N. M. et al. Programmable base editing of A>T to G>C in genomic DNA without DNA cleavage. *Nature* **551**, 464–471 (2017).

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature America, Inc. 2019

Methods

Design and construction of pTarget and pEditor plasmids. Plasmids and primers used in this report are listed in Supplementary Data Set 1. pcDNA3.1(+)-IRES-GFP was a gift from K. L. Collins (Addgene plasmid 51406)²². pCMV-BE3 was a gift from D. Liu (Addgene plasmid 73021)¹². pGH335_MS2-AID* Δ -Hygro was a gift from M. Bassik (Addgene plasmid 85406)⁴. pHR-EF1 α -puro-T2A-Tet-on 3G-TRE3G-Ascl1 was a gift from S. Qi (Addgene plasmid 118593)²³. Lenti_CMV_T-IR, Lenti_PAX2, and Lenti_VSVg were gifts from J. L. Marshall. pLX-TRC313 was obtained from the Broad Institute Genetic Perturbation Platform. T7 RNAP and T2A-tdTomato were ordered as gBlocks from Integrated DNA Technologies (IDT). T7 promoter-*MEK1* complementary DNA was ordered as a gene fragment from Genewiz. To generate pAPOBEC-T7 and pAPOBEC-T7-UGI, the Cas9 (D10A) in the pCMV-BE3 construct was replaced with T7 RNAP by Gibson assembly. The original T7 promoter was deleted to avoid self-editing. To generate pAID-T7 and pAID-T7-UGI, rat APOBEC1 in pAPOBEC-T7 and pAPOBEC-T7-UGI was replaced with AID* Δ amplified from pGH335_MS2-AID* Δ -Hygro. For the pTarget plasmid, the T7 promoter-GFP fragment was amplified from pcDNA3.1(+)-IRES-GFP and was subcloned into a pUC19 backbone. This fragment was also subcloned into Lenti_CMV-T-IR to generate Lenti_CMV_T7_GFP-T-IR. A pTarget plasmid without the T7 promoter was also cloned as a negative control. The BFP fragment was generated from the GFP sequence via site-directed mutagenesis and was inserted into the pLX-TRC313 backbone. pAID-T7G645A-UGI, pAID-T7P266L-UGI, pAID-T7P266LG645A-UGI, and pAID-T7G645AQ744R-UGI were cloned via site-directed mutagenesis using wild type pAID-T7-UGI as a template. The AID-T7G645AQ744R-UGI fragment was subcloned into the pcDNA3.1(+)-backbone via Gibson assembly to generate pcDNA3.1(+)-AID-T7G645AQ744R-UGI. The T2A-tdTomato sequence was inserted at the 3' end of the UGI to generate pcDNA3.1(+)-AID-T7G645AQ744R-UGI-T2A-tdTomato. The AID-T7 fragment was subcloned into pHR-EF1 α -puro-T2A-Tet-on 3G-TRE3G-Ascl1 via Gibson assembly to generate pHR-EF1 α -puro-T2A-Tet-on 3G-TRE3G-AID-T7. All plasmid sequences were verified using Sanger sequencing. All cloning primers were ordered from IDT. Plasmids were extracted using a Qiaprep Spin Miniprep Kit or Plasmid Plus Midi Kit (Qiagen).

Cell culture and plasmid transfection. HEK293T cells (ATCC) were grown in high-glucose (4.5 g l⁻¹) DMEM medium supplemented with GlutaMAX, 1 mM sodium pyruvate, 10% FBS, 100 units ml⁻¹ of penicillin and 100 μ g ml⁻¹ of streptomycin in a humidified chamber with 5% CO₂ at 37°C. Cells were maintained at ~80% confluence in 24-well plates on the day of transfection. For dual-plasmid transfection, 250 ng of pTarget and 250 ng of pEditor plasmids were mixed together with 1 μ l of *TransIT-X2* reagent (Mirus) and the mixture was incubated in 50 μ l of Opti-MEM (Thermo Fisher Scientific) for 30 min. The mixture was then added drop-wise to each well. Alternatively, target-integrated single cell clones were cultured in 12-well plates and were transfected with 1,000 ng of pTarget plasmids.

Lentivirus production and generation of single cell clones. Three million HEK293T cells were cultured in 10 ml of culture medium in a 10 cm dish. Cells were transfected with 12 μ g of transfer plasmids, 9 μ g of Lenti_PAX2 and 3 μ g of Lenti_VSVg. Culture medium was replaced 24 h after transfection with 6 ml of high-glucose (4.5 g l⁻¹) DMEM supplemented with GlutaMAX, 1 mM sodium pyruvate, 30% FBS, 100 units ml⁻¹ of penicillin and 100 μ g ml⁻¹ of streptomycin. Supernatant containing viral particles was filtered through 0.45 μ m filters 24 h later. To generate the target-integrated HEK293T single cell clone, 0.5 \times 10⁶ HEK293T cells in a 6-well plate with 2.5 ml of culture medium received, drop-wise, 500 μ l of viruses and 8 μ g ml⁻¹ of polybrene. Two days after transduction, successfully integrated cells were selected by puromycin at a concentration of 1.5 μ g ml⁻¹. Integrated cells were subjected to fluorescence-activated cell sorting (FACS) in single cell format into 96-well plates using a MoFlo Astrios EQ Cell Sorter (Beckman Coulter) and single cells were allowed to expand to form colonies. HEK293T cells were transduced with pLX-TRC313-BFP and pHR-EF1 α -puro-T2A-Tet-on 3G-TRE3G-AID-T7 virus and were selected by puromycin at a concentration of 1.5 μ g ml⁻¹ and hygromycin at a concentration of 200 μ g ml⁻¹ to generate fully TRACE-integrated cell lines. To generate T7 promoter-*MEK1*-integrated A375 cells, cells were selected by hygromycin at a concentration of 200 μ g ml⁻¹.

Long-term monitoring of the accumulation of mutations. For the experiment shown in Fig. 2a, pEditor variant pAID-T7-UGI in an SV40 origin of replication-containing pcDNA3.1(+)-backbone was used to enable HEK293T cells to replicate the plasmid. A T2A-tdTomato sequence was also fused to the 3' end of the editor sequence to monitor transfection efficiency and plasmid DNA replication. Target-integrated single cell clones were cultured in 12-well plates and were transfected with 1,000 ng of pTarget plasmids. Half of the cell population was collected at the indicated time points, and the other half was passaged to a new 12-well plate. Red fluorescence (tdTomato) was visually monitored daily. A second transfection was performed on day 11. Genomic DNA (gDNA) samples from day 1 (before diversification), days 5, 10, 15, and 20 were obtained and the T7 promoter-controlled target (~2,000 bp) was sequenced.

For the experiment shown in Supplementary Fig. 6, a tetracycline-inducible promoter-controlled pEditor variant pAID-T7 and the T7 promoter-controlled target were stably integrated into the HEK293T cell genome. Doxycycline (Dox, an analog of tetracycline) was added into the culture medium (1 μ M) to activate the TRACE system. The TRACE-integrated HEK293T cells were cultured in bulk with or without Dox for 20 d. gDNA samples from day 1 (before diversification), days 6, 10, 16 and 20 were obtained and the T7 promoter-controlled target (~800 bp) was sequenced. Cell viability was assessed by the dye exclusion test using 0.4% trypan blue solution (Thermo Fisher Scientific) and a Cellometer Auto T4 (Nexcelom Bioscience).

Fluorescence microscopy and image analysis. T7 promoter-BFP-integrated HEK293T cells were seeded in a 24-well glass-bottom plate and transfected with pEditor. Cells were imaged using an inverted Nikon CSU-W1 Yokogawa spinning disk confocal microscope with 488 nm (GFP) and 405 nm (BFP) lasers, an air objective (Plan Apo λ , 4 \times or 20 \times , Nikon, or Plan Fluor 40 \times , Nikon), and an Andor Zyla sCMOS camera. NIS-Elements AR software (v4.30.01, Nikon) was used for image capture. Images were processed using ImageJ (National Institutes of Health (NIH)). CellProfiler (version 3.1.5, Broad Institute)²⁴ was used for segmentation and for counting BFP- and GFP-positive cells. GFP-positive cells were further thresholded by Otsu's method using integrated intensity with the R package `autothresholdr`²⁵.

On-target mutation analysis. To sequence the targeted region (~2,000 bp) on pTarget, plasmids were extracted from ~1 million cells using a Qiaprep Spin Miniprep Kit. PCR was performed using Phusion U Hot Start PCR master mix (2 \times , Thermo Fisher Scientific). Approximately 100 ng of those plasmids was used as a template (primer sequences are shown in Supplementary Data Set 1) in a final volume of 40 μ l. The following program was used: 1 cycle of 98°C for 30 s, 25 cycles of 98°C for 10 s, 60°C for 30 s, 72°C for 1 min, and 1 cycle of 72°C for 5 min. Magnetic Ampure XP beads (Beckman Coulter) were added to samples at a 0.8:1 ratio to size select for the PCR fragments. Purified PCR products were eluted in 25 μ l of water. The concentration of each sample was measured by Qubit (Thermo Fisher Scientific). From each sample, 0.5 ng of DNA in a volume of 2.5 μ l was used as input for the subsequent library preparation. The sequencing library was prepared following the Nextera XT Kit protocol (Illumina), with half of the recommended amount of each reagent used. Sequencing was performed on a MiSeq (Illumina) with paired-end reads (read1, 100 bp; index1, 8 bp; index2, 8 bp; read2, 100 bp).

To sequence the targeted genomic loci, genomic DNA was extracted from ~1 million cells using the Quick-DNA Kit (Zymo Research). Approximately 75 ng of extracted genomic DNA was used as input. PCR was performed using Phusion U Hot Start PCR master mix with primers shown in Supplementary Data Set 1. DMSO (0.2 μ l) was added to the PCR mix to make a final volume of 20 μ l. The following program was used: 1 cycle of 98°C for 30 s, 25 cycles of 98°C for 10 s, 60°C for 30 s, 72°C for 2 min, and 1 cycle of 72°C for 5 min. Ampure XP beads were added to samples at a 0.8:1 ratio to size select for the PCR fragments. Purified PCR products were eluted in 10 μ l of water. The concentration of each sample was measured by Qubit and the same Nextera XT Kit protocol as mentioned above was used to prepare the sequencing library. Sequencing was performed on a MiSeq (Illumina) with paired-end reads (read1, 100 bp; index1, 8 bp; index2, 8 bp; read2, 100 bp).

To calculate the mutation rate, ~1 million reads were produced for each sample. Illumina sequencing adapters were trimmed during sample demultiplexing using `bc12fastq2` (v2.19.1). Bases in each read with an Illumina quality score lower than 25 were filtered. Alignment on the respective reference sequence was performed using `Bowtie 2` (v2.2.4.1)²⁶. Alignment files were generated in BAM format and were visualized in `Geneious` (v11.1.5). The mutation enrichment at each base was calculated with custom MATLAB scripts. The first and last 15 bases of each read aligned to the 5' and 3' ends of the reference sequence were trimmed to filter low coverage and low alignment quality reads due to the limits of Nextera tagmentation. Bases with a read count of less than 100 were also excluded from the analysis. We calculated the transitions, transversions, and indels observed at each position and plotted the C>T and G>A mutation profiles, respectively, for each sample. The mutation rate per base data were obtained by dividing the number of reads with mutations over the number of total reads at each base. The average mutation rate for each possible combination of base switching for each sample was calculated by averaging the mutation rate per base across the targeted region. The pT7 sample was used to estimate the measurement noise that was introduced through sample preparation and Illumina sequencing. To plot the mutations per read distribution shown in Fig. 2b, read 1 aligning to the targeted *EGFP* region was examined and the number of mutations for each read was counted.

Off-target mutation analysis. To test whether TRACE edits certain sites of the genome in preference to the target region²⁷, we studied several housekeeping genes (*GAPDH*, β -actin, and α -tubulin) and a non-coding RNA (*Malat1*). Gene-specific reverse primers were designed to target the above-mentioned genomic regions (Supplementary Data Set 1). The entire genome was first tagged into DNA fragments with varying lengths using custom Tn5 (see below). A read 1

adapter sequence and a 6-base UMI were appended to one end of all of the DNA fragments, which were then PCR amplified using a universal forward primer and gene-specific reverse primers. Because all gene-specific reverse primers are ~20 bp in length, regions in other chromosomes sharing sequence homology with the primers were also amplified. Thus, the final sequencing library covers additional genomic regions to the initial seven regions that were targeted. The details of the sequencing library preparation and analysis are described below.

First, a read 1 adaptor was generated by annealing a forward strand sequence with a complementary mosaic end sequence (Supplementary Data Set 1). This equimolar mixture was heated at 85 °C for 3 min and then cooled to 20 °C over approximately 1–1.5 h (1% ramp). The annealed read 1 adaptor was diluted with 100% glycerol at a 1:1 ratio. Second, 100 µl of Tn5 dilution buffer was prepared by mixing 50 µl of 100% glycerol, 5 µl of 1 M Tris pH 7.5, 2 µl of 5 M NaCl, 2 µl of 5 mM EDTA, 1 µl of 100 mM DTT, 1 µl of 10% NP-40 and 39 µl of water. Third, 25 µl of unloaded Tn5 enzyme (a gift from the F. Zhang Laboratory) was mixed with 25 µl of Tn5 dilution buffer and 50 µl of annealed read 1 adaptors. The mixture was then incubated at room temperature (22 °C) for 30 min. The loaded Tn5 can be stored at –20 °C and can be used for up to 2 weeks. Fourth, ~100 ng of genomic DNA, collected from HEK293T cells that had been incubated with pEditor variants for 6 d, was incubated with 2.5 µl of loaded Tn5 in 25 µl of Tagment DNA Buffer from the Nextera XT Kit. The final volume was 50 µl. The tagmentation mix was incubated at 55 °C for 5 min. The tagmentation reaction was stopped by performing a column purification using a NucleoSpin Gel and PCR Clean-Up Kit (Takara Bio). Purified DNA fragments were eluted in 40 µl of water. Fifth, 5 µl of tagmented DNAs were amplified by using barcoded universal forward primers and gene-specific reverse primers (Supplementary Data Set 1). Phusion U Hot Start PCR master mix (10 µl) was pre-heated to 85 °C and was then added into the PCR mix. The final volume for each PCR reaction was 20 µl. The following program was used: 1 cycle of 72 °C for 5 min, 98 °C for 30 s, 10 cycles of 98 °C for 10 s, 64 °C for 30 s, 72 °C for 30 s, and 1 cycle of 72 °C for 5 min. Ampure XP beads were added to samples at a 1.8:1 ratio to purify the PCR fragments. Purified PCR products were eluted in 10 µl of water. Sixth, the 10 µl elution products were used for a second PCR in which a barcoded universal reverse primer and the same barcoded universal forward primer were mixed with 20 µl of Q5 Hot Start High-Fidelity 2× Master Mix (NEB). The final volume of the PCR reaction was 20 µl. The following program was used: 1 cycle of 98 °C for 30 s, 15 cycles of 98 °C for 10 s, 64 °C for 30 s, 72 °C for 30 s, and 1 cycle of 72 °C for 5 min. Ampure XP beads were added to samples at a 1.8:1 ratio to purify the PCR fragments. The concentration of each sample was measured by Qubit. Sequencing was performed on a Nextseq (Illumina) with paired-end reads (read1, 75 bp; index1, 8 bp; index2, 8 bp; read2, 75 bp). Custom sequencing primers (Supplementary Data Set 1) were spiked-in according to the manufacturer's instructions.

To quantify the off-target mutation rate, each sample received 1 million paired-end reads. We trimmed 26 bases from the 5' end of read 1, which contained the UMI, the custom read1 adaptor sequence and the mosaic end sequence. We also trimmed 20 bases from the 5' end of read 2, which contained the gene-specific primer binding region, using custom MATLAB and python scripts. Trimmed paired-end reads were then aligned to hg19 using Bowtie2 with the parameter -X2000 to allow fragments of up to 2 kb to align. Reads that mapped to the mitochondria, chromosome Y, and unmapped contigs were removed. From the resulting BAM files, we used mpileup (samtools) to look at positions with reads that had minimum mapping qualities of 30 and minimum Q-scores of 32. Of those positions, any position that mapped to pericentromeric regions (on chromosome 1 and chromosome 7), as well as any position with fewer than 10 reads were removed. UMIs for each read were extracted and were used to correct for sequencing errors; the most common base call for UMIs with at least two matching reads was determined to be the consensus base for that UMI at that position. Positions with <5 bases after UMI deduplication were filtered, and any position with <80% of the deduplicated UMIs with the same base call was filtered out as a SNP. Finally, a mutation rate was calculated at each position by dividing the number of UMI-deduplicated bases that were different from the consensus base by the total number of UMI-deduplicated bases measured. Statistical significance was determined with a Dunnett's test.

Barcoded TRACE target library generation. Ultramers containing 20 bp barcodes (20 Ns; IDT) were annealed with equimolar annealing primers and phosphorylated with T4 polynucleotide kinase (95 °C for 5 min; cooled to 25 °C at ~5% ramp; oligonucleotide sequences are shown in Supplementary Data Set 1). Second strand synthesis of the annealed oligonucleotide was performed using Klenow fragment (NEB) at 25 °C for 30 min. The resulting double-stranded barcoded oligonucleotides were gel purified and inserted into a lentiviral backbone using Gibson assembly (primer sequences are shown in Supplementary Data Set 1). The Gibson assembly products were then purified by isopropanol precipitation and were transformed into Endura electrocompetent cells (Lucigen) using the following parameters: voltage, 1,800 V; capacitance, 10 µF; resistance, 600 Ω; cuvette, 1 mm. Approximately 0.8 × 10⁶ colonies were recovered and plasmids were then extracted. Barcoded lentiviruses were produced as described above and HEK293T cells were transduced at a multiplicity of infection such that no cells received the same barcoded target. Cells were diversified by TRACE for an amount of time indicated

in Supplementary Fig. 5 and gDNAs were extracted. A 900 bp region covering the 20 Ns barcode, the T7 promoter and the T7 promoter-controlled target was amplified by PCR with Phusion U Hot Start PCR master mix (2×, Thermo Fisher Scientific). The following program was used: 1 cycle of 98 °C for 30 s, 10 cycles of 98 °C for 10 s, 56.2 °C for 30 s, 72 °C for 1 min, and 1 cycle of 72 °C for 5 min. Exonuclease I (2 µl, *E. coli*; NEB) was added to the PCR products and the mixture was incubated at 37 °C for 15 min. Magnetic Ampure XP beads were then added to samples at a 0.8:1 ratio to size select for the PCR fragments. Purified PCR products were eluted in 10 µl of water. A second PCR was performed on the purified PCR products to attach the Illumina P5 and P7 handles using KAPA HiFi HotStart ReadyMix PCR Kit (2×, Kapa Biosystems). The following program was used: 1 cycle of 95 °C for 3 min, 15 cycles of 98 °C for 20 s, 64 °C for 15 s, 72 °C for 1 min, and 1 cycle of 72 °C for 1 min (primer sequences are shown in Supplementary Data Set 1). The final PCR products were gel purified. Sequencing was performed on a MiSeq (V3; Illumina) with paired-end reads (read1, 225 bp; index1, 8 bp; index2, 8 bp; read2, 150 bp).

MEK1 inhibitor-resistance screen. A375 cells were a gift from the F. Zhang Laboratory. T7-promoter-controlled *MEK1* cDNA was lentivirally integrated into the genome. Cells were diversified by transfections of pEditor variant pcDNA3.1(+)-AIDT7G645AQ744R-UGI for 3 d. Cells transfected with empty vectors were treated as an undiversified control. Approximately 500,000 cells per well in a 24-well plate were placed under selection with either 1 µM selumetinib or 1 µM trametinib for 2 weeks. After selection, cells were harvested and genomic DNA was extracted. The *MEK1* cDNA containing the mutations was PCR amplified using Phusion U Hot Start Master Mix (primer sequences were shown in Supplementary Data Set 1). Ampure XP beads were added to samples at a 0.8:1 ratio to size select for the PCR fragments. Purified PCR products were eluted in 10 µl of water. The concentration of each sample was measured by Qubit and the Nextera XT Kit protocol was used to prepare the sequencing library. Sequencing was performed on a MiSeq (Illumina) with paired-end reads (read1, 100 bp; index1, 8 bp; index2, 8 bp; read2, 100 bp). The mutation rate for each base of the *MEK1* sequence was calculated for both pre- and post-selection samples. Fold enrichment was calculated by dividing the post-selection mutation rate of a base by the pre-selection mutation rate. Mutations with a mutation rate of more than 1% and fold enrichment of more than 5 were identified. For Sanger sequencing on single *MEK1* molecules, *MEK1* molecules were PCR amplified using gDNA from trametinib-resistant cell colonies and were subsequently cloned into a pLX-TRC313 plasmid vector using Gibson assembly (primer sequences are shown in Supplementary Data Set 1). *MEK1*-containing plasmids were transformed into bacteria and single colonies were picked and mini-prepped.

SRE reporter assay. pcDNA3.1(+)-*MEK1* wild type, pcDNA3.1(+)-*MEK1E38K*, pcDNA3.1(+)-*MEK1V211D*, and pcDNA3.1(+)-*MEK1E38K*V211D were generated using Gibson assembly (primer sequences are shown in Supplementary Data Set 1). The SRE reporter assay was performed using the SRE reporter kit (BPS Bioscience) according to the manufacturer's instructions. In brief, ~40,000 cells in 100 µl of growth medium per well were seeded in 96-well white assay plates. HEK293T cells were transfected with 60 ng of reporter plasmids along with 50 ng of respective *MEK1* plasmids for 6 h before being washed and replenished with 50 µl of trametinib-containing culture medium with 0.5% FBS. After 12 h, cells were washed and incubated with 50 µl of 0.5% FBS-containing culture medium supplemented with recombinant human epidermal growth factor protein (final concentration 10 ng ml⁻¹) for 6 h. Reporter activity was then monitored using a dual luciferase (Firefly-Renilla) assay system (BPS Bioscience) according to the manufacturer's instructions using a Synergy H4 plate reader (BioTek). The ratio between Firefly luminescence intensity and Renilla luminescence intensity was calculated for each well after subtracting background luminescence.

Statistics. The relevant statistical test, sample size, replicate type and *P* values for each figure and table are found in the figure or table and/or the corresponding figure legends or table footnotes.

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

The sequencing data supporting the findings of this study are available in NCBI BioProject database with accession number [PRJNA555784](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA555784).

Code availability

All custom code will be available upon request.

References

- Schaefer, M. R. et al. A novel trafficking signal within the HLA-C cytoplasmic tail allows regulated expression upon differentiation of macrophages. *J. Immunol.* **180**, 7804–7817 (2008).

23. Liu, Y. et al. CRISPR activation screens systematically identify factors that drive neuronal fate and reprogramming. *Cell Stem Cell* **23**, 758–771 (2018).
24. Carpenter, A. E. et al. CellProfiler: image analysis software for identifying and quantifying cell phenotypes. *Genome Biol.* **7**, R100 (2006).
25. Landini, G., Randell, D. A., Fouad, S. & Galton, A. Automatic thresholding from the gradients of region boundaries. *J. Microsc.* **265**, 185–195 (2017).
26. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).
27. Martin, A. & Scharff, M. D. Somatic hypermutation of the AID transgene in B and non-B cells. *Proc. Natl Acad. Sci. USA* **99**, 12304–12308 (2002).

Acknowledgements

We thank Y. Jiang for help with generating figures, and J. Strecker for purified TN5 transposase. F.C. acknowledges support from Eric and Wendy Schmidt as funders of the Schmidt Fellows Program at the Broad Institute. This work was supported by the NIH Director's Early Independence Award (DP5-OD024583) to F.C. H.C. is supported by The Lalor Foundation. S.L. is supported by a Molecular Biophysics Training Grant (NIH/National Institute of General Medical Sciences T32 GM008313) and the National Science Foundation (NSF) Graduate Research Fellowship Program. K.G. is supported under Graduate Fellowships from the Fannie and John Hertz Foundation, and the Charles

Stark Draper Laboratory. A.L. is supported by a Paul and Daisy Soros Fellowship for New Americans and the NSF Graduate Research Fellowship Program.

Author contributions

H.C. and F.C. designed experiments. K.G. designed prototype constructs. H.C., S.L., S.P., D.L., and J.L.M. performed experiments. S.L., A.L., and F.C. contributed computational tools. H.C., S.L., D.L., A.L., T.Z., and F.C. analyzed the data. F.C. conceived of and supervised the study. H.C., S.L., and F.C. wrote the manuscript with input from S.P., D.L., K.G., A.L., T.Z., and J.L.M. All authors reviewed and approved the manuscript.

Competing interests

Some of the authors (H.C., S.L., S.P., K.G. and F.C.) have filed a patent related to this work.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41587-019-0331-8>.

Correspondence and requests for materials should be addressed to F.C.

Reprints and permissions information is available at www.nature.com/reprints.

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

NIS-Elements AR v4.30.01 was used for image capture and Geneious v11.1.5 was used for alignment files visualization.

Data analysis

A description of the software and code has been included in the Methods. Initial post-processing of raw data was performed using the bcl2fastq2 (v2.20.0) software package from Illumina, Bowtie2 (v2.3.4) software package, ImageJ2 and CellProfiler (v3.1.5). Subsequent computational analysis was performed using R package autothreshldr (v1.3.0) as well as custom code written in Matlab R2015a and Python (v3) which will be publicly available upon request. Plots were generated in RStudio (v1.1.463) using ggplot2 (v3.0.0) software package.

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

The raw sequencing data that support the findings of this study are openly available in NCBI BioProject database with an accession number PRJNA555784. . Figure 1-2 and Supplementary Figure S2-S4, S6 have associated raw data.

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](https://www.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	During imaging experiments, the number of cells counted were determined by the number of cells imaged per field of view.
Data exclusions	No data were excluded.
Replication	All experiments were reproduced to reliably support conclusions stated in the manuscript.
Randomization	Cells were assigned randomly to experimental and control groups. During imaging experiments, at least three fields of view were randomly selected for each condition.
Blinding	The investigators were not blinded during data collection. Computational analysis was not performed blinded.

Reporting for specific materials, systems and methods

Materials & experimental systems

n/a	Involvement in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> Unique biological materials
<input checked="" type="checkbox"/>	<input type="checkbox"/> Antibodies
<input type="checkbox"/>	<input checked="" type="checkbox"/> Eukaryotic cell lines
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology
<input checked="" type="checkbox"/>	<input type="checkbox"/> Animals and other organisms
<input checked="" type="checkbox"/>	<input type="checkbox"/> Human research participants

Methods

n/a	Involvement in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> ChIP-seq
<input checked="" type="checkbox"/>	<input type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging

Eukaryotic cell lines

Policy information about [cell lines](#)

Cell line source(s)	HEK293T (ATCC), A375 (ATCC)
Authentication	All cell lines authenticated by STR method.
Mycoplasma contamination	All cell lines were tested negative for mycoplasma contamination.
Commonly misidentified lines (See ICLAC register)	None of the used cell lines is listed in ICLAC database.