SYNTHETIC BIOLOGY

Multiplex recording of cellular events over time on CRISPR biological tape

Ravi U. Sheth,^{1,2} Sung Sun Yim,¹ Felix L. Wu,^{1,2} Harris H. Wang^{1,3*}

Although dynamics underlie many biological processes, our ability to robustly and accurately profile time-varying biological signals and regulatory programs remains limited. Here we describe a framework for storing temporal biological information directly in the genomes of a cell population. We developed a "biological tape recorder" in which biological signals trigger intracellular DNA production that is then recorded by the CRISPR-Cas adaptation system. This approach enables stable recording over multiple days and accurate reconstruction of temporal and lineage information by sequencing CRISPR arrays. We further demonstrate a multiplexing strategy to simultaneously record the temporal availability of three metabolites (copper, trehalose, and fucose) in the environment of a cell population over time. This work enables the temporal measurement of dynamic cellular states and environmental changes and suggests new applications for chronicling biological events on a large scale.

NA is the primary information storage medium in living organisms and can be used in synthetic cellular memory devices that convert biological signals into heritable changes in nucleotide sequences. For example, approaches using recombinases (1-6), single-stranded DNA recombineering (7), and CRISPR-Cas9 (8-12) have been developed to record the level of a biological signal or to track developmental lineage. However, a major outstanding challenge has been the robust recording of temporally varying biological states or signals (e.g., gene expression or metabolite fluctuations) in living cells. Such a biological recording system would have powerful applications in studying dynamic cellular processes, such as complex regulatory programs, or in engineering "sentinel" cells that track changing environmental signals over time.

The bacterial CRISPR-Cas adaptation process exemplifies a naturally occurring biological memory system. When foreign genetic elements such as plasmids and phages invade a cell, short fragments of these exogenous nucleic acids can be captured by CRISPR-Cas adaptation proteins and integrated into genomic CRISPR arrays as spacers (13-15). This spacer acquisition process occurs in a unidirectional manner; new spacers are inserted at the 5' of CRISPR arrays (16, 17) and subsequently can be used by CRISPR-Cas immunity proteins to repel invaders that exhibit matching sequence identity (18). The DNA-writing potential of the adaptation process was recently applied to record the sequence and ordering of chemically synthesized oligonucleotides that were serially electroporated into cell populations (19, 20). However, engineering the CRISPR-Cas adaptation system to directly record biological signals and their temporal context, and not simply sequence information of exogenous DNA, has not been achieved to date.

A tape recorder converts temporal signals such as analog audio into recordable data written onto a tape substrate as it is passed at a set rate across the recorder. Inspired by this temporal data storage scheme (Fig. 1A), we set out to develop a biological realization of the system, which we call temporal recording in arrays by CRISPR expansion (TRACE). In this framework, a biological input signal is first transformed into a change in the abundance of trigger DNA within living cells. The CRISPR-Cas spacer acquisition machinery is then used to record the amount of trigger DNA into CRISPR arrays in a unidirectional manner (Fig. 1B). Through this architecture, the presence of an input signal increases the frequency of trigger spacers incorporated into arrays, which constitutes recording of the positive signal. However, in the absence of a signal, reference spacers can still be acquired into arrays at a background rate from sources other than the trigger DNA, such as the genome (21). These reference spacers serve as pace-denoting markers that are embedded during the recording session, akin to the physical spacing on a tape substrate that represents time intervals.

We first explored an approach to convert the presence of a biological input into an increase in the abundance of trigger DNA within a population of *Escherichia coli* cells. We used a copy number–inducible trigger plasmid (pTrig), which contained a mini-F origin for stable maintenance and the phage P1 lytic replication protein RepL placed downstream of the Lac promoter. In the presence of the test input signal, isopropyl-β-D-1-thiogalactopyranoside (IPTG), transcription from the Lac promoter increases and results in expression of RepL. The RepL protein subsequently initiates plasmid replication from an origin located within the RepL coding sequence (*22*), which in turn increases pTrig copy number (Fig. 1C). Anal-

ysis of pTrig by quantitative polymerase chain reaction (qPCR) revealed a 653 ± 5 -fold increase in copy number in cells induced with IPTG for 6 hours, compared with copy number in cells with no induction (methods, Fig. 1D, and figs. S1 and S2). This demonstrates that a biological signal that elicits a transcriptional response can be coupled to the alteration of an intracellular DNA pool.

Next, we assessed whether an increase in pTrig copy number could be recorded in CRISPR arrays across a cell population. Expression of the CRISPR adaptation proteins Cas1 and Cas2 promotes unidirectional integration of ~33-base pair DNA spacers into genomic CRISPR arrays in E. coli (19, 21, 23). We constructed a recording plasmid (pRec) that expresses Cas1 and Cas2 upon addition of anhydrotetracycline (aTc), which results in spacer acquisition (Fig. 1E and fig. S3A). Cells with pRec or with pRec and pTrig were induced with aTc and with or without IPTG, and their CRISPR arrays were assessed by sequencing to determine the source of newly acquired spacers. either from pRec, pTrig, or the genome (methods; Fig. 1, F and G; and fig. S4). In cells with pRec, spacers were preferentially derived from the pRec plasmid, consistent with enriched spacer acquisition from plasmids in E. coli documented in the literature (21). Cells with pRec and pTrig, but without IPTG induction, resulted in similar spacer acquisitions and low pTrig spacer incorporation (0.23 ± 0.06% of spacers). However, IPTG induction of pTrig increased overall spacer acquisition (fig. S3B) and, more importantly, increased the percentage of pTrig-derived spacers (32.4 \pm 0.4% of spacers). This result demonstrates that an induced increase in trigger DNA abundance can be specifically recorded in CRISPR arrays. We further explored different input IPTG concentrations and observed an increasing relationship between pTrig copy number and the resulting percentage of pTrigderived spacers (fig. S5). Although increased pTrig spacer incorporation could be detected after 4 hours of induction, robust recording was best achieved when the signal persisted for at least 6 hours (fig. S6).

Having assessed the two main components of the system-(i) transformation of a biological signal to increase abundance of intracellular DNA and (ii) capture of the amplified pool into CRISPR arrays-we next tested whether TRACE could be used to record biological signals in the temporal domain. We performed a systematic time-course recording experiment in which cells experienced the presence or absence of IPTG across 4 sequential days, constituting 16 distinct temporal signal profiles (Fig. 2A). Sequencing the resulting CRISPR arrays confirmed an overall expansion of arrays over time (fig. S7), with 24.7 \pm 5.2% of all arrays having incorporated at least one new spacer by day 4. On average, about one in 15 arrays acquired a new spacer each day. As expected, arrays with increasing numbers of spacers were detected with decreasing frequency across the population (Fig. 2B). Because longer arrays contained more temporal information, we additionally

¹Department of Systems Biology, Columbia University, New York, NY, USA. ²Integrated Program in Cellular, Molecular, and Biomedical Studies, Columbia University, New York, NY, USA. ³Department of Pathology and Cell Biology, Columbia University, New York, NY, USA.

^{*}Corresponding author. Email: hw2429@columbia.edu

implemented a size enrichment protocol (methods) that facilitated the analysis of arrays with up to five new spacers (Fig. 2B).

For TRACE to function as a useful biological tape recorder, the spacer identity (reference or trigger) and ordering within CRISPR arrays should correlate with the actual temporal signal profile. We first noted that the system can act as a simple signal counter by observing that the total percentage of pTrig spacers increased proportionally with the number of times the signal was present in the signal profile (Fig. 2C). Next, we analyzed pTrig spacer incorporation and ordering in CRISPR arrays. For example, individual arrays from a sample receiving the IPTG profile [on, on, off, off] were variable but displayed an overall enrichment of pTrig spacers at distal positions in the array (Fig. 2D and fig. S8A). To visualize these incorporation patterns across each of the 16 signal profiles, for arrays of different lengths (L1 to L5), we calculated the population average of pTrig spacers at each spacer position (Fig. 2, D and E, and fig. S8B). These patterns of pTrig frequencies exhibited a high degree of correspondence to their respective temporal signal profiles when considered in reverse (i.e., oldest to newest acquired spacers; Fig. 2F), which suggested the successful recording of temporal biological signals.

To improve the interpretation of TRACE data, we explored a method for accurate and automated inference of the input temporal signal profiles from recorded CRISPR arrays. We hypothesized that the array expansion process could be mod-

eled to vield a useful classification scheme for matching an observed pattern of arrays to its corresponding signal profile. To test this approach, we first defined a cell population's repertoire of CRISPR arrays as a distribution of "array types." Array types constitute all possible array configurations across all array lengths with either reference or trigger spacers occupying each spacer position (Fig. 3A). We then developed a simple analytical model of the CRISPR expansion process for calculating the expected frequencies for all array types given a signal profile (methods). Only four constants are needed to parameterize the model for each array length: the rates of array expansion and pTrig incorporation per recording interval, in the presence or absence of a signal (fig. S9 and table S5). Using this model, we calculated the expected distributions of array types for all 16 temporal signal profiles and compared these distributions of array-type frequencies with those from experimentally recorded arrays. The predicted and observed array-type distributions matched closely (fig. S10). For example, for two signal profiles with an equal number of inductions but different temporal ordering, our models yielded distinctive array-type distributions that appeared to recapitulate the corresponding experimental data (Fig. 3B).

To quantitatively compare and classify the observed data with model array-type distributions, we calculated all pairwise Euclidean distances between them. An observed CRISPR array population was assigned to the most probable signal profile on the basis of the data-model pair with the shortest Euclidean distance (Fig. 3C). Using L1 arrays only, which do not contain any temporal information, only five of 16 signal profiles could be correctly classified. Using array types L2 to L4 individually resulted in much higher accuracy of assignments (13 to 14 of 16 correct). When array types L2 to L4 were used together, we could perfectly classify all 16 populations with their correct temporal signal profiles (methods and Fig. 3D). Only a few hundred arrays of a given length, corresponding to minimum populations of $\sim 10^5$ total arrays, were required to recapitulate reasonable classification accuracy (fig. S11). This demonstrates that temporal signals can be recorded and subsequently reconstructed with high accuracy from CRISPR arrays by using a simple model of the expansion process.

Beyond simply assigning spacer identity as reference or trigger, we hypothesized that spacer sequences themselves may additionally contain population lineage information, given the large pool of potential spacers. In the time-course recording experiment, cell populations were experimentally split into subpopulations each day, which resulted in a defined branching history of the 16 populations (Fig. 3E). By performing lineage reconstruction using a simple metric to assess spacer repertoire distance between populations (methods), we could reconstruct the entire experimental population lineage with nearly perfect accuracy (Fig. 3F).



measurement of pTrig relative copy number (log10 scale) in cells exposed to no IPTG or 1 mM IPTG for 6 hours. (**E**) The pRec recording plasmid includes an aTc-inducible *E. coli* Cas1 and Cas2 expression cassette. (**F**) Experimental induction scheme and CRISPR array sequencing approach. (o/n, overnight). (**G**) Cells with pRec or with pRec and pTrig were exposed to 100 ng/ μ L aTc and no or 1 mM IPTG and subjected to sequencing; resulting arrays with a single new spacer and identified source (genome, pRec, or pTrig) are plotted as a percentage of all measured CRISPR arrays. Error bars represent standard deviation of three biological replicates.

Fig. 2. Temporal recording of 4-day input profiles. (A) Cell populations were subjected to daily exposures over 4 sequential days (d1 to d4), constituting all

Α

Ε

В

С

D

16 possible temporal signal profiles. (B) Resulting CRISPR arrays were sequenced with (black) and without (gray) a size-enrichment method. The frequencies (log10 scale) of unexpanded (un) and expanded arrays of different lengths (L1 to maximum detectable L5) are plotted. (C) Input profiles are grouped by number of pTrig inductions, and the percentage of pTrig spacers in each profile is displayed; red lines indicate means and standard deviations. (D) On the left, 50 L4 arrays sampled from the full data set for the input profile [on, on, off, off] are shown (shaded, pTrig spacer; unshaded, reference spacer; positions p1 to p4, 5'-to-3' of array). Spacer incorporation can be analyzed across arrays of different lengths (L) and positions (p) as a heatmap displaying percentages of pTrig spacers detected at each location (right). (E) CRISPR arrays derived from recordings of all 16 temporal signal profiles. (F) The input signal profile (left) and corresponding L4 arrays (right, shown in reverse order to aid visual comparison) are displayed.

Fig. 3. Reconstructing temporal signal profiles and population lineages.

(A) CRISPR array populations can be described as a frequency distribution consisting of all permutations of reference (R, blue) and trigger (T, orange) spacers for a given array length (L); L3 arrays are depicted. (B) As an example, for two distinct profiles with an equal number of inductions, observed (black) and model-predicted (white) L3 array-type frequencies are plotted; L3 positional averages are shown for reference (inset). (C) Euclidean distances between observed (rows) and modelpredicted (columns) array-type distributions were calculated and normalized by row (L2, L3, and L4 array-type distributions are concatenated). The correct temporal signal profiles are indicated by white asterisks, and the models with minimum distance to the observed data are indicated by black outlines. (D) Number of profiles correctly classified using arrays L1 to L4 individually or

arrays L2 to L4 together as in (C); the gray dashed line indicates the expected random classification (one of 16 correct). (E) A defined branching history was used in the temporal recording experiment. (F) The mapping locations for genomic spacers within L1 arrays were used as the sequence identity of the spacer. The Jaccard distances





between all samples (1 minus the proportion of spacers shared between two samples) are displayed. Lineage reconstruction was performed using the Fitch-Margoliash method on this distance matrix and is displayed on the left; only one lineage is not fully differentiated (cells receiving induction on d1).

To further characterize the recording performance of TRACE, we assessed the stability of stored information and the potential for longerterm recordings. Propagation of recordings stored within cell populations over 8 days (~50 generations) did not appear to alter array-type distributions (fig. S12, A and B), and induction of recording showed negligible loss of previously acquired spacers (fig. S12C). These results demonstrate stable data storage. We repeated recording experiments on selected temporal signal profiles for 10 days, which showed reasonable reconstruction accuracy up to 6 days (four of seven correctly classified; fig. S13). In general, longer arrays increased the accuracy of signal profile reconstruction during longer recording sessions, which suggests that longer-read sequencing may further increase the performance of long-term recording analysis.

Last, we explored the possibility of using TRACE for multichannel temporal recording. We devised a multiplexing strategy wherein various pTrig sensor systems could be associated with uniquely barcoded CRISPR arrays within a cell population (Fig. 4A). Specifically, we chose to mutagenize the 3' direct repeat (DR) sequence, which should not affect spacer integration (24), as a barcode. This allowed for multiplexing with no modification to the sequencing protocol. More importantly, this enabled more stringent calling of barcodes because the DR sequence is duplicated during each spacer incorporation event (23, 25). Using MAGE [multiplex automated genome engineering (26)], we generated strains with new genomic DR barcodes. In distinct barcoded strains, we coupled different sensors to pTrig and screened their performance (fig. S14). Three orthogonal and robust biosensors that detected the biologically meaningful chemicals copper (heavy-metal contaminant), trehalose (dietary sugar metabolite), and fucose [associated with mammalian gut infection (27)] were selected for multiplex recording experiments. To assess the capacity for multichannel recording, we exposed cell populations containing a mix of all three strains to all eight combinations of the three input chemicals. The resulting CRISPR arrays were sequenced and demultiplexed using the DR barcodes. Each sen-



Fig. 4. Multiplex temporal recording with a barcoded sensor population. (A) The direct repeat (DR) of a CRISPR array can be barcoded to associate sensors with specific arrays; generated distal DR sequences with barcodes (bc) are shown. Sensors of copper, trehalose, and fucose were linked to the pTrig system and introduced into barcoded strains. The copper sensor uses a native promoter with endogenous transcription factor expression, whereas the trehalose and fucose sensors use an engineered transcription factor. (B) The three barcoded sensor strains were mixed and exposed to eight combinatorial inputs of the three chemicals; the resulting percentage of pTrig spacers for each barcoded sensor strain is displayed (average of three biological replicates). (C) The strain mixture was exposed to combinatorial inputs over 3 days. As an example, profile 5 is displayed, along with CRISPR arrays for each sensor (plotted as in Fig. 2, but the color map is rescaled for each sensor to aid visualization) and the resulting classification (correct, blue checkmark; incorrect, red X). (D) Of 512 (8³) possible profiles, 16 were tested (six defined and 10 randomly generated); the resulting classification is shown as in (C) (black arrows indicate the time course, d1 to d3). (\mathbf{E}) Single-channel classification accuracy. Profiles were classified for each sensor on the basis of L2 and L3 arrays; the gray dashed line indicates the expected random classification (two of 16 correct). (F) Multichannel classification accuracy. Predictions were considered across all three sensors, and the number classified correctly within a Hamming distance threshold is shown (black line) compared with the expected random classification (gray dashed line).

sor strain displayed a robust increase in pTrigderived spacers (>24-fold) only in the presence of their cognate input (Fig. 4B and fig. S15). Importantly, these results indicate modular compatibility of TRACE for multichannel recording with a variety of sensing systems, including engineered sensors or native promoters with endogenous transcription factor expression.

To explore multiplex temporal recording, we used the three-strain sensing system to perform a time-course exposure experiment over 3 days. Cell populations were exposed to 16 selected temporal signal profiles of 512 possible profiles, and resulting CRISPR arrays were sequenced. Sensor strains fluctuated in their final abundance but were maintained at sufficient levels to enable CRISPR array analysis (fig. S16). We parameterized models for each sensor individually as before and inferred the exposure history of each of the three inputs individually for all 16 populations by classification against model predictions (Fig. 4C). We were able to correctly classify 14, 13, and 12 of the 16 signal profiles for the copper, trehalose, and fucose sensors, respectively (Fig. 4, D and E). Classification accuracy for all three inputs simultaneously was assessed by the Hamming distance threshold to the actual temporal signal profiles; eight of 16 profiles were perfectly classified, and the rest were within a Hamming distance of 2 (Fig. 4F), implying that even incorrect predictions were close to actual signal profiles. Together, these results demonstrate accurate multichannel recording with the TRACE system.

Our work enables new applications in biological recording. TRACE could be used to record metabolite fluctuations, gene expression changes, and lineage-associated information across cell populations in difficult-to-study habitats, such as the mammalian gut, or in open settings, such as soil or marine environments. Applying inducible intracellular DNA production systems in parallel (28) and other CRISPR-Cas adaptation machinery (13, 29) could extend our system to other bacteria (or even eukaryotes) and increase the temporal resolution of recording beyond the levels demonstrated here (6 hours, ~45 µHz). The system could be further optimized by increasing the spacer incorporation rate (30), increasing the sequencing length (e.g., by nanopore sequencing), and improving reconstruction algorithms. These advances could further facilitate biological recording of inputs across many signal channels, with higher temporal resolution, and in smaller populations, possibly down to single cells. TRACE and future strategies for massively parallel recording of biological states should greatly advance our ability to delineate and understand complex cellular processes across time.

REFERENCES AND NOTES

- A. E. Friedland *et al.*, *Science* **324**, 1199–1202 (2009).
 J. Bonnet, P. Yin, M. E. Ortiz, P. Subsoontorn, D. Endy,
- Science **340**, 599–603 (2013). 3. L. Yang et al., Nat. Methods **11**, 1261–1266 (2014).
- L. Hang et al., Nat. Methods 11, 1201–1200 (2014).
 V. Hsiao, Y. Hori, P. W. Rothemund, R. M. Murray,
- Mol. Syst. Biol. 12, 869–14 (2016).

- N. Roquet, A. P. Soleimany, A. C. Ferris, S. Aaronson, T. K. Lu, Science 353, aad8559 (2016).
- 6. W. Pei et al., Nature 548, 456–460 (2017).
- 7. F. Farzadfard, T. K. Lu, Science 346, 1256272 (2014).
- 8. A. McKenna et al., Science 353, aaf7907 (2016).
- 9. K. L. Frieda et al., Nature 541, 107-111 (2017).
- 10. S. T. Schmidt, S. M. Zimmerman, J. Wang, S. K. Kim,
- S. R. Quake, ACS Synth. Biol. 6, 936–942 (2017).
 S. D. Perli, C. H. Cui, T. K. Lu, Science 353, aag0511 (2016).
- R. Kalhor, P. Mali, G. M. Church, *Nat. Methods* 14, 195–200 (2017).
- 13. S. A. Jackson et al., Science 356, eaal5056 (2017).
- S. H. Sternberg, H. Richter, E. Charpentier, U. Qimron, Mol. Cell 61, 797–808 (2016).
- 15. A. V. Wright et al., Science 357, 1113-1118 (2017).
- 16. R. Barrangou et al., Science 315, 1709-1712 (2007).
- 17. J. McGinn, L. A. Marraffini, Mol. Cell 64, 616-623 (2016).
- 18. L. A. Marraffini, Nature 526, 55-61 (2015).
- S. L. Shipman, J. Nivala, J. D. Macklis, G. M. Church, *Science* 353, aaf1175 (2016).
- S. L. Shipman, J. Nivala, J. D. Macklis, G. M. Church, *Nature* 547, 345–349 (2017).

- 21. A. Levy et al., Nature 520, 505–510 (2015).
- 22. M. B. Łobocka et al., J. Bacteriol. 186, 7032-7068 (2004).
- 23. I. Yosef, M. G. Goren, U. Qimron, Nucleic Acids Res. 40,
- 5569-5576 (2012).
- 24. J. K. Nuñez, L. Bai, L. B. Harrington, T. L. Hinder, J. A. Doudna, *Mol. Cell* 62, 824–833 (2016).
- J. K. Nuñez, A. S. Y. Lee, A. Engelman, J. A. Doudna, *Nature* 519, 193–198 (2015).
- 26. H. H. Wang et al., Nature 460, 894-898 (2009).
- 27. J. M. Pickard et al., Nature 514, 638-641 (2014).
- 28. J. Elbaz, P. Yin, C. A. Voigt, Nat. Commun. 7, 11179 (2016).
- 29. S. Silas et al., Science 351, aad4234 (2016).
- 30. R. Heler et al., Mol. Cell 65, 168-175 (2017).

ACKNOWLEDGMENTS

We thank T. Blazejewski, C. Munck, and members of the Wang laboratory for advice and comments on the manuscript. Sequencing data associated with this study are available in the National Center for Biotechnology Information Sequence Read Archive under PRJNA417866, and plasmids are available through Addgene. H.H.W. acknowledges specific funding support from the Department of Defense Office of Naval Research (N00014-17-1-2353 and N00014-15-1-2704), an NIH Director's Early Independence Award (1-

DP50D009172-02), and the Sloan Foundation (FR-2015-65795) for this work. R.U.S. is supported by a Fannie and John Hertz Foundation Graduate Fellowship and an NSF Graduate Research Fellowship (DGE 16-44869), F.L.W. is supported by an NIH training grant (T32CM008224), R.U.S., S.S.Y., and H.H.W. developed the initial concept; R.U.S., S.S.Y., and F.L.W. performed experiments; R.U.S. and F.L.W. analyzed the sequencing data; R.U.S. and H.H.W. wrote the manuscript; and all authors discussed results and commented on and approved the manuscript. H.H.W. and R.U.S. and R.U.S. and provisional patent application filed by the Trustees of Columbia University in the City of New York regarding this work.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/358/6369/1457/suppl/DC1 Materials and Methods Figs. S1 to S16 Tables S1 to S5 References (*31–44*)

16 June 2017; resubmitted 29 September 2017 Accepted 13 November 2017 Published online 23 November 2017 10.1126/science.aao0958



Multiplex recording of cellular events over time on CRISPR biological tape

Ravi U. Sheth, Sung Sun Yim, Felix L. Wu and Harris H. Wang

Science **358** (6369), 1457-1461. DOI: 10.1126/science.aao0958originally published online November 23, 2017

A CRISPR device to record time

The CRISPR adaptation system has been used to record the sequence and ordering of exogenous oligonucleotides that are electroporated into cell populations. Sheth *et al.* engineered a system bypassing the use of exogenous DNA to directly record temporal signals. An input biological signal is transformed into the ratio of the frequency of incorporating trigger DNA to that of incorporating reference DNA into the genomes of a bacterial population. A multiplexing strategy enables simultaneous recording of three environmental signals with high temporal resolution. *Science*, this issue p. 1457

ARTICLE TOOLS	http://science.sciencemag.org/content/358/6369/1457
SUPPLEMENTARY MATERIALS	http://science.sciencemag.org/content/suppl/2017/11/21/science.aao0958.DC1
REFERENCES	This article cites 44 articles, 18 of which you can access for free http://science.sciencemag.org/content/358/6369/1457#BIBL
PERMISSIONS	http://www.sciencemag.org/help/reprints-and-permissions

Use of this article is subject to the Terms of Service

Science (print ISSN 0036-8075; online ISSN 1095-9203) is published by the American Association for the Advancement of Science, 1200 New York Avenue NW, Washington, DC 20005. 2017 © The Authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original U.S. Government Works. The title *Science* is a registered trademark of AAAS.