

4. A. J. Wood *et al.*, *Science* **333**, 307 (2011).
5. M. Christian *et al.*, *Genetics* **186**, 757 (2010).
6. F. Zhang *et al.*, *Nat. Biotechnol.* **29**, 149 (2011).
7. J. C. Miller *et al.*, *Nat. Biotechnol.* **29**, 143 (2011).
8. D. Reyon *et al.*, *Nat. Biotechnol.* **30**, 460 (2012).
9. J. Boch *et al.*, *Science* **326**, 1509 (2009).
10. M. J. Moscou, A. J. Bogdanove, *Science* **326**, 1501 (2009).
11. B. L. Stoddard, *Q. Rev. Biophys.* **38**, 49 (2005).
12. M. Jinek *et al.*, *Science* **337**, 816 (2012).
13. G. Gasiunas, R. Barrangou, P. Horvath, V. Siksnys, *Proc. Natl. Acad. Sci. U.S.A.* **109**, E2579 (2012).
14. J. E. Garneau *et al.*, *Nature* **468**, 67 (2010).
15. H. Deveau, J. E. Garneau, S. Moineau, *Annu. Rev. Microbiol.* **64**, 475 (2010).
16. P. Horvath, R. Barrangou, *Science* **327**, 167 (2010).
17. K. S. Makarova *et al.*, *Nat. Rev. Microbiol.* **9**, 467 (2011).
18. D. Bhaya, M. Davison, R. Barrangou, *Annu. Rev. Genet.* **45**, 273 (2011).
19. E. Deltcheva *et al.*, *Nature* **471**, 602 (2011).
20. R. Sapranaukas *et al.*, *Nucleic Acids Res.* **39**, 9275 (2011).
21. A. H. Magadán, M. E. Dupuis, M. Villion, S. Moineau, *PLoS ONE* **7**, e40913 (2012).
22. H. Deveau *et al.*, *J. Bacteriol.* **190**, 1390 (2008).
23. F. J. Mojica, C. Díez-Villaseñor, J. García-Martínez, C. Almendros, *Microbiology* **155**, 733 (2009).
24. M. Jinek, J. A. Doudna, *Nature* **457**, 405 (2009).
25. C. D. Malone, G. J. Hannon, *Cell* **136**, 656 (2009).
26. G. Meister, T. Tuschl, *Nature* **431**, 343 (2004).
27. M. T. Certo *et al.*, *Nat. Methods* **8**, 671 (2011).
28. P. Mali *et al.*, *Science* **339**, 823 (2013).
29. P. A. Carr, G. M. Church, *Nat. Biotechnol.* **27**, 1151 (2009).

**Acknowledgments:** We thank the entire Zhang lab for their support and advice; P. A. Sharp for generous help with Northern blot analysis; C. Jennings, R. Desimone, and M. Kowalczyk for helpful comments; and X. Ye for help with confocal imaging. L.C. and X.W. are Howard Hughes Medical Institute International Student Research Fellows. D.C. is supported by the Medical Scientist Training Program. P.D.H. is a James Mills Pierce Fellow. X.W. is supported by NIH grants R01-GM34277 and R01-CA133404 to P. A. Sharp,

X.W.'s thesis adviser. L.A.M. is supported by Searle Scholars, R. Allen, an Irma T. Hirsch Award, and a NIH Director's New Innovator Award (DP2AI104556). F.Z. is supported by a NIH Director's Pioneer Award (DP1MH100706); the Keck, McKnight, Gates, Damon Runyon, Searle Scholars, Klingenstein, and Simons foundations; R. Metcalfe; M. Boylan; and J. Pauley. The authors have no conflicting financial interests. A patent application has been filed relating to this work, and the authors plan on making the reagents widely available to the academic community through Addgene and to provide software tools via the Zhang lab Web site ([www.genome-engineering.org](http://www.genome-engineering.org)).

#### Supplementary Materials

[www.sciencemag.org/cgi/content/full/science.1231143/DC1](http://www.sciencemag.org/cgi/content/full/science.1231143/DC1)  
Materials and Methods  
Figs. S1 to S8  
Tables S1 and S2  
References (30–32)

5 October 2012; accepted 12 December 2012  
Published online 3 January 2013;  
10.1126/science.1231143

# RNA-Guided Human Genome Engineering via Cas9

Prashant Mali,<sup>1\*</sup> Luhan Yang,<sup>1,3\*</sup> Kevin M. Esvelt,<sup>2</sup> John Aach,<sup>1</sup> Marc Guell,<sup>1</sup> James E. DiCarlo,<sup>4</sup> Julie E. Norville,<sup>1</sup> George M. Church<sup>1,2†</sup>

Bacteria and archaea have evolved adaptive immune defenses, termed clustered regularly interspaced short palindromic repeats (CRISPR)/CRISPR-associated (Cas) systems, that use short RNA to direct degradation of foreign nucleic acids. Here, we engineer the type II bacterial CRISPR system to function with custom guide RNA (gRNA) in human cells. For the endogenous AAVS1 locus, we obtained targeting rates of 10 to 25% in 293T cells, 13 to 8% in K562 cells, and 2 to 4% in induced pluripotent stem cells. We show that this process relies on CRISPR components; is sequence-specific; and, upon simultaneous introduction of multiple gRNAs, can effect multiplex editing of target loci. We also compute a genome-wide resource of ~190 K unique gRNAs targeting ~40.5% of human exons. Our results establish an RNA-guided editing tool for facile, robust, and multiplexable human genome engineering.

**B**acterial and archaeal clustered regularly interspaced short palindromic repeats (CRISPR) systems rely on CRISPR RNAs (crRNAs) in complex with CRISPR-associated (Cas) proteins to direct degradation of complementary sequences present within invading viral and plasmid DNA (1–3). A recent in vitro reconstitution of the *Streptococcus pyogenes* type II CRISPR system demonstrated that crRNA fused to a normally trans-encoded tracrRNA is sufficient to direct Cas9 protein to sequence-specifically cleave target DNA sequences matching the crRNA (4). The fully defined nature of this two-component system suggested that it might function in the cells of eukaryotic organisms such as yeast, plants,

and even mammals. By cleaving genomic sequences targeted by RNA sequences (4–6), such a system could greatly enhance the ease of genome engineering.

Here, we engineer the protein and RNA components of this bacterial type II CRISPR system in human cells. We began by synthesizing a human codon-optimized version of the Cas9 protein bearing a C-terminal SV40 nuclear localization signal and cloning it into a mammalian expression system (Fig. 1A and fig. S1A). To direct Cas9 to cleave sequences of interest, we expressed crRNA-tracrRNA fusion transcripts, hereafter referred to as guide RNAs (gRNAs), from the human U6 polymerase III promoter. Directly transcribing gRNAs allowed us to avoid reconstituting the RNA-processing machinery used by bacterial CRISPR systems (Fig. 1A and fig. S1B) (4, 7–9). Constrained only by U6 transcription initiating with G and the requirement for the PAM (protospacer-adjacent motif) sequence -NGG following the 20-base pair (bp) crRNA target, our highly versatile approach can, in principle, target any genomic site of the form GN<sub>20</sub>GG (fig.

S1C; see supplementary text S1 for a detailed discussion).

To test the functionality of our implementation for genome engineering, we developed a green fluorescent protein (GFP) reporter assay (Fig. 1B) in human embryonic kidney HEK 293T cells similar to one previously described (10). Specifically, we established a stable cell line bearing a genomically integrated GFP coding sequence disrupted by the insertion of a stop codon and a 68-bp genomic fragment from the AAVS1 locus that renders the expressed protein fragment non-fluorescent. Homologous recombination (HR) using an appropriate repair donor can restore the normal GFP sequence, which enabled us to quantify the resulting GFP<sup>+</sup> cells by flow-activated cell sorting (FACS).

To test the efficiency of our system at stimulating HR, we constructed two gRNAs, T1 and T2, that target the intervening AAVS1 fragment (Fig. 1B) and compared their activity to that of a previously described TAL effector nuclease heterodimer (TALEN) targeting the same region (11). We observed successful HR events using all three targeting reagents, with gene correction rates using the T1 and T2 gRNAs approaching 3% and 8%, respectively (Fig. 1C). This RNA-mediated editing process was notably rapid, with the first detectable GFP<sup>+</sup> cells appearing ~20 hours post transfection compared with ~40 hours for the AAVS1 TALENs. We observed HR only upon simultaneous introduction of the repair donor, Cas9 protein, and gRNA, which confirmed that all components are required for genome editing (fig. S2). Although we noted no apparent toxicity associated with Cas9/gRNA expression, work with zinc finger nucleases (ZFNs) and TALENs has shown that nicking only one strand further reduces toxicity. Accordingly, we also tested a Cas9D10A mutant that is known to function as a nickase in vitro, which yielded similar HR but lower nonhomologous end joining (NHEJ) rates (fig. S3) (4, 5). Consistent with (4), in which a related Cas9 protein is shown to cut both strands

<sup>1</sup>Department of Genetics, Harvard Medical School, Boston, MA 02115, USA. <sup>2</sup>Wyss Institute for Biologically Inspired Engineering, Harvard University, Cambridge, MA 02138, USA. <sup>3</sup>Biological and Biomedical Sciences Program, Harvard Medical School, Boston, MA 02115, USA. <sup>4</sup>Department of Biomedical Engineering, Boston University, Boston, MA 02215, USA.

\*These authors contributed equally to this work.

†To whom correspondence should be addressed. E-mail: [gchurch@genetics.med.harvard.edu](mailto:gchurch@genetics.med.harvard.edu)

3 bp upstream of the PAM, our NHEJ data confirmed that most deletions or insertions occurred at the 3' end of the target sequence (fig. S3B). We also confirmed that mutating the target genomic site prevents the gRNA from effecting HR at that locus, which demonstrates that CRISPR-mediated genome editing is sequence-specific (fig. S4). Finally, we showed that two gRNAs targeting sites in the GFP gene, and also three additional gRNAs targeting fragments from homologous regions of the DNA methyl transferase 3a (DNMT3a) and DNMT3b genes could sequence-specifically induce significant HR in the engineered reporter cell lines (figs. S5 and S6). Together, these results confirm that RNA-guided genome targeting in human cells is simple to execute and induces robust HR across multiple target sites.

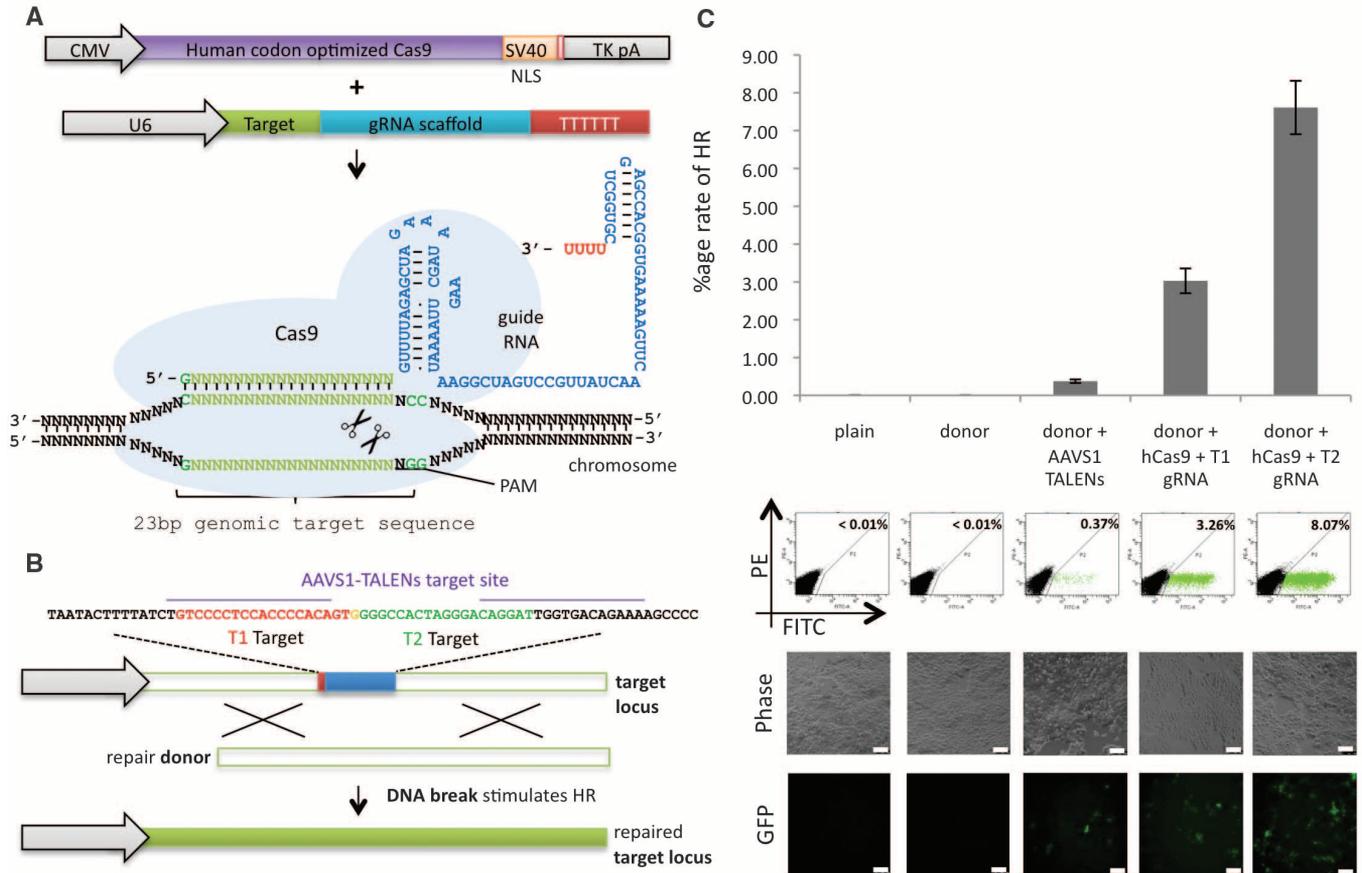
Having successfully targeted an integrated reporter, we next turned to modifying a native locus. We used the gRNAs described above to target the AAVS1 locus located in the PPP1R12C

gene on chromosome 19, which is ubiquitously expressed across most tissues (Fig. 2A). We targeted 293Ts, human chronic myelogenous leukemia K562 cells, and PGP1 human induced pluripotent stem (iPS) cells (12) and analyzed the results by next-generation sequencing of the targeted locus. Consistent with our results for the GFP reporter assay, we observed high numbers of NHEJ events at the endogenous locus for all three cell types. The two gRNAs T1 and T2 achieved NHEJ rates of 10 and 25% in 293Ts, 13 and 38% in K562s, and 2 and 4% in PGP1-iPS cells, respectively (Fig. 2B). We observed no overt toxicity from the Cas9 and gRNA expression required to induce NHEJ in any of these cell types. As expected, NHEJ-mediated deletions for T1 and T2 were centered around the target site positions, which further validated the sequence-specificity of this targeting process (figs. S7 to S9). Simultaneous introduction of both T1 and T2 gRNAs resulted in high-efficiency deletion of the intervening 19-bp fragment (fig. S8), which

demonstrated that multiplexed editing of genomic loci is feasible using this approach.

Last, we attempted to use HR to integrate either a double-stranded DNA donor construct (13) or an oligo donor into the native AAVS1 locus (Fig. 2C and fig. S10). We confirmed HR-mediated integration, using both approaches, by polymerase chain reaction (PCR) (Fig. 2D and fig. S10) and Sanger sequencing (Fig. 2E). We also readily derived 293T or iPS clones from the pool of modified cells using puromycin selection over 2 weeks (Fig. 2F and fig. S10). These results demonstrate that this approach enables efficient integration of foreign DNA at endogenous loci in human cells.

Our versatile RNA-guided genome-editing system can be readily adapted to modify other genomic sites by simply modifying the sequence of our gRNA expression vector to match a compatible sequence in the locus of interest. To facilitate this process, we bioinformatically generated ~190,000 specific gRNA-targetable sequences



**Fig. 1.** Genome editing in human cells using an engineered type II CRISPR system. (A) RNA-guided gene targeting in human cells involves coexpression of the Cas9 protein bearing a C-terminal SV40 nuclear localization signal (NLS) with one or more gRNAs expressed from the human U6 polymerase III promoter. Cas9 unwinds the DNA duplex and cleaves both strands upon recognition of a target sequence by the gRNA, but only if the correct PAM is present at the 3' end. Any genomic sequence of the form GN<sub>20</sub>GG can, in principle, be targeted. CMV, cytomegalovirus promoter; TK, thymidine kinase; pA, polyadenylation signal. (B) A genomically integrated GFP coding sequence is

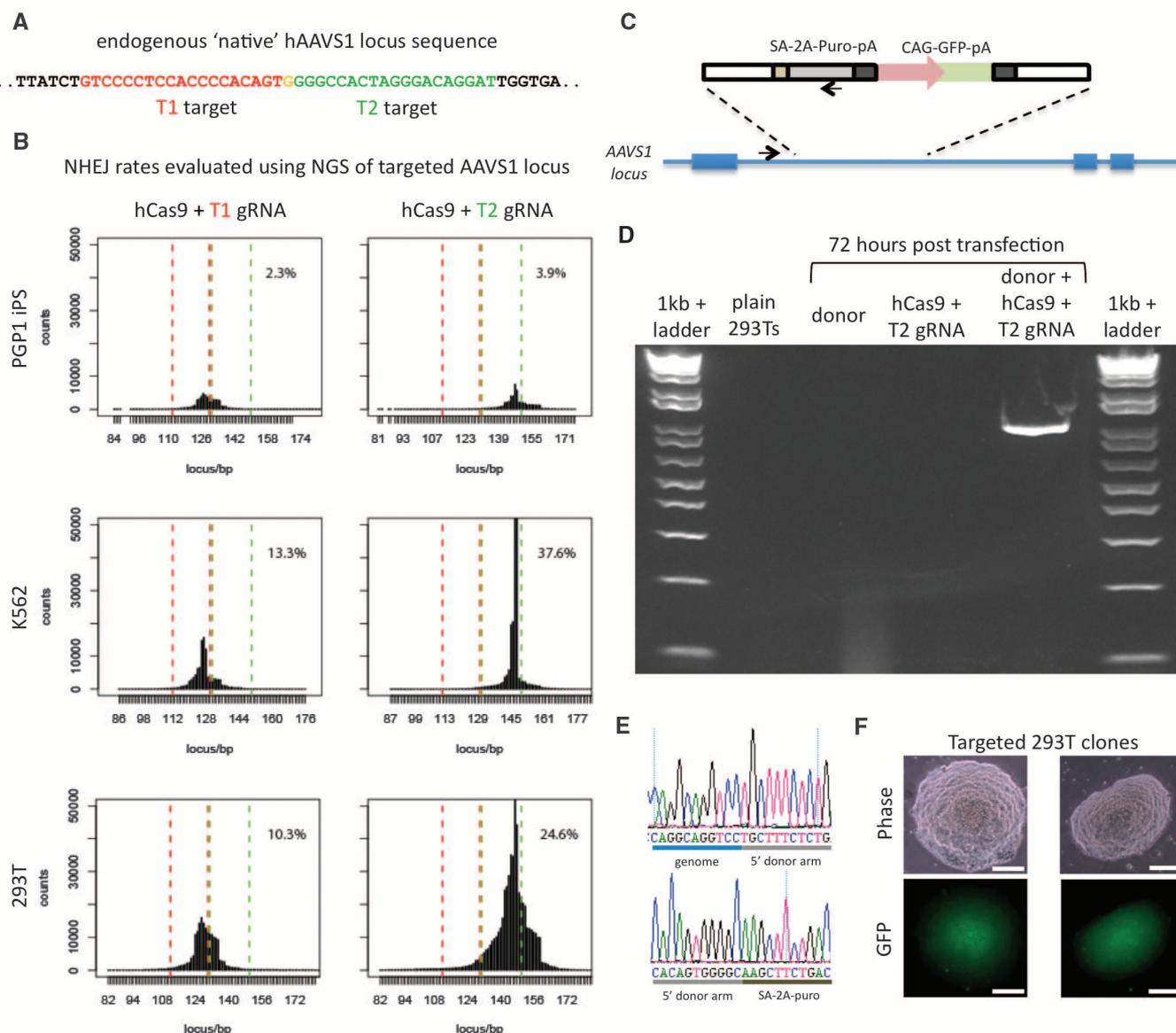
disrupted by the insertion of a stop codon and a 68-bp genomic fragment from the AAVS1 locus. Restoration of the GFP sequence by HR with an appropriate donor sequence results in GFP<sup>+</sup> cells that can be quantified by FACS. T1 and T2 gRNAs target sequences within the AAVS1 fragment. Binding sites for the two halves of the TALEN are underlined. (C) Bar graph depicting HR efficiencies induced by T1, T2, and TALEN-mediated nuclease activity at the target locus, as measured by FACS. Representative FACS plots and microscopy images of the targeted cells are depicted below. (Scale bar, 100 μm.) Data are shown as means ± SEM (N = 3).

targeting ~40.5% exons of genes in the human genome (refer to methods and table S1). We also incorporated these target sequences into a 200-bp format compatible with multiplex synthesis on DNA arrays (14) (fig. S11 and tables S2 and S3). This resource provides a ready genome-wide reference of potential target sites in the human genome and a methodology for multiplex gRNA synthesis.

Our results demonstrate the promise of CRISPR-mediated gene targeting for RNA-guided, robust, and multiplexable mammalian

genome engineering. The ease of retargeting our system to modify genomic sequences greatly exceeds that of comparable ZFNs and TALENs, while offering similar or greater efficiencies (4). Existing studies of type II CRISPR specificity (4) suggest that target sites must perfectly match the PAM sequence NGG and the 8- to 12-base “seed sequence” at the 3’ end of the gRNA. The importance of the remaining 8 to 12 bases is less well understood and may depend on the binding strength of the matching gRNAs or on the inherent tolerance of Cas9 itself. Indeed, Cas9 will

tolerate single mismatches at the 5’ end in bacteria and in vitro, which suggests that the 5’ G is not required. Moreover, it is likely that the target locus’s underlying chromatin structure and epigenetic state will also affect the efficiency of genome editing in eukaryotic cells (13), although we suspect that Cas9’s helicase activity may render it more robust to these factors, but this remains to be evaluated. Elucidating the frequency and underlying causes of off-target nuclease activity (15, 16) induced by CRISPR, ZFN (17, 18), and TALEN (19, 20) genome-engineering



**Fig. 2.** RNA-guided genome editing of the native AAVS1 locus in multiple cell types. **(A)** T1 (red) and T2 (green) gRNAs target sequences in an intron of the PPP1R12C gene within the chromosome 19 AAVS1 locus. **(B)** Total count and location of deletions caused by NHEJ in 293Ts, K562s, and PGP1 iPS cells after expression of Cas9 and either T1 or T2 gRNAs as quantified by next-generation sequencing. Red and green dashed lines demarcate the boundaries of the T1 and T2 gRNA targeting sites. NHEJ frequencies for T1 and T2 gRNAs were 10% and 25% in 293T, 13% and 38% in K562, and 2% and 4% in PGP1 iPS cells, respectively. **(C)** DNA donor architecture for HR at

the AAVS1 locus, and the locations of sequencing primers (arrows) for detecting successful targeted events, are depicted. **(D)** PCR assay 3 days after transfection demonstrates that only cells expressing the donor, Cas9 and T2 gRNA exhibit successful HR events. **(E)** Successful HR was confirmed by Sanger sequencing of the PCR amplicon, which showed that the expected DNA bases at both the genome-donor and donor-insert boundaries are present. **(F)** Successfully targeted clones of 293T cells were selected with puromycin for 2 weeks. Microscope images of two representative GFP+ clones is shown. (Scale bar, 100  $\mu$ m.)

tools will be of utmost importance for safe genome modification and perhaps for gene therapy. Potential avenues for improving CRISPR specificity include evaluating Cas9 homologs identified through bioinformatics and directed evolution of these nucleases toward higher specificity. Similarly, the range of CRISPR-targetable sequences could be expanded through the use of homologs with different PAM requirements (9) or by directed evolution. Finally, inactivating one of the Cas9 nuclease domains increases the ratio of HR to NHEJ and may reduce toxicity (figs. S1A and fig. S3) (4, 5), whereas inactivating both domains may enable Cas9 to function as a retargetable DNA binding protein. As we explore these areas, we note that another parallel study (21) has independently confirmed the high efficiency of CRISPR-mediated gene targeting in mammalian cell lines. We expect that RNA-guided genome targeting will have broad implications for synthetic biology (22, 23), the direct and multiplexed perturbation of gene networks (13, 24), and targeted ex vivo (25–27) and in vivo gene therapy (28).

#### References and Notes

1. B. Wiedenheft, S. H. Sternberg, J. A. Doudna, *Nature* **482**, 331 (2012).
2. D. Bhaya, M. Davison, R. Barrangou, *Annu. Rev. Genet.* **45**, 273 (2011).
3. M. P. Terns, R. M. Terns, *Curr. Opin. Microbiol.* **14**, 321 (2011).
4. M. Jinek *et al.*, *Science* **337**, 816 (2012).
5. G. Gasiunas, R. Barrangou, P. Horvath, V. Siksnys, *Proc. Natl. Acad. Sci. U.S.A.* **109**, E2579 (2012).
6. R. Sapranasuskas *et al.*, *Nucleic Acids Res.* **39**, 9275 (2011).
7. T. R. Brummelkamp, R. Bernards, R. Agami, *Science* **296**, 550 (2002).
8. M. Miyagishi, K. Taira, *Nat. Biotechnol.* **20**, 497 (2002).
9. E. Deltcheva *et al.*, *Nature* **471**, 602 (2011).
10. J. Zou, P. Mali, X. Huang, S. N. Doherty, L. Cheng, *Blood* **118**, 4599 (2011).
11. N. E. Sanjana *et al.*, *Nat. Protoc.* **7**, 171 (2012).
12. J. H. Lee *et al.*, *PLoS Genet.* **5**, e1000718 (2009).
13. D. Hockemeyer *et al.*, *Nat. Biotechnol.* **27**, 851 (2009).
14. S. Kosuri *et al.*, *Nat. Biotechnol.* **28**, 1295 (2010).
15. V. Pattanayak, C. L. Ramirez, J. K. Joung, D. R. Liu, *Nat. Methods* **8**, 765 (2011).
16. N. M. King, O. Cohen-Haguener, *Mol. Ther.* **16**, 432 (2008).
17. Y. G. Kim, J. Cha, S. Chandrasegaran, *Proc. Natl. Acad. Sci. U.S.A.* **93**, 1156 (1996).
18. E. J. Rebar, C. O. Pabo, *Science* **263**, 671 (1994).
19. J. Boch *et al.*, *Science* **326**, 1509 (2009).
20. M. J. Moscou, A. J. Bogdanove, *Science* **326**, 1501 (2009).
21. L. Cong *et al.*, *Science* **339**, 819 (2013).
22. A. S. Khalil, J. J. Collins, *Nat. Rev. Genet.* **11**, 367 (2010).
23. P. E. Purnick, R. Weiss, *Nat. Rev. Mol. Cell Biol.* **10**, 410 (2009).
24. J. Zou *et al.*, *Cell Stem Cell* **5**, 97 (2009).
25. N. Holt *et al.*, *Nat. Biotechnol.* **28**, 839 (2010).
26. F. D. Urnov *et al.*, *Nature* **435**, 646 (2005).
27. A. Lombardo *et al.*, *Nat. Biotechnol.* **25**, 1298 (2007).
28. H. Li *et al.*, *Nature* **475**, 217 (2011).

**Acknowledgments:** This work was supported by NIH grant P50 HG005550. We thank S. Kosuri for advice on the oligonucleotide pool designs and synthesis. G.M.C. and P.M. have filed a patent based on the findings of this study.

#### Supplementary Materials

www.sciencemag.org/cgi/content/full/science.1232033/DC1  
Materials and Methods

Supplementary Text

Figs. S1 to S11

Tables S1 to S3

References (29–46)

26 October 2012; accepted 12 December 2012

Published online 3 January 2013;

10.1126/science.1232033

## Cyclic GMP-AMP Is an Endogenous Second Messenger in Innate Immune Signaling by Cytosolic DNA

Jiayi Wu,<sup>1\*</sup> Lijun Sun,<sup>1,2\*</sup> Xiang Chen,<sup>1</sup> Fenghe Du,<sup>1</sup> Heping Shi,<sup>3</sup> Chuo Chen,<sup>3</sup> Zhijian J. Chen<sup>1,2,†</sup>

Cytosolic DNA induces type I interferons and other cytokines that are important for antimicrobial defense but can also result in autoimmunity. This DNA signaling pathway requires the adaptor protein STING and the transcription factor IRF3, but the mechanism of DNA sensing is unclear. We found that mammalian cytosolic extracts synthesized cyclic guanosine monophosphate–adenosine monophosphate (cyclic GMP-AMP, or cGAMP) in vitro from adenosine triphosphate and guanosine triphosphate in the presence of DNA but not RNA. DNA transfection or DNA virus infection of mammalian cells also triggered cGAMP production. cGAMP bound to STING, leading to the activation of IRF3 and induction of interferon- $\beta$ . Thus, cGAMP in metazoans and functions as an endogenous second messenger that triggers interferon production in response to cytosolic DNA.

Host defense against foreign genetic elements is one of the most fundamental functions of a living organism. The presence of self or foreign DNA in the cytoplasm is sensed by eukaryotic cells as a danger signal or a sign of foreign invasion (1). DNA can be introduced into the cytoplasm by bacterial or viral infection, transfection, or “leakage” from the nu-

cleus or mitochondria under some pathological conditions that cause autoimmune diseases such as lupus. In mammalian cells, cytosolic DNA triggers the production of type I interferons and other cytokines through the endoplasmic reticulum protein STING (also known as MITA, MPYS, or ERIS) (2). STING recruits and activates the cytosolic kinases IKK and TBK1, which activate the transcription factors NF- $\kappa$ B and IRF3, respectively. NF- $\kappa$ B and IRF3 then enter the nucleus and function together to induce interferons and other cytokines. DNA-dependent RNA polymerase III has been shown to be a sensor that detects and transcribes AT-rich DNAs such as poly(deoxyadenosine-deoxythymidine) [poly(dA:dT)] into an RNA ligand capable of stimulating the RIG-I pathway to induce interferons (3, 4). However, most DNA sequences do

not activate the RNA polymerase III–RIG-I pathway. Instead, cytosolic DNA activates the STING-dependent pathway in a sequence-independent manner. How cytosolic DNA activates the STING pathway remains elusive.

We hypothesized that DNA binds to and activates a putative cytosolic DNA sensor, which then directly or indirectly activates STING, leading to the activation of IRF3 and NF- $\kappa$ B (fig. S1A). To test this model, we developed an in vitro complementation assay using the murine fibrosarcoma cell line L929, which is known to induce interferon- $\beta$  (IFN- $\beta$ ) in a STING-dependent manner (5) (Fig. 1A). We used an L929 cell line stably expressing a short hairpin RNA (shRNA) against STING such that DNA transfection would only activate factors upstream of STING, including the putative DNA sensor (fig. S1, A and B). The L929-shSTING cells were transfected with different types of DNA, and then cytoplasmic extracts from these cells were mixed with the human monocytic cell line THP1 or murine macrophage cell line Raw264.7, which was permeabilized with perfringolysin O (PFO; Fig. 1A). PFO treatment pokes holes in the plasma membrane (6), allowing the cytoplasm to diffuse in and out of cells, while retaining organelles including the endoplasmic reticulum (which contains STING) and the Golgi apparatus inside the cells (7). If an upstream activator of STING is generated in the DNA-transfected cells, the cytoplasm containing such an activator is expected to activate STING in the PFO-permeabilized cells, leading to the phosphorylation and dimerization of IRF3.

Cytoplasmic extracts from L929-shSTING cells transfected with a DNA sequence known as interferon-stimulatory DNA (ISD; Fig. 1B, lane 2), poly(dA:dT), a GC-rich 50–base pair

<sup>1</sup>Department of Molecular Biology, University of Texas Southwestern Medical Center, Dallas, TX 75390, USA. <sup>2</sup>Howard Hughes Medical Institute, University of Texas Southwestern Medical Center, Dallas, TX 75390, USA. <sup>3</sup>Department of Biochemistry, University of Texas Southwestern Medical Center, Dallas, TX 75390, USA.

\*These authors contributed equally to this work.

†To whom correspondence should be addressed. E-mail: zhijian.chen@utsouthwestern.edu