Multiplex Genome Engineering Using CRISPR/Cas Systems

Le Cong *et al.*

Science **339**, 819 (2013);
DOI: 10.1126/science.1231143

Science (print ISSN 0036-8075; online ISSN 1095-9203) is published weekly, except the last week in December, by the American Association for the Advancement of Science, 1200 New York Avenue NW, Washington, DC 20005.

because the efficiency of transcription initiation correlates with the binding affinity of T7 RNAp (*25*). We therefore inserted DNA templates used in vitro (fig. S15) into the chromosome of *Escherichia coli* and measured the expression level of *lacZ* using the Miller assay (Fig. 4A) (*26*). Indeed, the gene expression level oscillates as a function of L with a periodicity of ~10 bp (Fig. 4B). Similar oscillations of T7 RNAp activity were observed on plasmids in *E. coli* cells by using a yellow fluorescent protein as a reporter (fig. S16). The oscillation of gene expression levels with a 10-bp periodicity was also seen in a classic experiment on *lac* operon with a DNA loop formed by two operators (*27*). However, our T7 RNAp result illustrates that DNA allostery results in such an oscillatory phenomenon even without a DNA loop, which is consistent with a recent study in which *E. coli* RNA polymerase was used (*10*).

Pertinent to eukaryotic gene expression, DNA allostery may affect the binding affinity of transcription factors near nucleosomes that are closely positioned (*28*, *29*). We placed GRE downstream of a nucleosome (Fig. 4C) and observed a similar DNA allosteric effect in the k_{off} of GRDBD (Fig. 4D and fig. S17). To evaluate DNA allostery in an internucleosomal space, we used two nucleosomes to flank a GRE (Fig. 4C). At the same separation L, GRDBD resides on GRE for a relatively longer time with a single nucleosome nearby than it does with a pair of nucleosomes on both sides of GRE (Fig. 4D). Nonetheless, the fold change between the maximal and minimal k_{off} is larger for GRDBD with two nucleosomes (approximately sevenfold). This indicates moderately large cooperativity between the two flanking nucleosomes in modifying the binding affinity of GRDBD, which is in line with previous in vivo experiments (*30*, *31*). The fact that histones modify a neighboring transcription factor's binding suggests that allostery through DNA might be physiologically important in affecting gene regulation.

References and Notes

1. J. Monod, J. Wyman, J. P. Changeux, *J. Mol. Biol.* **12**, 88 (1965).
2. D. E. Koshland Jr., G. Némethy, D. Filmer, *Biochemistry* **5**, 365 (1966).
3. F. M. Pohl, T. M. Jovin, W. Baehr, J. J. Holbrook, *Proc. Natl. Acad. Sci. U.S.A.* **69**, 3805 (1972).
4. M. Hogan, N. Dattagupta, D. M. Crothers, *Nature* **278**, 521 (1979).
5. B. S. Parekh, G. W. Hatfield, *Proc. Natl. Acad. Sci. U.S.A.* **93**, 1173 (1996).
6. J. Rudnick, R. Bruinsma, *Biophys. J.* **76**, 1725 (1999).
7. D. Panne, T. Maniatis, S. C. Harrison, *Cell* **129**, 1111 (2007).
8. R. Moretti *et al.*, *ACS Chem. Biol.* **3**, 220 (2008).
9. E. F. Koslover, A. J. Spakowitz, *Phys. Rev. Lett.* **102**, 178102 (2009).
10. H. G. Garcia *et al.*, *Cell Reports* **2**, 150 (2012).
11. S. Kim, P. C. Blainey, C. M. Schroeder, X. S. Xie, *Nat. Methods* **4**, 397 (2007).
12. Materials and methods are available as supplementary materials on *Science* Online.
13. B. F. Luisi *et al.*, *Nature* **352**, 497 (1991).
14. M. Newman, T. Strzelecka, L. F. Dorner, I. Schildkraut, A. K. Aggarwal, *Science* **269**, 656 (1995).
15. F. K. Winkler *et al.*, *EMBO J.* **12**, 1781 (1993).
16. M. Lewis *et al.*, *Science* **271**, 1247 (1996).
17. C. G. Kalodimos *et al.*, *EMBO J.* **21**, 2866 (2002).
18. K. J. Durniak, S. Bailey, T. A. Steitz, *Science* **322**, 553 (2008).
19. S. B. Smith, L. Finzi, C. Bustamante, *Science* **258**, 1122 (1992).
20. Z. Bryant *et al.*, *Nature* **424**, 338 (2003).
21. A. A. Travers, *Annu. Rev. Biochem.* **58**, 427 (1989).
22. R. Rohs *et al.*, *Annu. Rev. Biochem.* **79**, 233 (2010).
23. A. Újvári, C. T. Martin, *J. Mol. Biol.* **295**, 1173 (2000).
24. G.-Q. Tang, S. S. Patel, *Biochemistry* **45**, 4936 (2006).
25. Y. Jia, A. Kumar, S. S. Patel, *J. Biol. Chem.* **271**, 30451 (1996).
26. J. H. Miller, *Experiments in Molecular Genetics* (Cold Spring Harbor Laboratory, 1972).
27. J. Müller, S. Oehler, B. Müller-Hill, *J. Mol. Biol.* **257**, 21 (1996).
28. W. Lee *et al.*, *Nat. Genet.* **39**, 1235 (2007).
29. E. Sharon *et al.*, *Nat. Biotechnol.* **30**, 521 (2012).
30. S. John *et al.*, *Nat. Genet.* **43**, 264 (2011).
31. S. H. Meijsing *et al.*, *Science* **324**, 407 (2009).
32. H. Viadiu, A. K. Aggarwal, *Nat. Struct. Biol.* **5**, 910 (1998).
33. R. Daber, S. Stayrook, A. Rosenberg, M. Lewis, *J. Mol. Biol.* **370**, 609 (2007).
34. P. T. Lowary, J. Widom, *J. Mol. Biol.* **276**, 19 (1998).

Acknowledgments: We thank K. Wood for his early involvement and J. Hynes, A. Szabo, C. Bustamante, and J. Gelles for helpful discussions. This work is supported by NIH Director's Pioneer Award to X.S.X., Peking University for BIOPIC, Thousand Youth Talents Program for Y.S., as well as the Major State Basic Research Development Program (2011CB809100), National Natural Science Foundation of China (31170710, 31271423, 21125311).

Supplementary Materials

www.sciencemag.org/cgi/content/full/339/6121/816/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S18
Tables S1 to S11
References (*35*–*75*)

23 August 2012; accepted 7 November 2012
10.1126/science.1229223

Multiplex Genome Engineering Using CRISPR/Cas Systems

Le Cong,[1,2]* F. Ann Ran,[1,4]* David Cox,[1,3] Shuailiang Lin,[1,5] Robert Barretto,[6] Naomi Habib,[1] Patrick D. Hsu,[1,4] Xuebing Wu,[7] Wenyan Jiang,[8] Luciano A. Marraffini,[8] Feng Zhang[1]†

Functional elucidation of causal genetic variants and elements requires precise genome editing technologies. The type II prokaryotic CRISPR (clustered regularly interspaced short palindromic repeats)/Cas adaptive immune system has been shown to facilitate RNA-guided site-specific DNA cleavage. We engineered two different type II CRISPR/Cas systems and demonstrate that Cas9 nucleases can be directed by short RNAs to induce precise cleavage at endogenous genomic loci in human and mouse cells. Cas9 can also be converted into a nicking enzyme to facilitate homology-directed repair with minimal mutagenic activity. Lastly, multiple guide sequences can be encoded into a single CRISPR array to enable simultaneous editing of several sites within the mammalian genome, demonstrating easy programmability and wide applicability of the RNA-guided nuclease technology.

Precise and efficient genome-targeting technologies are needed to enable systematic reverse engineering of causal genetic variations by allowing selective perturbation of individual genetic elements. Although genome-editing technologies such as designer zinc fingers (ZFs) (*1*–*4*), transcription activator–like effectors (TALEs) (*4*–*10*), and homing meganucleases (*11*) have begun to enable targeted genome modifications, there remains a need for new technologies that are scalable, affordable, and easy to engineer. Here, we report the development of a class of precision genome-engineering tools based on the RNA-guided Cas9 nuclease (*12*–*14*) from the type II prokaryotic clustered regularly interspaced short palindromic repeats (CRISPR) adaptive immune system (*15*–*18*).

The *Streptococcus pyogenes* SF370 type II CRISPR locus consists of four genes, including the Cas9 nuclease, as well as two noncoding CRISPR RNAs (crRNAs): trans-activating crRNA (tracrRNA) and a precursor crRNA (pre-crRNA) array containing nuclease guide sequences (spacers) interspaced by identical direct repeats (DRs) (fig. S1) (*19*). We sought to harness this prokaryotic

[1]Broad Institute of MIT and Harvard, 7 Cambridge Center, Cambridge, MA 02142, USA, and McGovern Institute for Brain Research, Department of Brain and Cognitive Sciences, Department of Biological Engineering, Massachusetts Institute of Technology (MIT), Cambridge, MA 02139, USA. [2]Program in Biological and Biomedical Sciences, Harvard Medical School, Boston, MA 02115, USA. [3]Harvard-MIT Health Sciences and Technology, Harvard Medical School, Boston, MA 02115, USA. [4]Department of Molecular and Cellular Biology, Harvard University, Cambridge, MA 02138, USA. [5]School of Life Sciences, Tsinghua University, Beijing 100084, China. [6]Department of Biochemistry and Molecular Biophysics, College of Physicians and Surgeons, Columbia University, New York, NY 10032, USA. [7]Computational and Systems Biology Graduate Program and Koch Institute for Integrative Cancer Research, Massachusetts Institute of Technology, Cambridge, MA 02139, USA. [8]Laboratory of Bacteriology, The Rockefeller University, 1230 York Avenue, New York, NY 10065, USA.

*These authors contributed equally to this work.
†To whom correspondence should be addressed. E-mail: zhang@broadinstitute.org

RNA-programmable nuclease system to introduce targeted double-stranded breaks (DSBs) in mammalian chromosomes through heterologous expression of the key components. It has been previously shown that expression of tracrRNA, pre-crRNA, host factor ribonuclease (RNase) III, and Cas9 nuclease is necessary and sufficient for cleavage of DNA in vitro (*12*, *13*) and in prokaryotic cells (*20*, *21*). We codon-optimized the *S. pyogenes Cas9* (*SpCas9*) and *RNase III* (*SpRNase III*) genes and attached nuclear localization signals (NLSs) to ensure nuclear compartmentalization in mammalian cells. Expression of these constructs in human 293FT cells revealed that two NLSs are most efficient at targeting SpCas9 to the nucleus (Fig. 1A). To reconstitute the noncoding RNA components of the *S. pyogenes* type II CRISPR/Cas system, we expressed an 89-nucleotide (nt) tracrRNA (fig. S2) under the RNA polymerase III U6 promoter (Fig. 1B). Similarly, we used the U6 promoter to drive the expression of a pre-crRNA array comprising a single guide spacer flanked by DRs (Fig. 1B). We designed our initial spacer to target a 30–base pair (bp) site (protospacer) in the human *EMX1* locus that precedes an NGG trinucleotide, the requisite protospacer-adjacent motif (PAM) (Fig. 1C and fig. S1) (*22*, *23*).

To test whether heterologous expression of the CRISPR system (SpCas9, SpRNase III, tracrRNA, and pre-crRNA) can achieve targeted cleavage of mammalian chromosomes, we transfected 293FT cells with different combinations of CRISPR/Cas components. Because DSBs in mammalian DNA are partially repaired by the indel-forming nonhomologous end joining (NHEJ) pathway, we used the SURVEYOR assay (fig. S3) to detect endogenous target cleavage (Fig. 1D and fig. S2B). Cotransfection of all four required CRISPR components resulted in efficient cleavage of the protospacer (Fig. 1D and fig. S2B), which was subsequently verified by Sanger sequencing (Fig. 1E). SpRNase III was not necessary for cleavage of the protospacer (Fig. 1D), and the 89-nt tracrRNA is processed in its absence (fig. S2C). Similarly, maturation of pre-crRNA does not require RNase III (Fig. 1D and fig. S4), suggesting that there may be endogenous mammalian RNases that assist in pre-crRNA maturation (*24*–*26*). Removing any of the remaining RNA or Cas9 components abolished the genome cleavage activity of the CRISPR/Cas system (Fig. 1D). These results define a minimal three-component system for efficient RNA-guided genome modification in mammalian cells.

Next, we explored the generalizability of RNA-guided genome editing in eukaryotic cells by targeting additional protospacers within the *EMX1* locus (Fig. 2A). To improve codelivery, we designed an expression vector to drive both pre-crRNA and SpCas9 (fig. S5). In parallel, we adapted a chimeric crRNA-tracrRNA hybrid (Fig. 2B, top) design recently validated in vitro (*12*), where a mature crRNA is fused to a partial tracrRNA via a synthetic stem loop to mimic the natural crRNA:tracrRNA duplex (Fig. 2B, bottom). We observed cleavage of all protospacer targets when SpCas9 is coexpressed with pre-crRNA (DR-spacer-DR) and tracrRNA. However, not all chimeric RNA designs could facilitate cleavage of their genomic targets (Fig. 2C and table S1). We then tested targeting of additional genomic loci in both human and mouse cells by designing pre-crRNAs and chimeric RNAs targeting the human *PVALB* and the mouse *Th* loci (fig. S6). We achieved efficient modification at all three mouse *Th* and one *PVALB* targets by using the crRNA:tracrRNA duplex, thus demonstrating the broad applicability of the CRISPR/Cas system in modifying different loci across multiple organisms (table S1). For the same protospacer targets, cleavage efficiencies of chimeric RNAs were either lower than those of crRNA:tracrRNA duplexes or undetectable. This may be due to differences in the expression and stability of RNAs, degradation by endogenous RNA interference machinery, or secondary structures leading to inefficient Cas9 loading or target recognition.

Effective genome editing requires that nucleases target specific genomic loci with both high

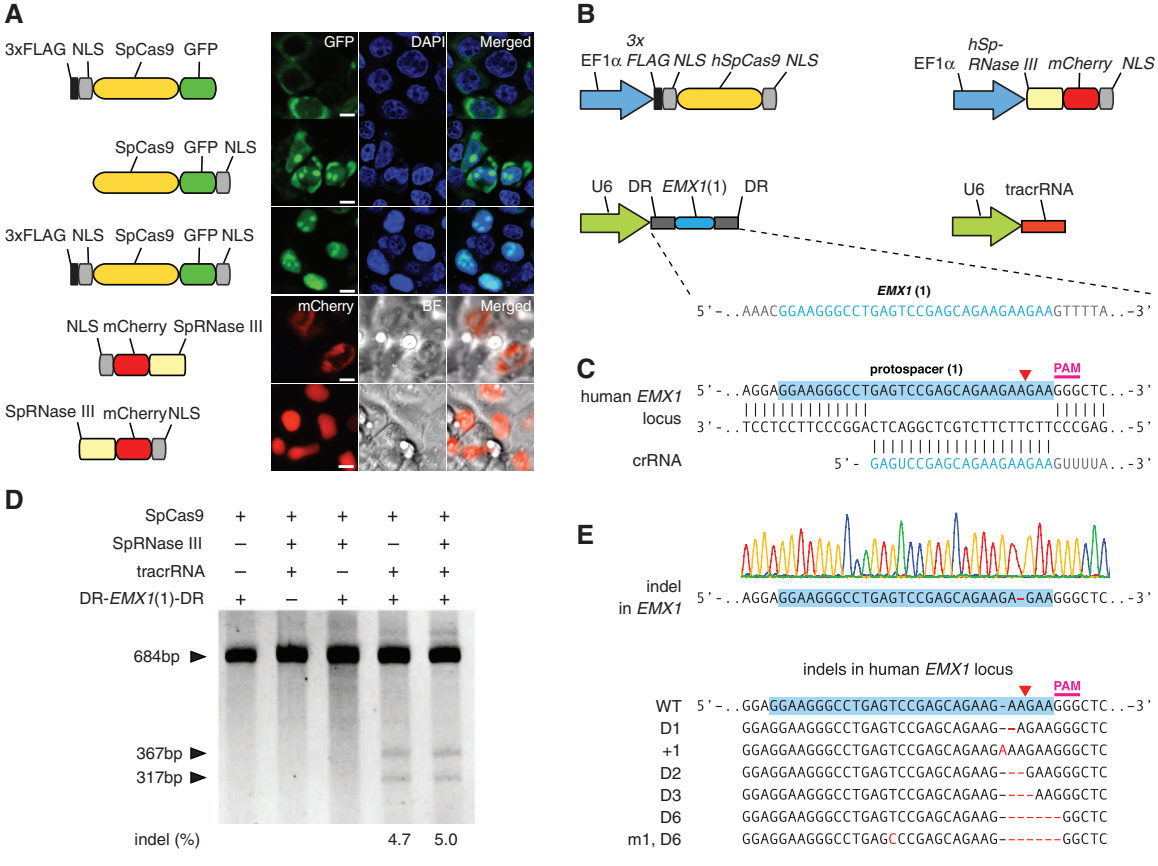

Fig. 1. The type II CRISPR locus from *S. pyogenes* SF370 can be reconstituted in mammalian cells to facilitate targeted DSBs of DNA. (**A**) Engineering of SpCas9 and SpRNase III with NLSs enables import into the mammalian nucleus. GFP indicates green fluorescent protein; scale bars, 10 μm. (**B**) Mammalian expression of human codon–optimized SpCas9 (hSpCas9) and SpRNase III (hSpRNase III) genes were driven by the elongation factor 1α (EF1α) promoter, whereas tracrRNA and pre-crRNA array (DR-Spacer-DR) were driven by the U6 promoter. A protospacer (blue highlight) from the human *EMX1* locus with PAM was used as template for the spacer in the pre-crRNA array. (**C**) Schematic representation of base pairing between target locus and *EMX1*-targeting crRNA. Red arrow indicates putative cleavage site. (**D**) SURVEYOR assay for SpCas9-mediated indels. (**E**) An example chromatogram showing a microdeletion, as well as representative sequences of mutated alleles identified from 187 clonal amplicons. Red dashes, deleted bases; red bases, insertions or mutations.

precision and efficiency. To investigate the specificity of RNA-guided genome modification, we analyzed single-nucleotide mismatches between the spacer and its mammalian protospacer target (Fig. 3A). We observed that single-base mismatch up to 11 bp 5′ of the PAM completely abolished genomic cleavage by SpCas9, whereas spacers with mutations farther upstream retained activity against the protospacer target (Fig. 3B). This is consistent with previous bacterial and in vitro studies of Cas9 specificity (*12*, *20*). Furthermore, SpCas9 is able to mediate genomic cleavage as efficiently as a pair of TALE nucleases (TALENs) targeting the same *EMX1* protospacer (Fig. 3, C and D).

Targeted modification of genomes ideally avoids mutations arising from the error-prone NHEJ mechanism. The wild-type SpCas9 is able to mediate site-specific DSBs, which can be repaired through either NHEJ or homology-directed repair (HDR). We engineered an aspartate-to-alanine substitution (D10A) in the RuvC I domain of SpCas9 to convert the nuclease into a DNA nickase (SpCas9n, Fig. 4A) (*12*, *13*, *20*), because nicked genomic DNA is typically repaired either seamlessly or through high-fidelity HDR. SURVEYOR (Fig. 4B) and sequencing of

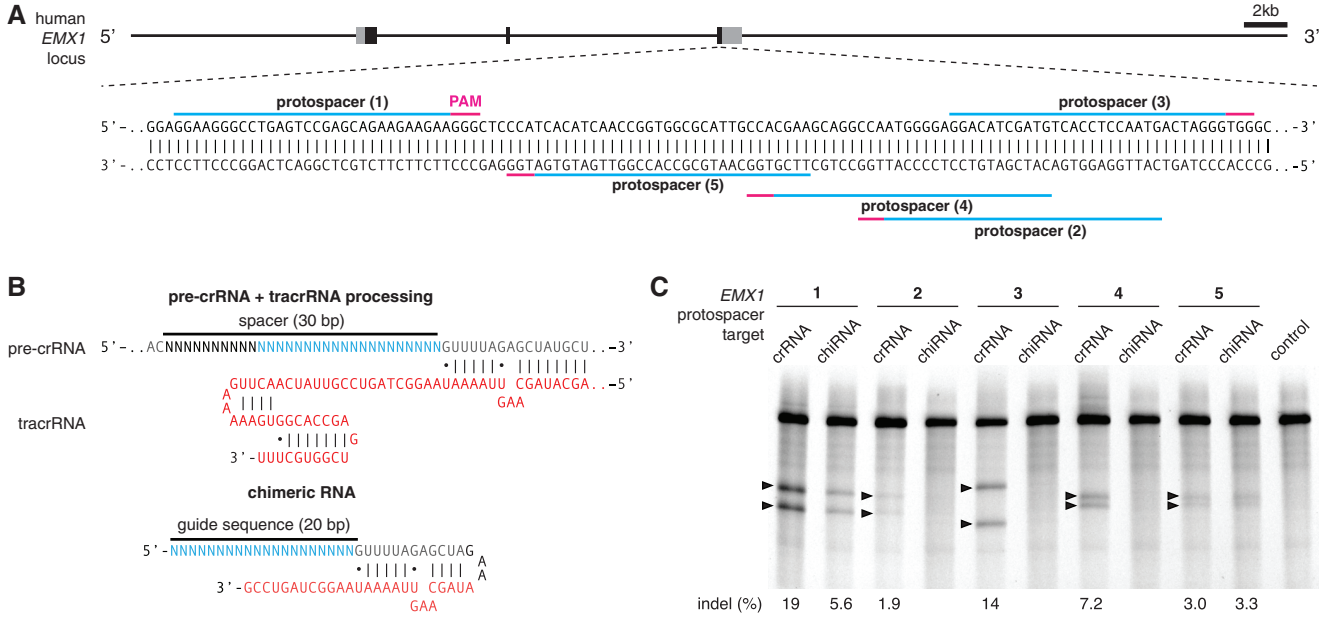

Fig. 2. SpCas9 can be reprogrammed to target multiple genomic loci in mammalian cells. (**A**) Schematic of the human *EMX1* locus showing the location of five protospacers indicated by blue lines with corresponding PAM in magenta. (**B**) Schematic of the pre-crRNA:tracrRNA complex (top) showing hybridization between the direct repeat (gray) region of the pre-crRNA and tracrRNA. Schematic of a chimeric RNA design (*12*) (bottom). tracrRNA sequence is shown in red and the 20-bp spacer sequence in blue. (**C**) SURVEYOR assay comparing the efficacy of Cas9-mediated cleavage at five protospacers in the human *EMX1* locus. Each protospacer was targeted by using either processed pre-crRNA:tracrRNA complex (crRNA) or chimeric RNA (chiRNA). Arrowheads indicate cleavage products for each protospacer target.

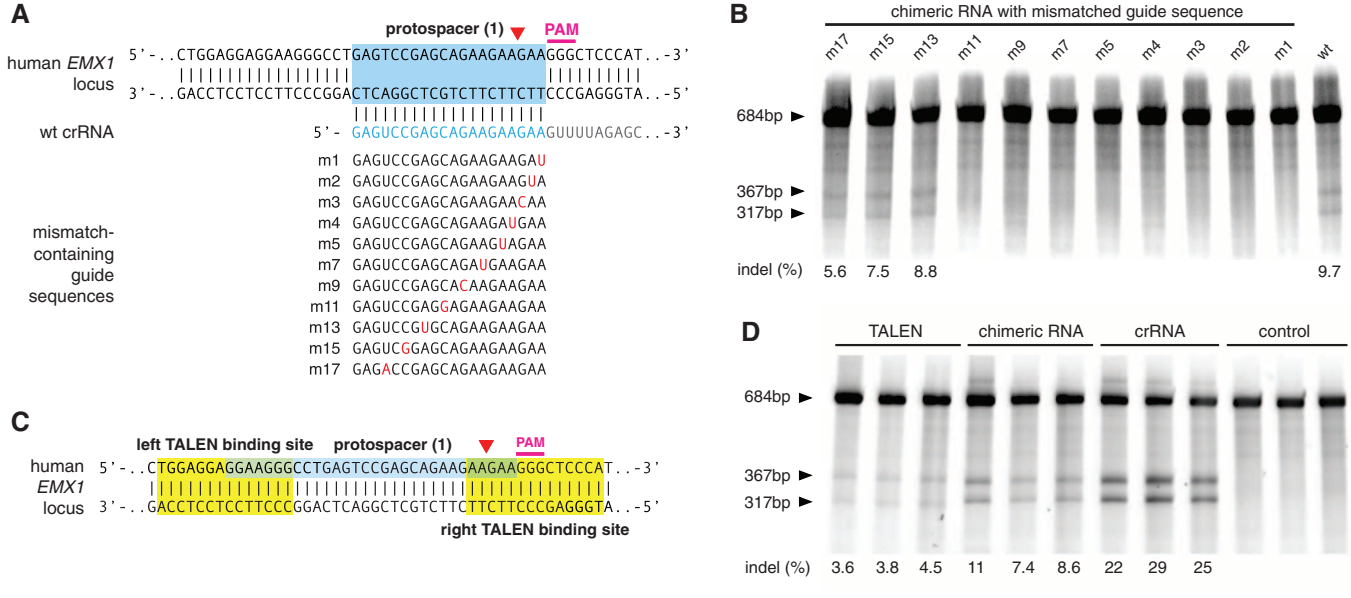

Fig. 3. Evaluation of the SpCas9 specificity and comparison of efficiency with TALENs. (**A**) *EMX1*-targeting chimeric crRNAs with single point mutations were generated to evaluate the effects of spacer-protospacer mismatches. (**B**) SURVEYOR assay comparing the cleavage efficiency of different mutant chimeric RNAs. (**C**) Schematic showing the design of TALENs that target *EMX1*. (**D**) SURVEYOR gel comparing the efficiency of TALEN and SpCas9 ($N = 3$).

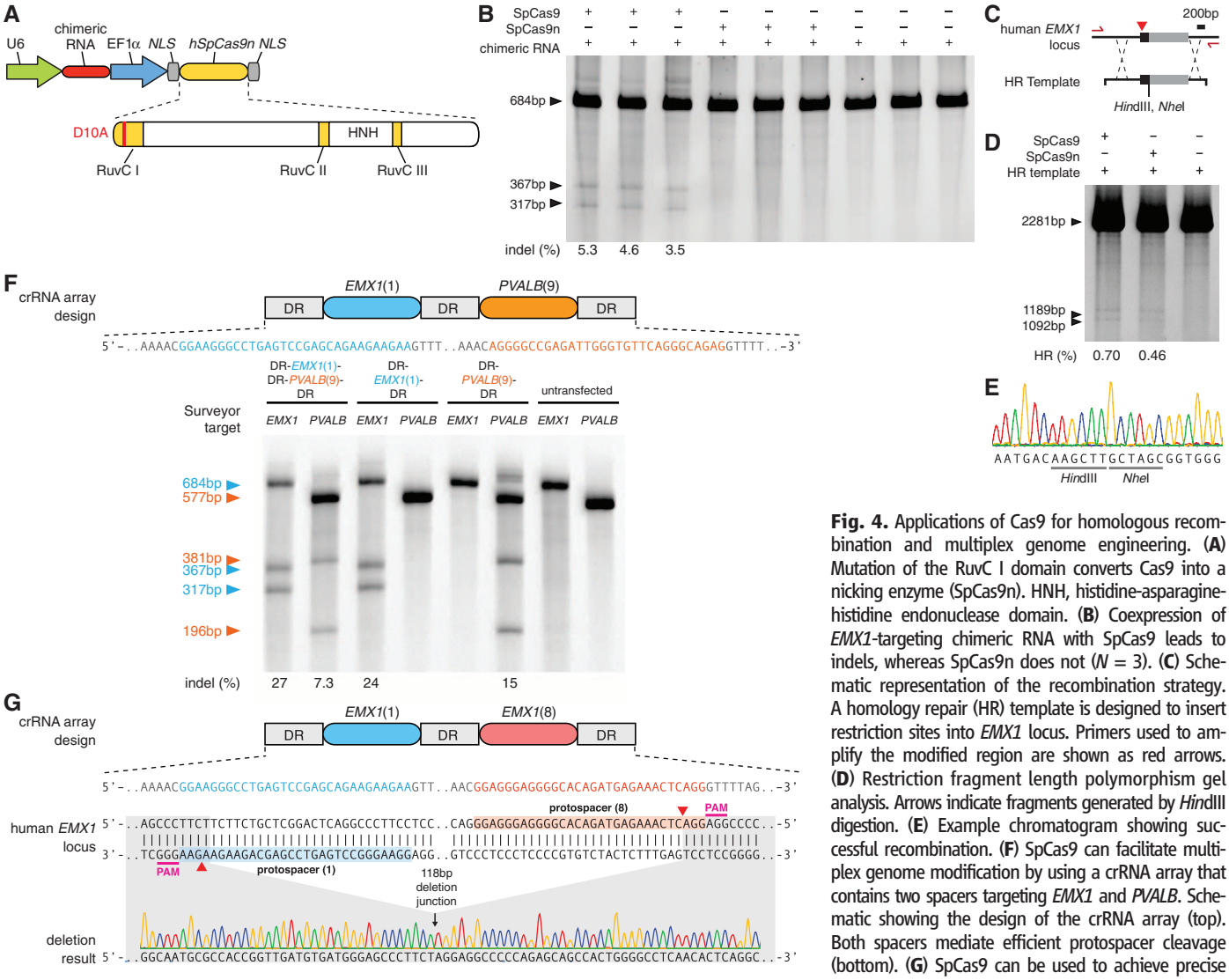

Fig. 4. Applications of Cas9 for homologous recombination and multiplex genome engineering. (**A**) Mutation of the RuvC I domain converts Cas9 into a nicking enzyme (SpCas9n). HNH, histidine-asparagine-histidine endonuclease domain. (**B**) Coexpression of *EMX1*-targeting chimeric RNA with SpCas9 leads to indels, whereas SpCas9n does not ($N = 3$). (**C**) Schematic representation of the recombination strategy. A homology repair (HR) template is designed to insert restriction sites into *EMX1* locus. Primers used to amplify the modified region are shown as red arrows. (**D**) Restriction fragment length polymorphism gel analysis. Arrows indicate fragments generated by *Hin*dIII digestion. (**E**) Example chromatogram showing successful recombination. (**F**) SpCas9 can facilitate multiplex genome modification by using a crRNA array that contains two spacers targeting *EMX1* and *PVALB*. Schematic showing the design of the crRNA array (top). Both spacers mediate efficient protospacer cleavage (bottom). (**G**) SpCas9 can be used to achieve precise genomic deletion. Two spacers targeting *EMX1* (top) mediated a 118-bp genomic deletion (bottom).

327 amplicons did not detect any indels induced by SpCas9n. However, nicked DNA can in rare cases be processed via a DSB intermediate and result in a NHEJ event (*27*). We then tested Cas9-mediated HDR at the same *EMX1* locus with a homology repair template to introduce a pair of restriction sites near the protospacer (Fig. 4C). SpCas9 and SpCas9n catalyzed integration of the repair template into *EMX1* locus at similar levels (Fig. 4D), which we further verified via Sanger sequencing (Fig. 4E). These results demonstrate the utility of CRISPR for facilitating targeted genomic insertions. Given the 14-bp (12 bp from the seed sequence and 2 bp from PAM) target specificity (Fig. 3B) of the wild-type SpCas9, the use of a nickase may reduce off-target mutations.

Lastly, the natural architecture of CRISPR loci with arrayed spacers (fig. S1) suggests the possibility of multiplexed genome engineering. By using a single CRISPR array encoding a pair of *EMX1*- and *PVALB*-targeting spacers, we detected efficient cleavage at both loci (Fig. 4F). We further tested targeted deletion of larger genomic regions through concurrent DSBs by using spacers against two targets within *EMX1* spaced by 119 bp and observed a 1.6% deletion efficacy (3 out of 182 amplicons, Fig. 4G), thus demonstrating the CRISPR/Cas system can mediate multiplexed editing within a single genome.

The ability to use RNA to program sequence-specific DNA cleavage defines a new class of genome engineering tools. Here, we have shown that the *S. pyogenes* CRISPR system can be heterologously reconstituted in mammalian cells to facilitate efficient genome editing; an accompanying study has independently confirmed high-efficiency RNA-guided genome targeting in several human cell lines (*28*). However, several aspects of the CRISPR/Cas system can be further improved to increase its efficiency and versatility. The requirement for an NGG PAM restricts the target space of SpCas9 to every 8 bp on average in the human genome (fig. S7), not accounting for potential constraints posed by crRNA secondary structure or genomic accessibility resulting from chromatin and DNA methylation states. Some of these restrictions may be overcome by exploiting the family of Cas9 enzymes and its differing PAM requirements (*22*, *23*) across the microbial diversity (*17*). Indeed, other CRISPR loci are likely to be transplantable into mammalian cells; for example, the *Streptococcus thermophilus* LMD-9 CRISPR1 system can also mediate mammalian genome cleavage (fig. S8). Lastly, the ability to carry out multiplex genome editing in mammalian cells enables powerful applications across basic science, biotechnology, and medicine (*29*).

References and Notes

1. M. H. Porteus, D. Baltimore, *Science* **300**, 763 (2003).
2. J. C. Miller *et al.*, *Nat. Biotechnol.* **25**, 778 (2007).
3. J. D. Sander *et al.*, *Nat. Methods* **8**, 67 (2011).

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. Sapranauskas *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. Hirschl 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).

Supplementary Materials

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,[1]* Luhan Yang,[1,3]* Kevin M. Esvelt,[2] John Aach,[1] Marc Guell,[1] James E. DiCarlo,[4] Julie E. Norville,[1] George M. Church[1,2]†

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.

Bacterial 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 $GN_{20}GG$ (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 nonfluorescent. Homologous recombination (HR) using an appropriate repair donor can restore the normal GFP sequence, which enabled us to quantify the resulting GFP^+ 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^+ 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

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