

# Targeted genome engineering in human cells with the Cas9 RNA-guided endonuclease

Seung Woo Cho<sup>1-3</sup>, Sojung Kim<sup>1-3</sup>, Jong Min Kim<sup>1,2</sup> & Jin-Soo Kim<sup>1,2</sup>

**We employ the CRISPR-Cas system of *Streptococcus pyogenes* as programmable RNA-guided endonucleases (RGENs) to cleave DNA in a targeted manner for genome editing in human cells. We show that complexes of the Cas9 protein and artificial chimeric RNAs efficiently cleave two genomic sites and induce indels with frequencies of up to 33%.**

The clustered, regularly interspaced, short palindromic repeats (CRISPR)–CRISPR-associated protein (Cas) system provides prokaryotes with adaptive immunity to viruses and plasmids<sup>1</sup>. When Cas9, a protein component in the type II CRISPR-Cas system, is complexed with two RNAs called CRISPR RNA (crRNA) and trans-activating crRNA (tracrRNA), it forms a sequence-specific endonuclease that cleaves foreign genetic sequences to protect host cells. The crRNA is transcribed from a host genomic CRISPR element previously captured from such foreign invaders. A single-chain chimeric RNA produced by fusing crRNA and tracrRNA sequences can replace the two RNAs in the Cas9-RNA complex to form a single-guide-RNA:Cas9 endonuclease (sgRNA:Cas9)<sup>2</sup>. Thus, in contrast to the widely used genome-editing technologies based on zinc finger nucleases (ZFNs) and transcription activator–like effector nucleases (TALENs), the specificity of RNA-guided endonucleases (RGENs) can be customized by replacing a short synthetic RNA molecule without changing the protein component. Here we show that sgRNA:Cas9 can induce site-specific genome modifications in human cells at high frequencies.

We first tested the DNA cleavage activity of the Cas9 derived from *Streptococcus pyogenes* (Supplementary Fig. 1) in the presence or absence of a chimeric RNA *in vitro* (Supplementary Methods and Supplementary Table 1). To this end, we used recombinant Cas9 protein to cleave plasmid DNA that contained a 23-bp target sequence from human *CCR5*, which encodes an essential co-receptor of HIV and is a potential target for the treatment for AIDS<sup>3</sup>. A Cas9 target sequence consists of a 20-bp DNA sequence complementary to the crRNA or the chimeric RNA and the trinucleotide (5′-NGG-3′) protospacer adjacent motif (PAM) recognized by Cas9 itself (Fig. 1a)<sup>2</sup>. Cas9 cleaved the plasmid efficiently at the expected position only in the presence of the synthetic RNA and did not cleave a control plasmid that lacked the target sequence (Fig. 1b).

Next, we used an RFP-GFP reporter<sup>4</sup> to investigate whether the sgRNA:Cas9 can cleave the target sequence incorporated between the

RFP and GFP sequences in human embryonic kidney (HEK) 293T cells (Fig. 1c). In this reporter, the GFP sequence is fused to the RFP sequence out-of-frame. GFP is expressed only when the target sequence is cleaved by site-specific nucleases, which causes small, frameshifting insertions or deletions around the target sequence by means of error-prone nonhomologous end joining repair of the double-strand breaks. We found that GFP-expressing cells were obtained at frequencies ranging from 5% to 7% in three independent experiments only when the cells were co-transfected with the *in vitro* transcribed chimeric RNA and a plasmid encoding Cas9, demonstrating that sgRNA:Cas9 can recognize and cleave the target DNA sequence in eukaryotic cells.

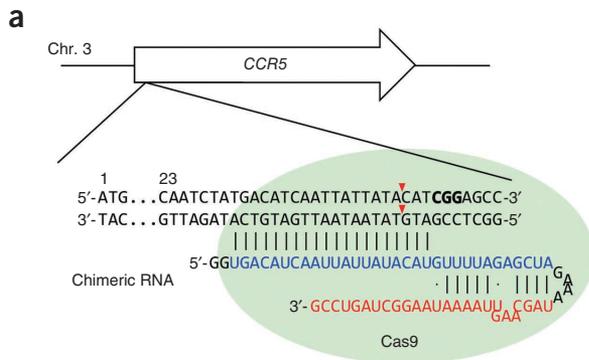
To test whether the RGENs could be used for targeted disruption of endogenous genes in human cells, we analyzed genomic DNA isolated from transfected cells using T7 endonuclease I (T7EI), an endonuclease that specifically cleaves heteroduplexes formed by the hybridization of wild-type and mutant DNA sequences<sup>5</sup>. We found that mutations were induced only when the cells were co-transfected with both Cas9 and RNA. Mutation frequencies estimated from the relative DNA band intensities were RNA-dosage dependent, ranging from 2.0% to 18% in three independent experiments (Fig. 2a and Supplementary Fig. 2a), on par with those obtained with ZFNs or TALENs at the *CCR5* locus<sup>5-8</sup>. DNA sequencing analysis of the PCR amplicons corroborated the induction of sgRNA:Cas9-mediated mutations at the endogenous site (Fig. 2a).

Both ZFNs and TALENs are associated with off-target effects<sup>7,9</sup>, which may limit the utility and safety profile of these enzymes in gene or cell therapy. Furthermore, the repair of two concurrent double-strand breaks induced by these enzymes at on-target and off-target sites can give rise to genome rearrangements<sup>10,11</sup>. The most important off-target sites associated with *CCR5*-specific ZFNs and TALENs reside in *CCR2*, a close homolog of *CCR5*, located 15 kbp upstream of *CCR5*. To avoid off-target mutations in *CCR2* and unwanted chromosomal rearrangements, we intentionally chose the target site of our RGEN to recognize a region within the *CCR5* sequence that has no apparent homology with the *CCR2* sequence.

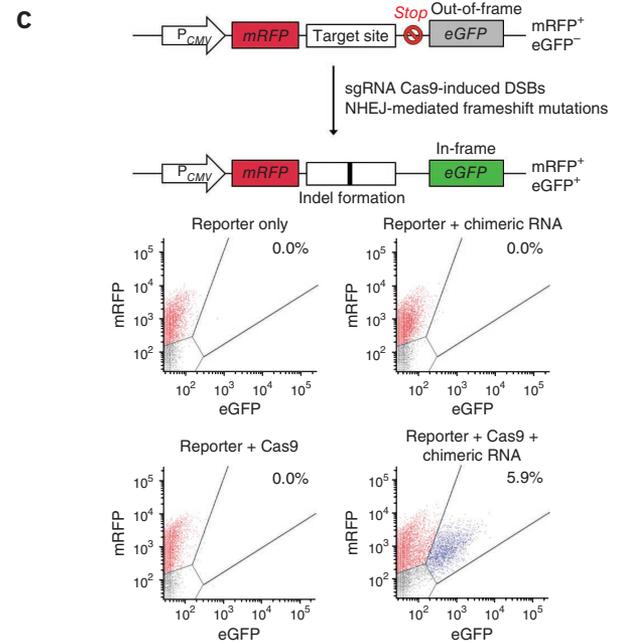
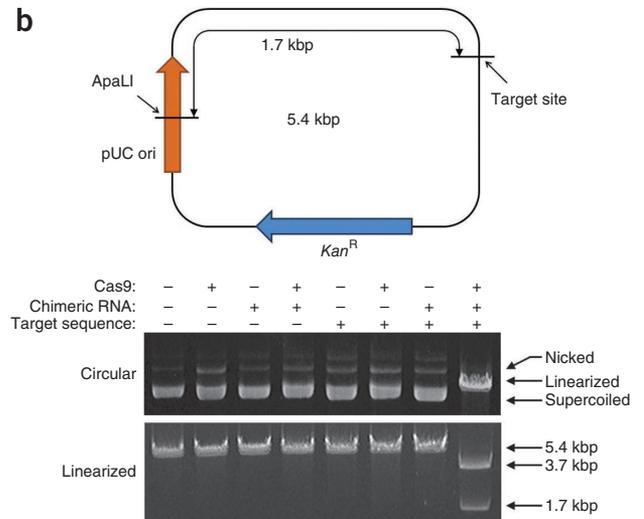
To investigate whether the *CCR5*-specific RGEN was associated with off-target effects, we searched for potential off-target sites in the human genome that are most homologous to the 23-bp target sequence. As expected, no such sites were found in *CCR2*. Instead, we found four sites elsewhere in the human genome, each of which carries 3-base mismatches to the on-target site (Fig. 2b). The T7EI assays showed that mutations were not induced at these sites (assay sensitivity, ~0.5%). Furthermore, we used PCR to detect the induction of chromosomal deletions in cells separately transfected with plasmids encoding either the ZFN or the chimeric RNA specific to *CCR5* and a plasmid encoding Cas9. As expected, because of the difference in target sites, the ZFN induced deletions, whereas gRNA-Cas9 did not (Supplementary Fig. 3). Although we did not detect any off-target effects with the *CCR5*-specific RGEN in this study,

<sup>1</sup>National Creative Research Initiatives Center for Genome Engineering, Seoul National University, Seoul, South Korea. <sup>2</sup>Department of Chemistry, Seoul National University, Seoul, South Korea. <sup>3</sup>These authors contributed equally to this work. Correspondence should be addressed to J.-S.K. (jskim01@snu.ac.kr).

Received 20 November 2012; accepted 14 January 2013; published online 29 January 2013; doi:10.1038/nbt.2507



**Figure 1** RGEN-catalyzed cleavage of plasmid DNA *in vitro* and *in cellula*. (a) Schematic representation of target DNA and chimeric RNA sequences. Red triangles indicate cleavage sites. The PAM sequence recognized by Cas9 is shown in bold. The sequences in the chimeric RNA derived from crRNA and tracrRNA are shown in red and blue, respectively. (b) *In vitro* cleavage of plasmid DNA by Cas9. An intact circular plasmid or ApaI-digested plasmid was incubated with Cas9 and chimeric RNA. (c) RGEN-induced mutagenesis at an episomal target site. (Top) Schematic overview of cell-based assays using a RFP-GFP reporter. (Bottom) Flow cytometry of cells transfected with Cas9 and chimeric RNA. The percentage of cells that express RFP and GFP is indicated. DBSs, double-strand breaks; NHEJ, nonhomologous end joining.



it is possible that RGENs in general can induce off-target mutations at sites that carry mismatches distal from the PAM sequence<sup>2</sup>. In addition, deep sequencing and whole-genome sequencing may reveal off-target mutations induced by Cas9-based endonucleases. Further studies are warranted to clarify this issue.

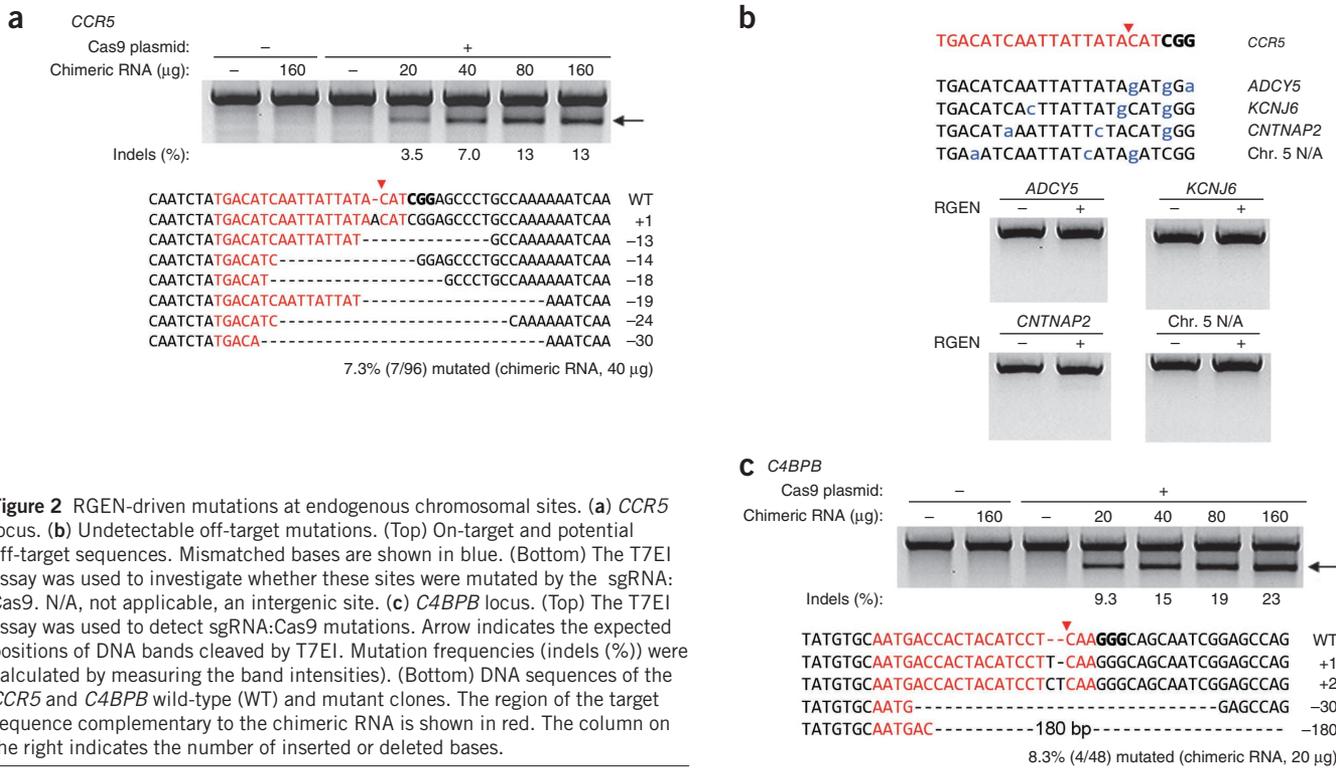
Next, we reprogrammed Cas9 by replacing the *CCR5*-specific chimeric RNA with a newly synthesized RNA designed to target human *C4BPB*, which encodes the beta chain of the C4b-binding protein, a transcription factor. This enzyme induced mutations at the chromosomal target site in human K562 cells at high, dose-dependent frequencies that ranged from 5% to 33% in three independent experiments (Fig. 2c and Supplementary Fig. 2b). Sequencing analysis of PCR amplicons revealed that, out of four mutant sequences, two clones contained a single-base or two-base insertion precisely at the cleavage site, a pattern that was also observed at the *CCR5* site. These results indicate that sgRNA:Cas9 complexes cleave chromosomal DNA at expected positions.

The RGENs used in this study induced no obvious cytotoxicity. Mutation frequencies were stably maintained even 8 d after transfection (Supplementary Fig. 4), which allowed us to isolate clonal populations of mutant cells by limiting dilution (Supplementary Fig. 5). As expected, the sgRNA-Cas9 endonucleases produced ~1 double-strand break in each cell, which was detected by TP53BP1 staining (Supplementary Fig. 6).

ZFNs and TALENs enable targeted mutagenesis in a variety of eukaryotic cells, plants and animals, but the mutation frequencies obtained with individual nucleases vary widely. Furthermore, some ZFNs and TALENs fail to show any genome editing activities<sup>12,13</sup> and it is technically challenging and time consuming to make custom nucleases. In this regard, RGENs could provide useful options for genome editing. Compared to ZFNs and TALENs, RGENs can be more readily customized because, as only the RNA component has to be replaced to make a new genome-editing nuclease, no subcloning steps are involved to make Cas9-based nucleases with new specificities. (While this paper was under review,

two other groups independently reported genome editing in mammalian cells using the gRNA-Cas9 system, but their methods require a subcloning step to make plasmids that encode the gRNA<sup>14,15</sup>.) The relatively small size of the Cas9 gene (4.1 kbp) compared with a pair of TALEN genes (~6 kbp) provides an advantage for this system in gene delivery. These features will make RGENs scalable and convenient tools for genome engineering in cells and organisms.

Unlike ZFNs and TALENs, whose specificities of DNA recognition are tunable by changing the number of DNA-binding modules, Cas9-based RGENs recognize target sequences of a fixed length. The specificity of Cas9-based RGENs is further limited by the requirement for a 5'-GG-3' dinucleotide in the PAM sequence. Thus, RGENs can be designed to cleave DNA once per 8 bp (= 4 × 4/2) on average. This limitation might be relieved by engineering Cas9. Whereas FokI-based ZFNs and TALENs produce 5' overhangs at cleavage sites, Cas9-based RGENs yield blunt ends rather than cohesive ends<sup>2</sup>. Our results show that double-strand breaks with blunt ends can also be readily repaired in mammalian cells. It would be interesting to investigate how and whether blunt, double-strand-break ends are differentially repaired by the endogenous end-joining processes.



**Figure 2** RGEN-driven mutations at endogenous chromosomal sites. **(a)** *CCR5* locus. **(b)** Undetectable off-target mutations. (Top) On-target and potential off-target sequences. Mismatched bases are shown in blue. (Bottom) The T7E1 assay was used to investigate whether these sites were mutated by the sgRNA: Cas9. N/A, not applicable, an intergenic site. **(c)** *C4BPB* locus. (Top) The T7E1 assay was used to detect sgRNA:Cas9 mutations. Arrow indicates the expected positions of DNA bands cleaved by T7E1. Mutation frequencies (indels (%)) were calculated by measuring the band intensities. (Bottom) DNA sequences of the *CCR5* and *C4BPB* wild-type (WT) and mutant clones. The region of the target sequence complementary to the chimeric RNA is shown in red. The column on the right indicates the number of inserted or deleted bases.

Two independent studies in this issue report RGEN-mediated genome engineering in microbial organisms<sup>16</sup> and in zebrafish<sup>17</sup>, demonstrating broad applications of this new technology.

Note: Supplementary information is available in the online version of the paper.

**ACKNOWLEDGMENTS**

This work was supported by the National Research Foundation of Korea (2012-0001225) and ToolGen, Inc. We thank Jae Kyung Chon for bioinformatic analysis.

**AUTHOR CONTRIBUTIONS**

S.W.C., S.K. and J.M.K. performed the experiments. J.-S.K. wrote the manuscript.

**COMPETING FINANCIAL INTERESTS**

The authors declare competing financial interests: details are available in the online version of the paper.

Published online at <http://www.nature.com/doi/10.1038/nbt.2507>.

Reprints and permissions information is available online at <http://www.nature.com/reprints/index.html>.

1. Wiedenheft, B., Sternberg, S.H. & Doudna, J.A. *Nature* **482**, 331–338 (2012).
2. Jinek, M. *et al. Science* **337**, 816–821 (2012).
3. Cohen, J. *Science* **332**, 784–789 (2011).
4. Kim, H., Um, E., Cho, S.R., Jung, C. & Kim, J.S. *Nat. Methods* **8**, 941–943 (2011).
5. Kim, H.J., Lee, H.J., Kim, H., Cho, S.W. & Kim, J.S. *Genome Res.* **19**, 1279–1288 (2009).
6. Miller, J.C. *et al. Nat. Biotechnol.* **29**, 143–148 (2011).
7. Mussolino, C. *et al. Nucleic Acids Res.* **39**, 9283–9293 (2011).
8. Perez, E.E. *et al. Nat. Biotechnol.* **26**, 808–816 (2008).
9. Kim, E. *et al. Genome Res.* **22**, 1327–1333 (2012).
10. Lee, H.J., Kweon, J., Kim, E., Kim, S. & Kim, J.S. *Genome Res.* **22**, 539–548 (2012).
11. Lee, H.J., Kim, E. & Kim, J.S. *Genome Res.* **20**, 81–89 (2010).
12. Kim, J.S., Lee, H.J. & Carroll, D. *Nat. Methods* **7**, 91, author reply 91–92 (2010).
13. Reyon, D. *et al. Nat. Biotechnol.* **30**, 460–465 (2012).
14. Cong, L. *et al. Science* doi:10.1126/science.1231143 (3 January 2013).
15. Mali, P. *et al. Science* doi:10.1126/science.1232033 (3 January 2013).
16. Jiang, W., Bikard, D., Cox, D., Zhang, F. & Marraffini, L.A. RNA-guided editing of bacterial genomes using CRISPR-Cas systems. *Nat. Biotechnol.* advance online publication, doi:10.1038/nbt.2508 (29 January 2013).
17. Hwang, W.Y. *et al.* Efficient genome editing in zebrafish using a CRISPR-Cas system. *Nat. Biotechnol.* advance online publication, doi:10.1038/nbt.2501 (29 January 2013).