Precise gene deletion and replacement using the CRISPR/Cas9 system in human cells

Qiupeng Zheng\textsuperscript{1,†}, Xiaohong Cai\textsuperscript{2,†}, Meng How Tan\textsuperscript{3,4}, Steven Schaffert\textsuperscript{2}, Christopher P. Arnold\textsuperscript{2}, Xue Gong\textsuperscript{2}, Chang-Zheng Chen\textsuperscript{2,5}, and Shenglin Huang\textsuperscript{1,2}

\textsuperscript{1}Fudan University Shanghai Cancer Center, Institute of Biomedical Sciences, and Department of Oncology, Shanghai Medical School, Fudan University, Shanghai, China, \textsuperscript{2}Baxter Laboratory for Stem Cell Biology, and Department of Microbiology and Immunology, Stanford University School of Medicine, Stanford, CA, \textsuperscript{3}School of Chemical and Biomedical Engineering, Nanyang Technological University, Singapore, \textsuperscript{4}Genome Institute of Singapore, Agency for Science Technology and Research, Singapore, and \textsuperscript{5}Achelois Pharmaceuticals, Inc., San Francisco, CA.

\textsuperscript{†}Q.Z. and X.C. contributed equally to this work

The prokaryotic type II CRISPR/Cas9 system has been adapted to perform targeted genome editing in cells and model organisms. Here, we describe targeted gene deletion and replacement in human cells via the CRISPR/Cas9 system using two guide RNAs. The system effectively generated targeted deletions of varied length, regardless of the transcriptional status of the target gene. It is notable that targeted gene deletions generated via CRISPR/Cas9 and two guide RNAs resulted in the formation of correct junctions at high efficiency. Moreover, in the presence of a homology repair donor, the CRISPR/Cas9 system could guide precise gene replacement. Our results illustrate that the CRISPR/Cas9 system can be used to precisely and effectively generate targeted deletions or gene replacement in human cells, which will facilitate characterization of functional domains in protein-coding genes as well as noncoding regulatory sequences in animal genomes.

METHOD SUMMARY

Here we show that two guide RNAs coupled with Cas9 efficiently generate DNA deletions of up to 10 kb in human cells in a process where repair of the deletion is largely accomplished by precise end joining. In addition, we provide data showing that the CRISPR/Cas9 system can replace large genomic fragments in the presence of a linear homologous repair donor.
mutations generated from repairing a single DSB may not be useful in experiments aimed at characterizing the functional domains of protein-coding genes or for inactivation of genomic elements such as intergenic or intronic regulatory sequences or noncoding RNA genes. DNA fragment deletions in target loci would provide an avenue to study these functional elements. To this end, multiple DSBs have been introduced to generate deletions in Drosophila (12,13), zebrafish (14), and human cells (8), albeit with low efficiency. Targeted genomic DNA deletions have also been achieved using zinc finger nuclease (ZFN) or transcription activator-like effector nuclease (TALEN) in human cells (15–17). However, the efficiencies of these approaches are generally low. In addition, ZFNs and TALENs remain somewhat difficult and expensive to design, develop, and empirically test in the cellular context.

Here, we examined the generation of fragment deletions in human cells catalyzed by the CRISPR/Cas9 system. We show that 2 gRNAs coupled with Cas9 can efficiently create DNA deletions of up to 10 kb. Of interest, we found that repair of this deletion process is largely accomplished by precise end joining. Moreover, targeted deletion with CRISPR/Cas9 appears to be independent of the transcriptional status of the targeted locus. Finally, we show that the CRISPR/Cas9 system can be used to replace large genomic fragments in the presence of a linear homologous repair donor.

Material and Methods

Plasmid construction

The basic H1 promoter was amplified from pLVTHM plasmid (Addgene, #12247, Cambridge, MA). Oligonucleotides containing the modified H1 promoter and backbone of desired gRNA sequences with two BsaI sites were synthesized (PAN Facility, Stanford University). The resulting full-length products were amplified by PCR and cloned into the pUC19 vector. The ampicillin gene (amp) and the H1 promoter in the pUC19 vector contain BsaI restriction enzyme sites; these were mutated (the amp gene was changed from G1601C, which does not alter the amino acid sequence; the H1 promoter was changed from GAGACC to GAGGACC) to eliminate the BsaI sites. The protocol for gRNA cloning is presented in the Supplementary Material. All targeting sites sequences are presented in Supplementary Table S1.

Cell culture

HEK 293T, SK-Hep1, and HeLa cells were cultured in Dulbecco’s modified Eagle’s medium (DMEM) supplemented with 10% fetal bovine serum (FBS) (Hyclone, Logan, UT) and penicillin/streptomycin (pen/strep) (Invitrogen, Carlsbad, CA). PC3 cells were cultured in RPMI-1640 medium supplemented with 10% FBS and pen/strep. For tumor necrosis factor α (TNF-α) stimulation, 293T cells were treated with indicated concentrations of TNF-α (R&D Systems, Minneapolis, MN). Cells were maintained at 37°C and 5% CO2 in a humidified incubator.

Targeted DNA deletion

HEK 293T cells were seeded in 12-well plates at a density of 100,000 cells per well. After 24 h, the cells were transiently transfected with 1 µg Cas9 plasmid (Addgene, #41815), 0.5 µg gRNA T1, and 0.5 µg gRNA T2 plasmids using Lipofectamine 2000 (Invitrogen) as per the manufacturer’s protocols.

Figure 1. An optimized system for targeted gene deletion in human cells with CRISPR/Cas9. (A) Design of the guide RNA (gRNA) expression vector. The vector was designed to produce gRNA transcripts with a synthetic gRNA fused to the trans-activating RNA/CRISPR RNA (tracrRNA). The H1 promoter was modified to eliminate the internal type IIS restriction enzyme BsaI site by changing GAGACC in the H1 promoter to GAGGACC. A BsaI site was introduced to create cloning sites for gRNA and tracrRNA fusions by insertion of synthetic oligonucleotide duplexes with compatible overhangs. The 3’ end of the H1 promoter was modified to be CCACAGATCCCC to facilitate transcription of gRNAs with any nucleotide at the 5’ end. (B) The steps of targeted gene deletion with CRISPR/Cas9.
Genomic DNA was extracted 48 h after transfection using QuickExtract DNA Extraction Solution (Epicentre Biotechnologies, Madison, WI). Common PCR was conducted to amplify the targeted region using primers flanking the targeted regions. Wild-type and truncated genomic fragments were resolved by gel electrophoresis. Real-time PCR (RT-PCR) was performed to quantify the percent of deletion using primers across the junction or within the deletion region. The comparative C_{r} method was used to calculate the expression level of the target region relative to a reference region (ACTB locus). Percent of deletion in the target region was further calculated by the ratio of target cells relative to control cells. All primer sequences are listed in Supplementary Table S2.

**Target sequencing**
Cells were harvested two days after transfection, and the genomic DNA was extracted using QuickExtract DNA Extraction Solution (Epicentre Biotechnologies). PCR was conducted to amplify the targeting region with genomic DNA derived from the cells, and amplicons were deep sequenced by MiSeq Personal Sequencer (Illumina, San Diego, CA).

**Targeted DNA replacement**
The linear donor was generated by PCR from pG3-GFP-SV40pA plasmid, created by replacing the Renilla gene with the GFP gene in pRL-TK (Promega, Madison, WI). The primer sequences used for PCR were:

\[
\begin{align*}
\text{CCL2-donor-F} \quad \text{A*CGACGCCAGGAGAACGCAGG} \\
\text{GCTGAGACTAACCAGAAAACATCCAC} \\
\text{ATGCTTTTACGCAGGCTCTAGG} \\
\text{CCL2-donor-R} \quad \text{C*AAATATTTATTTGGTGTAATAG} \\
\text{TTACAAAATATTTCATTTCCACAA} \\
\text{CTGGATCCTTATCGA}
\end{align*}
\]

The underlined regions indicate the termini of analogous oligonucleotides with 50 bp of CCL2 homology. The donor sequences are presented in the Supplementary Material. The two 5’-most linkages are phosphorothioate (indicated by asterisks). Cells in 6-well plates were transiently transfected with 2.0 µg Cas9 plasmid, 0.8 µg gRNA T1 plasmid, 0.8 µg gRNA T2 plasmid, and 0.4 µg linear donor using Lipofectamine 2000 (Invitrogen). At 48 h after transfection, cells were treated with 1 ng/mL TNF-α for 24 h, and then GFP-positive cells were sorted.

**Luciferase assay**
For the luciferase assay, HEK 293T cells were seeded in 96-well plates at a density of 5000 cells per well. After 24 h, the cells were transiently transfected with 5 ng of pRL-TK Renilla luciferase reporter and 100 ng of luciferase reporter with cytomegalovirus (CMV), SV40 (Simian virus 40), or basic promoter. After 48 h, luciferase activity was measured with the dual luciferase reporter assay system (Promega).

**Western blot**
Proteins were separated by sodium dodecyl sulfate–PAGE (SDS-PAGE) and transferred to nitrocellulose membranes. The membranes were blocked with 5% nonfat milk and incubated with GFP antibody (CST, #2555S, Danvers, MA). The antigen-antibody complex was detected with enhanced chemiluminescence reagents.

**Results and discussion**
We adapted the bacterial type II CRISPR/Cas9 system to mutagenize genomic DNA in human cells. The human codon-optimized version of \textit{S. pyogenes} Cas9 protein bearing a C-terminus SV40 nuclear localization signal was expressed using a previously described system (6). To direct Cas9 cleavage to the desired sequence, we expressed crRNA-tracrRNA fusion transcripts, hereafter referred to as guide RNAs (gRNAs), from a modified human H1 polymerase III promoter. The 3’ end of the H1 promoter was modified to permit transcription of gRNAs that begin with any nucleotide. Constrained only by the requirement that the 20 bp crRNA target be followed by

Figure 2. Targeted deletion of human CDC42 with CRISPR/Cas9. (A) Schematic diagram depicting the locations of guide RNAs (gRNAs) targeting the \textit{CDC42} locus. (B–D) The efficiency of targeted deletion with CRISPR/Cas9 was determined by PCR in HEK 293T cells. Primers outside expected deletion regions were used. The percent of deletion was quantified by RT-PCR using primers across the junction or within the deletion region. (E–F) The efficiency and precision of targeted deletion in \textit{CDC42} was confirmed by (E) Sanger sequencing and (F) high-throughput sequencing analyses. The PCR product containing only the deletion amplicon was enriched for sequencing.
the PAM sequence, NGG (where N is any nucleotide), this approach can in principle be used to target any genomic location that has the form N_{n-1}NGG. To facilitate the cloning of the gRNA expression vector, we used a Type II restriction enzyme, BsaI. This required the synthesis of a 24 bp oligonucleotide containing a region of complementarity to the target site on the DNA. The simple and efficient protocol for the cloning of the gRNA expression vector (Figure 1A) is described in detail in the Supplementary Material.

For the deletion of a large segment of genomic DNA, we used a pair of gRNAs against the targeted locus (Figure 1B). Two target sites with the pattern N_{n-1}NGG were selected at the boundary of the target region. The efficiencies of targeted deletion guided by various combinations of gRNA pairs were determined by PCR analyses using primers flanking the targeted regions. Wild-type and truncated genomic fragments were resolved by gel electrophoresis. To avoid PCR amplification bias, percent deletion was quantified by RT-PCR using one primer pair. The primers were designed across the deletion junction (one primer outside the deletion region, the other primer within the deletion region) or within the deletion regions (both primers locate within the deletion region). Thus, only a single band is amplified with the primer pair for both the targeted cells and control cells. We calculated percent deletion by comparing the relative amount of PCR products (target cells versus control cells) amplified by the same primer pair. Targeted deletions were further verified by sequencing. To assess how gRNA pairs might affect subsequent repair and generation of deletions, we first designed sets of gRNAs targeted against the human CDC42 genomic locus and separated by distances ranging from approximately 200 to 10,000 bp (Figure 2A and Supplementary Table S1). We then assessed the ability of each gRNA pair to generate deletions in human HEK 293T cells in the presence of Cas9. Robust efficiencies of NHEJ-based deletions (up to 68%) were confirmed by qPCR (Figure 2B–2D). Even for deletion of a 10 kb genomic region, we obtained targeting rates of 16% to 28%, depending on the gRNA pair. This RNA-mediated editing process was rapid, with the first detectable deletion appearing approximately 12 h post-transfection (Supplementary Figure S1). The system was effective in a variety of cell types, including: PC3, SK-Hep1, and HeLa cells (Supplementary Figure S2).

Deletions were further confirmed by sequencing PCR products spanning the expected cleavage sites. Sanger sequencing showed deletion junctions resulted from the precise ligation of the blunt-ended DSBs created by Cas9; each DSB occurred exactly 3 bp upstream of the PAM sequence (Figure 2E and Supplementary Figure S3). We also used deep sequencing of deletion amplicons to assess the accuracy of deletion efficiency; in about 80% of reads, the targeted DSBs were perfectly repaired (Figure 2F).

We recapitulated these findings by testing gRNA pairs designed to delete fragments from a genomic locus containing the microRNA miR-21 gene in HEK 293T cells. Two gRNAs were designed to target the boundaries of the miR-21 hairpin (Supplementary Figure S4). Deletion efficiency was 38% following transfection with the two gRNAs and Cas9, as measured using a PCR assay (Supplementary Figure S4B). Deep sequencing confirmed the deletion occurred precisely as expected (Supplementary Figure S4C).

To investigate whether the CRISPR/Cas9-mediated gene deletion system is influenced by the transcriptional state of targeted genes, the gene encoding the chemokine (C-C motif) ligand 2 (CCL2) was targeted. CCL2 is a small cytokine belonging to the CC chemokine family; the CCL2 gene is a target of NF-κB signaling. We randomly selected eight target sites located in the 5’ and 3’ ends of the CCL2 gene locus (Figure 3A). We

![Figure 3. Deletion with CRISPR/Cas9 was independent of the transcriptional activity of the targeted gene.](image)

(A) Schematic diagram depicting the locations of guide RNAs (gRNAs) targeting the CCL2 locus. (B–D) The efficiency of targeted deletion of CCL2 with CRISPR/Cas9 in HEK 293T cells. (E) CCL2 mRNA levels were determined upon tumor necrosis factor α (TNF-α) addition by quantitative RT-PCR analyses in HEK 293T cells. Data were shown with the means ± SEM in triplicate. (F) The efficiency of targeted deletion of CCL2 with CRISPR/Cas9 after treatment with TNF-α for 24 h in HEK 293T cells. (G) The level of luciferase activity from cytomegalovirus (CMV), or SV40, or a basic promoter in HEK 293T cells. Data were shown with the means ± SEM in triplicate. (H) Efficiency of targeted deletion of a luciferase gene controlled by CMV, or SV40, or a basic promoter in HEK 293T cells. (B–D,F,H) The percent of deletion was quantified by RT-PCR using primers across the junction or within the deletion region.
achieved robust and efficient deletion of different regions of the gene using Cas9 and different pairs of gRNAs in HEK 293T cells (Figure 3B–3D).

Expression of the CCL2 gene was dramatically induced by TNF-α (up to ~300-fold increase) in HEK 293T cells (Figure 3E), providing a good model to investigate whether the CRISPR/Cas9-mediated targeted genome editing is affected by transcriptional activity. Of interest, the efficiency of targeted deletion of the CCL2 gene locus was not affected by treatment of cells with TNF-α (Figure 3F), suggesting that transcription did not alter CRISPR/Cas9-mediated deletion. To further confirm this result, we targeted an exogenous reporter gene driven by various promoters with differing strengths, where the different transcriptional activities could be evaluated using a luciferase assay (Figure 3G). PCR assays revealed similar efficiency in targeted deletions in HEK 293T cells after co-transfection of the reporter gene together with Cas9 and gRNA pairs (Figure 3H). This result indicates NHEJ-mediated repair can occur despite the occurrence of varying degrees of transcriptional activity.

DSBs can stimulate HDR to enable highly precise replacement of the damaged region with a homologous donor. To obtain targeted genomic DNA replacement, we introduced a pair of gRNAs, Cas9, and a linear donor with homology to the targeted region into cells (Figure 4A). The linear donor was obtained by PCR amplification with primers bearing a 50 bp of homologous sequence. This same donor was successfully inserted using a ZFN-based HDR repair system (18). To test the feasibility of CRISPR/Cas9-mediated replacement by HDR, we targeted the CCL2 locus with a pair of gRNAs (#39 and #1854 shown in Figure 3) and a donor bearing the enhanced green fluorescent protein (EGFP) coding sequence and SV40 poly(A) site (Figure 4A; sequences and positions are presented in the Supplementary Material). Using this system, approximately 0.5% of targeted cells were EGFP-positive, whereas only 0.023% were EGFP-positive in mock transfection cells (just transfected with donor), which was similar to control cells (0.021%, without transfection). The EGFP-positive cells were then sorted by flow cytometry.

Figure 4. Targeted gene replacement with CRISPR/Cas9 using a PCR-generated linear short-homology donor. (A) Schematic diagrams depicting the procedure of targeted gene replacement using CRISPR/Cas9 in human cells. To test the efficacy of targeted gene replacement, guide RNAs (gRNAs) were designed to delete the indicated region (marked by sites 1 and 2) of the CCL2 gene and to replace the deleted region with the EGFP-polyA cassette donor with arms with short regions of homology. The targeted sites of site 1 and site 2 are #39 and #1854 within the CCL2 gene shown in Figure 3. The homologous sequences (50 bp) are just upstream and downstream of the deletion sites. The positions and sequences in detail are presented in the Supplementary Material. (B) The efficiency of targeted gene replacement with CRISPR/Cas9 was determined by PCR in HEK 293T cells. Primers spanning the junctions between CCL2 and EGFP were used for PCR amplification. (C) PCR assay for the targeted gene replacement of the single clones. (D,E) Expression of EGFP protein upon TNF-α addition in HEK 293T cells was determined by (D) Western blot and (E) FACS.
can be potentially used to target any genomic loci.

There has been concern regarding specificity of the CRISPR/Cas9 system (19–21). To rule out unwanted phenotypes due to off-target mutations, we suggest that at least two different pairs of gRNAs be used for each target region. In our study, the need for multiple gRNA pairs was not a major limitation, given the simplicity and highly efficiency of this system. It is notable that different pairs of gRNAs targeted to the same region worked with high efficiency (Figures 2 and 4). Another approach to avoid unwanted mutations is use of the double nickase method (22,23). We also successfully applied the double nickase method to generate deletion of genomic DNA, but the efficiency was considerably lower.

It is known that the repair of DNA DSBs is largely mediated by error-prone NHEJ, in which the two ends are processed and ligated together in a way that is frequently accompanied by nucleotide insertions and deletions. Such error-prone end joining was observed in the repair of DSBs created by ZFNs or TALENs. In contrast, the repair of DSBs generated by Cas9 and two gRNAs was very precise. Our results suggest that the breaks are directly ligated without end processing, revealing a previously unappreciated advantage of the NHEJ pathway. The mechanism that results in the precise ligation remains to be determined. One possibility is that targeted deletion using Cas9 and two gRNAs results in a junction that is not recognized by either of the original gRNAs. We also analyzed the efficiency in generating indel mutations for the individual gRNA and gRNA pair (#39 and #224 gRNA in Figure 3A) by Sanger sequencing of PCR amplicons (TA cloning). Of interest, we observed that the efficiency of generating indel mutations for single gRNA was quite low (9.5%, 2 of 21 clones for #39 gRNA; 5%, 1 of 20 clones for #224 gRNA). However, the gRNA pair generated high efficiency of indel mutations (50%, 10 of 20 clones for #39 and #224), which was similar to the assay using qPCR (52%, Figure 3B). We propose that a single gRNA often results in one blunt end of cleavage site, which will be precisely repaired by NHEJ. Thus, the efficiency of generating mutation is much lower with use of a single gRNA than a pair of gRNAs.

Author contributions

The conception and development of the experiments: S.H., and C.Z.C.; the execution and analysis of the experiments: Q.Z., X.C., S.S., C.P.A., X.G., and S.H.; and the writing and editing of the manuscript: S.H., C.Z.C., and M.H.T.

Acknowledgments

This work was supported by grants from National Institute of Health (No. DP1CA174421) and the W. M. Keck foundation to C.-Z. C. and National Natural Science Foundation of China (No. 81101481) and Shanghai Medical Talent Training Program (No. XYQ2011048) to S.L.H. This paper is subject to the NIH Public Access Policy.

Competing interests

The authors declare no competing interests.

References


Received 26 March 2014; accepted 28 July 2014.

Address correspondence to Shenglin Huang, Fudan University Shanghai Cancer Center, Institute of Biomedical Sciences, and Department of Oncology, Shanghai Medical School, Fudan University, Shanghai, China, or Chang-Zhen Chen, Achelos Pharmaceuticals, Inc., San Francisco, CA. E-mail: shuang@fudan.edu.cn or czchen@achelospharm.com

To purchase reprints of this article, contact: biotechniques@fosterprinting.com