Video Article

Generation of Genomic Deletions in Mammalian Cell Lines via CRISPR/Cas9

Daniel E. Bauer*[1,2,3], Matthew C. Canver*[1], Stuart H. Orkin[1,2,3,4]

[1]Harvard Medical School

[2]Division of Hematology/Oncology, Boston Children's Hospital

[3]Department of Pediatric Oncology, Dana-Farber Cancer Institute

[4]Howard Hughes Medical Institute

*These authors contributed equally

Correspondence to: Daniel E. Bauer at Daniel.Bauer@childrens.harvard.edu, Stuart H. Orkin at stuart_orkin@dfci.harvard.edu

Abstract

The prokaryotic clustered regularly interspaced short palindromic repeats (CRISPR)/CRISPR-associated (Cas) 9 system may be re-purposed for site-specific eukaryotic genome engineering. CRISPR/Cas9 is an inexpensive, facile, and efficient genome editing tool that allows genetic perturbation of genes and genetic elements. Here we present a simple methodology for CRISPR design, cloning, and delivery for the production of genomic deletions. In addition, we describe techniques for deletion, identification, and characterization. This strategy relies on cellular delivery of a pair of chimeric single guide RNAs (sgRNAs) to create two double strand breaks (DSBs) at a locus in order to delete the intervening DNA segment by non-homologous end joining (NHEJ) repair. Deletions have potential advantages as compared to single-site small indels given the efficiency of biallelic modification, ease of rapid identification by PCR, predictability of loss-of-function, and utility for the study of non-coding elements. This approach can be used for efficient loss-of-function studies of genes and genetic elements in mammalian cell lines.

Video Link

The video component of this article can be found at http://www.jove.com/video/52118/

Introduction

Recent advances in genome engineering technology have allowed for unprecedented opportunities for site-specific modification of the genome. This technology may be utilized to investigate the function of genes and regulatory elements via prospective genetic perturbation. Zinc finger nucleases (ZFNs), transcription-activator like (TAL) effector nucleases (TALENs), and CRISPR/Cas9 RNA-guided nucleases each leverage customizable DNA specificity to localize a nuclease for the introduction of DSBs[1–3]. The resulting DSBs can be repaired by indel-forming NHEJ or by homology-directed repair (HDR) using a donor template[4].

The CRISPR/Cas9 nuclease pathway, an adaptive immune system in prokaryotic cells[5], has been recently adapted for mammalian genome engineering[2,3]. This tool has been demonstrated to be an inexpensive, efficient, and reliable genome engineering technique[6]. Briefly, a complex of *Streptococcus pyogenes*-derived Cas9 nuclease and a sgRNA achieve target recognition via Watson-Crick base-pairing with cognate genomic DNA sequences. sgRNAs include 20-mer sequences complementary to genomic sequences adjacent to an obligate protospacer adjacent motif (PAM) NGG. Cas9 induces a DSB at a predictable position within the target site. Additionally, variants of Cas9 with single-strand cleavage capacity or catalytic inactivity may be used to facilitate "nicking" or transcriptional regulation respectively[7–9]. CRISPR/Cas9 has been used for a wide range of applications including both knock-in and knockout[10,11], large-scale genomic deletions[12–14], pooled library screening for gene discovery[15,16], genetic engineering of numerous model organisms[10,11,17–21], as well as gene therapy[22,23].

Here we describe a protocol for efficient deletion of desired genomic regions. The protocol includes CRISPR design, cloning, and delivery, as well as deletion, identification, and characterization. Genomic deletions can be generated by the introduction of two CRISPR sgRNAs with Cas9 to induce repair of the resultant two DSBs by NHEJ with deletion of the intervening segment. This strategy has been used to create deletions ranging from one kilobase to over one megabase[12]. Deletions can be informative for the study of genes and other genetic elements, either in isolation or in combination. There are several potential advantages of genomic deletions as compared to HDR or single-site small indel production. First, this method capitalizes on the high efficiency of NHEJ in many cellular contexts[7]. The high frequency of deletion limits the number of clones needed to be screened to identify informative clones. Deletion frequency is inversely related to deletion size. Biallelic deletion clones may be retrieved at frequencies at least as great as of probabilistic expectation[12]. Second, both monoallelic and biallelic deletions may be easily identified and distinguished by conventional PCR, simplifying the screening process. Strategies relying on small indels or point mutations may require RFLP, allele-specific PCR, T7EN1 cleavage assay, Sanger sequencing, RT-qPCR, or immunoblotting, which may be more laborious. Third, by removing a substantial portion of a gene or element of interest, a reliable loss-of-function allele may be obtained. In contrast, frameshift mutations in protein-coding sequences may not always induce nonsense-mediated decay, may produce a hypomorphic or neomorphic allele,

or target an exon excluded from an alternate isoform[24]. Finally, deletions may be particularly revealing for the study of non-coding DNA such as regulatory elements since frameshift mutations as produced by single-site indels would not be relevant[25].

Protocol

1. CRISPR Design

1. Design sgRNAs manually or using freely available online tools[7]. Use these tools to help identify guide sequences that minimize identical genomic matches or near-matches to reduce risk of cleavage away from target sites (off-target effects). Ensure that the guide sequences consist of a 20-mer ("protospacer sequence") upstream of an "NGG" sequence ("protospacer adjacent motif" or PAM) at the genomic recognition site.
   NOTE: **Figure 1** describes possible deletion strategies for genes and non-coding elements. For creating a gene knockout, two sgRNA located within exons will enrich even monoallelic deletion clones for loss of function. This is due to the high frequency of indels formed on non-deleted alleles[12], which are likely to cause frameshift mutations leading to nonsense mediated decay of the mRNA transcript (**Figure 1B**).

2. Use the example guides for the intended deletion of *Pim1* in mouse (*Mus musculus*; **Table 1**, **Figure 2A**).
   NOTE: In this example, sgRNA-A's protospacer sequence and PAM happen to fall on the bottom (Crick) strand while sgRNA-B's protospacer sequence and PAM fall on the top (Watson) strand (**Figure 2A**). However, DSB will occur independent of orientation of the protospacer sequence/PAM relative to the top or bottom strand.

3. Determine the reverse complement (rc) of each guide sequence. Example reverse complement sequences of the Pim1 sgRNA from **Table 1** are found in **Table 2**.

4. Obtain 24- or 25-mer oligos for each guide and its associated reverse complement including additional nucleotides for cloning and expression purposes.
   NOTE: Here we discuss the usage of the pSpCas9(BB) plasmid (pX330) (Addgene plasmid ID 42230). This plasmid allows for the simultaneous expression of sgRNA and SpCas9, but does not contain markers for selection[7]. Other constructs may be utilized, such as pX458 (Addgene plasmid ID 48138) or pX459 (Addgene plasmid ID 48139), which include GFP and puromycin as selectable markers, respectively, or constructs in which multiple sgRNAs may be expressed from a single plasmid.
   1. Add "CACC" before the 20-mer guide sequence and "AAAC" before the guide's reverse complement for cloning into the pX330 vector using BbsI restriction enzyme (**Table 3**).
   2. Add a G nucleotide after the CACC sequence and before the 20-mer if the first position of the 20-mer is not G. sgRNA expression from the U6 promoter of the pX330 vector is enhanced by the inclusion of a G nucleotide after the CACC sequence. Add a C at the 3' end of the reverse complement oligo (*e.g.,* sgRNA-A in **Table 4**). The resultant oligos would be 25-mer oligos.
   3. However, if the first position of the 20-mer (protospacer sequence) is G, do not add another G (*e.g.,* sgRNA-B in **Table 4**) and do not add C to the final position of the reverse complement oligo. In this case, the resultant oligos would be 24-mer oligos (**Table 4**).

2. Design Deletion Screening Primers

1. Design one set of primers internal to the sequence to be deleted ("non-deletion band") and another set of primers upstream and downstream of the sgRNA cleavage sites ("deletion band"; **Figures 1 - 2**). In the absence of deletion, the "deletion band" is often too large to efficiently amplify. Typically use primers at least 100 bp from the predicted cleavage site to ensure detection would not be impacted by a small indel at the sgRNA target site.
   1. Design additional primers to analyze for scarring (small indels produced at the sgRNA cleavage site without the intended deletion). Use a pair of forward and reverse primers flanking each sgRNA target site (within 150 - 350 bp) to amplify the sgRNA target site to examine for scarring. This may be useful to characterize the non-deleted allele in monoallelic deletion clones.

2. For small deletions (as the "deletion band" may still amplify), resolve amplicons on an agarose gel to determine if size is consistent with the presence or absence of deletion. For this approach, the internal primers described in **step 2.1** may be omitted.

3. CRISPR Cloning

1. Anneal and phosphorylate oligos.
   1. Resuspend oligos at a concentration of 100 µM in ddH$_2$O.
   2. Prepare a 10 µl reaction mix for each guide and its reverse complement: 1.0 µl sgRNA 24- or 25-mer oligo (100 µM; see **step 1.4**), 1.0 µl sgRNA 24- or 25-mer reverse complement oligo (100 µM; see **step 1.4**), 1.0 µl 10x T4 Ligation Buffer, 6.5 µl ddH$_2$O, and 0.5 µl T4 Polynucleotide Kinase (PNK) (10,000 U/ml).
      NOTE: Phosphorylated oligos may be ordered instead. For this approach, the use of T4 PNK is omitted.
   3. Anneal in a thermocycler using the following parameters: 37 °C for 30 min; 95 °C for 5 min and then ramp down to 25 °C at 5 °C/min.
   4. Dilute oligos 1:10 in ddH$_2$O (*e.g.,* 1.0 µl annealed oligos + 9.0 µl ddH$_2$O to yield a concentration of 1 µM).

2. Ligate annealed oligos into pX330 using a Golden Gate assembly cloning strategy[26].
   1. Prepare a 50 µl reaction mix: 100 ng circular pX330 vector, 1.0 µl annealed oligos (1 µM; see **step 3.1.4**), 5.0 µl restriction enzyme buffer (10x), 4.0 µl (20 U) BbsI restriction enzyme (5,000 U/ml), 5.0 µl ATP (10 mM), 0.25 µl (5 µg) BSA (20 mg/ml), 0.375 µl (750 U) T4 DNA ligase (2,000,000 U/ml), and H$_2$O to final volume of 50 µl. This reaction may be scaled down to a smaller final volume if necessary.
   2. Run samples in a thermocycler using the following parameters: Cycles 1-20 (37 °C for 5 min, 20 °C for 5 min); Cycle 21 (80 °C for 20 min). These cycling conditions allow for digestion and ligation to occur in one reaction (see **step 3.2**).

3. Transform 10 µl of DH5α *E. coli* cells with 1 µl of reaction (from **3.2.1 - 3.2.2**).
4. Plate onto a lysogeny broth (LB) agar plate with 100 µg/ml ampicillin and incubate O/N at 37 °C.

3. Pick 2 - 3 colonies and inoculate into a mini-prep culture.
    1. Perform mini-prep for each sample and sequence each colony using a U6 promoter forward primer: CGTAACTTGAAAGTATTTCGATTTCTTGGC. This is a representative sequencing primer; other flanking primers may be utilized.

4. Choose a sequence-verified colony and inoculate into a maxi-prep culture. Prep size may be scaled based on the required DNA yield.
5. Perform maxi-prep for each CRISPR/Cas9 construct.

4. Transfecting CRISPRs into Cells of Interest

NOTE: This protocol involves the delivery of CRISPR/Cas9 plasmids by electroporation[27]. This protocol is described in detail for murine erythroleukemia (MEL) cells, a suspension cell line. The culture medium in all steps consists of DMEM supplemented with 2% penicillin/ streptomycin and 1% L-glutamine, which is used for MEL cells. However, transient transfection of CRISPR/Cas9 plasmids may be successfully adapted to numerous cell types using preferred culture conditions and transfection strategies for each cell type. While MEL cells are suspension cells, instructions for adherent cells have also been included.

1. Ensure there are 2×10^6 cells per CRISPR pair. Resuspend 2×10^6 cells in 100 µl of electroporation solution and add to electroporation cuvette.
    1. Add 5 µg of each CRISPR/Cas9 construct (10 µg total). Add 0.5 µg of GFP expression construct.

2. Electroporate cells with 250 volts for 5 msec in a 2 mm cuvette using an electroporation system.
   NOTE: Alternatively use another transfection method such as cationic liposome-based transfection. Optimize transfection conditions for each cell line with a reporter construct to ensure robust plasmid delivery before attempting genome editing.
3. Immediately transfer solution from cuvette into 1 ml of culture media after electroporation. Minimize the time between electroporation and transferring the solution into media to enhance cell viability.
4. Incubate at 30 - 37 °C for 24 - 72 hr. 30 °C may enhance genome editing efficiency, but 37 °C is acceptable.

5. Fluorescence Activated Cell Sorting (FACS) of Transfected Cells

1. Prepare cells for FACS by filtering them through a 50 µm filter into a FACS tube.
2. FACS sort the top ~3% of GFP positive cells in order to enrich for cells that received high levels of the CRISPR/Cas9 constructs.
3. Plate sorted cells individually into 96-well round-bottom plates using sorter or by using limiting dilution at 30 cells per 96-well round-bottom plate. Optimize plating the cell type used at limiting dilution prior to performing this step to reliably obtain approximately 30 cells per 96-well plate.
4. Include 100 µl per well of cell culture media.
5. For the remaining sorted cells ("bulk") that were not plated, freeze half of the cells for future plating. Plate the other half for screening and primer validation (see **step 6**).
   NOTE: This protocol is for suspension cells. Adherent cells can either grow as individual cells in 96-well flat bottom plate or in a 10 cm dish at low concentration so that individual single-cell derived clones can be picked and moved to a flat bottom 96-well plate.
6. Allow the bulk cells to incubate at 37 °C for 3 - 7 days and allow the clones to incubate at 37 °C for 7 - 14 days. Vary these times depending on the doubling time of the cell line used.
   NOTE: This incubation time allows for sufficient cell proliferation for screening genomic DNA (gDNA) for the intended deletion by PCR (see **steps 6.1 and 7.1**). The bulk cells have sufficiently proliferated when the concentration exceeds ~100,000 cells/ml or for adherent cells, the cells have reached ~80% confluence. The clones have sufficiently proliferated once macroscopically visible with ~2 mm diameter.

6. Primer Validation and Screening for CRISPR/Cas9-Mediated Deletion

1. Isolate gDNA from parental and bulk sorted cells by resuspending parental and bulk cell pellets in 50 µl of DNA extraction solution.
   NOTE: Generally ~100,000 cells are used for DNA extraction, although a wide range of cell numbers is acceptable. The bulk sorted cells are composed of a polyclonal population exposed to sgRNA-A and sgRNA-B (see **step 5**). The purpose of the following PCR is to validate primers and verify the presence of intended genomic deletion.
2. Run sample in thermocycler and run the following program: 65 °C for 6 min, 98 °C for 2 min to extract gDNA. Measure the DNA concentration.
   NOTE: While **steps 6.1 and 6.2** recommend an efficient method for DNA extraction, any method for genomic DNA isolation may be utilized to be able to perform PCR in **step 6.3**.
3. Assemble a 20 µl PCR with the following components: 10 µl 2x PCR mix, 0.5 µl forward primer (10 µM), 0.5 µl reverse primer (10 µM), 50-100 ng gDNA, and H$_2$O up to 20 µl. Use the primers designed in **step 2** above. Conduct PCR for "non-deletion band" and "deletion band" in separate reactions.
   NOTE: Numerous polymerases may be used for **step 6.3**.
    1. Run samples in a thermocycler using the following parameters: 95 °C for 15 min, 35 cycles of (95 °C for 30 sec, 60 °C for 1 min, 72 °C for 1 min), and 72 °C for 10 min. Optimize PCR conditions for each primer pair designed based on testing the bulk sorted cells.

4. Run samples on 2% agarose gel at 10 V/cm using 1x Tris-acetate-EDTA (TAE) buffer.
5. Examine samples for the presence/absence of non-deletion and deletion bands (**Figure 2**). Consider multiplexing the "deletion" and "non-deletion" PCR primer pairs in a single reaction. Optimize multiplexing in a polyclonal population (*i.e.,* bulk sorted cells) before screening individual clones. It is critical that the deletion and non-deletion amplicons be easily resolved on an agarose gel for multiplexing.

7. Screening CRISPR/Cas9 Clones for Deletions and Clone Selection

1. For suspension cells, transfer all clones to a single 96-well plate that already contains 50 µl cell culture media per well for a final volume of 150 µl. This facilitates screening by allowing a multichannel pipette to be used for the remainder of the steps in **step 7**.
    1. Transfer 50 µl from each well (leaving 100 µl in each well) to a 96-well PCR plate using a multichannel pipette.
    2. Centrifuge PCR plate at 400 x g for 5 min and remove supernatant by flicking the PCR plate over a sink. Add 50 µl of DNA extraction solution per well and resuspend. Continue to **step 7.3** for suspension cells.

2. For adherent cells, aspirate media. Add 20 µl of 0.05% trypsin-EDTA to each well with a clone present.
    1. Resuspend cells in 200 µl of media. Pipette mix to detach cells.
    2. Plate 100 µl each into two separate 96-well flat-bottom plates. Keep one plate to allow for clones to grow and use the other plate to screen each clone for deletions.
    3. Add an additional 100 µl to each well for a total volume of 200 µl. Wait 24 - 72 hr to allow cells to grow.
    4. Aspirate media. Add 50 µl DNA extraction solution per well, resuspend and transfer to 96-well PCR plate. Continue to **step 7.3**.

3. Extract the gDNA from clones. Run sample in thermocycler: 65 °C for 6 min and 98 °C for 2 min to extract gDNA.
4. Screen each clone using the same PCR primers and reaction conditions optimized on the bulk cells (see **step 6**).
5. Select the clones identified with the desired deletion and move to larger plate or flask for growth.

8. Validation of Biallelic Deletion Clones

1. In order to characterize obtained clones and validate a successful knockout, evaluate clones at the DNA as well as RNA and/or protein levels.
    1. To evaluate the DNA, amplify deletion bands from biallelic deletion clones with a proofreading polymerase and clone the amplicons (*e.g.,* with a PCR cloning kit) into a plasmid vector. Transform the plasmid into DH5α *E. coli* cells and plate onto LB agar plates with the relevant antibiotic. Select multiple colonies, mini-prep each one, and subject each clone to Sanger sequencing to characterize each deletion allele[28–30]. Repeating the PCR test for deletion after the initial screen ensures that the correct clone was selected and reproducibility of results.
    2. To evaluate the RNA, perform RT-qPCR for gene expression of the relevant gene[30,31].
    3. To evaluate the protein, perform an immunoblot using an antibody against the relevant protein[33].

Representative Results

The goal of this experiment was the deletion of *Pim1* in MEL cells. Use of multiple non-overlapping sgRNA pairs (*i.e.,* independent protospacer sequences) may help to control for off-target effects (**Figure 1A**). A consistent phenotype would be more likely to result from an on-target effect as opposed to a common off-target effect shared by multiple independent protospacer sequences. Each pair would lead to production of a unique deletion breakpoint. If close together (*i.e., n* less than ~150 bp), the same screening primers could be used to detect deletions produced by each set of sgRNAs. Genomic deletions may disrupt genes by using sgRNA pairs in various locations with respect to the gene (**Figure 1B**). For example, the sgRNA pair may flank a gene for deletion of the entire gene body; the pair could be located within two exons, with the potential to create frameshift indels even if one or both alleles were not deleted; or the pair may flank a specific exon to allow disruption of a particular isoform.

The deletion strategy used for *Pim1* was to design flanking sgRNAs to delete the entire gene body, an 8 kb deletion (**Figure 2**). This strategy was chosen in part due to the relatively small size of the *Pim1* gene. This example shows one PAM (green) on the top (Watson) strand and one PAM on the bottom (Crick) strand; however, DSB is independent of PAM sequence localization to the top or bottom strand. sgRNA pairs can have both PAM sequences on the top strand, both on the bottom strand, or one of each. Using the protocol described above, two sgRNA were designed, cloned into the pX330 expression vector, and delivered to MEL cells by electroporation along with a GFP reporter (**Figure 2A**). The top 3% of GFP⁺ cells were sorted two days post-electroporation and plated clonally at limiting dilution. Screening primers were designed as described in **step 2** and as shown in **Figure 2**. PCR conditions were optimized using gDNA isolated from parental MEL cells and from "bulk" sorted cells.

10 days after plating, gDNA was isolated from all clones and screened for deletion via PCR, which identified non-deletion, monoallelic and biallelic deletion clones according to the patterns of non-deletion (ND) and deletion (D) amplicons (**Figure 3**). Non-deletion clones were identified as having the presence of the non-deletion amplicon and the absence of the deletion amplicon. Monoallelic clones were identified as having the presence of both the non-deletion and deletion amplicons. Biallelic clones were identified as having the absence of the non-deletion amplicon and the presence of the deletion amplicon. For this deletion, 400 clones were screened which identified 126 monoallelic deletion clones and 32 biallelic deletion clones (it is important to note that deletion frequency varies with deletion size[12]). Biallelic deletion clones were selected and moved to flasks with 8 ml of media. After allowing 5 days for expansion, each clone was retested by PCR of gDNA to confirm biallelic deletion and deletion amplicons were subjected to Sanger sequencing to identify the precise deletion (**Figure 2B**). Heterogeneity within the deletion amplicons reflects imperfect indel-forming NHEJ repair. Sequencing of the non-deletion allele in monoallelic deletion clones uncovered indels in the majority of cases, demonstrating that even the non-deleted allele is frequently edited by CRISPR/Cas9, which may be important for applications where both monoallelic and biallelic deletion clones are required. RNA was isolated from biallelic deletion clones and analyzed by RT-qPCR to confirm loss of Pim1 expression (**Figure 4**).
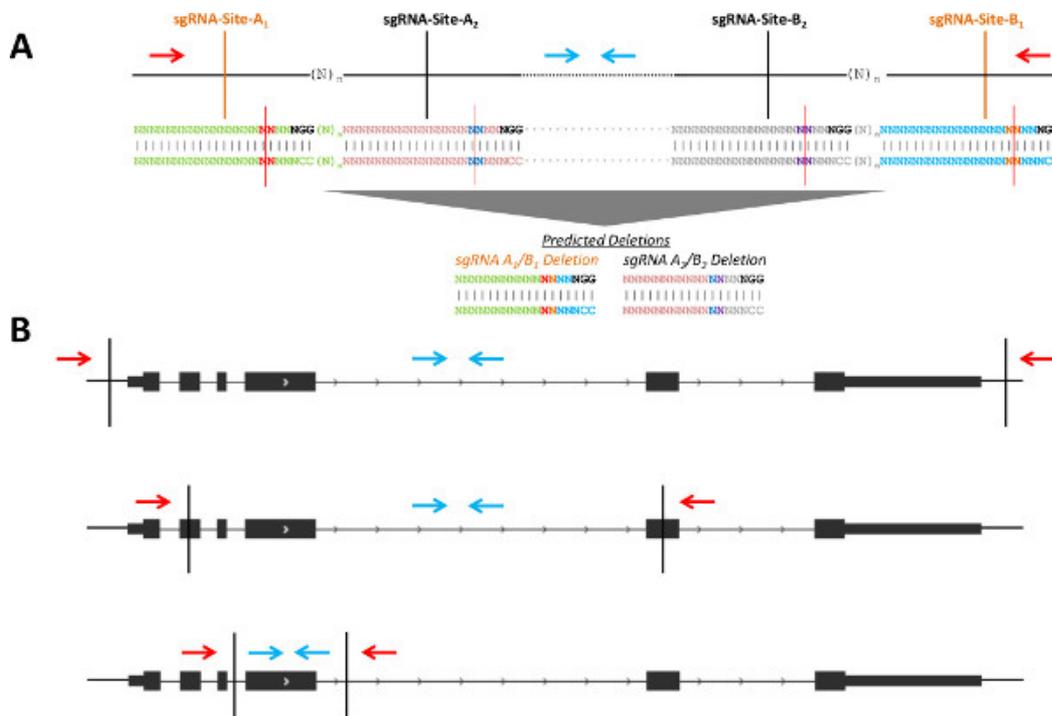
Figure 1. Schematic of possible deletion strategies. (A) Two example sgRNA pairs for genomic deletions (shown in black and orange, respectively). The blue arrows indicate primers to detect the non-deletion amplicon and red arrows indicate primers to detect the deletion amplicon. sgRNA positions 17 and 18 are highlighted in red and blue at sgRNA-Sites-A_1/A_2 and are highlighted in purple and orange at sgRNA-Sites-B_1/B_2 with a red line indicating the predicted Cas9 cleavage between positions 17 and 18. **(B)** CRISPR-directed cleavages are shown as vertical black lines. The blue arrows indicate primers to detect the non-deletion amplicon and red arrows indicate primers to detect the deletion amplicon. Please click here to view a larger version of this figure.
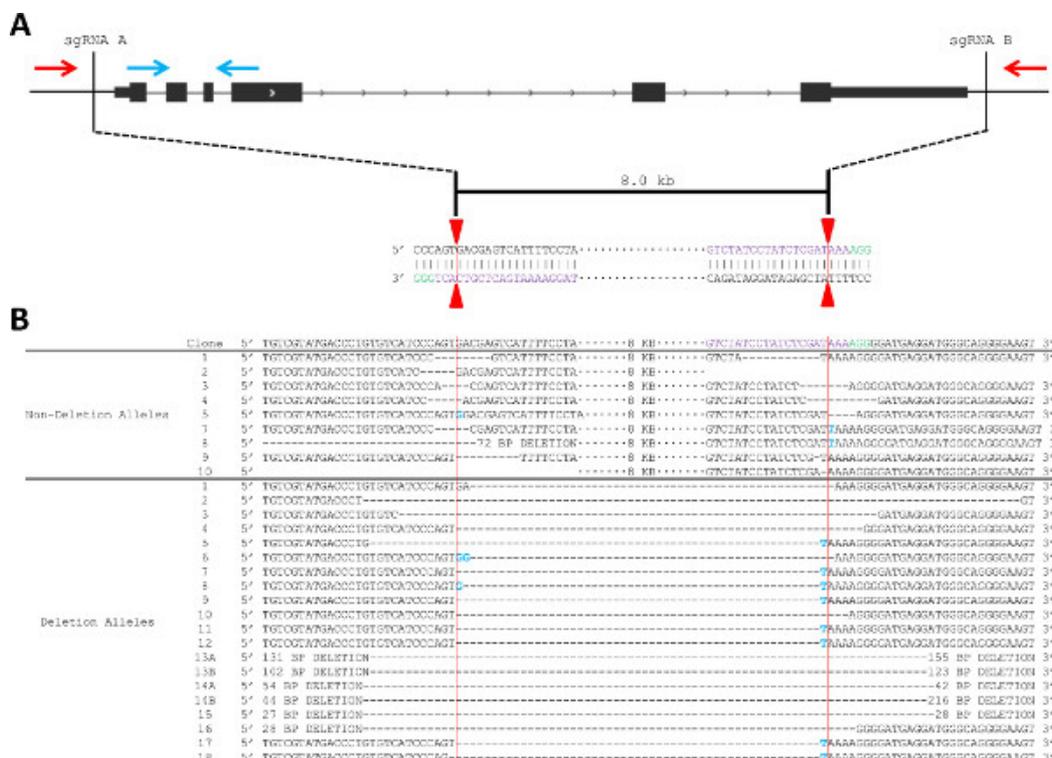
Figure 2. Strategy to produce and detect deletion of *Pim1*, including sequencing *deletion and non-deletion alleles.* (A) Schematic of the PCR-based screening strategy to identify *Pim1* deletion clones. One primer pair is located internal to the deletion (blue arrows) and one primer pair is localized external to the deletion (red arrows). sgRNA sequences (protospacer sequences) are shown in purple. The vertical red lines indicate the predicted Cas9 cleavage between positions 17 and 18 of the sgRNA sequence. **(B)** Sanger sequencing reveals indel formation at the sgRNA recognition site. sgRNA sequences are shown in purple and PAM sequences in green. Deletion events are shown by an equivalent number of dash marks and insertions are highlighted in blue. Vertical red lines indicate predicted cleavage site, between positions 17 and 18 of the sgRNA. Please click here to view a larger version of this figure.
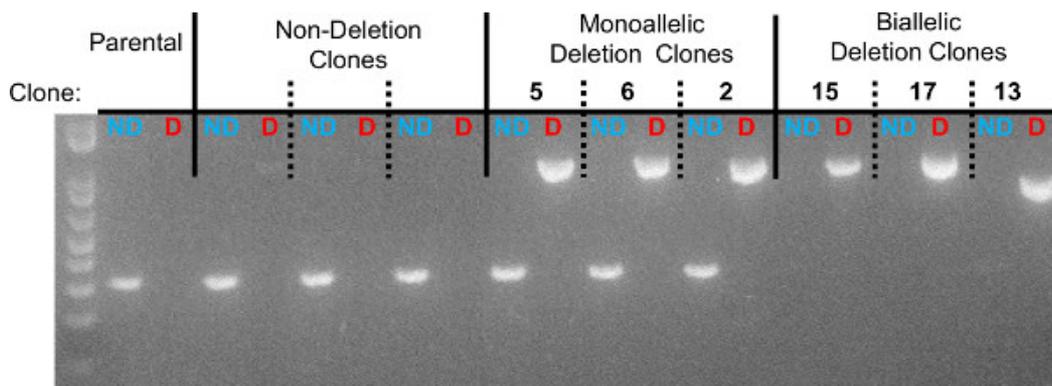
Figure 3. **Representative gel identifying non-deletion, monoallelic, and biallelic clones.** For each clone, the left lane represents the non-deletion amplicon ("ND" in blue) and the right lane represents the deletion amplicon ("D" in red). Individual clones are separated by dotted lines. The three monoallelic deletion clones are clones 5, 6, and 2. The three biallelic deletion clones are clones 15, 17, and 13. As seen in the sequencing data, the deletion band for clone 13 has a smaller size due to larger deletions at the deletion junction.
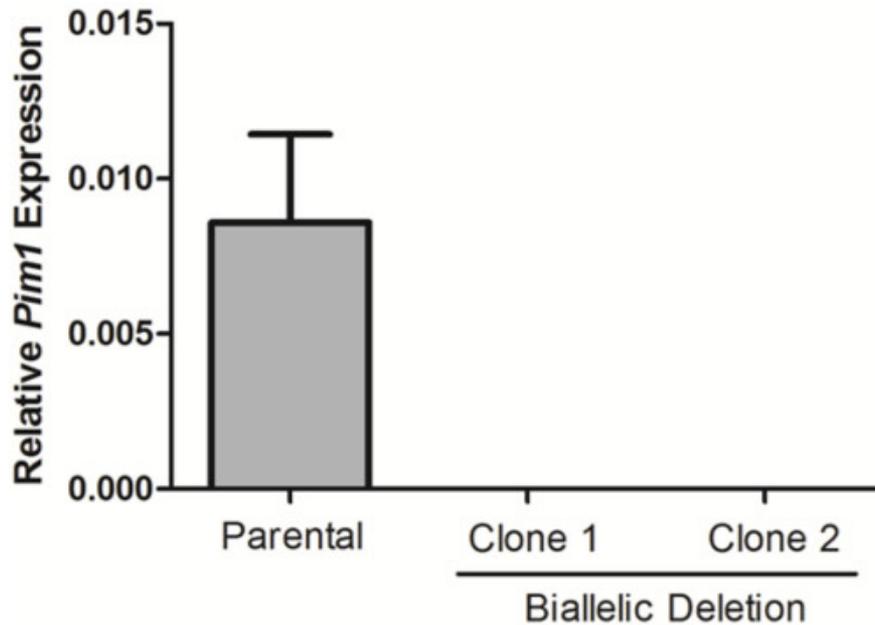
Figure 4. Loss of Pim1 expression in biallelic deletion clones. Pim1 expression was calculated for two biallelic deletion clones by RT-qPCR. Data was normalized to Gapdh using the $2^{-\Delta Ct}$ method.

	Protospacer Sequence
sgRNA-A	5'-TAGGAAAATGACTCGTCACT-3'
sgRNA-B	5'-GTCTATCCTATCTCGATAAA-3'

Table 1. 20-mer protospacer sequences for two sgRNA for the deletion of *Pim1*.

	Reverse Complement of Protospacer Sequence
sgRNA-A-rc	5'-AGTGACGAGTCATTTTCCTA-3'
sgRNA-B-rc	5'-TTTATCGAGATAGGATAGAC-3'

Table 2. Reverse complement of the protospacer sequences for the two sgRNA from Table 1.

	Sequences
sgRNA-A	5'-CACCTAGGAAAATGACTCGTCACT-3'
sgRNA-A-rc	5'-AAACAGTGACGAGTCATTTTCCTA-3'
sgRNA-B	5'-CACCGTCTATCCTATCTCGATAAA-3'
sgRNA-B-rc	5'-AAACTTTATCGAGATAGGATAGAC-3'

Table 3. Protospacer sequences and their reverse complements with "CACC" and "AAAC" added for cloning into the pX330 vector using BbsI restriction enzyme.

	Sequences
sgRNA-A	5'- CACCGTAGGAAAATGACTCGTCACT-3'
sgRNA-A-rc	5'- AAACAGTGACGAGTCATTTTCCTAC-3'
sgRNA-B	5'- CACCGTCTATCCTATCTCGATAAA-3'
sgRNA-B-rc	5'- AAACTTTATCGAGATAGGATAGAC-3'

Table 4. sgRNA expression from the U6 promoter of the pX330 vector is enhanced by the addition of a G nucleotide after the CACC sequence and before the 20-mer. The addition of an extra G nucleotide requires the addition of a C nucleotide at the 3' end of the reverse complement oligo (*e.g.,* sgRNA-A). However, if the first position of the 20-mer (protospacer sequence) is already a G nucleotide, there is no need to add another G (*e.g.,* sgRNA-B) and no need to add C to the final position of the reverse complement oligo.

Discussion

The CRISPR/Cas9 system may be used to generate genomic deletions of a range of sizes. Although we have observed that the frequency of deletion varies inversely with respect to intended deletion size, we have been able to recover deletions of up to 1 Mb, and deletions up to 100 kb routinely yield multiple biallelic deleted clones. We have observed no loss in efficiency of sequentially introducing deletions into a cell line. This strategy can be used for creation of combinatorial deletion of numerous genes and elements. The process of obtaining biallelic deletion clones can be expedited by estimating the minimum number of clones needed to be screened based on deletion size to obtain the desired number of clones with biallelic deletion[12].

The ability to obtain monoallelic deletion with the absence of biallelic deletion at probabilistic distribution could indicate cell lethality associated with complete loss of function. Low frequency or absent deletions could reflect a number of scenarios including poor transfection, inefficient sgRNAs or inefficient PCR screening primers (due to lack of a positive control to validate PCR primers to screen for deletion). GFP+ cells can be used as a surrogate for transfection efficiency (see **step 5.2**), so a reduction in GFP+ cells likely reflects poor transfection and a resulting decreased deletion efficiency. Using two different sgRNA pairs with independent screening primers can help control for inefficient sgRNA and screening PCR primers and maximize chances of obtaining biallelic deletion clones. Cell sorting for GFP+ cells enriches for deletion alleles. While this step may be omitted, omission will likely necessitate screening more clones to identify those with monoallelic or biallelic deletions. To the degree that transfection efficiency may be optimized, we would expect genome editing efficiency to be enhanced.

The NHEJ events that underlie deletions and local repair result in a series of alleles with a variety of indels at the target sites. The predominant outcome is small ~1 - 10 bp insertions or more commonly deletions at the site of sgRNA-directed cleavage (**Figure 2B**). Often these alleles appear to be the result of microhomology-based repair[34,35]. It should be noted that the PCR-based detection strategy we describe will not identify larger or more complicated insertions, deletions, inversions, or rearrangements. Although these events are less common, we have observed clones in which neither deletion nor non-deletion amplicons could be detected, and upon further investigation reflect these more complex outcomes.

We have observed extensive CRISPR/Cas9-mediated "scarring" of non-deletion alleles from monoallelic and non-deletion clones (see **Figure 2B**). These "scars" consist of small indels produced at the sgRNA cleavage site without the intended deletion (*i.e.,* deletion of the intervening segment between sgRNAs A and B). These scars often interrupt target recognition by the sgRNA. Therefore we would urge caution in retargeting alleles in cells previously exposed to sgRNAs using the same sgRNAs. A more successful retargeting strategy would utilize unique sgRNA sequences distinct from previously "scarred" recognition sites. In cases when a pair of sgRNAs recognizes exonic sequences (**Figure 1B**, bottom), frameshift alleles may be produced even in the absence of deletion. Therefore, monoallelic deletion clones can be enriched for loss-of-function due to the high frequency of frameshift mutations on the non-deleted allele[12].

One concern with the CRISPR/Cas9 system is off-target effects, *i.e.,* genomic modification at unintended sites[36–38]. Recent reports have suggested that shorter guide RNAs with 17 - 19 nucleotides can reduce the frequency of CRISPR/Cas9-based off-target effects[39]. Additionally, a double-nicking strategy using two guides per target site with a nickase can be used to create DSBs while minimizing off-target effects[7]. Alternatively, analogous to strategies used for RNAi, we suggest that different pairs of sgRNAs with non-overlapping protospacer sequences be used to demonstrate that the observed phenotype is the result of the on-target CRISPR/Cas9 modification as opposed to a potential off-target effect. A convenient approach would be to design at least two adjacent but non-overlapping sgRNA pairs so that a single set of screening primers (see **step 2**) may be used for multiple sgRNA pairs (**Figure 1A**). Furthermore, complementing a deletion cell line by reintroducing the missing sequence and/or disrupted gene can substantiate a causal relationship between a given genomic deletion and phenotype.

For biologists working with cellular model systems, RNAi has represented a powerful tool for functional genomics. However, limitations of this approach have included incomplete reduction in target mRNA transcript levels, heterogeneity of effect of independent reagents targeting the same gene, and known off-target effects including seed-based and non-seed effects[40–42]. Genome editing strategies promise to address many of these concerns and represent an exciting, complementary approach for prospective genetic perturbation[8,36,37]. Furthermore, genome editing allows for the study of non-coding genetic elements in a way not possible by RNAi and challenging by conventional targeting approaches[25]. We encourage generation of genomic deletions by CRISPR/Cas9 as a robust and specific method to produce and characterize loss-of-function alleles.

Disclosures

We have no conflicts of interest related to this report.

Acknowledgements

References

1. Jinek, M. *et al.* A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science.* **337** (6096), 816–21, doi: 10.1126/science.1225829 (2012).
2. Cong, L. *et al.* Multiplex genome engineering using CRISPR/Cas systems. *Science.* **339** (6121), 819–23, doi: 10.1126/science.1231143 (2013).

3. Mali, P. *et al.* RNA-guided human genome engineering via Cas9. *Science.* **339** (6121), 823–6, doi: 10.1126/science.1232033 (2013).
4. Gaj, T., Gersbach, C. A., & Barbas, C. F. ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. *Trends in Biotechnology.* **31** (7), 397–405, 10.1016/j.tibtech.2013.04.004 (2013).
5. Barrangou, R. *et al.* CRISPR provides acquired resistance against viruses in prokaryotes. *Science.* **315**, 1709–1712 (2007).
6. Ding, Q. *et al.* Enhanced efficiency of human pluripotent stem cell genome editing through replacing TALENs with CRISPRs. *Cell Stem Cell.* **12** (4), 393–4, doi: 10.1016/j.stem.2013.03.006 (2013).
7. Ran, F. A. *et al.* Genome engineering using the CRISPR-Cas9 system. *Nature Protocols.* **8** (11), 2281–308, doi: 10.1038/nprot.2013.143 (2013).
8. Ran, F. A. *et al.* Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. *Cell.* **154** (6), 1380–9, doi: 10.1016/j.cell.2013.08.021 (2013).
9. Larson, M. H. *et al.* CRISPR interference (CRISPRi) for sequence-specific control of gene expression. *Nature Protocols.* **8** (11), 2180–96, doi: 10.1038/nprot.2013.132 (2013).
10. Yang, H. *et al.* One-step generation of mice carrying reporter and conditional alleles by CRISPR/Cas-mediated genome engineering. *Cell.* **154** (6), 1370–9, doi: 10.1016/j.cell.2013.08.022 (2013).
11. Wang, H. *et al.* One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering. *Cell.* **153** (4), 910–8, doi: 10.1016/j.cell.2013.04.025 (2013).
12. Canver, M. C. *et al.* Characterization of Genomic Deletion Efficiency Mediated by CRISPR/Cas9 in Mammalian Cells. *Journal of Biological Chemistry.* **289** (31), 21312-24, doi: 10.1074/jbc.M14.564625 (2014).
13. Xiao, A. *et al.* Chromosomal deletions and inversions mediated by TALENs and CRISPR/Cas in zebrafish. *Nucleic Acids Research.* **41** (14), e141, doi:10.1093/nar/gkt464 (2013).
14. Zhou, J. *et al.* Dual sgRNAs facilitate CRISPR/Cas9 mediated mouse genome targeting. *The FEBS Journal.* doi:10.1111/febs.12735 (2014).
15. Wang, T. *et al.* Genetic screens in human cells using the CRISPR-Cas9 system. *Science.* **343** (6166), 80–4, doi: 10.1126/science.1246981 (2014).
16. Shalem, O. *et al.* Genome-scale CRISPR-Cas9 knockout screening in human cells. *Science.* **343** (6166), 84–7, doi: 10.1126/science.1247005 (2014).
17. Niu, Y. *et al.* Generation of Gene-Modified Cynomolgus Monkey via Cas9/RNA-Mediated Gene Targeting in One-Cell Embryos. *Cell.* **156** (4), 836–843, doi: 10.1016/j.cell.2014.01.027 (2014).
18. Guo, X. *et al.* Efficient RNA/Cas9-mediated genome editing in Xenopus tropicalis. *Development.* **141** (3), 707-14, doi: 10.1242/dev.099853 (2014).
19. Xie, K., & Yang, Y. RNA-guided genome editing in plants using a CRISPR-Cas system. *Molecular Plant.* **6** (6), 1975–83, doi: 10.1093/mp/sst119 (2013).
20. Waaijers, S. *et al.* CRISPR/Cas9-targeted mutagenesis in Caenorhabditis elegans. *Genetics.* **195** (3), 1187–91, doi: 10.1534/genetics.113.156299 (2013).
21. Bassett, A. R., & Liu, J. L. CRISPR/Cas9 and Genome Editing in Drosophila. *Journal of Genetics and Genomics.* **41** (1), 7–19, doi: 10.1016/j.jgg.2013.12.004 (2014).
22. Schwank, G. *et al.* Functional Repair of CFTR by CRISPR/Cas9 in Intestinal Stem Cell Organoids of Cystic Fibrosis Patients. *Cell Stem Cell.* **13** (6), 653–8, doi: 10.1016/j.stem.2013.11.002 (2013).
23. Wu, Y. *et al.* Correction of a Genetic Disease in Mouse via Use of CRISPR-Cas9. *Cell Stem Cell.* **13** (6), 659–62, doi: 10.1016/j.stem.2013.10.016 (2013).
24. Lappalainen, T. M. *et al.* Transcriptome and genome sequencing uncovers functional variation in humans. *Nature.* **501** (7468), 506–11, doi: 10.1038/nature12531 (2013).
25. Bauer, D. E. *et al.* An Erythroid Enhancer of BCL11A Subject to Genetic Variation Determines Fetal Hemoglobin Level. *Science.* **342** (6155), 253–257, doi: 10.1126/science.1242088 (2013).
26. Engler, C., Kandzia, R.., & Marillonnet, S. A one pot, one step, precision cloning method with high throughput capability. *PloS One.* **3**, e3647, doi:10.1371/journal.pone.0003647 (2008).
27. Gehl, J. Electroporation: theory and methods, perspectives for drug delivery, gene therapy and research. *Acta Physiologica Scandinavica.* **177**, 437–47, doi:10.1046/j.1365-201X.2003.01093.x (2003).
28. Zhang, S., Cahalan, M. D. Purifying Plasmid DNA from Bacterial Colonies Using the Qiagen Miniprep Kit. *J. Vis. Exp.* (6), e247, doi:10.3791/247 (2007).
29. Froger, A., Hall, J. E. Transformation of Plasmid DNA into E. coli Using the Heat Shock Method. *J. Vis. Exp.* (6), e253, doi:10.3791/253 (2007).
30. Finney, M., Nisson, P. E., & Rashtchian, A. *Molecular cloning of PCR products. Current Protocols in Molecular Biology.* Chapter 15, Unit 15.4, doi: 10.1002/0471142727.mb1504s56 (2001).
31. Gordanpour, A., Nam, R. K., Sugar, L., Bacopulos, S., Seth, A. MicroRNA Detection in Prostate Tumors by Quantitative Real-time PCR (qPCR). *J. Vis. Exp.* (63), e3874, doi:10.3791/3874 (2012).
32. Schularick, N. M., Clark, J. J., Hansen, M. R. Primary Culture of Human Vestibular Schwannomas. *J. Vis. Exp.* (89), e51093, doi:10.3791/51093 (2014).
33. Eslami, A., Lujan, J. Western Blotting: Sample Preparation to Detection. *J. Vis. Exp.* (44), e2359, doi:10.3791/2359 (2010).
34. McVey, M., & Lee, S. E. MMEJ repair of double-strand breaks (director's cut): deleted sequences and alternative endings. *Trends in Genetics.* **24** (11), 529–38, doi: 10.1016/j.tig.2008.08.007 (2008).
35. Symington, L. S., & Gautier, J. Double-strand break end resection and repair pathway choice. *Annual Review of Genetics.* **45**, 247–71, doi: 10.1146/annurev-genet-110410-132435 (2011).
36. Fu, Y. *et al.* High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. *Nature Biotechnology.* **31** (9), 822–826, doi: 10.1038/nbt.2623 (2013).
37. Wu, X. *et al.* Genome-wide binding of the CRISPR endonuclease Cas9 in mammalian cells. *Nature Biotechnology.* doi:10.1038/nbt.2889 (2014).
38. Hsu, P. D. *et al.* DNA targeting specificity of RNA-guided Cas9 nucleases. *Nature Biotechnology.* **31** (9), 827–32, doi: 10.1038/nbt.2647 (2013).
39. Fu, Y. *et al.* Improving CRISPR-Cas nuclease specificity using truncated guide RNAs. *Nature Biotechnology.* **32**, 279–284, doi: 10.1038/nbt.2808 (2014).

40. Alemán, L. M., Doench, J., & Sharp, P. a Comparison of siRNA-induced off-target RNA and protein effects. *RNA.* **13** (3), 385–95 (2007).
41. Anderson, E. M. *et al.* Experimental validation of the importance of seed complement frequency to siRNA specificity Experimental validation of the importance of seed complement frequency to siRNA specificity. *RNA.* **14**, 853–861, doi: 10.1261/rna.704708 (2008).
42. Marine, S. *et al.* Common seed analysis to identify off-target effects in siRNA screens. *Journal of Biomolecular Screening.* **17** (3), 370–8, doi: 10.1177/1087057111427348 (2012).