

Application Note: Gene Knock-out of *GAPDH* using SBI's CRISPR/Cas9 SmartNuclease™ system and PrecisionX™ HR Targeting Vectors

Table of Contents:

I. Background	Pg. 1
II. Analysis of Gene Target	Pg. 2
III. Guide RNA Design	Pg. 4
IV. Designing Homology Arms and HR Targeting Vector Construction	Pg. 7
V. Protocol for Co-Transfection of HR Targeting Vector with Cas9/gRNA Plasmids and Characterization of Cells	Pg. 10



I. Background:

The recent discovery of the CRISPR/Cas9 complex has provided researchers an invaluable tool to target and modify any genomic sequence with high levels of efficacy and specificity. The system, consisting of an RNA-guided nuclease (Cas9) and guide RNA (gRNA) complementary to a target sequence, allows for sequence-specific cleavage of a target locus containing a protospacer adaptor motif (PAM) of sequence “NGG” (Fig. 1). By changing the gRNA sequence, virtually any gene sequence with an adjacent PAM can be targeted by the CRISPR/Cas9 system, enabling the possibility of systematic targeting of sequences on a genomic scale.

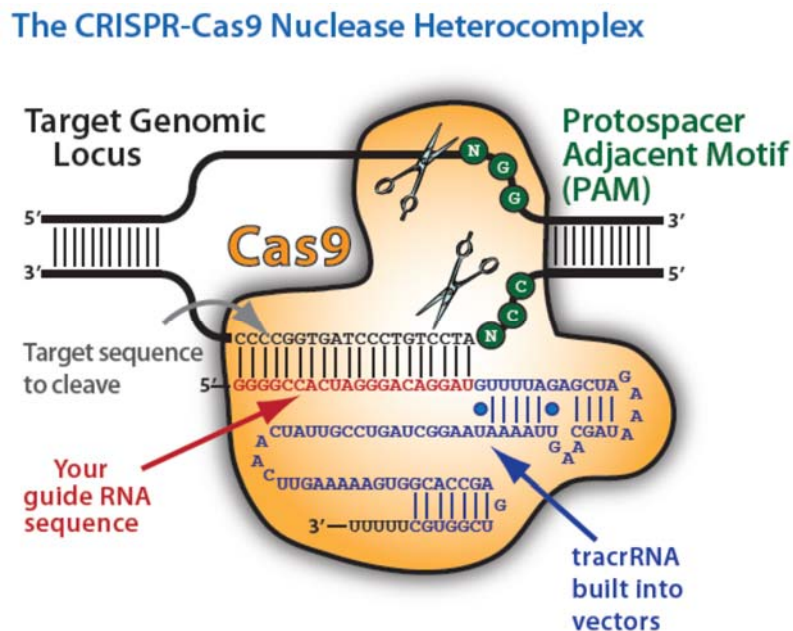


Fig. 1. Illustration of the CRISPR/Cas9 Heterocomplex

This application note is designed for first-time and experienced users of the CRISPR/Cas9 system to learn how to Knock-Out (KO) the human *GAPDH* gene using a combination of SBI's PrecisionX Cas9 SmartNuclease system and a Homologous Recombination (HR) Targeting vector. The protocol is designed for use of the CAS940A-1 [CMV-hspCas9-H1-gRNA] cloning vector and HR110PA-1 [MCS-LoxP-EF1a-RFP-T2A-Puro-LoxP-MCS] HR Knock-out vector. Other combinations of Cas9 SmartNuclease Vectors and HR Knock-out vectors may be substituted as well.

II. Analysis of Gene Target

Wild-type GAPDH (glyceraldehyde-3-phosphate dehydrogenase) genomic sequence:

```

ATCGGGAAGG TGAAGGTCGG AGTCAACGGg tgagttcgcg ggtggctggg 6644048
gggcccctggg ctgcgaccgc ccccgaaccg cgtctacgag ccttgccggc 6644098
tccgggtcctt tgcagtcgta tgggggcagg gtagctgttc cccgcaagga 6644148
gagctcaagg tcagcgctcg gacctggcgg agccccgcac ccaggctgtg 6644198
gcgcccctgtg cagctccgcc cttgcggcgc catctgcccg gagcctcctt 6644248
cccctagtc cccagaaacag gaggtcccta ctcccgcctg agatcccgcac 6644298
ccggaccctt aggtggggga cgctttcttt cctttcgcgc tctgcccggg 6644348
cacgtgtcgc agaggagccc ctccccacg gcctccggca ccgcaggccc 6644398
cgggatgcta gtgcgcagcg ggtgcatccc tgtccggatg ctgcccctgc 6644448
ggtagagcgg ccgccatggt gcaaccggga aggaaatgaa tgggcagccg 6644498
ttaggaaagc ctgcccgtga ctaaccctgc gctcctgcct cgatgggtgg 6644548
agtgcgctgt ggcggggaag tcaggtggag cgaggctagc tggcccgatt 6644598
tctcctccgg gtgatgcttt tcctagatta ttctctggta aatcaaagaa 6644648
gtgggtttat ggaggctctc ttgtgtcccc tccccgcaga ggtgtggtgg 6644698
ctgtggcatg gtgccaaacc gggagaagct gactcatggg tagttggaaa 6644748
aggacatttc caccgcaaaa tggcccctct ggtggtggcc ccttccctgca 6644798
gcgcccggctc acctcacggc cccgcccttc ccctgccagc ctagcgttga 6644848
cccgacccca aaggccaggc tgtaaatgtc accgggagga ttgggtgtct 6644898
gggcccctcg gggaaacctgc ccttctcccc attccgtctt ccggaaacca 6644948
gatctcccac cgcaccctgg tctgaggtta aatatagctg ctgacctttc 6644998
tgtagctggg ggccctgggt ggggctctct cccatccctt ctccccacac 6645048
acatgcactt acctgtgctc ccactcctga tttctggaaa agagctagga 6645098
aggacaggca acttggcaaa tcaaagccct gggactaggg ggttaaataa 6645148
cagcttcccc tcttcccacc cgccccagtc tctgtccctt ttgtaggagg 6645198
gacttagaga aggggtgggc ttgccctgtc cagttaattt ctgaccttta 6645248
ctcctggcct ttgagtttga tgatgctgag tgtacaagcg ttttctccct 6645298
aaagggtgca gctgagctag gcagcagcaa gcattcctgg ggtggcatag 6645348
tgggggtggtg aataccatgt acaaagcttg tgcccagact gtgggtggca 6645398
gtgccccaca tggccgcttc tcctggaagg gcttcgtatg actgggggtg 6645448
ttgggcagcc ctggagcctt cagttgcagc catgccttaa gccaggccag 6645498
cctggcaggg aagctcaagg gagataaaat tcaacctctt gggccctcct 6645548
gggggtaagg agatgctgca ttgcacctct taatggggag gtggcctagg 6645598
gctgctcaca tattctggag gaggctcccc tctcatgccc ttcttgctc 6645648
ttgtctctta gATTTGGTGC TATTGGGCGC CTGGTCACCA GGGCTGCTTT 6645698
TAACTCTGGT AAAGTGGATA TTGTTGCCAT CAATGACCCC TTCATTGACC 6645748
TCAACTACAT Ggtgagtgct acatggtgag ccccaaagct ggtgtggggag 6645798
gagccacctg gctgatgggc agccccctca taccctcacg tattccccca 6645848
gGTTTACATG TTCCAATATG ATTCCACCCA TGGCAAATTC CATGGCACCG 6645898
TCAAGGCTGA GAACGGGAAG CTTGTTCATCA ATGGAAATCC CATCACCATC 6645948
TTCCAGGAgt gagtgggaaga cagaatggaa gaaatgtgct ttggggaggc 6645998
aactaggatg gtgtggctcc cttgggtata tggtaacctt gtgtccctca 6646048
atatggtcct gtccccatct cccccccacc cccatagGCG AGATCCCTCC 6646098
AAAATCAAGT GGGCGCATGC TGGCGTGTAG TACGTCGTGG AGTCCACTGG 6646148
CGTCTTCACC ACCATGGAGA AGGCTGGGgt gagtgcagga gggcccgcgg 6646198
gaggggaagc tgactcagcc ctgcaaagcc aggaccggg ttcataactg 6646248
tctgcttctc tgctgtagGC TCATTTGCAG GGGGGAGCCA AAAGGGTCAT 6646298
CATCTCTGCC CCCTCTGCTG ATGCCCCCAT GTTCGTCATG GGTGTGAACC 6646348
ATGAGAAGTA TGACAACAGC CTCAAGATCA TCAGgtgagg aaggcagggc 6646398
ccgtggagaa gcggccagcc tggcacccta tggacacgct cccctgactt 6646448

```

```

gcgccccget cctcttttct ttgcagCAAT GCCTCCTGCA CCACCAACTG 6646498
CTTAGCACCC CTGGCCAAGG TCATCCATGA CAACTTTGGT ATCGTGGAAG 6646548
GACTCATGgt atgagagctg gggaatggga ctgaggctcc cacctttctc 6646598
atccaagact ggctcctccc tgccggggct gcgtgcaacc ctggggttgg 6646648
gggttctggg gactggcttt cccataatth cttttcaagg tggggagggg 6646698
ggtagagggg tgatgtgggg agtacgctgc agggcctcac tccttttgca 6646748
gACCACAGTC CATGCCATCA CTGCCACCCA GAAGACTGTG GATGGCCCCT 6646798
CCGGGAAACT GTGGCGTGAT GGCCGCGGGG CTCTCCAGAA CATCATCCCT 6646848
GCCTCTACTG GCGCTGCCAA GGCTGTGGGC AAGGTCATCC CTGAGCTGAA 6646898
CGGGAAGCTC ACTGGCATGG CCTTCCGTGT CCCCCTGCC AACGTGTCAG 6646948
TGGTGGACCT GACCTGCCGT CTAGAAAAAC CTGCCAAATA TGATGACATC 6646998
AAGAAGGTGG TGAAGCAGGC GTCGGAGGGC CCCCTCAAGG GCATCCTGGG 6647048
CTACACTGAG CACCAGGTGG TCTCCTCTGA CTTCAACAGC GACACCCACT 6647098
CCTCCACCTT TGACGCTGGG GCTGGCATTG CCCTCAACGA CCACTTTGTC 6647148
AAGCTCATTT CCTGgtatgt ggctggggcc agagactggc tcttaaaaag 6647198
tgcagggctc ggcgcctct ggtggctggc tcagaaaaag ggccctgaca 6647248
actcttttca tcttctagGT ATGACAACGA ATTTGGCTAC AGCAACAGGG 6647298
TGGTGGACCT CATGGCCCAC ATGGCCTCCA AGGAGTAAga cccctggacc 6647348
accagcccca gcaagagcac aagaggaaga gagagaccct cactgctggg 6647398
gagtccctgc cacactcagt cccccaccac actgaatc

```

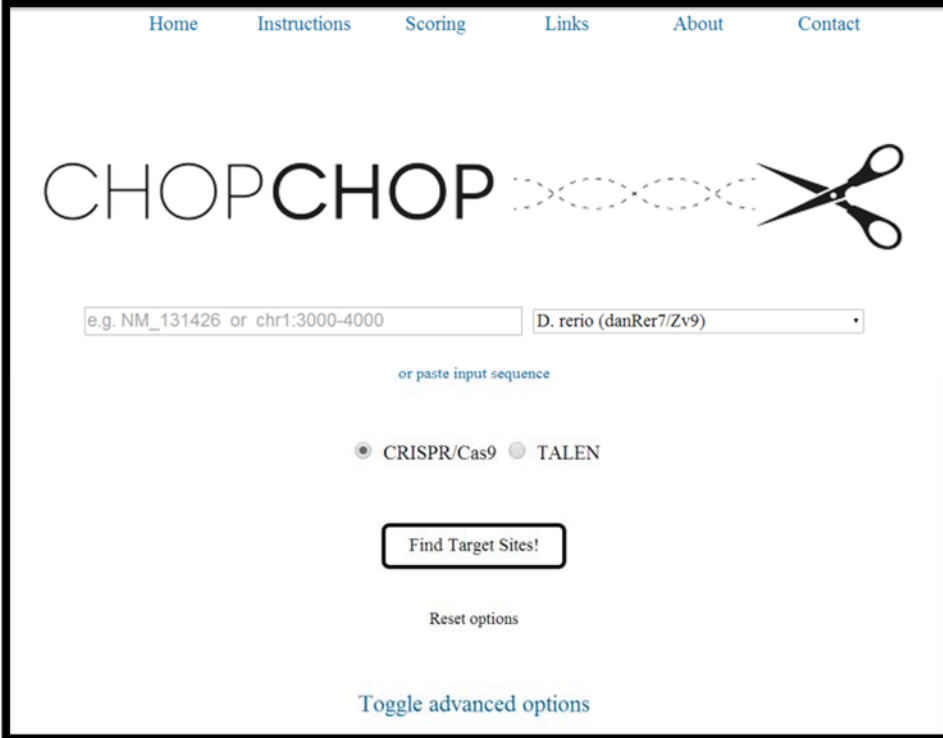
Exons colored blue, introns in black

Translation Start Site

Stop Site

III. Guide RNA Design

A single gRNA targeting an exon – intron junction has been selected for the KO of human *GAPDH* gene using the CHOPCHOP CRISPR/gRNA algorithm (<https://chopchop.rc.fas.harvard.edu/>) which identifies and ranks gRNA targets using an evidence-based scoring algorithm incorporating the # of potential off-target hits as well as GC-content.



The screenshot shows the CHOPCHOP web interface. At the top, there is a navigation menu with links for Home, Instructions, Scoring, Links, About, and Contact. The main heading is "CHOPCHOP" followed by a graphic of a DNA double helix being cut by a pair of scissors. Below this, there is a text input field containing "e.g. NM_131426 or chr1:3000-4000" and a dropdown menu set to "D. rerio (danRer7/Zv9)". A link "or paste input sequence" is positioned below the input field. There are two radio buttons for "CRISPR/Cas9" (selected) and "TALEN". A "Find Target Sites!" button is centered below the radio buttons, with a "Reset options" link underneath. At the bottom, there is a "Toggle advanced options" link.

Fig. 1. Screenshot of CHOPCHOP CRISPR/gRNA design utility

1. Open the CHOPCHOP CRISPR/gRNA design tool (<https://chopchop.rc.fas.harvard.edu/>)
2. Obtain the genomic DNA sequence of *GAPDH* starting from the first coding exon to the stop codon (include introns) using genome browser utilities such as the UCSC Genome Browser (<http://genome.ucsc.edu/>). Gene ID or RefSeq accession numbers may also be used for the next step, but we find it is helpful to have a copy of the reference genomic sequence available for gRNA and HR donor vector design.

3. Paste the sequence into a plasmid utility program (such as the ApE plasmid editor, <http://biologylabs.utah.edu/jorgensen/wayned/ape/>) to remove any non-nucleotide characters and copy the resulting sequence file as input into the CHOPCHOP program by clicking the “or paste input sequence” link, or enter the target Gene ID or mRNA accession number.

Note: The CHOPCHOP program provides options for changing parameters to specify general sequence and algorithm restrictions for identifying gRNAs under the “Toggle Advanced Options” link. For the purposes of this application note, the default settings can be used.



The screenshot shows the CHOPCHOP web interface. At the top, the text "CHOPCHOP" is displayed next to a DNA double helix icon and a pair of scissors icon. Below this, there is a text input field containing "e.g. NM_131426 or chr1:3000-4000" and a dropdown menu set to "H. sapiens (hg19/GRCh37)". A link "or paste input sequence" is positioned below the input field. A large text area contains a pasted genomic DNA sequence:

```
CATCCTGGGCTACACTGAGCACCAGGTGGTCTCCTCTGACTTCAACAGCGACACCCACT
CCTCCACCTTTGACGCTGGGGCTGGCATTGCCCTCAACGACCACTTTGTCAAGCTCATT
CCTGgtatgtggctggggccagagactggctctfaaaaagtgcagggctgcccctctggctggctcagaaaaaggccctgacaact
ctttcatcttagGTATGACAACGAATTTGGCTACAGCAACAGGGTGGTGGACCTCATGGCCCA
CATGGCCTCCAAGGAGTAAgaccctggaccaccagccccagcaagagcacaagaggaagagagaccctcactgctgg
ggagtccctgccacactcagtccccaccacactgaatc
```

Below the text area, there are two radio buttons: "CRISPR/Cas9" (selected) and "TALEN". A "Find Target Sites!" button is located below the radio buttons, and a "Reset options" link is at the bottom of the interface.

Fig. 2. Screenshot of pasting genomic DNA sequence into CHOPCHOP program

4. Click on “Find Target Sites” to get a ranked list of suitable gRNA targets (Fig. 3 on next page). We chose the first gRNA on the list to target a sequence closest to the ATG start site of *GAPDH* to maximize disruption of the wild-type protein, while minimizing the chance of targeting other sequences in the genome. For other gene targets, one can also consider targeting key functional domains, such as those involved in protein/DNA binding or enzymatic activity (e.g. phosphorylation domain for kinases) for effective knockouts.

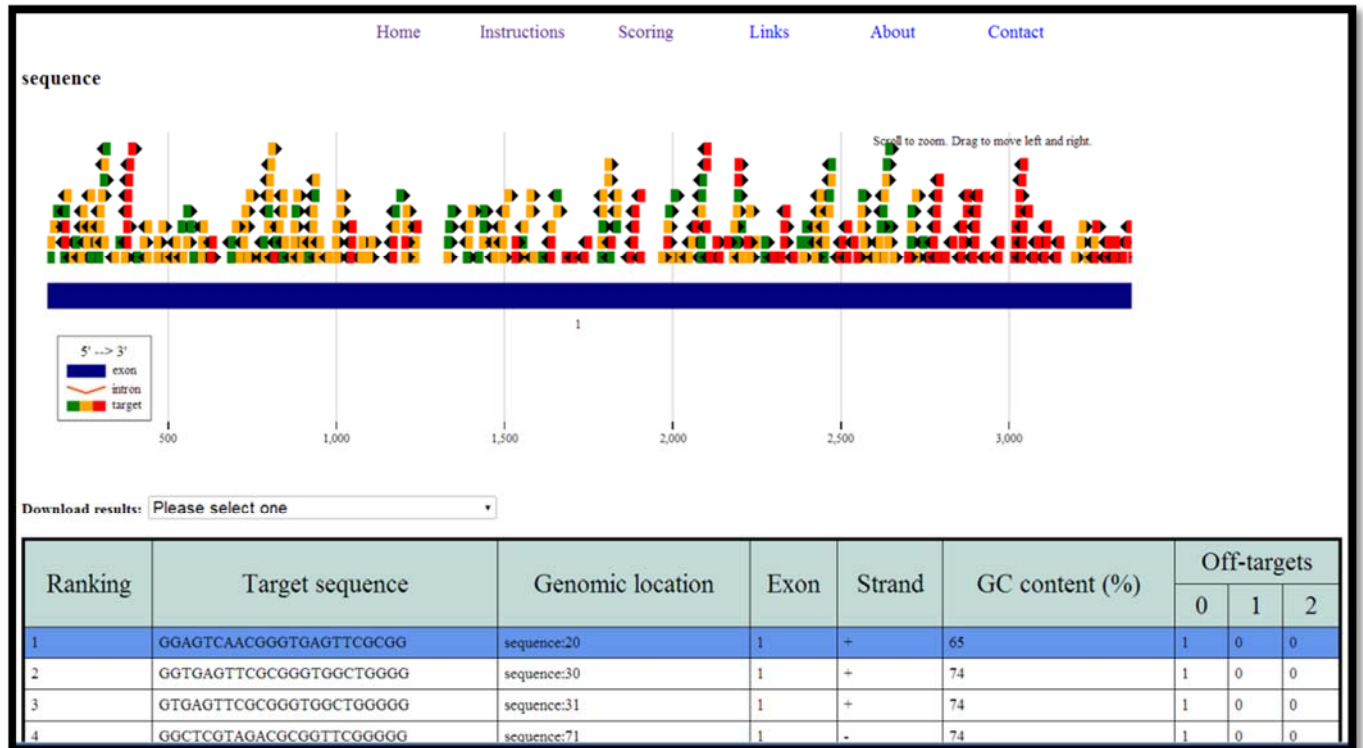


Fig. 3 – Screenshot of CHOPCHOP gRNA output

<u>gRNA ID</u>	<u>Location of gRNA</u>	<u>Target Sequence</u>	<u>Protospacer Adaptor Motif</u>
gRNA-GAPDH	Chr12: 6644017	5' GGAGTCAACGGGTGAGTTCG 3'	CGG

The gRNA will target the corresponding sequence in the genomic DNA as shown below:

ATGGGGAAGGTGAAGGTCGGAGTCAACGGGgtgagttcgCGGgtggctggggggcctgggctgcaccgccccgaaccggtctac
 gagcctgcgggctccgggtcttgcagtcgtatgggggcagggtagctgttcccccaaggagagctcaaggtcagcgtcggacctggcggagccc
 cgaccaggctgtggcgcctgtgcagctccgacctgc

Exon in upper case, intron in lower case

Translation Start Site

gRNA sequence

PAM motif

↓ Nuclease cut site

5. Clone the gRNA insert into SBI's All-in-One PrecisionX SmartNuclease cloning vector (CMV-hspCas9-H1-gRNA, catalog #CAS940A-1) following the recommended protocols in the user manual. (<http://www.systembio.com/downloads/Cas9-SmartNuclease-user-manual.pdf>)

IV. Designing Homology Arms and HR Targeting Vector Construction

While gene targeting using the CRISPR/Cas9 system alone will likely induce formation of random mutations by the NHEJ repair process, leading to functional knockouts, in practice it is difficult to determine exactly which cells in a heterogeneous population have been mutated by CRISPR/Cas9 in the absence of a selection marker. In order to facilitate the generation of stable KO cell lines, SBI has built a line of HR Targeting Vectors (<http://www.systembio.com/genome-engineering-precisionx-HR-vectors>) that leverage the cell's ability to incorporate large exogenous DNA sequences via homologous recombination (HR) at double-stranded DNA breaks (DSB). The incorporation of fluorescent and/or selection markers via HR allows researchers to easily generate and identify KO clones.

HR Targeting Vectors serve as the donor template to induce HR in cells that have been targeted at a specific locus using CRISPR/Cas9 or TALEN nucleases. A general targeting vector will contain the following features:

- 1) Homologous sequences (to the template with DSBs) at 5' and 3' ends of the DSB (cutting) site
- 2) An expression cassette bearing a promoter, insert (cDNA, microRNA, non-coding RNA, etc.), fluorescent marker, and/or selection agent to select cells that have undergone HR

The 5' and 3' homologous sequences (termed "homology arms") should be exactly homologous to the template (genomic sequence) with the DSB, and preferably directly adjacent to the actual DSB site. An example of homology arm design is shown in Figure 4:

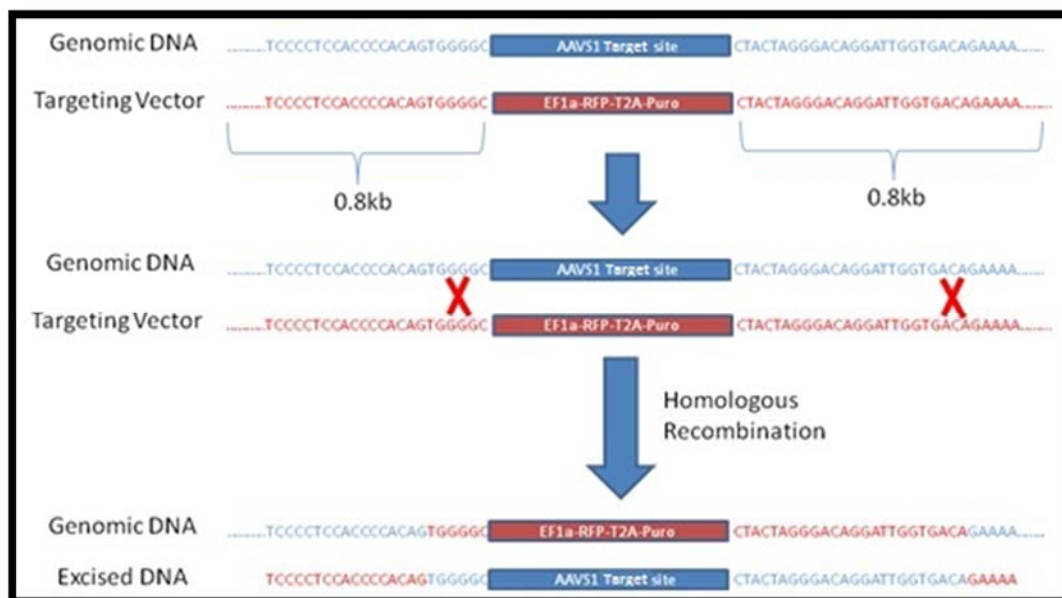


Figure 4. Schematic Diagram of Homologous Recombination (HR) Process

In Figure 4, a region of the human AAVS1 safe harbor locus (in blue) is targeted by a gRNA + Cas9 in order to insert (knock-in) an EF1 α -RFP-T2A-Puro expression cassette (in red) present in the HR Targeting Vector through homologous recombination (HR). The HR targeting vector contains homology arms at the 5' and 3' end of the expression cassette which each include ~0.8kb of sequence homologous to the genomic DNA surrounding the targeted AAVS1 locus. This region of homology is crucial for the success of the homologous recombination reaction, as it serves as the guide template for specifically targeting the exogenous cassette into this genomic locus. The typical size range for homology arms varies by the application, but it should be anywhere from 0.5kb to 1kb for each arm for efficient recombination to occur. Please note that the actual regions of recombination (Red "Xs") at the 5' and 3' of the target site can vary widely, thus it is difficult to predict the actual sites as this is determined by the cell.

With the CRISPR/Cas9 system, the cleavage site is 2-3bp upstream of the protospacer adaptor motif (PAM) immediately following the guide RNA sequence; therefore, homology arms should be designed to be as close as possible to this cut site (<10bp) at both ends for efficient HR reaction. (For TALEN-mediated cleavage, homology arms are designed to be adjacent to the spacer region between TALEN binding sites, which spans 15-30bp and is the site of the DSB.)

Important Caution when Designing Homology Arms for Use with CRISPR/Cas9: It is best practice to avoid including the full target sequence (gRNA sequence + PAM) in the HR Targeting Vector, to ensure that the donor vector is not targeted for cleavage by the Cas9/gRNA complex. As such, homology arms can be designed to either:

- a) Design the arms so that the gRNA spans across both 5' and 3' arms, so that neither arm contains the full gRNA + PAM sequence; or
- b) Delete some or all of the gRNA or PAM sequence for a particular gRNA from the homology arm, so that the Cas9/gRNA complex does not cut the Targeting Vector. Deleting the PAM sequence – or even the full gRNA sequence – from a homology arm will introduce this deletion into the targeted genome upon successful homologous recombination, but for Gene Knock-Out applications, this should not be a concern.

Homology Arm Design for the GAPDH Knock-Out HR Vector:

5' Arm Design

5' EcoRI RE site-----1.0kb of upstream sequence----**ATG**GGGAAGGTGAAGGTC*----KpnI RE site 3'

3' Arm Design

5' BamHI RE site----**cgg**gtggctggggggccctgggctgcga----1.0kb of downstream sequence----SphI RE Site 3'

*- Note: In this case, the entire gRNA target sequence was excluded from both the 5' and 3' homology arms, to ensure that the Cas9/gRNA complex will not target the HR donor vector.

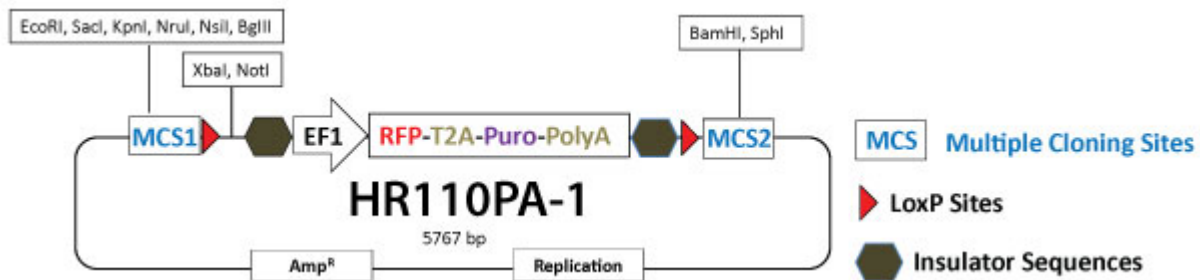
The final HR targeting vector would look similar to the following:

[Rest of 5' Arm] ATCGGGAAGGTGAAGGTCGGTACC [LoxP-Ins1-EF1a-RFP-T2A-Puro-Ins2-LoxP]
 GGATCCcgggtggctggggggccctgggctgcgaccgccccgaaccgctctacgagccttgccgggctccgggtctttgagtcgtatgggggca
 gggtagctgttccccgaaggagagctcaaggtcagcgcctcgacactggcggagccccgcaccaggctgtggcgccctgtgcagctccgccccttgcc
[Rest of 3' Arm]
Restriction Sites

By introducing a selection cassette into the first coding exon of GAPDH, this would disrupt the normal splicing of the GAPDH mRNA, and would create a functional knock-out of the GAPDH gene. The inserted selection cassette would be expressed normally as it is under the control of an autonomous EF1a promoter, enabling selection of successfully targeted cells.

Cloning of the Homology Arms into the HR Targeting Vector

The 5' and 3' homology arms can be created by PCR amplification of the genomic DNA sequence or through gene synthesis. The resulting homology arms can be readily cloned into our HR Targeting Vector (e.g. HR110PA-1) using the available restriction sites present in the vector:



General cloning strategy of homology arms into HR110PA-1 Vector Using Traditional Ligation Cloning:

1. Clone in the 5' homology arm by digesting the vector using two different enzymes in the 5' MCS that are compatible in the same buffer (e.g. EcoRI and KpnI). We suggest using NEB High-Fidelity (HF) enzymes or equivalent for best results. Please note that the homology arms must contain compatible restriction sites for directional cloning of the arm into the vector.
2. Transform competent cells, pick colonies, propagate the plasmid, and check via agarose gel analysis/sequencing to confirm insertion of the 5' arm.
3. Using the correct clone from Step 2, clone in the 3' homology arm using BamHI and SphI HF enzymes.

For fast, efficient, no-hassle cloning of homology arms into HR donor vectors, we suggest using SBI's Cold Fusion Cloning Kit, with no restriction enzymes or ligation required:

<http://www.systembio.com/molecular-tools/cold-fusion-cloning/overview>

We suggest sequential cloning of the homology arms into the MCS for best results.

V. Protocol for Co-Transfection of HR Targeting Vector with Cas9/gRNA Plasmids and Characterization of Cells

1) Plate 200,000 to 300,000 cells (e.g. 293T cells) into a single well of a 12-well plate in 1 ml of appropriate growth medium. Include a single well of cells as negative control (which can be non-relevant plasmid DNA).

2) Next day, or when cells are 50-60% confluent, co-transfect target cells with Cas9/gRNA plasmid(s) and the HR targeting vector using a suitable transfection reagent following the manufacturer's recommended protocol for 12-well plates. The use of reduced or serum-free media containing no antibiotics to dilute the vector/transfection complex is highly recommended.

Note: For 293T cells, we suggest 0.5 µg of SBI's Cas9 SmartNuclease vector in conjunction with 0.5 µg of the HR targeting vector into cells for efficient cleavage and HR reaction. For other cell types, we suggest optimizing the amounts and ratios of Cas9 plasmid to targeting vector for optimal results.

3) Allow at least 12 hours before changing transfection media to complete growth media.

4) Assay for positive HR events 48-96 hours after co-transfection. Select cells with insertion of targeting vector using fluorescent or antibiotic selection. If using selection by Puromycin, select cells for a minimum of 5 days prior to further characterization. Cutting efficiency of Cas9 can be measured by Surveyor Nuclease Assay and HR efficiency by % of fluorescence signal via FACS sorting.

5) After selection of cells in Puro, remaining colonies can be characterized by PCR genotyping or Sanger sequencing to confirm knockout of one or both alleles. SBI offers the EZ-Genotyping kit for fast characterization of engineered cells:

<http://www.systembio.com/genome-engineering-ez-genotyping-kit>

6) Colonies possessing the desired genotype(s) can be clonally propagated for further phenotypic characterization and experimentation.

Copyright © System Biosciences, 2014