

# Application Note: Single Nucleotide Gene Correction Of Mutant (E6V) Beta-globin (HBB) Gene To Wild-Type 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 against Mutant HBB Locus	Pg. 3
IV. Designing Homology Arms and HR Targeting Vector Construction	Pg. 7
V. Protocol for Co-Transfection of HR210PA-1 Targeting Vector with Cas9/gRNA Plasmids and Characterization of Cells	Pg. 11



## 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.

### The CRISPR-Cas9 Nuclease Heterocomplex

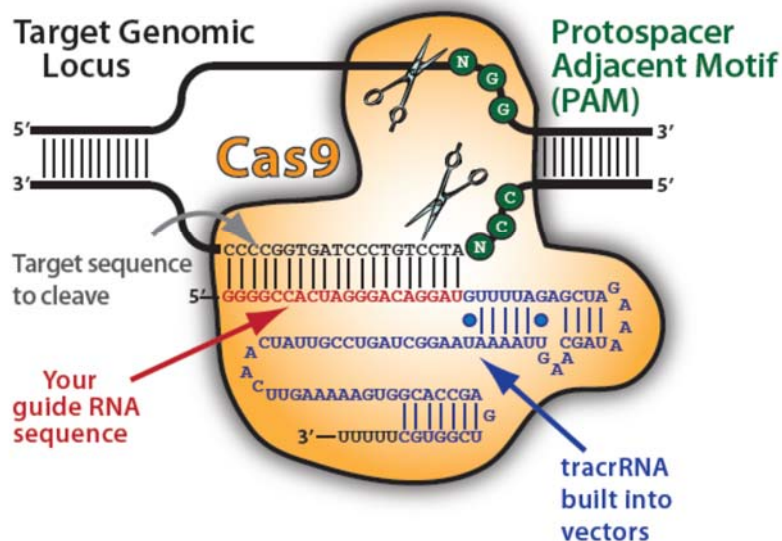


Fig. 1. Illustration of the CRISPR/Cas9 Heterocomplex

This application note is designed for first-time and experienced users of the CRISPR/Cas9 system to learn how to perform single nucleotide gene correction of mutant human beta-globin gene (*HBB*, *E6V mutation*) to the wild-type sequence using a combination of SBI's PrecisionX Cas9 SmartNuclease and Homologous Recombination (HR) Targeting vectors. The protocol is designed for use of the CAS940A-1 [CMV-hspCas9-H1-gRNA] cloning vector and HR210PA-1 [MCS-LoxP-EF1a-GFP-T2A-Puro-P2A-hsvTK-LoxP-MCS] HR Targeting vector. Other combinations of Cas9 SmartNuclease Vectors and HR Targeting vectors may be substituted as well.

## II. Analysis of Gene Target

Mutant beta-globin (HBB) genomic sequence containing E6V mutation

```

tgtggagcca caccctaggg ttggccaatc tactcccagg agcagggagg 5248352
gcaggagcca gggctgggca taaaagtcag ggcagagcca tctattgctt 5248302
ACATTTGCTT CTGACACAAC TGTGTTCACT AGCAACCTCA AACAGACACC 5248252
ATCGTGCATC TGA CTCTCTGT GGAGAAGTCT GCCGTTACTG CCCTGTGGGG 5248202
CAAGGTGAAC GTGGATGAAG TTGGTGGTGA GGCCCTGGGC AGgttgggat 5248152
caaggttaca agacagggtt aaggagacca atagaaactg ggcattgtgga 5248102
gacagagaag actcttgggt ttctgatagg cactgactct ctctgcctat 5248052
tggctctattt tcccaccctt agGCTGCTGG TGGTCTACCC TTGGACCCAG 5248002
AGGTTCTTTG AGTCCTTTGG GGATCTGTCC ACTCCTGATG CTGTTATGGG 5247952
CAACCCTAAG GTGAAGGCTC ATGGCAAGAA AGTGCTCGGT GCCTTTAGTG 5247902
ATGGCCTGGC TCACCTGGAC AACCTCAAGG GCACCTTTGC CACACTGAGT 5247852
GAGCTGCACT GTGACAAGCT GCACGTGGAT CCTGAGAACT TCAGGgtgag 5247802
tctatgggac gcttgatggt ttctttcccc ttcttttcta tggttaagtt 5247752
catgtcatag gaaggggata agtaacaggg tacagttagg aatgggaaac 5247702
agacgaatga ttgcatcagt gtggaagtct caggatcgtt ttagtttctt 5247652
ttatttgctg ttcataacaa ttgttttctt ttgtttaatt ctgctttct 5247602
ttttttttct tctccgcaat ttttactatt atacttaatg ccttaacatt 5247552
gtgtataaca aaaggaata tctctgagat acattaagta acttaaaaaa 5247502
aaactttaca cagtctgcct agtacattac tatttggaat atatgtgtgc 5247452
ttatttgcat attcataatc tccctacttt attttctttt atttttaatt 5247402
gatacataat cattatacat atttatgggt taaagtgtaa tgttttaata 5247352
tgtgtacaca tattgaccaa atcagggtaa ttttgcaatt gtaattttaa 5247302
aaaatgcttt ctctctttta tatacttttt tgtttatctt atttctaata 5247252
ctttccctaa tctcttttct tcagggcaat aatgatacaa tgtatcatgc 5247202
ctctttgac cattctaaag aataacagtg ataatttctg ggttaaggca 5247152
atagcaatat tctctgatat aaatatttct gcatataaat tgtaactgat 5247102
gtaagagggt tcatattgct aatagcagct acaatccagc taccattctg 5247052
cttttatatt atgggtggga taaggctgga ttattctgag tccaagctag 5247002
gcccttttgc taatcatggt catacctctt atcttctctc cacagCTCCT 5246952
GGGCAACGTG CTGGTCTGTG TGCTGGCCCA TCACTTTGGC AAAGAATTCA 5246902
CCCCACCGT GCAGGCTGCC TATCAGAAAG TGGTGGCTGG TGTGGCTAAT 5246852
GCCCTGGCCC ACAAGTATCA CTAAAGCTCGC TTTCTTGCTG TCCAATTTCT 5246802
ATTAAAGGTT CCTTTGTTCC CTAAGTCCAA CTACTAACT GGGGGATATT 5246752
ATGAAGGGCC TTGAGCATCT GGATTCTGCC TAATAAAAAA CATTATTATT 5246702
CATTGCaatg atgtatttaa attatttctg aatattttac taaaaggga 5246652
atgtgggagg tcagtgcat taaaacataa agaaatgaag agctagttca 5246602
aacctt

```

Exons colored blue, introns in black

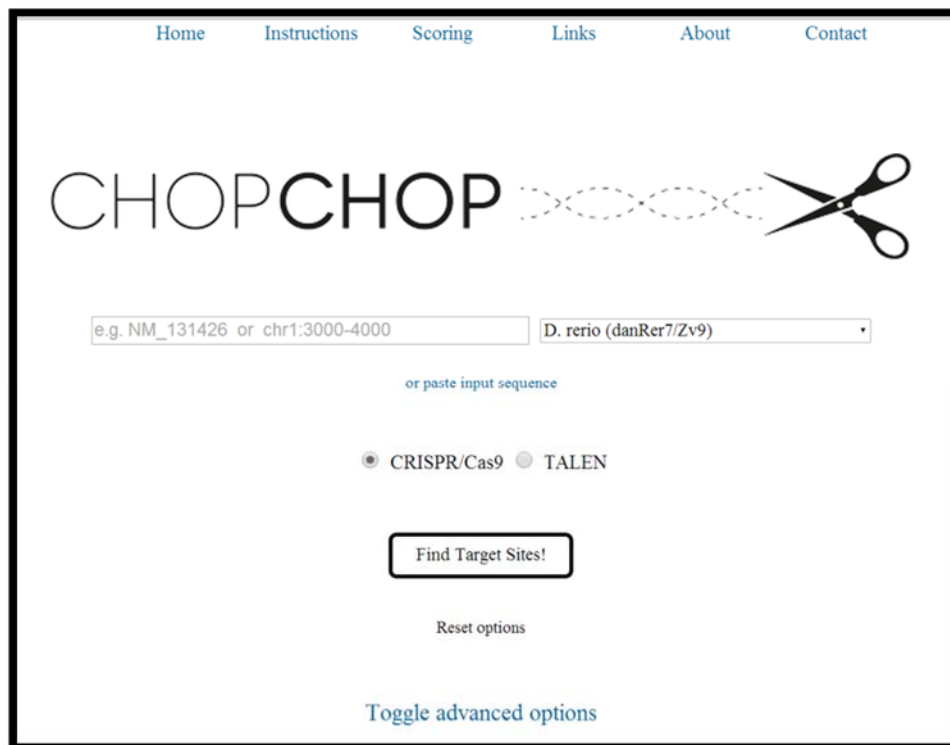
Translation Start Site

Stop Site

E6V mutation (GAG->GTG)

### III. Guide RNA Design against Mutant *HBB* Locus

A single gRNA targeting an intronic region has been selected for correction of the mutant *HBB* 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.

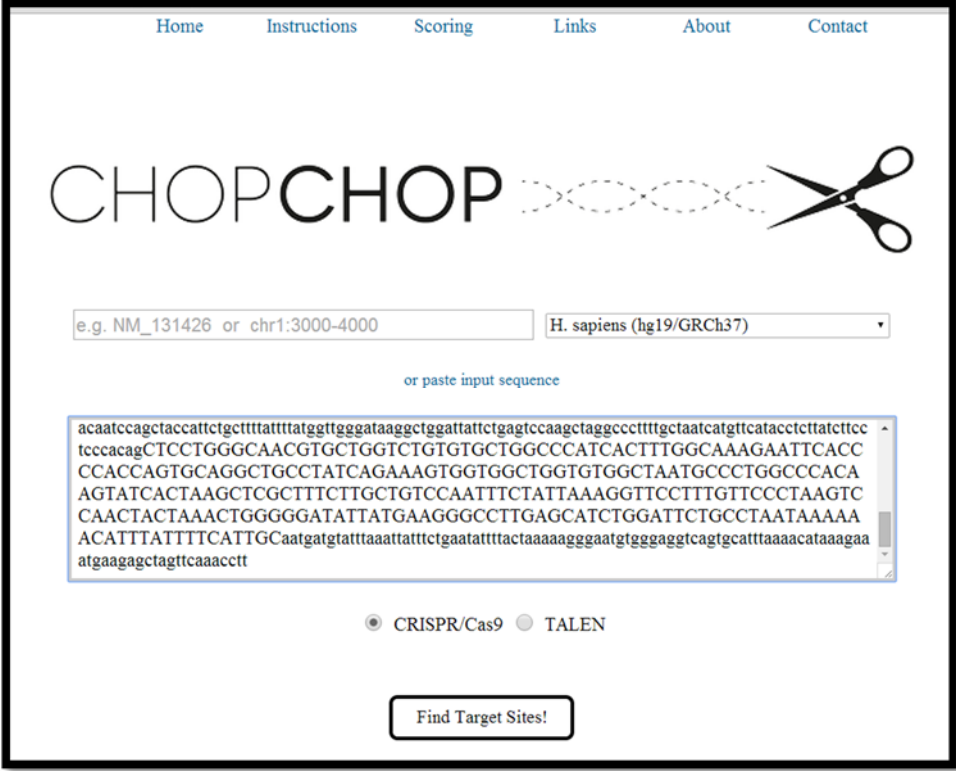


**Fig. 1. Screenshot of CHOPCHOP CRISPR/gRNA design utility**

1. Open the CHOPCHOP CRISPR/gRNA design tool (<https://chopchop.rc.fas.harvard.edu/>)
2. Obtain the genomic DNA sequence of mutant *HBB* gene (or manually change the nucleotide sequence corresponding to amino acid position 6 from GAG->GTG) 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, there are navigation links: Home, Instructions, Scoring, Links, About, and Contact. The main heading is "CHOPCHOP" with a pair of scissors icon to its right. Below the heading, 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)". Below these fields is a link that says "or paste input sequence". A large text area contains a long, multi-line genomic DNA sequence. At the bottom of the interface, there are two radio buttons: "CRISPR/Cas9" (which is selected) and "TALEN". A "Find Target Sites!" button is located at the bottom center.

**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 a guide RNA targeting an intron close to the desired correction site with the highest possible ranking to limit potential off-target cutting as well as minimize disruption of gene function when a selection cassette is inserted (more information in Section IV on HR Donor design).

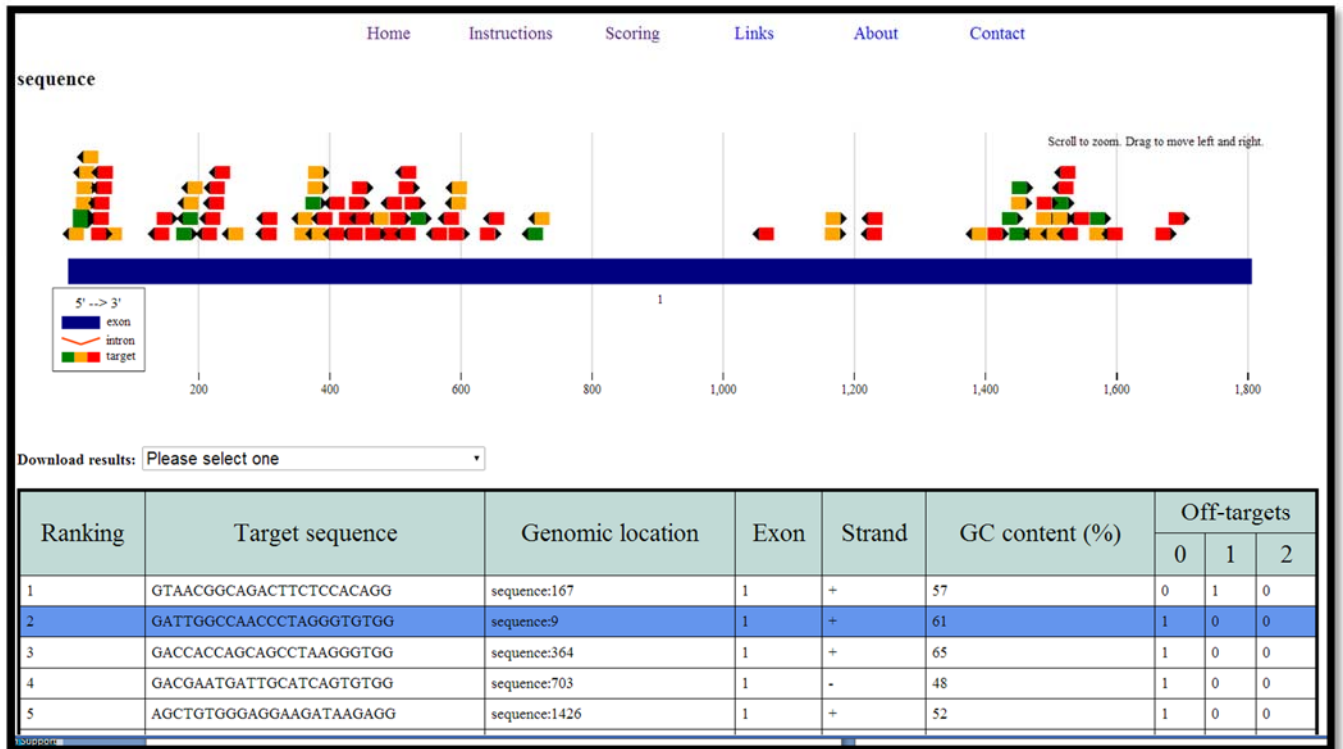


Fig. 3. Screenshot of CHOPCHOP gRNA output

<u>gRNA ID</u>	<u>Location of gRNA</u>	<u>Target Sequence</u>	<u>Protospacer Adaptor Motif</u>
gRNA-HBB-int	Chr11: 5248372	5' GATTGGCCAACCCTAGGGT 3'	TGG

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

↓

tgtggagccacaccctagggttgcccaatctactcccaggagcagggagggcaggagccagggtgggcataaaagt  
cagggcagagccatctattgcttATCGTGCATCTGACTCCTGTGGAGAAGTCTGCCGTTACTGCCCTGTGGGGCAAG  
GTGAACGTGGATGAAGTTGGTGGTGAGGCCCTGGGCAGgttggatcaaggtacaagacaggtttaaggagaccaa  
tagaaactgggcatgtggagacagagaagactcttgggttctgataggcactgactctctctgcctattggctcat  
tttcccacccttagGCTGCTGGTGGTCTACCCTTGGACCCAGAGGTTCTTTGAGTCCTTTGGGGATCTGTCCACTCC  
TGATGCTGTTATGGGCAACCCTAAGGTGAAGGCTCATGGCAAGAAAGTGCTCGGTGCCTTTAGTGATGGCCTGGCTC  
ACCTGGACAACCTCAAGGGCACCTTTGCCACACTGAGTGAGCTGCACTGTGACAAGCTGCACGTGGATCCTGAGAAC  
TTCAGGgtgagtctatgggacgcttgatgttttctttcccttcttttctatggttaagttcatgtcataggaaggg  
gataagtaacagggtagagtttagaatgggaaacagacgaatgattgcatcagtggtggaagtctcaggatcgttta  
gtttcttttatttgctgttcataacaattgttttcttttgtttaattcttgctttcttttttttcttctccgcaat  
tttactattatacttaatgccttaacattgtgtataacaaaaggaaatatctctgagatacattaagtaacttaaa  
aaaaaactttacacagctcgcctagtagtactatttggaaatatatgtgtgcttatttgcatattcataatctccc  
tactttattttcttttatttttaattgatacataatcattatacatatttatggggttaaagtgtaatgttttaatat

gtgtacacatattgaccaaatacagggaatgttgcatttgtaatgttaaaaaatgctttcttcttttaataacttt  
tttgtttatcttatttctaatactttccctaatactcttcttccagggaataatgatacaatgtatcatgcctctt  
tgcaccattctaaagaataacagtgataatgtctgggttaaggcaatagcaatatctctgcatataaatatgtctgc  
atataaattgtaactgatgtaagagggttcatattgctaatagcagctacaatccagctaccattctgcttttattt  
tatgggtgggataaggctggattattctgagctcaagctaggcccttttgctaatacatgttcatacctcttatcttc  
ctcccacagCTCCTGGGCAACGTGCTGGTCTGTGTGCTGGCCCATCACTTTGGCAAAGAATTCACCCCACCAGTGCA  
GGCTGCCTATCAGAAAGTGGTGGCTGGTGTGGCTAATGCCCTGGCCCAAGTATCACTAA

### Exons in upper case, introns in lower case

**Note:** gRNA is complimentary to antisense strand

Translation Start Site

gRNA sequence

PAM motif

E6V mutation (GAG->GTG)

Stop Codon

↓ 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>)

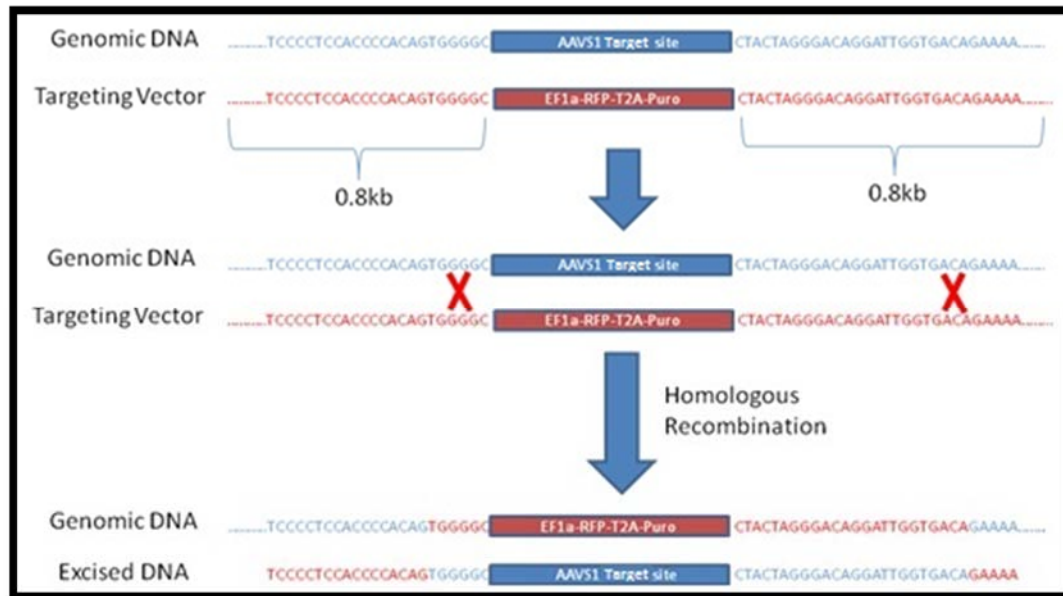
## IV. Designing Homology Arms and HR Targeting Vector Construction

In order to facilitate the generation of cell lines with single nucleotide gene corrections, 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 clones that have the desired gene correction of interest.

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 TALE 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 $\alpha$ -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.)

#### **Special Rules for Designing Homology Arms for Gene Editing Applications:**

- Care must be taken to avoid disrupting the gene coding sequence being targeted for a single base-pair modification.
- Most importantly, the HR Targeting Vector must be designed so that the exogenous selection cassette (used to identify positively engineered clones) is not inserted into the middle of the coding region (as intentionally done for Gene Knock-Out applications).
- We recommend choosing a gRNA (in Section III) to cut in an intronic region adjacent to the desired modification site, which allows the homology arms to be designed in such a way as to insert the selection cassette into this intron, avoiding disruption of the coding sequence.
- When targeting a selection cassette into an intron sequence, care should be taken to not disrupt the native splice donor ("GT") and splice acceptor ("AG") sites flanking the intron.
- The single base-pair mutation or correction is then included in one of the homology arms, to introduce this base change into the target genome.

**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 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 (which may be acceptable if targeting an intronic sequence); or
- c) Generate a silent mutation of the PAM sequence adjacent to the gRNA target when targeting an exonic sequence.

**Note:** For gene editing applications where the gRNA target or PAM site is located in an intron, option A or B may be employed to prevent the donor vector from being cut; however, if the gRNA target or PAM lies in a coding region, care must be taken to avoid introducing deletions in the sequence. In this case, either option A or C is recommended to ensure that the donor vector is not cut without affecting the coding sequence.

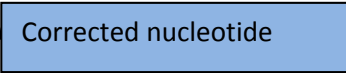
Homology Arm Design for Mutant *HBB* Gene Correction:

### 5' Arm Design

5' [SpeI RE site]-----[1.0kb of Homology Arm Sequence]---tgtggagcca\*---[BsaI RE site] 3'

### 3' Arm Design

5' [NotI RE site]  
 caccctaggggtggccaatc tactcccaggagcaggggaggcaggagccagggctggggcataaaaagt cagggcagag  
 ccatctattgctt ATCGTGCATCTGACTCCTGAGGAGAAGTCTGCCGTTACTGCCCTGTGGGGCAAGGTGAACGTGG  
 ATGAAGTTGGTGGTGGGCCCCTGGGCAGggttggt-- [1.0kb of Homology Arm Sequence]---AgeI RE site 3'



\*Note: In this case, the PAM site and gRNA target sequence span the 5' and 3' homology arms, to ensure that the Cas9/gRNA complex will not target the HR donor vector.

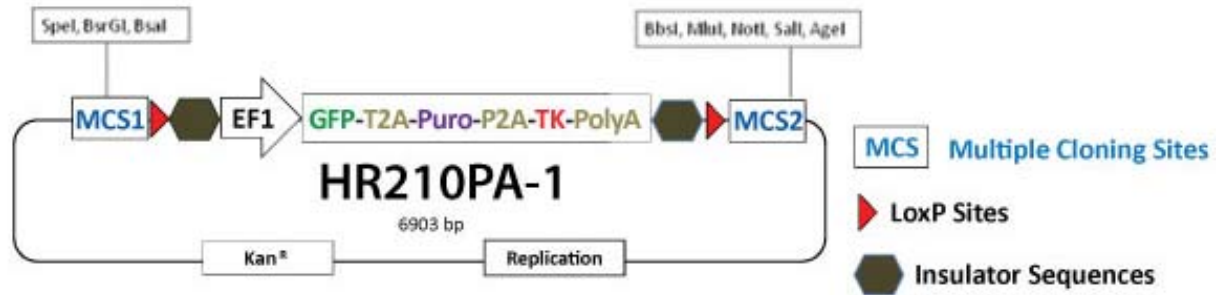
The final HR targeting vector would look similar to the following:

**[Rest of 5' Arm]** tgtggagccaGGTCTC **[LoxP-Ins1-EF1a-GFP-T2A-Puro-P2A-hsvTK-Ins2-LoxP]**  
 GCGGCGC caccctaggggtggccaatc tactcccaggagcaggggaggcaggagccagggctggggcataaaaagtca  
 gggcagagccatctattgctt ATCGTGCATCTGACTCCTGAGGAGAAGTCTGCCGTTACTGCCCTGTGGGGCAAGGT  
 GAACGTGGATGAAGTTGGTGGTGGGCCCCTGGGCAGggttggt **[Rest of the 3' Arm Sequence]**  
**Restriction Sites**

Note that the single nucleotide change to correct the mutant *HBB* sequence is contained in the 3' homology arm, which enables correction of the coding sequence of the mutant *HBB* to wild-type, allowing the cell to produce wild-type *HBB* protein. The selection cassette is inserted into the intron upstream of the first exon and will be expressed independently as it is under the control of an autonomous EF1a promoter, enabling selection of successfully targeted cells. The selection cassette can then be left in this intron, where it should be spliced out via post-transcriptional splicing, OR it can be excised using Cre Recombinase to leave behind only a single LoxP site in the intron (See Step 7 in Section V below).

### Cloning of the Homology Arms into 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 homology arms can be readily cloned into our HR Targeting Vector (e.g. HR210PA-1) using the available restriction sites present in the vector:



### General cloning strategy of homology arms into HR210PA-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. SpeI and BsaI). 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 remaining 3' homology arm using NotI and AgeI 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 HR210PA-1 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 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 propagated for further characterization by PCR genotyping and Sanger sequencing to confirm correction 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 corrections (either heterozygous or homozygous for wild-type *HBB*) can be subsequently passaged and clonally propagated.

7) If desired, the remaining cells can be transfected with the Cre recombinase to excise the selection cassette inserted into the genome. After 48-72hrs post-transfection with Cre recombinase, the cells can be subject to TK negative selection using ganciclovir (2-5ug/ml) or FIAU (20ug/ml). Those cells that still contain the selection cassette after Cre excision will be killed by TK selection, while those remaining will survive.

**Copyright © System Biosciences, 2014**