

Generating plasmids for CRISPR/CAS9 insertion of frameshifting mutations

[This protocol has been adapted from one sent to us by Mike East while in Lee Graves lab at UNC-Chapel Hill]. KO of mouse ELMOD1 is used as an example. There is one splice variant in RefSeq for mouse ELMOD1: [NM_177769.4](#)]

The goal of this protocol is to generate at least two clones each from at least two sequence-independent guide RNAs in which both alleles of your gene of interest contain frameshifting indels in the hopes of generating null lines, which are confirmed by DNA sequencing around the targeted cut sites. Before starting to design oligos for guides, look at NCBI Gene for your gene of interest to see how many transcripts are known. If there are known alternative exons, you want to avoid those to get a complete KO. You might need to do some alignments of the different transcripts to clarify how they differ and keep that alignment handy when picking guide sequences, to further ensure that you are targeting exon sequence that is present in all transcripts. You can ignore the XM/XP transcripts/proteins.

Step 1: Selecting Guide RNA sequences from Benchling and Broad Analyses:

- 1) You will need to create a Benchling account but it is free and a great resource. Login and look for the button that is a plus sign (+) and says "create" when you hover over (left side of screen).
- 2) From the Create button use it to get to drop down menu, under CRISPR choose "CRISPR guides"
- 3) Type your gene name into the box that comes up and choose the appropriate genome (mouse).
- 4) Verify that the accession number that comes up matches the one from PubMed Gene (don't take this step for granted)
- 5) Clicking next will bring up another window for the design type. Select "Single guides" if not already selected (guide length = 20, PAM = NGG). Use all the default settings otherwise.
- 6) Click "Finish" and you'll go to the entire gene (genomic DNA) sequence. The middle of three panels displays both strands of the genomic sequence and will indicate the open reading frame. This coding region is annotated so just scroll down to find it, look for lots of colors. The genomic sequence is shown with the gene name in a tan box underneath. Scroll down to find exons, shown in different colored boxes for each exon, right under the gene box. Alternatively, in the upper right part of the screen there is a list of options that include "Linear map, Design CRISPR, Description, and Metadata". Click on Linear map to see the whole gene and from here you can hover over and click on any exon. Then click on Design Crispr to get back to the screen that will show sequences of guides with scores. The ORF is shown in different colors for each residues, based upon the amino acid. Click on a residue and it will highlight that residue and the three nucleotides that encode it. Click on the exon box and it will highlight the entire exon.
- 7) We want to target indels near the start codon (N-terminus of the protein) to minimize the likelihood of a functional, truncated protein being made, rather than a true "null". Thus, start by clicking on the first exon that includes ORF. Alternatively, you can click on the box in the panel on the right that says "Target Region." You can manually input target regions here. You can also look for a button that looks like a participation ribbon at the top right hand corner of the screen (just above the scissor icon). This button will provide a shortcut to all of the exons within your gene.
- 8) Once you define the region of interest, click the green button with the plus sign in the middle, there should be a blue circle hovering nearby it. There should be one line with start and end residues listed. Under "genome region" click on set region, if more than one option, click the appropriate one, then click "set genome region".
- 9) Repeat steps above for all exons you're interested in. You will need do this (set the genome region) for each exon you click on.
- 10) As you load exons, the right panel will populate guide candidates, with "off-target" and "On-target" scores. The important one of these is the **off-target** score because that defines how specific your guide is for your region. Higher is better or fewer off-targets. Don't use anything below 50. I try to only use those with off-target scores >70. You can order the candidate guides by clicking on the "Off-target Score" button. The On-target score should tell you how likely your region of the genome is to be modified. From my experience these numbers are meaningless, but higher is supposed to be better.

- 11) You can click each guide RNA sequence and it will highlight that region of the gene. You want ideally 4 gRNAs that are sequence independent, meaning not overlapping. Thus, look at your top scoring guides and pick the number you want that or not overlapping or nearly so.
- 12) Click the "Save" button and Export All (.tsv). This copies all the guide RNAs to your clipboard which can then be pasted into an excel file. (Just open an Excel and click Paste) (Note that the On-target score has been changed to "Efficiency Score" and that the Off-target score has been changed to "Specificity Score."
- 13) Use the "Specificity Score" in descending order and you get a rank order of the best guide RNAs for your target. I have not found a way to export the exon number with the guides so you'll want to manually check back with the Benchling site to get specific guide locations. It is a good idea to record not only the exon targeted by guides chosen but also the residue number in the protein, for later consideration and documentation.
- 14) The Broad Institute has their own algorithms for finding sequences to target for CRISPR. Go to: <https://portals.broadinstitute.org/gpp/public/analysis-tools/sgrna-design>
 Select CRISPR enzyme, should be *S. pyogenes* (default). Select your target taxon (mouse). Type in your gene name in the box, using first letter capital, rest lower case for mouse genes. Or use other identifiers as listed on the website. Set the Quota at 5 (default) to generate its top 5 picks to target (I recently switched to 10). Click on report unpicked sequences, confirm you are not a robot, and submit. When analysis is completed, click on the "sgRNA picking results". The data comes in a .txt file. I open this in my Text editor, then cut and paste it all into Excel. It is tab delimited so should go right in. There's lots of data here but focus on the last few columns, which show its first 5 choices and reasons why it excluded others. They don't give you the same visual picture that Benchling does so I tend to use both; Benchling and Broad to find good targets, and then Benchling to compare targets and locate the Broad targets in the gene/exons. Broad is more concerned about on-target scores than is Benchling. Off-target score is still the most important but it doesn't hurt to pay attention to the on-target scores as well. Off-target score is more about specificity and you want this as high as possible to lower chances of off-target effects. On-target is more about efficiency of cutting, which hasn't been an issue in my lab, hence we tend to minimize, compared to off-target. You want to use off-target scores of as high as possible, preferably above 70 but never below 50. I start by looking for guides that score well in both.

Step 2: Designing oligos and cloning guides into pX459v2 (available from Addgene and the full name is pSpCas9(BB)-2A-Puro V2)

Below is an example of four sequences identified using Benchling for the generation of indels in exon 3 of the mouse ELMOD1 gene. Benchling focuses on the off-target score, with the higher the better.

Name	Sequence	Strand	Exon	ON-Target	OFF-Target
MmELMOD1_1	TGATGCGGAAACTCACCGGA	1	3	67.9	90.8
MmELMOD1_2	ATTTGCTACGGCACCAAACC	1	3	54.8	83.7
MmELMOD1_3	CCTTTGCAGTTCACACCTTC	-1	3	38.4	66.5
MmELMOD1_4	GCTGCATGAAGTTTGTGATG	1	3	65.5	60.6

1. Order forward primers based upon the Benchling designed targets (20-mers) by adding CACCG 5' of each forward primer. The G at the end of this five nucleotide addition actually **replaces** whatever nucleotide is at the 5' end of the 20-mer sequence from Benchling (see table below if confused). The result is a 24-mer.

2. Reverse primers are determined by making the reverse complement of the forward primer (the original 20-mer, before adding the CACCG), adding AAAC at the 5' end and **replacing** the last nucleotide with a C at the 3' end, hence also a 24-mer. When annealed there should be 4 nt overhang at each end:

Forward: **CACCG**NNNNNNNNNNNNNNNNNNNNNN (Change first NT of original 20-mer guide to g)
 Reverse: **AAAC**REVERSECOMPLEMENTNN**c** (Change last NT of guide to c)

Guide	Direction	Sequence	Primer ID
-------	-----------	----------	-----------

MmELMOD1_1	Forward	CACCgGATGCGGAAACTCACCGGA	ME01
MmELMOD1_1	Reverse	AAACTCCGGTGAGTTTCCGCATCc	ME02
MmELMOD1_2	Forward	CACCgTTTGCTACGGCACCAAACC	ME03
MmELMOD1_2	Reverse	AAACGGTTTGGTGCCGTAGCAAACc	ME04
MmELMOD1_3	Forward	CACCgCTTTGCAGTTCACACCTTC	ME05
MmELMOD1_3	Reverse	AAACGAAGGTGTGAACTGCAAAGc	ME06
MmELMOD1_4	Forward	CACCgCTGCATGAAGTTTGTGATG	ME07
MmELMOD1_4	Reverse	AAACCATCACAAACTTCATGCAGc	ME08

Note: both primers are 24-mers and there are 4bp overhang on each after annealing, e.g.: 5'- CACCGGATGCGGAAACTCACCGGA -3'
3'- CCTACGCCTTTGAGTGGCCTCAAA -5'

The resulting overhangs are used for cloning into the BbsI sites of pX459v2. Note that there are two BbsI sites in pX4592, near to each other and on opposite strands, this is what yields the two different 4nt overhangs and makes cloning a snap. Primers are made up in water to a concentration of 100µM.

While doing this designing and oligo ordering, it is a good idea to go ahead and design your primers for PCR and DNA sequencing that will happen later. The PCR primers should be roughly 200-400 bp away from your cut site and you can try using the same primers in DNA sequencing. They just need to be at least 20-mers that are ~50% GC content ($T_m \geq \sim 55^\circ\text{C}$). Note that if your guides are targeting different exons or sites that are far apart you will likely have to design more than one set of primers for PCR/sequencing. Because these primers will be used over a much longer period of time it is a good idea to aliquot and freeze, to limit degradation. Primers are made up in sterile water to 100 µM.

1. Annealing primers: In separate tubes, mix 1 µl forward and reverse guide primers and add 8 µl water. In a PCR machine with temperature ramping abilities, heat these tubes to 95°C for 5 min then ramp down to 20°C, ramping down 5°C per minute. Can store at 4°C when done.

2. Digesting vector (pX459v2): In one tube per gRNA (primer pair) add the following to a total of 20 µl:
2 µl NEB Buffer 2.1
16 µl H₂O
1 µl pX459v2 (0.5 – 1 µg/µl)
1 µl BbsI (10,000 U/ml)

Note: if using HF-BbsI, use the appropriate NEB buffer (e.g., CutSmart). Incubate at 37°C for one hour, then heat inactivate at 65°C for 20 min, then store at 4°C until needed.

3. Ligation: Add to each tube of digested plasmid (20µL):
2.5 µl 10X T4 DNA ligase buffer
1.0 µl annealed gRNAs (primers)
1.5 µl T4 DNA ligase

Total should now be 25 µl . Incubate at 37°C for one hour (or colder and longer), can then heat inactivate the ligase, though not required. I would allow reaction to continue by just putting tubes at 4°C.

4. Transformation:

Use 3 µl in 50uL cells to transform competent DH5alpha and plate on LB/Amp O/N.

5. Checking for inserts with Mini-prep or Colony PCR.

This cloning typically yields vector with your insert well over 50% of the time. So I think it is easiest to pick a few colonies (no more than 1-2 needed), grow up O/N and do a mini-prep. These can be used for DNA sequencing as definitive evidence of correct insertion and sequence.

Alternatively, colony PCR can be used to check for inserts in colonies from ligated vector transformation. This is based upon the use of one primer that is specific to the insert (reverse primers

from above) and the other from the vector (U6F: GAGGGCCTATTTCCCATGATTCC). The protocol below works, though so should any your lab has used and optimized.

For a 20 μ l final volume PCR reaction to test each colony to be screened, combine in a PCR tube the following (this scales linearly to generate a master mix for multiple samples):

- 2 μ l 10X Taq buffer (we use OneTaq these days and the buffer that comes with it)
- 1 μ l 10 mM dNTPs (equal amounts of dGATC)
- 1 μ l 25 mM MgCl₂ (might want to check and see if this is required for each Taq)
- 0.5 μ l U6F primer (100 μ M)
- 0.5 μ l Reverse primer from above (100 μ M)
- 0.2 μ l Taq polymerase
- 4.8 μ l water

Total at this point is 10 μ l. Add 10 μ l of the cells in water to each PCR tube. This will bring up the mix to a total of 20 μ l. Run under PCR program: Colony PCR2. Run PCR reactions when completed in a 1.5% agarose gel. An appropriately sized product (~300 bp) means the construct has an insert because the associated reverse primer from the construct was used in PCR. Inoculate LB-Amp cultures of each construct with insert per guide for midi-prep and freezer stocks from LB-Agar plate.

I am more comfortable after confirming inserts by DNA sequencing. We then use Qiagen midi-prep kit to generate from 100 ml cultures DNA preps of 100-200 μ l of ~ 1 μ g/ μ l DNA.

Step 3: Transfection of MEFs and selection with puromycin

Transfect (Day1) WT MEFs with guide RNAs (we use 4 μ g DNA for MEFs) in one well of a 6 well dish according to standard lab protocols for Lipofectamine 2000, and the next day (day 2) re-plate into 10 cm dish. The following day (day 3), replace with fresh medium containing 3 μ g/mL puromycin in DMEM to select for transfected cells. Put on fresh puromycin medium on day 5. On day 7 remove puromycin with fresh DMEM, grow cells to confluence. Collect cells with 1 mL of trypsin/EDTA for ~ 5 min, add 9 mL fresh DMEM, triturate to resuspend all cells. Take 1 mL of this, pellet cells and freeze for later use in T7 endonuclease assay. Pellet remaining 9 mL cells, resuspend in 1 mL DMEM, freeze down two vials of 450 μ L cells plus 50 μ l DMSO for cell stocks.

Step 4: T7 Endonuclease assay to check for NHEJ

Thaw the cell pellet from previous step. These cells will be used to generate genomic DNA prep, which will then be used in PCR reaction followed by T7 endonuclease treatment (which only cuts mis-matched DNA) to look for cutting as evidence that CRISPR generated some indels in the region of interest, which is what was amplified in the PCR reaction.

Mix 2 μ L NEBuffer 2, 200 ng purified PCR product and molecular biology grade water to a total of 19 μ L in a PCR tube. Run a hybridization reaction in a PCR cycler:

- a. 95°C, 5min
- b. Ramp down to 85°C at -2°C/s
- c. Ramp down to 25°C at -0.1°C/s
- d. Hold at 4°C

Add 1 μ L T7 endonuclease (NEB; 10 U/ μ l) and incubate at 37°C for 15 minutes. Immediately run a 1.5% agarose gel on the product. Run an equivalent amount of PCR product side by side with T7 endonuclease product. T7 endonuclease cleaves mis-matched DNA. Thus, if you see a cleavage product, the guide RNA is working. The intensity of the cleavage product can serve as a crude, semi-quantitative measure of how efficient your guide is at modifying genomic DNA. If you compare four different guides, use the one that cuts the most first, etc.

Step 5: Clonal isolation of transfected cells.

Thaw out one of the two vials of cells that has been stored in DMSO. We do this in a 37°C water bath to minimize time cells are in DMSO. Put cells into 10 cm plate with 9 mL fresh medium. Next day, take up cells, count and use to seed two 96-well plates, one each at 3 and 5 cells/well (for MEFs, may

need to optimize for other cell types). Two days later start checking each well every day by visual inspection to ensure single colonies arising. If more than one, mark the well and ignore in the future. We typically get ~40 clones from these two plates. As each well becomes confluent, expand each line by moving into 6-well plate and later into a 10 cm plates (for freezing down 2 vials of cells) plus 1 or 2 wells of 6 well plate. One well will be used to generate genomic DNA for PCR and sequencing and the other can be used to screen by immunoblot if you would like (and have an antibody capable of detecting the endogenous MEF protein). We use the QIAamp DNA blood kit to make genomic DNA. Use the protocol in the appendix as it applies to cultured cells.

Step 6: DNA sequencing to confirm two frameshifting alleles of your gene

This part can be challenging but with a little practice gets easier. Also, given that you may have generated ~40 clonal lines per guide, which is far more than is needed, you might want to sequence more than the minimum number in order to find indels that are readily identified. As soon as you start generating clones and freezing them down you can start screening them. Use the frozen cell pellet described in the section above to make genomic DNA, using the Qiagen blood kit. This will serve as template in a PCR reaction. Use the PCR protocol below (40 cycles) and it is a good idea to run some of the PCR product on a gel to confirm it is a single band (though could be a doublet). Use DNA clean up and determine the concentration of DNA before sending off for sequencing. You can use either of the two primers used for PCR as primer for DNA sequencing.

The primers were designed to be ≥ 200 bp away from the CRISPR cut site so expect ≥ 200 bp of high quality DNA sequence before all of a sudden you see double peaks. This is what you want. If you see triple peaks consistently discard that clone as not a pure clone. How you deconvolve the DNA sequence data is typically empirically determined by each person. I like to find the cut site in several clones sequenced at the same time (they should be all the same for the same guide). I then look downstream in the region with double peaks in hopes of finding short homopolymeric stretches. These are useful landmarks to then build out from to reconstruct the two reads. Occasionally you will get both alleles with the same indel. This is great because it makes deciphering the sequence much easier so don't discard sequence that is all one set of nice looking peaks. BLAST that sequence against the database to look for an indel. If anyone has any good advice or ways to make this easier please share.

Below is the protocol we use for lipofectamine based MEF transfections:

Transfection of WT MEFs with Cas9 vector + guide:

1. Seed WT MEFs to be ~90% dense in wells of a 6-well dish (0.5mL from 10mL suspension from a confluent 10 cm plate should be sufficient).
2. In two separate, sterile Eppendorf tubes, add 250 μ L OptiMEM (no FBS).
3. To one of these, add 4 μ g DNA (in this case, pX459 vector with your guide inserted).
4. For MEFs, we've optimized our transfection protocol so that we have a 1:3 ratio of DNA:lipofectamine 2000. Therefore, add 12 μ L of lipofectamine 2000, which is 1 μ g/ μ L, to the second tube of OptiMEM (still no FBS). Briefly vortex and incubate at RT for 5 minutes.
5. Immediately (no more than 5 min incubation above) transfer contents of lipofectamine tube to DNA tube, pipetting up and down 1-2 times. The final volume should be around 500 μ L. Incubate for 20 minutes at room temperature.
6. Change medium on cells to 1mL OptiMEM (no FBS).
7. After 20 min incubation, using 1 mL pipette add complexes of lipofectamine+DNA dropwise to respective well, reaching final volume of 1.5mL.
8. Incubate cells for 4 hrs at 37°C.
9. Aspirate medium, rinse once with PBS, trypsinize cells, and transfer entire contents of well to a 10 cm dish. Let grow up overnight at 37°C before begin puromycin selection.

PCR Programs

Colony PCR2, for screening of inserts into pX459v2

Temperature	Time	
95 °C	5 min	1 cycle
95 °C	30 s	10 cycles
52 °C	30 s	
72 °C	45 s	
95 °C	30 s	30 cycles
56 °C	30 s	
72 °C	45 s	
72 °C	5 min	1 cycle
4 °C	Hold	

Indel PCR, for screening of changes in genomic DNA

Temperature	Time	
95 °C	10 min	1 cycle
95 °C	45 s	40 cycles
59	120 s	
72 °C	90 s	
72 °C	10 min	1 cycle
4 °C	Hold	