

JIN LAB CRISPR-Cas9 gRNA editing efficiency analysis through NGS

2023.8.21 / Boli Wu / v1

NOTES

This protocol is for indel analysis to assess gRNA targeting efficiency using wild-type Cas9.

Low-bind tubes and vortexing

Use DNA low-bind tubes (e.g. Eppendorf™ 022431021) to ensure maximal recovery of your library. Remember to vortex throughout your experiment process (after making a mix, before starting the reaction, before making an aliquot, etc.). These are necessary for a successful sequencing.

METHODS

Prepare materials and reagents:

gRNA cloning

Name	Catlog #	Notes
gRNA cloning oligos	customized	See "Oligo design and ordering" for details
Quick Ligation Reaction Buffer (5x)	NEB B6058S	
Recombinant Albumin	NEB B9200S	Molecular Biology Grade, 20mg/mL
T7 ligase	Enzymatics L6020L	
FastDigest Esp3I	Thermo Scientific FD0454	Isoschizomers of BsmBI
T4 Polynucleotide Kinase (PNK)	NEB M0201S	Come with 10x ligation buffer

Cell culture and genome DNA extraction

Name	Catlog #
Fetal Bovine Serum (FBS)	Corning™ 35011CV
DMEM	Gibco™ 11965118
Opti-MEM	Gibco™ 31985088

Penicillin-Streptomycin (PS)	Gibco™ 15140122
Polyethylenimine Hydrochloride	24765-1
QuickExtract DNA Extraction Solution (QE)	QE09050

NGS library preparation

Name	Catlog #
NGS library amplification primers	See “Oligo design and ordering” for details
NEBNext High-Fidelity 2X PCR Master Mix	NEB M0541S
KAPA Library Quantification Kits	KK4824
E-Gel Agarose Gels, 1%	A42100

gRNA design

This step is to use the gRNA design wizard on benchling.com to design gRNA.

1. Log in to benchling. Click on the “+” button on the left panel, then choose “CRISPR” and “CRISPR guides”.
2. **Import sequence:** search for the genome sequence you want to target by gene names or chromosome coordinates.
 - *Pay attention to choosing the correct genome version from GRC and the correct species. Choose the functional protein-coding transcripts.*
 - *Otherwise, you can also customize the targeted sequence by uploading files/raw sequences.*

Use the gene *Scn2a* as an example:

Design CRISPR Guides: Import sequence ×

Import From

Gene ID or Name ▼

scn2a × [Example](#)

Genome ▼
GRCm39_EnsemblRelease_104 (Mus mu)
Don't see the genome you're looking for? We may be able to import it. [Just ask.](#)

Gene ▼
Scn2a (ENSMUSG00000075318)

Location ▼
chr2 65,451,115—65,597,791 (+)

Transcript ▼
Scn2a-201 (ENSMUSG00000075318_EN)

Sequence Name ▼
Scn2a (ENSMUSG00000075318)

3. **Adjust parameters:** choose single gRNA, 20bp guide length, and the same genome reference as in 2.

In the advanced settings, to be consistent with the on-target score algorithm, the CFD off-target score from Doench, Fusi et al. 2016 is recommended.

CRISPR Analysis Settings ×

Design Type

Single guide
Wild-type Cas9, single gRNA (higher efficiency)

Paired guides
Double Cas9 nickase, two gRNAs (lower off-target effects)

Guides for "base editing" (Komor et al., 2016)
C -> T (or G -> A) substitution, no dsDNA breaks

Guide Length

20

Only guide lengths of 20bp are supported for the CFD off-target score.

Genome

GRCm39_EnsemblRelease_104 (Mus musculus) ▾

Don't see the genome you're looking for? We may be able to import it, [just ask](#).

PAM

NGG (SpCas9, 3' side) ▾

Only NGG PAMs can be used for the CFD off-target score.

[Hide Advanced Settings](#)

Masked Regions

Include masked regions in off-target search ?

Off-Target Score

Use the scoring method from [Hsu et al., 2013](#)

Use the Cutting Frequency Determination (CFD) score from [Doench, Fusi et al. \(2016\)](#). Only available for NGG PAM sites and 20mer guides.

On-Target Score

Use Optimized Score from Doench, Fusi et al., 2016. Only available for SpCas9. ?

4. Click on the "Finish" button to enter the gRNA design interface.
5. In the "sequence map" tab, select the coding exons to be targeted, ideally the first few exons in the transcripts to ensure complete loss-of-function editing.
6. With the ideal exon selected, go to the "design CRISPR" tab and add it to the targeted regions by clicking on the "+" button. Multiple exons can be added.
7. The gRNA candidates will be automatically generated after adding the targeted regions. Here are some criteria we are using for choosing good gRNA candidates:

- On-target score ≥ 60
- Off-target score ≥ 20 (double check the off-targeting by clicking on the score number, make sure that no protein-coding genes are listed in the top-10 hits)
- If picking multiple gRNAs, make sure that ≥ 50 bp spacings between every 2 gRNAs, ideally ≥ 100 bp.

Oligo design and ordering

Order gRNA cloning oligos:

These oligos will be used to clone gRNA into gRNA backbones with BsmBI cutting sites.

1. Design oligos of the original sequences of the gRNA candidates, as well as oligos of the reverse complementary sequences of the gRNA candidates.
The 2 oligos for a given gRNA candidate should be able to anneal with each other in order to perform the following Goldengate Annealing procedure.
2. Add “caccg” at the 5’ of forward oligo, “aaac” at the 5’ of reverse oligo, and “c” at the 3’ of reverse oligo.
The modifications serve as adaptors to insert gRNAs into the BsmBI site. If using different restriction enzymes, the adaptors could be different.

Order NGS library amplification oligos:

These oligos will be used to amplify gRNA targeting sites to check for potential insertion/deletion (indel analysis) using NGS. To construct a library for Illumina NGS, adaptor sequences need to be added in the amplification primers to build specific sequencing structures. Here we are using 2 rounds of PCR to build the Truseq library structure.

1. Locate the gRNA targeting site in the genome. Design primers (forward and reverse) that can amplify fragments containing this site. Use a primer design wizard (e.g. Primer3Plus) to ensure the primers are of good quality. The length of the amplicons is ideal to be 100-300bp.
2. Add the following adaptors to the designed primers in 2. This pair of primers will be used in the 1st round of NGS PCR.
 - Add CTTTCCCTACACGACGCTCTTCCGATCT to 5’ site of the forward primer
 - Add GACTGGAGTTCAGACGTGTGCTCTTCCGATCT to 5’ site of the reverse primer*These adaptor sequences are part of Read1 and Read2 sequences of the Truseq library structure from Illumina.*

Optional: check the amplification efficiency of these primers by performing a PCR using these primers to amplify genomic DNA sample (from a wildtype cell lysate is fine, see section **Test**

gRNA targeting efficiency *in vitro* for details). Run a gel to check if you can get a bright band with the expected length.

- Design primers for the 2nd round of PCR to complete the Truseq library structures. These primers should contain the rest part of Truseq **Read1** and **Read2**, **i5/i7 index**, and **P5/P7**. These primers can be used repeatedly for different NGS library prep. For example:

- Forward:

AATGATACGGCGACCACCGAGATCTACACAAGTAGAGACTCTTTCCCTACACGACGCTC
TTCC

- Reverse:

CAAGCAGAAGACGGCATACGAGATCATGATCGGTGACTGGAGTTCAGACGTGTGCTCTTC

Critical: change i5/i7 index to design more 2nd round primers for amplification of multiple samples. Avoid using primer pairs with the same indexes to amplify different samples in one round of sequencing.

gRNA cloning

This step is to clone gRNA into your desired vectors with BsmBI cutting sites using Golden Gate Assembly.

- Anneal gRNA cloning oligos.
 - Prepare the reaction mix. For each reaction, make a 10uL mix:

Reagents	Volume (uL)
Oligo 1 (100uM)	1
Oligo 2 (100uM)	1
T4 PNK	0.5
10x ligation buffer	1
ddH2O	6.5

- Briefly vortex and then spin down. Incubate in a thermocycler:

Temperature (°C)	Time (min)
37	30 min
95	5 min
Drop to 4	1°C / 5 s
4	hold

- c. The oligos are now annealed as dsDNA. Take 1 uL of the annealing products and mix with 19uL H₂O to make a 1:20 dilution.
2. Insert gRNA into vectors by Golden Gate Assembly.
 - a. Prepare the reaction mix. For each reaction, prepare a 15uL mix:

Reagents	Volume (uL)
Diluted dsDNA (1:20)	1.2
Vector plasmid	Add 25 ng
Quick Ligation Reaction Buffer (5x)	3
Recombinant Albumin	0.08
T7 ligase	0.08
Esp3I	0.3
ddH ₂ O	Fill up to 15 uL

- b. Briefly vortex and then spin down. Incubate in a thermocycler:

Step	Temperature (°C)	Time
1	37	5 min
2	20	5 min
Repeat step1 and 2 for 15 cycles		
3	4	hold

The cycle number can be reduced to ~10.

- c. The plasmid vector with gRNA inserts are now ready to use for transformation. Transform the competent cells and get the plasmid ready.

Test gRNA targeting efficiency *in vitro*

This step is to perform gRNA editing in vitro and prepare genome DNA to check indels.

1. Prepare plasmids expressing gRNA and Cas9 (ideally with different reporters).
2. Choose the correct cell line according to your gRNA targets. For example, use HEK293 for gRNA targeting human genome, N2A/HT22 for gRNA targeting mouse genome.
3. Prepare medium and solutions:

- **Cell growth medium:** made by adding 50 mL FBS and 5 mL PS to a bottle of DMEM medium (500 mL), mix well, filter and store at 4 °C.
 - **PEI transfection reagent:** dissolve the powder in water, adjust the pH to 7.0 using HCl, adjust the volume to make a final concentration of 1 mg/mL. Aliquot into 1 mL and store at -20 °C.
4. Plate the cells in a 96-well plate on Day 0. Adjust the seeding density according to the growing speed of your cell line to get a desired density for transfection on Day 1 (after overnight growth). For example:
- Using HEK293 cell line with a doubling time of 14 h, seed ~30 k cells in each 96-well plate to get ~70% confluency after 18 h growth.
 - Using HT22 cell line with a doubling time of 12 h, seed ~15 k cells in each 96-well plate to get ~70% confluency after 18 h growth.
5. Transfection.
- a. Change cell growth medium before transfection.
***Critical:** using fresh medium ensures the culture stays healthy in the following days waiting for Cas9 and gRNA to express.*
 - b. Take 2 96-well PCR plates (or a few 8-strip PCR tubes according to your reaction number). For each well of transfection, mix 200 ng of Cas9 plasmid, 200 ng of gRNA plasmid, and appropriate amount of Opti-MEM to make a 10 uL mix. In the other 96-well PCR plate, for each well of transfection, mix 1.2 uL PEI transfection reagent and appropriate amount of Opti-MEM to make a 10 uL mix.
***Critical:** scale up accordingly if using a different experimental format. Test the conditions before running a big experiment.*
 - c. Use a multichannel pipette to mix the plasmid solution and PEI solution.
 - d. Seal the plate using a film. Vortex thoroughly for 30 s. Briefly spin down.
***Critical:** this step ensures the complete mixing of plasmid and PEI.*
 - e. Incubate at RT for 15 min.
***Critical:** this step allows PEI and plasmid to form complexes.*
 - f. Add the transfection reagent to the cell growth medium.
 - g. Wait for ≥ 2 days before harvesting the cells to extract genome DNA to ensure the Cas9 and gRNA to express and be effective.
 - h. Remove the cell growth medium, wash with clean PBS, add 50 uL of QE for each well. Use pipette tips to scrape the cells from the well plate into the QE medium. (Use a multichannel pipette if you have multiple conditions)

Critical: check under the microscope to make sure most cells have fallen off from the well bottom.

- i. Transfer the QE to a 96-well PCR plate (or 8-strip PCR tube). Briefly spin down. Set up a thermocycler and run:

Temperature (°C)	Time
65	15 min
68	15 min
98	10 min
4	hold

- j. The genomic DNA in the 96-well PCR plate is now ready to use. Store at 4 °C for up to 1 week. - 20 °C for long-term storage.

Prepare NGS sequencing library

This step is to use 2 rounds of PCR to prepare a library for sequencing, as well as the following purification and quantification steps.

6. Prepare the reaction mix for the 1st round PCR. For each reaction make a 20uL reaction mix:

Reagents	Volume (uL)
NEBNext High-Fidelity 2X PCR Master Mix	10
Primers for 1st round NGS PCR (10uM)	1 for each
Genomic DNA	2
H2O	6

7. Incubate in a thermocycler:

Step	Temperature (°C)	Time
1	98	30 s
2	98	10 s
3	60 (adjust according to your primers)	15 s
4	72	20 s

Repeat step 2-4 for 15 cycles		
5	72	2 min
6	4	hold

8. Prepare the reaction mix for the 2nd round PCR. For each reaction make a 20uL reaction mix:

Reagents	Volume (uL)
NEBNext High-Fidelity 2X PCR Master Mix	10
Primers for 2nd round NGS PCR (10uM)	1 for each
PCR products from 1st round NGS PCR	2
H2O	6

9. Incubate in a thermocycler:

Step	Temperature (°C)	Time
1	98	30 s
2	98	10 s
3	60	15 s
4	72	20 s
Repeat step 2-4 for 20 cycles		
5	72	2 min
6	4	hold

Ideally the total cycle number of 1st and 2nd round NGS PCR should not exceed 50 cycles to avoid overamplification and bias.

10. Check the PCR product quality by running an E-Gel. Ideally there will be a single band with the expected length.

If only smear is visible on the gel, double-check the primer design and PCR setting.

11. If you are testing multiple gRNAs, pool PCR products of different gRNA amplicons together. Take a photo of your gel from 10. Use the brightness of the band of each gRNA amplicon as an indication of its relative concentration. According to this estimated concentration, add different volumes of PCR

products of different gRNA amplicons in the pooled library to make sure all gRNAs are represented equally.

- Run the pooled library on an E-Gel, cut the expected band, and perform a gel extraction to purify the library. Take a low-bind tube to collect the library.
- Measure the concentration of the final (pooled) library by nanodrop. Calculate the mole concentration using <https://nebiocalculator.neb.com/#!/dsdnaamt>.
- Take a few low-bind tubes. Make a diluted aliquot of the library at the concentration of 10-15 nM (use ddH₂O to dilute, also save the high concentration version). Further dilute it to 10-15 pM for qPCR quantification by mixing 1 uL of 15 nM sample with 999 uL H₂O.

Critical: vortex thoroughly before making the dilution. The 10-15 nM dilution will be used for final sequencing analysis so make sure it is mixed well.

- Quantify the library concentration using a KAPA qPCR quantification kit. Take a qPCR plate (Bio-rad 96-well PCR plates). Calculate the total number of reactions (including samples and standards). Prepare the master mix according to the reaction number. Make sure samples/standards were well-vortexed before adding to mastermix. For each reaction make a 15uL mix:

Reagents	Volume (uL)
2X KAPA SYBR FAST qPCR Master Mix	7.5
10X Illumina Primer Premix	0.5
H ₂ O	6
15pM sample/standards	1

Usually using standard 1-3 (20 - 0.2 pM) should be enough to cover the concentration range of your samples. Do not use less than 3 standards as you will not be able to plot a standard curve.

- Seal the plate using a seal film (e.g. Bio-rad Microseal). Vortex thoroughly and briefly spin down.
- Incubate the plate in a Bio-rad qPCR machine using the following protocol. Select SYBR or FAM as your readout channel:

Step	Temperature (°C)	Time
1	95	5 min
2	95	30 s
3	60	45 s
Repeat step 2-3 for 30-35 cycles		

18. Quantify the library based on the Cq number.

- a. Input Cq values in an Excel sheet for all standards and samples.
- b. Select all the standard Cq values to create a scatterplot with Cq values on x-axis and log concentration on y-axis. (standard1: 20pM, standard2: 2pM, standard3: 0.2pM, etc.)
- c. Right click on a plot point, select "add trendline" and "display R² value and equation on plot". Do not select "set intercept".
If a standard point is very far off from the rest, leave it out.
- d. Use the equation to calculate the log concentration for your samples. The real concentration of your sample will be 10^{logconc} in pM.

Sequencing by iSeq kit

This step is to make the final sequencing library and start the sequencing. Use iSeq as an example.

Prepare for iSeq: the iSeq usually takes 17.5 h. The cartridge of iSeq needs to be thawed in a RT water bath at least 6 hours prior to use. We recommend to run an overnight sequencing and thaw the cartridge earlier. The flow cell of iSeq needs to be warmed up at RT 15 min prior to the sequencing run. See [this](#) for more information.

1. Take the cartridge to RT to thaw 5-6 h prior to use.
2. Remove Resuspension Buffer (RSB) from -20 °C storage. Alternatively, use 10 mM Tris-HCl, pH 8.5 or Qiagen EB buffer in place of RSB.
Optional: remove 10 nM PhiX stock from -20 °C storage. PhiX is needed only for an optional spike-in if the sequencing library is not heterogeneous enough. See [PhiX Illumina](#) for more information.
3. According to your qPCR quantification results, re-determine the concentration of your 10-15 nM library aliquot.
 - a. Take a low-bind tube, use the 10-15 nM aliquot to prepare a 1 nM library. Vortex and briefly spin down. The 1 nM library can be stored at -20 °C.
Critical: use the exact aliquot you used for qPCR quantification to prepare this library to make sure the concentration is right.
 - b. Take a low-bind tube, take 8 µl of this 1nM library pool to 92 µl RSB buffer to make 100 µl of an 80 pM pool. Vortex and briefly spin down.
Critical: the optimal loading concentration varies and needs to be determined empirically for different samples. See [this](#) for more information.
4. Open the package of the cartridge, invert the reagent 5 times to mix the reagents.

5. Tap the cartridge (label facing up) on the bench or other hard surface 5 times to ensure reagent aspiration.
6. Using a new pipette tip, pierce the library reservoir and push the foil to the edges to enlarge the hole. Discard the pipette tip to prevent contamination.
7. Add 20 μ l 100pM library to the bottom of the reservoir. Avoid touching the foil.
Optional: *If you need PhiX control, this is the place to mix your library with the PhiX.*
8. Take the flow cell from 4 °C to RT 15 min prior to use. Open the white foil flow cell package. Use within 24 hours of opening. Pull the flow cell out of the package.
Critical: *touch **only the plastic** when handling the flow cell. Avoid touching the electrical interface, CMOS sensor, glass, and gaskets on either side of the glass.*
9. Hold the flow cell by the grip points with the label facing up. Insert the flow cell into the slot on the front of the cartridge.
10. Load the cartridge (with 100pM library and flow cell) and follow the interface to start the run.

Indel analysis

1. Use the Illumina [Basespace Sequence Hub](#) to monitor the run, store, and transfer data.
2. Create an account for Illumina Basespace Sequence Hub, make sure you have an account and log in on the sequencer. The data will be linked to your account after the sequencing is finished.
If you are using a shared sequencer and not able to log in your own account, make sure to ask for data sharing after the sequencing is finished.
3. Download the data.
4. Do demultiplexing to sort the raw sequencing reads into .fastq files for different libraries by their distinct indexes. Here we introduce how to use our customized codes based on SLURM to do this as an example:
 - a. Prepare a sample sheet using the template [SampleSheet template](#). Change content in red according to your own design. Remember to save this file as 'SampleSheet.csv' and put it in the same folder as your data.
 - b. Run the following codes by using `sbatch` command.

```
#!/bin/bash
#SBATCH --mem=12gb
#SBATCH --nodes=1
#SBATCH --cpus-per-task=8
#SBATCH --time=1:00:00
#SBATCH --output=myoutput.out
#SBATCH --error=myerror.err
```

Configured to work with Slurm

```
mkdir fastq_files
```

```
module load bcl2fastq
```

```
bcl2fastq -p 12 --barcode-mismatches 1 --output-dir ./fastq_files_10x --no-lane-splitting
```

c. The .fastq files will be generated in a subfolder named 'fastq_files' under your current folder.

5. Align the sequence in the .fastq files to the original sequence (from your reference genome) to identify potential insertions/deletions (indels). Here we introduce how to use our customized MATLAB codes (on [Lab Github](#) page) to do this as an example:

a. Have all .fastq files ready. Make sure all .fastq files are unzipped and under the same folder (referred to as 'working folder' below).

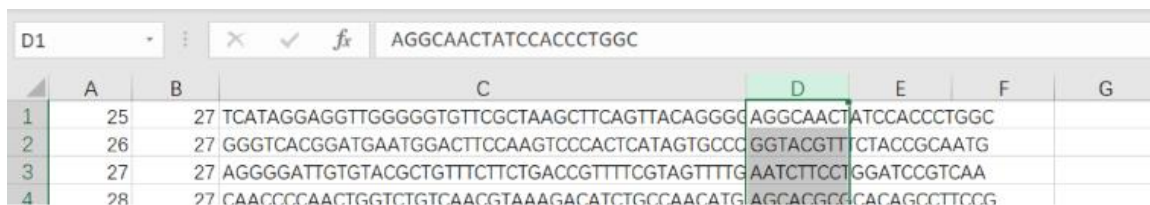
b. Prepare a .xlsx spreadsheet listing the information of all your samples. Each row describes one sample using 4 columns of data:

- i. the order number of i5 index
- ii. the order number of i7 index,

The correspondence between order number and index sequences were determined priorly, see [NGS2_plate \(32F 32R\).xlsx](#)

- iii. the amplicon sequence (excluding the common handles for NGS, only for reference genome)
- iv. the gRNA sequence.

Column or row title should not be included. See example below:



	A	B	C	D	E	F	G
1	25	27	TCATAGGAGGTTGGGGGTGTTTCGCTAAGCTTCAGTTACAGGGG	AGGCAACTATCCACCCTGGC			
2	26	27	GGGTACGGATGAATGGACTTCCAAGTCCCACTCATAGTGCCG	GGTACGTTCTACCGCAATG			
3	27	27	AGGGGATTGTACGCTGTTTCTTCTGACCGTTTTTCGTAGTTTTG	AATCTTCTGGATCCGCAA			
4	28	27	CAACCCCAACTGGTCTGTCAACGTAAAGACATCTGCCAACATG	AGCACGGCCACAGCCTTCG			

Save the spreadsheet as "processlist.xlsx". Put it under your working folder.

c. In MATLAB, set the working directory to your working folder.

d. Run the main.m. The program will give outputs of the analysis status for each sample (each row in the "processlist.xlsx").

If errors occur, it is likely your input format is wrong.

e. After the program is finished (usually in a few minutes), you will have a "_indelcounts.xlsx" file in your working folder with the following measurements.

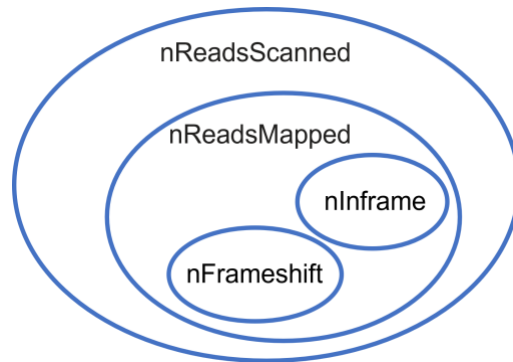
nInframe: the number of reads that have in-frame indels

nFrameshift: the number of reads that have frameshift indels

nReadsMapped: the number of reads that have the given barcode and successfully mapped to your amplicon

indel pct: the percentage of reads with indels that mapped to your amplicon.

nReadsScanned: the number of reads that have the given barcode.



- f. Add another column to calculate the percentage of frameshift indels only ($\frac{nFrameshift}{nReadsMapped}$). This percentage generally represents your gRNA targeting efficiency of silencing the gene.
6. Another way to quickly check a few .fastq files at a time is to use [Outknocker](#).
 - a. In step 1, specify the species, experiment type (select “monoclonal” for indel analysis), nuclease type, gene name, and gRNA sequence as needed. Click on “ANALYZE DATA (proceed to step 2)”
 - b. In step 2, click on “Select FASTQ files” to upload the .fastq files corresponding to the gRNA seq (can do multiple at a time). Click on “START ANALYSIS (proceed to step 3)”
 - c. In step 3, the website will generate a summary of the indels (percentages and locations) found in your .fastq files.