Apr 20, 2021

High throughput quantification of CRISPR gRNA efficiency based on surrogate lentivirus libraries

Xi Xiang[1], Yonglun Luo[1]

[1]Aarhus University

1 Works for me dx.doi.org/10.17504/protocols.io.bt9jnr4n

Yonglun Luo
Aarhus University

ABSTRACT

This protocol is developed for massively capture CRISPR gRNA efficiency in cells. The protocol is based on lentivirus vectors which carry a CRISPR gRNA expression cassette and a surrogate target site. The surrogate target site is used to captured indels introduced by CRISPR, which are quantified by next generation sequencing. A lentivirus plasmid vector is generated to streamline Golden-Gate Assembly based plasmid library generation and quantification of lentivirus titer.

DOI

dx.doi.org/10.17504/protocols.io.bt9jnr4n

PROTOCOL CITATION

Xi Xiang, Yonglun Luo 2021. High throughput quantification of CRISPR gRNA efficiency based on surrogate lentivirus libraries. **protocols.io** https://dx.doi.org/10.17504/protocols.io.bt9jnr4n

KEYWORDS

CRISPR, gRNA efficiency, lentivirus library, surrogate vector

CREATED

Apr 15, 2021

LAST MODIFIED

Apr 20, 2021

PROTOCOL INTEGER ID

49163

Design of surrogate oligonucleotides

1 Each oligo (170 bp) consists of the BsmBI recognition site “cgtctc” with 4 bp specific nucleotides “acca” upstream, following the GGA cloning linker “aCACC”, one bp “g” for initiating transcription from U6 promoter, 20 bp gRNA sequences of “gN20”, 82bp gRNA scaffold sequence, 37bp surrogate target sequences (10bp upstream sequences, 20 bp protospacer and 3 bp PAM sequences, 4 bp downstream sequence), the downstream linker “GTTTg” and another BsmBI binding site and its downstream flanking sequences “acgg”.

An example of the surrogate oligo is showed below:

5'/accacgtctc**aCACC**gCTGCAGCTGCAGGCAGCTCCGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCTTTTTTAGAAGTGGAGCTGCAGCTGCAGGCAGCTCCCGGATCC**GTTTg**gagacgacgg/3'

Amplification of surrogate oligonucleotides from the oligo pool

2 This step is to amplify the synthetic single stranded surrogate oligonucleotides into double stranded DNA.

Primers used for PCR are:

Oligo (BsmBI GGA)-F	TACAGCTaccacgtctcaCACC
Oligo (BsmBI GGA)-R	AGCACAAccgtcgtctccAAAC

Surrogate PCR1, the amplicon size is 184bp

The PCR reaction was carried out using PrimeSTAR HS DNA Polymerase (Takara, Japan) following the manufacturer's instruction. Briefly, each PCR reaction contained 1 µl oligo template, 0.2 µl PrimeSTAR polymerase, 1.6 µl dNTP mixture, 4 µl PrimeSTAR buffer,1 µl forward primer (10 uM)and 1 µl reverse primer (10 uM) and ddH2O to a final volume of 20 µl.

The thermocycle program was one cycle at 98°C for 2min, then 21cycles [98°C for 10s, 55°C for 10s, and 72°C for 30s), followed by one cycle at 72°C for 7min, and lastly hold at 4°C.

To avoid amplification bias of oligos introduced by PCR, we conducted gradient thermocycles and performed PCR products gray-intensity analysis to determine the optimal PCR cycles of 21. The best thermocycles should be in the middle of an amplification curve. In this study, the PCR cycle as 21 for oligos amplification. The final PCR product length was 184 bp. We performed 72parallel PCR reactions for 12K oligos amplification, then these PCR products were pooled, and gel purified by 2% agarose gel.

3 The purified PCR product can be saved at -20 °C, or used directly for cloning into the Lentiviral surrogate reporter vector: LentiU6-LacZ-GFP-Puro (BB)

Addgene ID: 170459

Golden Gate Assembly (GGA)

4 This step is to clone the PCR product into the LentiU6-LacZ-GFP-Puro plasmid by Golden Gate Assembly.

For each GGA reaction, the reaction mixture contained:

100 ng lentiviral backbone vector, LentiU6-LacZ-GFP-Puro,
10 ng purified 12K oligos-PCR products,
1 µlT4 ligase (NEB),
2 µlT4 ligase buffer (NEB),
1 µlBsmBI restriction enzyme (ThermoFisher Scientific, FastDigestion)
Add double distilled water ddH2O to a final volume of 20 µl.

The GGA reactions were performed:
1 cycle at 37°C for 5 min,
10 cycles at 22°C for 10 min
1 cycle at 37°C for 30 min
1 cycle at 75°C for 15 min.

For the 12K oligo pool library, prepare 36 parallel GGA reactions and the ligation products were pooled into one tube.

Transformation was then carried out using chemically competent DH5a cells following standard protocols. For each reaction, 10µlGGA ligation product was transformed in to 50 µlcompetent cells and all the transformed cells were spread on one LB plate (15 cm dish in diameter) with Xgal, IPTG and Ampicillin.

Citation: Xi Xiang, Yonglun Luo (04/20/2021). High throughput quantification of CRISPR gRNA efficiency based on surrogate lentivirus libraries. https://dx.doi.org/10.17504/protocols.io.bt9jnr4n

The ligation efficiency will be determined by first the presence of very few blue colonies. Second, to ensure that there is sufficient coverage of each gRNA of the 12K library, 42 parallel transformations were performed, and all the bacterial colonies were scraped off and pooled together for plasmids midi-prep.

For NGS-based quality quantification of the library coverage, midi-prep plasmids were used as DNA templates for surrogate PCR2, followed by gel purification and NGS sequencing.

primers for surrogate PCR2:

NGS-F1	GGACTATCATATGCTTACCGTA
NGS-R1	ACTCCTTTCAAGACCTAGCTAG

Surrogate PCR2, amplicon size is 252bp.

PCR conditions:
PCR reaction mixture:
2 ng plasmid DNA,
0.5 µg PrimeSTAR polymerase,
4 µ ldNTP mixture,
10 µl PrimeSTAR buffer,
2.5 µl NGS-F1 primer (10 uM)
2.5 µl NGS-R1 primer (10 uM)
Add ddH2O to a final volume of 50 µl.

The thermocycle program:
1 cycle at 98°C for 2min,
25 cycles at 98°C for 10s, 55°C for 10s, 72°C for 30s,
1 cycle at 72°C for 7min
hold at 4°C.

PCR product are gel purified at save at -20 °C or proceeded with deep sequencing.

Lentivirus packaging

5 Lentiviral packaging can be carried out using a standard protocol.

In our method, HEK293T cells were used for lentivirus packaging. All cells were cultured in Dulbecco's modified Eagle's medium (DMEM) (LONZA) supplemented with 10 % fetal bovine serum (FBS) (Gibco), 1% GlutaMAX (Gibco), and penicillin/streptomycin (100 units penicillin and 0.1 mg streptomycin/mL) in a 37°C incubator with 5% CO2 atmosphere and maximum humidity. Cells were passaged every 2-3 days when the confluence was approximately 80-90%.

For lentivirus packaging: (Day 1) Wild-type HEK293T cells were seeded to a 10 cm culture dish, $4x10^6$ cells per dish (10 dishes in total); (Day 2) Transfection. Briefly, we refreshed the medium with 7 mL fresh culture medium to 1 hour before transfection (gently, as the HEK293T cells are easy to be detached from the bottom of dish); Next, we performed transfection with the PEI 40000 transfection method. For 10 cm dish transfection, the DNA/PEI mixture contains 13 µg lentiviral 12K plasmid DNA, 3 µg pRSV-REV, 3.75 µg pMD.2G, 13 µg pMDGP-Lg/p-RRE, 100 µgPEI 40000 solution (1 µg/µl in sterilized ddH2O) and supplemented by serum-free opti-MEM without phenol red (Invitrogen) to a final volume of 1 mL. The transfection mixture was pipetted up and down several times gently, then kept at room temperature (RT) for 20 min, then added into cells in a dropwise manner and mix by swirling gently. (Day 3) Changed to fresh medium; (Day 4) Harvest and filter all the culture medium of the 10 cm dish through a 0.45 µmfilter, pool the filtered media into one bottle. Each 10 cm dish generated approximately 7~8 mL lentivirus crude. Add polybrene solution (Sigma-Aldrich) into the crude virus to a final concentration of 8 µg/mL.Aliquotthe crude virus into 15 mL tubes (5 mL/tube) and store in -80 °C freezer.

Quantification of lentivirus titer by FCM analysis

6 This section is for quantification of transduction titer of the crude lentivirus.

As the 12K lentiviral vector expresses an EGFP gene, the functional titer of our lentivirus prep was assayed by flow cytometry (FCM).

Citation: Xi Xiang, Yonglun Luo (04/20/2021). High throughput quantification of CRISPR gRNA efficiency based on surrogate lentivirus libraries. https://dx.doi.org/10.17504/protocols.io.bt9jnr4n

1) split and seed HEK293T cells to 24-well plate on day 1, 5 x10^4cells per well. Generally, 18 wells were used to perform the titter detection, a gradient volume of the crude lentivirus was added into the cells and each volume was tested by replicates. In this experiment, the crude virus gradients were 10 µl, 20 µl, 40 µl, 80 µl, and 160 µlfor each well (**Fig. S5**). Another 2 wells of cells were used for cell counting before transduction;

2) Conduct lentivirus transduction when cells reach up to 60~80% confluence on day 2. Before transduction, detach the last two wells of cells using 0.05% EDTA-Trypsin to determine the total number of cells in one well ($N_{initial}$). Then change the remaining wells with fresh culture medium containing 8µg/mL polybrene, then add the gradient volume of crude virus into each well and swirling gently to mix;

3) On day 3, change to fresh medium without polybrene;

4) On day 4, harvest all the cells and wash them twice in PBS. Fix the cells in 4% formalin solution at RT for 20 min, then spin down the cell pellet at 2,000 rpm for 5 min. Discard the supernatant and re-suspend the cell pellet carefully in 600 µlPBS, and conduct FCM analysis immediately. FCM was performed using a BD LSRFortessa^{TM}cell analyzer with at least 30,000 events collected for each sample in replicates.

The FCM output data was analyzed by the software Flowjo vX.0.7. Percentage of GFP-positive cells was calculated as:

$$\mathbf{Y\% = N_{GFP\text{-}positive\ cells} / N_{total\ cells} x100\ \%.}$$

Calculate the GFP percentage of all samples. For accurate titter determination, there should be a linear relationship between the GFP positive percentages and crude volume. The titter (Transducing Units (TU/mL) calculation according to this formula:

$$\mathbf{TU/mL = (N_{initial} x Y\% \times 1000)/V.}$$

V represents the crude volume (µl) used for initial transduction.

Lentivirus transduction

7 HEK293T-SpCas9 cells were cultured in growth medium with 50 µg/ml hygromycin throughout the whole experiment.

For 12K lentivirus library transduction:
Day -1: 2.5 x10^6cells per 10 cm dish were seeded (in 12 dishes). For each group, one dish was used for cell number determination before transduction and one dish for drug-resistance (puromycin) test control and the remaining 10 dishes were used for the 12K lentivirus library transduction (transduction coverage per gRNA exceeds 4000X);

2) Day 0: We first determined the approximate cell number per dish. This was used to determine the volume of crude lentivirus used for transduction using a multiplicity of infection (MOI) of 0.3. The low MOI (0.3) ensures that most infected cells receive only 1 copy of the lentivirus construct with high probability [41].

The calculation formula is: **V = N x0.3 / TU**. V = volume of crude lentivirus used for infection (ml); N = cell number in the dish before infection; TU = the titter of crude lentivirus (IFU/mL).The infected cells were cultured in a 37°Cincubator;

3) Day 1:24 hours after transduction, split the transduced cells of each dish to 3 dishes equally;

4) Day 2: For the 3 dishes of split (30 dishes in total, 3 divided into sub-groups), sub-group1 (10 dishes) were harvested and labeled as the Day 2 after the 12K lentivirus library transduction. All cells from this sub-group werepooled into one tube and stored in -20°Cfreezer for genomic DNA extraction; The sub-group 2 (10 dishes) was changed to fresh D10 medium contains 50 µg/mlhygromycin + 1 µg/mLpuromycin (Dox-free group); The sub-group 3 (10 dishes) was changed to D10 medium contains 50µg/mlhygromycin + 1 µg/mL puromycin + 1 µg/mL doxycycline (Dox-addition group).

5) The transduced cells were spitted every 2~3 days when cell confluence reaches up to 90%. Cells from Day 2, 8 and 10 were harvested and stored in -20 °C for further genomic DNA extraction. Parallel experiments were performed using wildtype HEK293T cells.

PCR application of surrogate sites from cells

8 Genomic DNA was extracted using the phenol-chloroform method.

Citation: Xi Xiang, Yonglun Luo (04/20/2021). High throughput quantification of CRISPR gRNA efficiency based on surrogate lentivirus libraries. https://dx.doi.org/10.17504/protocols.io.bt9jnr4n

The genomic DNA were digested with RNase A(OMEGA) to remove RNA contamination (In this study, 50 μg RNase A worked well to digest the RNA contamination in 100–200 μg genomic DNA after incubating in 37 °C for 30 min).

Then the genomic DNA was purified and subjected to surrogate PCR2 (Table S1). In this study, 5 ug genomic DNA was used as temperate in one PCR reaction which contained approximately 7.6x10^5copies of surrogate construct (assuming 1 x10^6cells contain 6.6 μg genomic DNA), which covered about 63 times coverage of the 12K library. In total, 32 parallel PCR reactions were performed to achieve approximately 2,016 times coverage of each gRNA and surrogate site.

For each PCR reaction, briefly, 50 μl PCR reaction system consists of 5μg genomic DNA, 0.5 μg PrimeSTAR polymerase, 4 μl dNTP mixture, 10 μl PrimeSTAR buffer, 2.5 μl forward primer (10 uM) and 2.5 μl reverse primer (10 uM) and supplemented with ddH2O to a final volume of 50 μl.

The thermocycle program was
1 cycle at 98°C for 2 min
25 cycles at 98°C for 10s, 55°C for 10s, 72°C for 30s
1 cycle at 72°C for 7min and held at 4°C hold.

Then purify all the PCR products by 2% gel, pool the products together and conduct deep amplicon sequencing.

NGS-F1	GGACTATCATATGCTTACCGTA
NGS-R1	ACTCCTTTCAAGACCTAGCTAG

Surrogate PCR2

NGS sequencing

9 NGS sequencing can be performed with NGS facilities. PE150 should be used for targeted sequencing.

In this study, MGISEQ-500 (MGI of BGI in China) was used to perform the amplicons deep sequencing following the standard operation protocol. First, PCR-free library was prepared using MGIeasy FS PCR-free DNA library Prep kit following the manufacturer's instruction. Briefly, measure the concentration of purified PCR products using Qubit4 ™ fluorometer (Invitrogen) and dilute the concentration of each sample to 10 ng/μl. 10 μldiluted PCR product was mixed with anA-Tailing reaction which contained A-Tailing enzyme and buffer, incubated at 37°C for 30 minutes then 65°Cfor 15 min to inactive the enzyme. Then the A-Tailed sample was mixed with PCR Free index adapters (MGI.), T4 DNA Ligase and T4 ligase buffer to add index adapter at both 3' and 5' ends of PCR products.The reaction was incubated at 23°C for 30 min and then purified with XP beads. Then denature the PCR products to be single-strand DNA (ssDNA) by incubating at 95°C for 3 min and keep on 4 °C for the subsequent step. Transform the ssDNA to be circles using cyclase (MGI) at 37 °C for 30 min and then digested to remove linear DNA using Exo enzyme at 37 °C for 30 min. Purify the products again by XP beads and assay the concentration of library by Qubit 4™fluorometer. The amplicons libraries were subjected to deep sequencing on the MGISEQ-2000 platform. In this study, for each lane 4 samples (6 ng each) were pooled together for deep sequencing. To avoid sequencing bias induced by base unbalance of surrogate PCR products, 12 ng whole-genome DNA library (balance library) was mixed with the 4 PCR samples in a final concentration of 1.5 ng/μland sequenced in one lane. All the samples were subjected to pair-ended 150 bp deep-sequencing on MGISEQ-500 platform.

Data analysis

10 In order to evaluate the sequencing quality of amplicons and filter the low-quality sequencing data, Fastqc-0.11.3 and fastp-0.19.6 were used with default parameters for each sample. The clean sequencing reads of pair-ended segments were merged using FLASh-1.2.1 to obtain full-length reads. In order to obtain the amplified fragment reads of each surrogate reference sequence, BsmBI Linker was removed from the surrogate reference sequence. The BWA-MEMalgorithm of bwa was used for local alignment, and the reads of all samples were divided into 12,000 independent libraries. Due to the existence of sequencing or oligo synthesis introduced errors, each library was then filtered. As SpCas9 mainly causes insertions and deletions, the length of surrogate sequence is expected to change from its original 37bp. We adopt the following steps for data processing and filtering: 1) Obtain the sequence containing gRNA

Citation: Xi Xiang, Yonglun Luo (04/20/2021). High throughput quantification of CRISPR gRNA efficiency based on surrogate lentivirus libraries. https://dx.doi.org/10.17504/protocols.io.bt9jnr4n

+ scaffold fragment as dataset1. 2) Obtain the sequence containing **GTTTGAAT** in dataset1 as dataset2 (BsmBI linker fragments changed in orientation (GTTTGGAG-> GTTTGAAT)). 3) Extract the intermediate surrogate sequence from dataset2, which removed the length limit. In order to eliminate the interference of background noise before analyzing editing efficiency, all mutations or indels found in WT HEK293T cells group were removed.

The total editing efficiency for each gRNA was calculated according to the following formula: total efficiency = number of reads with length not equal to 37 bp/total number of reads.

The average fraction of indels from 30bp deletion to 10bp insertion was calculated according to the following formula: average indel fraction = number of reads with length rang from 7 to 47 bp / total number of reads of the 12K library.

Citation: Xi Xiang, Yonglun Luo (04/20/2021). High throughput quantification of CRISPR gRNA efficiency based on surrogate lentivirus libraries. https://dx.doi.org/10.17504/protocols.io.bt9jnr4n