Chapter 11

Genome-Wide CRISPR/Cas9 Screening for High-Throughput Functional Genomics in Human Cells

Shiyou Zhu, Yuexin Zhou, and Wensheng Wei

Abstract

It is highly desirable to identify gene's function in a high-throughput fashion, and the CRISPR/Cas9 system has been harnessed to meet such a need. Here, we describe a general method to generate genome-scale lentiviral single-guide RNA (sgRNA) library and conduct a pooled function-based screening in human cells. This protocol would be of interest to researchers to rapidly identify genes in a variety of biological processes.

Key words CRISPR-Cas9 system, High-throughput, Knockout, Screening, sgRNA

1 Introduction

The CRISPR/Cas9 system is commonly used as a defense system in archaea and bacteria [1] The most widely used engineered CRISPR/Cas system is composed of the Cas9 nuclease and a single guide RNA (sgRNA). A~20-bp sequence at the 5′ terminus of the sgRNA recognizes the targeted sequence via Watson-Crick base pairing, which recruits Cas9 protein to the targeted locus to produce double-stranded DNA breaks (DSBs) and modulate endogenous gene expression [2–4] The high efficiency and easy programmability of the CRISPR/Cas9 system make it possible to create large-scale loss-of-function mutations in mammalian cells. The CRISPR/Cas9 system has been harnessed to produce pooled gene knockout libraries through lentivirus infection for the functional screening of genes in a certain biologic process, aided by deep-sequencing technology [5–8] Here, we describe the design of sgRNAs targeting all human genes, with six different sgRNAs targeting each gene. sgRNA-coding sequences were generated by array-based synthesis and cloned into a lentiviral vector. An

Shiyou Zhu and Yuexin Zhou contributed equally to this work.

Karen Mossman (ed.), *Innate Antiviral Immunity: Methods and Protocols*, Methods in Molecular Biology, vol. 1656, DOI 10.1007/978-1-4939-7237-1_11,

sgRNA-expressing cartridge was introduced into a cell line stably expressing Cas9 by lentiviral infection at a low multiplicity of infection (MOI $\leq$ 0.3), followed by selection with antibiotics or fluorescence-activated cell sorting (FACS). After library screening, sgRNA-coding sequences were PCR-amplified from genomes, followed by deep-sequencing analysis. Candidate gene targets were then selected for further validation.

2 Materials

2.1 sgRNA Library Synthesis, Primers, and Plasmids

1. Oligo B3 Synthesizer (CustomArray, Inc.).
2. SpeedVac (Thermo Fisher Scientific).
3. Primers for the amplification of oligos for sgRNA library construction (*see* Table 1).
4. Primers for PCR amplification of sgRNA-coding sequences in genome for deep-sequencing analysis (*see* Table 1).
5. sgRNA expressing vector: pLenti-sgRNA-Lib [5].
6. Virus packaging plasmid: pVSVG, pR8.74.

2.2 Enzymes, Chemicals, and Kits

1. TransTaq DNA polymerase High Fidelity.
2. Phusion Hot Start Flex DNA polymerase.
3. dNTP mix (2.5 mM each).
4. PCR product purification kit.
5. BsmBI restriction enzyme.
6. Tango buffer: 33 mM Tris-acetate pH 7.9, 10 mM magnesium acetate, 66 mM potassium acetate, 0.1 mg/mL BSA (or buffer compatible with BsmBI restriction enzyme).
7. T4 DNA ligase.
8. 10 mM ATP.
9. 50 mM DTT.

Table 1
Primers for PCR amplification of synthesized oligos and sgRNA-coding sequences in the genome

Primer	Sequence
Oligo-F	5′-TTGTGGAAAGGACGAAACCG-3′
Oligo-R	5′-TGCTGTCTCTAGCTCTACGT-3′
Lib-F	5′-TATCTTGTGGAAAGGACGAAACACC-3′
Lib-R	5′-AATACGGTTATCCACGCGGC-3′

10. Trans1-T1 competent cells.
11. LB broth: to 800 mL dH_2O add 10 g tryptone, 5 g yeast extract and 5 g NaCl. Adjust pH to 7.5 with NaOH. Adjust volume to 1 L with dH_2O; autoclave to sterilize.
12. Endo free plasmid Maxi kit.
13. TIANamp genomic DNA kit (TIANGEN).
14. NGS Fast DNA Library Prep Set for Illumina.

2.3 Cell Culture and Transfection

1. HEK293T cell line.
2. Mammalian cell line of interest (HeLa in this protocol).
3. Complete culture medium, e.g., Dulbecco's modified Eagle medium (DMEM) with 10% FBS and 1% Penicillin-Streptomycin solution.
4. 0.25% Trypsin-EDTA solution.
5. Phosphate-buffered saline (PBS).
6. X-tremeGENE HP or other DNA transfection reagent.

3 Methods

3.1 sgRNA Library Design

1. For each targeting gene, 5′-end coding sequence is preferred for the design of sgRNAs (*see* **Note 1**).
2. Six sgRNAs are designed for each gene (*see* **Note 2**).
3. For genome-scale sgRNA library, at least 1000 nontarget sgRNAs are included as negative controls (*see* **Note 3**).

3.2 PCR Amplification for Synthesized DNA Oligo

1. For each PCR reaction, mix the followings into a 0.2 mL PCR tube: synthesized oligo template, forward primer (10 μM, 2.5 μL), reverse primer (10 μM, 2.5 μL), Phusion Hot Start Flex DNA polymerase (0.5 μL), HF buffer (5×), dNTP mix (1 μL), ddH_2O to a total volume of 50 μL. No less than 24 tubes are needed (*see* **Note 4**).
2. Perform PCR reaction as follows: 98 °C, 30 s for hot start; thermal cycling (98 °C, 10 s; 58 °C, 20 s; 72 °C, 10 s; 26 cycles); 72 °C, 10 min; and hold at 4 °C.
3. Optional step: DNA electrophoresis with 1 ~ 2 μL PCR product for quality check. The size of PCR product is about 80 bp (*see* **Note 5**).
4. Purify the PCR products using kit of choice.

3.3 BsmBI Digestion, DNA Ligation, and Transformation

1. Mix the purified PCR fragments with the following into one tube: BsmBI (7.5 units), T4 DNA ligase (100 units), ATP (10 nmol), DTT (10 nmol), Tango buffer (10×), sgRNA expressing vector (~20 ng), and ddH_2O to a total volume of 10 μL (*see* **Note 6**).

2. Perform thermo cycles (37 °C, 5 min; 16 °C, 5 min; 16 cycles); 37 °C, 5 min; and hold at 4 °C.
3. Transform 2 μL product into each tube of 50 μL Trans1-T1 competent cells, then add 1 mL liquid LB without antibiotics and culture the mixture overnight at 37 °C (*see* **Note 7**).
4. Mix all the bacterial liquid together, and extract the plasmids using Endo free plasmid Maxi kit.

3.4 Lentivirus Package

1. Culture HEK293T cells in complete culture medium at 37 °C and 5% CO_2 in a humidified incubator. Seed 4×10^6 HEK293T cells onto 10 cm plates 24 h before transfection.
2. Co-transfect 0.4 μg pVSVG plasmid, 4 μg of pR8.74, and 4 μg of Cas9 expressed plasmid or sgRNA library plasmids into HEK293T cells.
3. Collect the media and centrifuge at 395 RCF for 10 min to pellet cell debris 72 h posttransfection.
4. Calculate the virus titer.

3.5 Cas9 Stably Expressed Cell Line Construction

1. Culture HeLa cells in complete culture media at 37 °C and 5% CO_2 in a humidified incubator. Seed 2×10^6 cells into 10 cm plates 24 h before viral infection.
2. Add polybrene into DMEM at a final concentration of 8 μg/mL. Infect cells with Cas9-producing virus.
3. Add 5 μg/mL of Blasticidin onto cells 48 h after virus infection to enrich Cas9-expressing cells. Isolate the best single clones that show high efficiency in the indel analysis (*see* **Note 8**).

3.6 sgRNAs Delivery and Cell Library Construction

1. Culture HeLa cells stably expressing Cas9 in complete culture medium at 37 °C and 5% CO_2 in a humidified incubator. For each repeat, seed at least 3.6×10^7 cells (15 cm plate, total of 9 plates) 24 h before viral infection.
2. Add polybrene into DMEM at a final concentration of 8 μg/mL, and then add sgRNA library virus into cells with MOI ≤ 0.3 (*see* **Note 9**).
3. Collect cells expressing sgRNA by FACS or antibiotic selection 48 h after virus infection. Keep culturing these cells for 14 days before splitting them into four cell libraries (at least 1.2×10^7 cells per library), one for control and three replicates for the screening (*see* **Notes 10** and **11**).

3.7 Library Screening and Deep-Sequencing Analysis

1. Perform library screen.
2. Extract genomic DNA from 1.2×10^7 of control library cells and experimental library cells.
3. PCR-amplify sgRNA-coding regions from cell genome. For each PCR reaction, mix the following material into 0.2 mL

PCR tube: genomic DNA (4 μg), primer Lib-F (10 μM, 2 μL), primer Lib-R (10 μM, 2 μL), dNTP mix (2.5 nM, 8 μL), Taq DNA polymerase buffer (10×, 10 μL), Taq DNA polymerase (5 units/μL, 2 μL), and ddH_2O to a total volume of 50 μL.

4. Perform PCR reaction as follows: 94 °C 5 min for hot start; thermal cycling (95 °C 30 s; 62 °C 30 s; 72 °C 30 s; 26 cycles); 72 °C 10 min; hold at 4 °C. For each sample, perform 20 separate 100-μL reactions with 4 μg genomic DNA in each.
5. Pool and purify PCR products (20 tubes total) of each replicate, followed by high-throughput sequencing analysis. Index control library and different experimental library replicates with barcodes.
6. Use DESeq2 (R software package from Bioconductor) to perform a statistical analysis of the sequencing data. Rank the enrichment of sgRNA by the average fold change of normalized counts ($reads_{Exp}/reads_{Ctrl}$) and the adjusted *P* value ≤ 0.05. Select candidate sgRNAs and their targeted genes for further validation (*see* **Note 12**).

4 Notes

1. The basic design principles are consistent across different libraries. First, the reading frame could be disrupted by indels mediated by DSBs at target sites of sgRNAs. One should design sgRNAs targeting 5′ end of coding sequences for gene knockout as much as possible to maximize the chance of gene knockouts. Second, one should select sgRNAs with the best efficiency based on certain sequence features and criteria to minimize the off-target effect and maximize on-target activity. For instance, the GC content should be in the range of 20% ~ 70%, and sequences containing homopolymer stretches (e.g., TTTT, GGGG) should be avoided [9, 10].
2. We suggest designing six sgRNAs for each gene in this protocol. Although one report has shown that only one sgRNA for each gene is enough through the optimized design [9], designing 4 ~ 6 sgRNAs for each gene would have a better chance at target identification and the statistical analysis of screening data.
3. For negative controls in the library, one could also design sgRNAs targeting safe locus on genome, such as *AAVS1*.
4. To minimize both the mutation rate and the amplification bias, fewer PCR cycles and more PCR reactions are recommended. We usually conduct 26 cycles with no less than 24 tubes for each library sample.

5. If there are other nonspecific amplified bands, one should perform gel purification on the PCR products.
6. Make sure that the BsmBI enzyme works at 37 °C.
7. Before culturing bacteria overnight, plate 1 μL mixture onto solid LB medium with 25 μg/mL of ampicillin to count the colony number for each reaction. We perform multiple tubes of transformation to ensure that the total colony numbers exceed 200-fold coverage of the sgRNA library size.
8. To maximize the gene knockout efficiency, one should select the best Cas9-expressing single clones to construct the sgRNA library, those that have the highest efficiency in generating Cas9-mediated DSBs.
9. Low MOI is used to lower the odd that more than one sgRNA enters the same cell.
10. One should keep culturing cells for 7 ~ 14 days after lentiviral infection in order to maximize the gene knockouts in the cell library [5, 11].
11. To ensure library complexity, one should maintain the number of cells in each library at least 100-fold. For the whole-genome sgRNA library containing ~10^5 sgRNAs, at least 1.2×10^7 cells should be harvested for each passage.
12. Other bioinformatics tools could also be used for the analysis of deep-sequencing data [12].

Acknowledgments

We thank J. Xi for providing the Cas9-encoding construct, and D. Trono for aid with the lentivirus packaging system. We thank Peking University High-throughput Sequencing Center operated by BIOPIC. This work was supported by funds from the National Basic Research Program of China (2010CB911800), the National Science Foundation of China (NSFC31170126, NSFC31070115), and the Peking-Tsinghua Centre for Life Sciences. Shiyou Zhu and Yuexin Zhou contributed equally to this work.

References

1. Barrangou R et al (2007) CRISPR provides acquired resistance against viruses in prokaryotes. Science 315:1709–1712
2. Jinek M et al (2012) A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science 337:816–821
3. Cong L et al (2013) Multiplex genome engineering using CRISPR/Cas systems. Science 339:819–823
4. Mali P et al (2013) RNA-guided human genome engineering via Cas9. Science 339:823–826
5. Zhou Y et al (2014) High-throughput screening of a CRISPR/Cas9 library for functional genomics in human cells. Nature 509:487–491
6. Wang T, Wei JJ, Sabatini DM, Lander ES (2014) Genetic screens in human cells using the CRISPR-Cas9 system. Science 343:80–84

7. Shalem O et al (2014) Genome-scale CRISPR-Cas9 knockout screening in human cells. Science 343:84–87
8. Koike-Yusa H, Li Y, Tan EP, Velasco-Herrera Mdel C, Yusa K (2014) Genome-wide recessive genetic screening in mammalian cells with a lentiviral CRISPR-guide RNA library. Nat Biotechnol 32:267–273
9. Doench JG et al (2016) Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. Nat Biotechnol 34:184–191
10. Doench JG et al (2014) Rational design of highly active sgRNAs for CRISPR-Cas9-mediated gene inactivation. Nat Biotechnol 32:1262–1267
11. Peng J, Zhou Y, Zhu S, Wei W (2015) High-throughput screens in mammalian cells using the CRISPR-Cas9 system. FEBS J 282:2089–2096
12. Li W et al (2014) MAGeCK enables robust identification of essential genes from genome-scale CRISPR/Cas9 knockout screens. Genome Biol 15:554