# sgRNA Library Generation using the Molecular Chipper Approach

**Jijun Cheng** ( j.cheng@yale.edu )
Department of genetics, Yale Stem Cell Center, Yale University School of Medicine, New Haven, Connecticut 06510, USA

**Jun Lu** ( jun.lu @yale.edu )
Department of genetics, Yale Stem Cell Center, Yale University School of Medicine, Yale Cooperative Center of Excellence in Hematology, Yale Center for RNA Science and Medicine, New Haven, Connecticut 06510, USA

**Wen Pan**
Department of genetics, Yale Stem Cell Center, Yale University School of Medicine, New Haven, Connecticut 06510, USA

**Method Article**

**Keywords:** sgRNA library, Molecular Chipper, noncoding regions, fragmentation, Type III restriction enzyme

**Posted Date:** April 1st, 2016

**DOI:** https://doi.org/10.1038/protex.2016.023

# Abstract

Genetic screens using single-guide-RNA \(sgRNA) libraries and CRISPR technology have been powerful to identify genetic regulators for both coding and noncoding regions of the genome. Interrogate functional elements in noncoding regions requires sgRNA libraries that are densely covering, and ideally inexpensive, easy to implement and flexible for customization. We present a Molecular Chipper protocol for generating dense sgRNA libraries from genomic regions of interest. This approach utilizes a combination of random fragmentation and a Type III restriction enzyme to derive a densely covering sgRNA library from input DNA.


# Introduction

Genome editing using Streptococcus pyogenes \(sp) Cas9 and sgRNA libraries is a powerful tool to screen for functional genetic regulators in mammalian cells by generating biallelic loss-of-function sequence alterations 1–6. Cas9 binds sgRNA, which can be designed to target Cas9 toward a defined locus in the genome. The nuclease activity of Cas9 cuts target DNA locus, leading to double-stranded DNA breaks, which upon DNA repair through non-homologous end-joining pathway frequently results in short deletions at the locus of interest. The powerful genomic editing capacity of the CRISPR system has led to the use of sgRNA libraries to interrogate protein-coding genes as well as noncoding regions. Indeed, Canver et al. recently demonstrated that cis-regulatory elements for BCL11A can be identified using computationally designed CRISPR libraries of ~1300 sgRNAs for selected enhancer regions 7. Several sgRNA libraries for protein-coding genes and/or limited numbers of non-coding genes have been reported 2–5,7,8, which were produced by careful bioinformatics design, oligonucleotide synthesis on microarray, and cloning of oligonucleotide pool\(s) into vectors. This synthetic approach has been very useful, but requires computational expertise for genome-wide sgRNA design and expensive microarray synthesis, and thus is challenging for most laboratories. Importantly, without prior knowledge of the locations of critical noncoding-element-containing regions, functional mapping of noncoding genomic regions requires sgRNA libraries that densely populate regions of interest, and the ideal method requires flexibility for adjusting the scale of sgRNA production to easily cope with this need. We describe here a detailed protocol of the Molecular Chipper approach that processes any input DNA piece\(s) to generate a near base-resolution sgRNA library densely covering the input DNA of interest.

# Reagents

• T4 DNA ligase \(M0202T, NEB) • T4 DNA ligase \(M0202S, NEB) • Agarose \(AB00972, AmericanBIO) • Agarose, low melting point \(AB00981, AmericanBIO) • 1X TAE gel running buffer \(50X stock: 242 g Tris base, 18.61 g EDTA and 57.1 ml glacial acetic acid) • Ethidium bromide, 10 mg/ml \(E1510, Sigma) • 10,000X CYBR Safe DNA Gel Stain \(S33102, Invitrogen) • 3M sodium acetate, pH 5.2 • 100% ethanol • 70% ethanol • Phenol, saturated with Tris, pH7.5 \(15593-031, Invitrogen) • Chloroform • PCI \(Phenol:chloroform:isoamyl alcohol 25:24:1), Tris saturated \(03 117 944 001, Roche) • NEBNext End Repair Module \(E6050S, NEB) • QIAEX II Gel Extraction Kit \(20021, Qiagen) • QIAquick PCR Purification

Kit \(28104, Qiagen), • QIAquick Gel Extraction kit \(28704, Qiagen) • QIAprep Spin Miniprep Kit \(27104, Qiagen) • MiniElute Gel Extraction kit \(28604, Qiagen) • EcoP15I enzyme \(R0646L, NEB) • BamHI-HF enzyme \(R3136S, NEB) • HindIII-HF enzyme \(R3104S, NEB) • NEB 5-alpha Electrocompetent E. coli \(C2989K, NEB) • Petri dishes • LB medium • Agar • Ampicillin • Distilled water

# Equipment

• 37 ºC water bath • IncuBlock heating block \(Danville Scientific) • Nanodrop 2000 \(Thermo Scientific) • Gel electrophoresis system • UltraSlim LED Illuminator \(Maestrogen) • S220 Focused-ultrasonicator and sonication vials \(Covaris) • GenePulser Xcell \(BIO-RAD) • Microcentrifuge

# Procedure

1. For large DNA, such as BAC clones, start from step 5. If input DNA are small and composed of multiple pieces, such as individual PCR products, purified DNA pieces are quantified by a Nanodrop 2000 and pooled in an equal molar ratio. 2. If using an input pool of multiple pieces of DNA, randomly ligated larger products are then generated using excessive T4 DNA ligase. This step has been added to avoid biasing against regions located near the ends of input DNA pieces \(because we perform a size selection after this fragmentation step, and sequences close to the ends would be represented by very small fragments after fragmentation, and thus would be under-represented in the final library). Specifically, for each ug of DNA, the DNA pool are ligated with 4,000 units of T4 DNA ligase in a 50 ul reaction for 3 hours at 37ºC. We start with ~20 ug total DNA that have 5’ phosphate groups. If directly using PCR products, perform a kinase reaction prior to ligation using T4 Polynucleotide Kinase. 3. Run 10 ul or 200 ng of ligation products on 1% agarose gel with 0.5 ug/ml ethidium bromide in 1X TAE buffer, in order to check sizes of ligation products \(usually >10 kb on average). 4. The ligated large amount of DNA is purified by ethanol precipitation. Add 10% volume of 3M sodium acetate, pH 5.2, and 2 volumes of 100% ethanol; mix and then precipitate in -20ºC for one hour; spin down using a pre-cooled microcentrifuge for 10 minutes, wash with 70% ethanol; air dry for one minute; and re-suspend in 150 ul water. 5. To generate random DNA fragments, ~14 ug of the input DNA \(ligated if originated from multiple pieces) in 120 ul of water is sonicated in a S220 Focused-ultrasonicator for 90 seconds to result in fragments peaking at sizes of 400-450bp \(Peak Power = 140V, Duty Factor = 5, Cycle/Burst = 200 and Average Power =7). Run at least 2-4 ul \(250-500 ng) DNA on 2% agarose gel with ethidium bromide to check peak sizes of fragmented DNA. 6. Sonicated fragments are repaired in a 150 ul End Repair reaction with 15 ul of the NEBnext End Repairing Enzyme Mix, followed by 1% agarose gel purification of the 400-450 bp DNA fragments \(visualized with 1X CYBR Safe in the gel by a blue-light gel dock, such as the Maestrogen Illuminator) by using QIAEX II Gel Extraction Kit. 7. To obtain fragment ends from both ends of the random DNA fragments, an EcoP15I-adaptor is first prepared by annealing two oligonucleotides aaaactcgagcagcagtggatccG and /5phos/Cggatccactgctgctcgag \(IDT) in equal molar ration in 1X ligation buffer \(NEB). The annealed DNA adaptor contains an EcoP15I site \(in bold) followed by a total 8-bp spacer, including a BamHI site \(underlined) and a G \(capitalized) at the end for later sgRNA cloning

to serve as transcription start site from the U6 promoter. 8. 12 ug of the DNA fragments are ligated, at a 1:10 molar ratio to 6.0 ug of the above annealed EcoP15I-adaptor with 20,000 units T4 DNA ligase in 300 ul reaction for 3 hours at 37ºC. 9. The adaptor-ligated DNA fragments are purified from adaptor monomer and other non-specific bands by running on 1% agarose gels and by using QIAEX II Gel Extraction Kit. 10. 5 ug of the EcoP15I-adaptor-ligated gel-purified DNA is digested by 100 units of EcoP15I enzyme \(NEB) in 300 ul reaction for 1 hour at 37 ºC. To check completeness of digestion and efficiency of adaptor ligation in step 8, run 15 ul of the digestion on 4% low-melting-point agarose gel with ethidium bromide to visualize the 38 bp \(expecting 25-50 ng of the 38 bp DNA fragment). 11. After digestion, EcoP15I digestion reaction is cleaned by phenol/PCI/chloroform extraction, ethanol precipitation and resuspension in 50 ul water. 12. Precipitated digestion products are gel-purified \(on 4% low-melting-point agarose gel) to obtain a ~38-bp DNA fragment pool \(EcoP15I-adaptor + 19/17 bases from ends of random DNA fragments) by diluting the gel slice with 2 volumes of 1X TAE buffer, melting the gel slice at 70 ºC on a heating block, extracting by phenol/PCI/chloroform and dissolving DNA in 20 ul water. 13. To ligate to the rest of sgRNA backbone, 280 ng of the purified 38-bp DNA pool is ligated in 50 ul reaction with 4,000 units of T4 DNA ligase for 3 hours at 37ºC, at a 1:5 molar ratio to 2.75 ug of an sgRNA-backbone-adaptor. The sgRNA-backbone-adaptor contains two Ns for binding to overhangs from EcoP15I digestion products, the remaining sgRNA sequence \(without the target-recognition domain), a polyT stretch for polymerase III transcriptional termination, and a HindIII site for cloning. This sgRNA-backbone adaptor was prepared by annealing two oligonucleotides below \(IDT), followed by 4% low-melting-point agarose gel-purification, by using QIAEX II Gel Extraction Kit, to eliminate improperly annealed products. 14. The ligated sgRNA DNA pool is cleaned by QIAquick PCR Purification Kit, digested in 50 ul with 20 units each of BamHI and HindIII \(NEB) overnight at 37ºC, and gel-purified by using 4% low-melting-point agarose and the MiniElute Gel Extraction kit to obtain a ~115-bp sgRNA pool. 15. This sgRNA pool is quantified by SYBR Safe Gel Stain \(Invitrogen) on a fluorometer, and ligated at 3:1 molar ration with 40 units of T4 DNA ligase per 100 ng total DNA per 10ul reaction at room temperature overnight into BamHI-HindIII sites of a retroviral vector pSUPER-CRISPR which contains a U6 promoter and a puromycin selection marker. 16. The ligated DNA is ethanol precipitated \(see step 4 for details) and dissolved in 10 ul water. 17. Ligation products are transformed into NEB5 alpha competent cells in aliquots by electroporation using a GenePulser Xcell. The transformed cells are then pool together. Several small fractions of transformation are plated to estimate total transformed clones. To estimate the % cloning efficiency, 10-20 colonies are grown up to miniprep DNA using the QIAprep Spin Miniprep Kit, and the cloned sgRNA inserts are confirmed by Sanger sequencing using oligo nucleotide ctccctttatccagccctca. 18. The transformation culture is grown overnight in 100 ml of LB medium containing 100 ug ml-1 of ampicillin for plasmid DNA preparation.

# Timing

Two-four weeks.

# Troubleshooting

• No 38 bp band after EcoP15I digestion: 1. Check if the adaptor is annealed well: ligate only adaptors and should see major band of the monomer and minor non-specific bands on 4% low melting point gel. 2. Check if sonicated DNA are repaired well and if ligation goes well: treat a small fraction of repaired DNA by Taq DNA polymerase in the presence of dNTPs at 72 ºC followed by TA cloning \(Invitrogen). Many white colonies should appear if end repairing works well. Sequencing of the cloned insert should have adaptors at both end of the cloned fragments. 3. Purify the adaptors from impurities to reduce background for gel purification of the ligated DNA. • Not enough colonies/coverage after transforming the final ligation: 1. Identify a commercial source of competent cell with high transformation efficiency. We use the electroporation competent cells from NEB. 2. Alternatively, make your own electro-competent cells of higher efficiency. 3. Perform a test transformation to calculate total colony number/complexity/coverage before transforming the rest of ligation to make the DNA library. • Cloning efficiency is not high: 1. Digest the full-length ligated sgRNA fragment with BamHI and HindIII for 6 hours instead of overnight to minimize star activity. 2. Test the vector quality by cloning a BamHI-HindIII fragment and if necessary, a stuffer can be cloned first to prepare the vector.

## Anticipated Results

When properly prepared, a library of >1 million clones can be easily achieved.

## References

1. Wiedenheft, B., Sternberg, S. H. & Doudna, J. a. RNA-guided genetic silencing systems in bacteria and archaea. Nature 482, 331–8 \(2012). 2. Shalem, O. et al. Genome-scale CRISPR-Cas9 knockout screening in human cells. Science 343, 84–7 \(2014). 3. Wang, T., Wei, J. J., Sabatini, D. M. & Lander, E. S. Genetic screens in human cells using the CRISPR-Cas9 system. Science 343, 80–4 \(2014). 4. Zhou, Y. et al. High-throughput screening of a CRISPR/Cas9 library for functional genomics in human cells. Nature 509, 487–91 \(2014). 5. Koike-Yusa, H., Li, Y., Tan, E.-P., Velasco-Herrera, M. D. C. & Yusa, K. Genome-wide recessive genetic screening in mammalian cells with a lentiviral CRISPR-guide RNA library. Nat. Biotechnol. \(2013). 6. Mali, P. et al. RNA-guided human genome engineering via Cas9. Science 339, 823–6 \(2013). 7. Canver, M. C. et al. BCL11A enhancer dissection by Cas9-mediated in situ saturating mutagenesis. Nature 527, 192–7 \(2015). 8. Gilbert, L. et al. Genome-Scale CRISPR-Mediated Control of Gene Repression and Activation. Cell 1–15 \(2014).

## Acknowledgements

We thank the Lu lab members for constructive discussions and technical assistance.