

Combinatorial CRISPR-Cas9 Knockout Screen

Dongxin Zhao (✉ doz006@eng.ucsd.edu)

Mali Lab

John Paul Shen (✉ jpshen@ucsd.edu)

Ideker Lab

Roman Sasik (✉ rsasik@ucsd.edu)

UCSD Center for Computational Biology & Bioinformatics

Trey Ideker

Ideker Lab

Prashant Mali


Mali Lab

Method Article

Keywords: CRISPR, Genetic Interaction, Synthetic Lethal, Cancer

Posted Date: March 20th, 2017

DOI: <https://doi.org/10.1038/protex.2017.063>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Abstract

Simultaneous mutation of two genes can produce a phenotype that is unexpected in light of each mutation's individual effect. This phenomenon, known as genetic interaction, identifies an underlying functional relationship between the genes, such as contributions to the same protein complex or pathway. Mapping these functional relationships has broad applicability for advancing our fundamental understanding of biological systems and also has direct implications for cancer therapeutic development. Here we describe a method using the CRISPR-Cas9 system to knockout pairs of genes, enabling high-throughput, systematic mapping of these genetic interaction networks in human cells. This protocol accompanies Shen et al (Nature Methods, publishing online 20 March, 2017; 10.038/nmeth.4225); it was added to the manuscript after formal peer review, as an aid to users.

Introduction

In the CRISPR-Cas9 system, a guide-RNA (gRNA), in complex with the Cas9 protein, targets genomic sequences homologous to the gRNA1-7. Targeting new genomic elements entails modifying the gRNA sequence, thus enabling many targeted genome editing and regulation capabilities¹. Notably, Cas9 also enables facile multiplex targeting via delivery of multiple gRNAs per cell⁸. As described in the associated publication, we have combined multiplex targeting with array-based oligonucleotide synthesis⁹⁻¹² to create dual-gRNA libraries covering up to 105 defined gene pairs. In these libraries, each construct bears two gRNAs, with each gRNA designed to target either a gene or a scrambled non-targeting sequence absent from the genome. Thus, all combinations of gene-gene (double gene perturbation) and gene-scramble (single gene perturbation) are exhaustively assayed for effects on cell growth. Notably, by enabling determination of single-gene and dual-gene perturbation effects in the same assay, this approach generates data that allows for the systematic quantification of genetic interactions in humans.

Reagents

Plasmids & DNA 1. gRNA-LGP vector 2. step 3 inserts, synthesized commercially (Integrated DNA Technologies) Kits & Chemicals 1. One Shot Stbl3 Chemically Competent E. coli (Invitrogen) 2. ElectroMAX Stbl4 Competent Cells (Invitrogen) 3. Kapa Hifi enzyme (Kapa Biosystems) 4. QIAquick gel extraction kit (Qiagen) 5. BsmBI (New England Biosciences) 6. Calf Intestinal Alkaline Phosphatase (New England Biosciences) 7. Carbenicillin 8. HiSpeed Plasmid Maxi Kit (Qiagen) 9. Centricon Plus-20 centrifugal ultrafilters with a cutoff 100,000 NMWL (Millipore) 10. Agencourt AMPure XP (Beckman Coulter) 11. QIAquick PCR Purification (Qiagen) 12. Next® Multiplex Oligos for Illumina (New England Biosciences) 13. Illumina Library Quantification kit (Kapa Biosystems)

Equipment

1. Tissue Culture Hood/Biosafety Cabinet 2. CO₂ incubator at 37°C and 5% CO₂ for mammalian cell culture 3. Cell counting system 4. Manual Pipettes (P2, P20, P200, P1000) & tips 5. 245 mm x 245 mm Bioassay dishes, tissue culture treated 5. Thermocycler for PCR 6. Electroporator 7. Bioanalyzer (Agilent) 8. Next-generation Sequencer (such as Illumina HiSeq)

Procedure

****Preparation of dual-gRNA library**** **_Library design_** 1) Identify genes and gene pairs to be tested in assay. In the associated publication we describe the testing of a square matrix of all pairwise interactions between 73 genes (73 x 73). However, the both gRNA for a dual-gRNA construct are directly specified during oligonucleotide synthesis thereby enabling the library constituents to be exactly defined to facilitate custom gRNA-pairing. Thus asymmetric formats (eg. 10 x 250, 25 x 100, ect.) are also possible. 2) Design 20 bp gRNA to target each gene. Multiple on-line resources to design gRNA are now available¹³⁻¹⁵. Poly-T sequences (i.e., more than two consecutive T's) should be avoided and, to avoid off-target editing, gRNAs should require at least three substitutions in order to match any other sequence in the genome. We selected three gRNA per gene, it is possible to select more gRNA per gene. The total number of unique elements in the library, including non-targeting controls, should not exceed ~50,000 in order to keep tissue culture manageable. Ideally all steps of the process should be performed with a minimum of 200X coverage, meaning 200 unique copies of each gRNA. For a library of 50,000 elements this would require 10 million cells. **_Cloning Step 1: paired gRNA cloning_** 1) Purchase pooled oligonucleotide libraries from commercial vendor such as CustomArray (Bothell, WA). 2) Amplify full-length oligonucleotides with dual gRNA spacers (i.e., 20-bp sequences used for targeting desired genes) using PCR using Kapa Hifi enzyme (Kapa Biosystems). Set up PCR reactions according to the manufacturer's protocol, using 1 µL of synthesized oligonucleotide template (typically around 20 ng), an annealing temperature of 55 °C, and an extension time of 15 s. The numbers of cycles of amplification should fall within the linear phase of amplification Primer sequences: OLS_gRNA-SP_F: TATATATCTTGTGGAAAGGACGAAACACCG OLS_gRNA-SP_R: CTTATTTAACTTGCTATTTCTAGCTCT 3) Repeat ten 50 µL PCR reactions for each library, this is to obtain high yield coverage of the PCR products. Too much genomic starting material can interfere with PCR reaction, which is why to parallel reactions are needed. 4) Run PCR product on 2% agarose gel electrophoresis, separate the 144 bp amplicon and purify with the QIAquick gel extraction kit (QIAGEN). 5) Digest gRNA-LGP vector with BsmBI (New England Biosciences) via the following reaction at 55°C for 3 hours: gRNA-LGP vector 4 µg Buffer 3.1 5 µL 10× BSA 5 µL BsmBI 3 µL H₂O Up to 50 µL 6) After digestion, treat vector with 2 µL of Calf Intestinal Alkaline Phosphatase (New England Biosciences) at 37°C for 30 minutes, 7) Purify with QIAquick PCR Purification Kit (Qiagen) 8) Assemble the paired gRNAs into the vector using 10 Gibson assembly reactions in parallel with reactions as follows: Linearized gRNA-LGP vector 200 ng Dual gRNA inserts 36 ng (molar ratio 1:10) 2× Gibson Assembly Master Mix (NEB) 10 µL H₂O Up to 20 µL 9) Incubate at 50°C for 1 hour 10) Purify with QIAquick PCR Purification Kit (Qiagen) 11) Transform into One Shot Stbl3 Chemically Competent E. coli (Invitrogen). Approximately twenty parallel transformations are needed to ensure adequate library representation. 12) Calculate library coverage: remove a small fraction (20-100 µL) of cultures and spread on carbenicillin (50 µg/ml) LB plates 13) Amplify rest of cultures overnight in 150 ml LB medium, maintaining ~100X library coverage 14) Extract plasmid DNA using the HiSpeed Plasmid Maxi Kit (Qiagen) 15) Pick twenty independent clones for Sanger-sequencing to estimate the overall quality of the library **_Cloning Step 2: Insertion of the gRNA scaffold and the mouse U6 promoter_** 1) Use BsmBI to digest the step 2 library plasmids as per the following reaction at 55°C for 3 hours: Cloning Step 1

library 4 µg Buffer 3.1 5 µL 10× BSA 5 µL BsmBI 3 µL H₂O Up to 50 µL 2) Treat the linearized plasmids with 2 µL of Calf Intestinal Alkaline Phosphatase at 37°C for 30 minutes; 3) Gel-purified cut plasmids using via 0.6% agarose gel electrophoresis and QIAquick gel extraction 4) Concurrently, digested the step 3 inserts with BsmBI per the following reaction at 55 °C for 3 hours. The inserts should be synthesized commercially (Integrated DNA Technologies or similar) and cloned into a TOPO vector. See sequence of insert below. Purified step 2 insert PCR product 0.8 µg Buffer 3.1 5 µL 10× BSA 5 µL BsmBI 3 µL H₂O Up to 50 µL Sequence of the step 2 insert, with left-gRNA scaffold (*italicized*) and mU6 promoters (****bold****): TATGAGGACGAATCTCCCCTTATACGCTCTCT_GTTTCAGAGCTATGCTGGAACTGCATAGCAAGTTGAAATAAGGCTAGTCCGTTATCAACTGAAAAAGTGGCACCGAGTCGGTGCTTTTTT_GTACTGAGTCGCCCA****GTCTCAGATAGATCCGACGCCCATCTCTAGGCCCGCGCCGCCCTCGCACAGACTGTGGGAGAAGCTCGGCTACTCCCCTGCCCGGTTAATTTGCATATAATA**GAGAGACGGTACAAGCACAGCTTTGTCAAGACC 5) Ligate with following reaction, involving overnight incubation at 16°C followed by heat inactivation at 65°C for 10 minutes: 10X T4 DNA Ligase Buffer 2 µL Step 1 library, digested 100 ng Step 2 insert, digested 100 ng T4 DNA Ligase (high concentration) 1 µL H₂O Up to 20 µL 6) Transform 4 µL of reaction from step (5) into 100 µL of ElectroMAX Stbl4 Competent Cells (Invitrogen) according to the manufacturer's protocol using an Eppendorf Electroporator. Spread a small fraction (1-10 µL) of cultures on carbenicillin (50 µg/ml) LB plates to calculate the library coverage. Use rest of the cultures to plate on ten 15 cm LB- carbenicillin plates and grow overnight at 37°C for amplification. Likely two parallel transformations were required to get ~100X library coverage. 7) Extract plasmid DNA using HiSpeed Plasmid Maxi Kit (Qiagen). Library diversity is determined by deep sequencing. ****Viral production**** _Note: it is important to check your HEK cells for contamination with mycoplasma before using them for the production of lentiviral particles._ 1) To produce lentivirus, seed HEK293T cells in 15 cm tissue culture dishes one day before transfection so that they are 70-80% confluent at the time of transfection. HEK293T cells to be maintained in DMEM medium supplemented with 10% fetal bovine serum. 2) On day of transfection, change to fresh culture media. 3) For each 15 cm dish, dilute 36 µL of Lipofectamine 3000 (Life Technologies) in 1.2 mL OptiMEM (Life Technologies). 4) Separately, dilute 3 µg pMD2.G (Addgene #12259), 12 µg of pCMV delta R8.2 (Addgene #12263), 9 µg of lentiviral vector and 48 µL of P3000 Reagent in 1.2 mL OptiMEM. Incubate for 5 min at room temperature. 5) After incubation, combine the Lipofectamine 3000 mixture and DNA mixture and incubate at room temperature for 30 min. 6) Add the mixture dropwise to HEK293T cells. 7) Return HEK293T cells to incubator, remove and collect media 48 hours after transfection to harvest virus. Replace with fresh media. 8) Remove and collect media again 72 hours after transfection. 9) Combine 48 and 72 hr harvests and further concentrate virus using Centricon Plus-20 centrifugal ultrafilters with a cutoff 100,000 NMWL (Millipore) to a final volume of 450 µL, then aliquot and freeze at -80°C. ****Cas9 cloning**** Prior to screening cell lines must be engineered to express the CRISPR Cas9 nuclease. This can be done by stable integration into the human AAVS1 site, or using a lentiviral vector to package a Cas9 expression plasmid such as Addgene #52962, followed by antibiotic selection to for cells with the Cas9 vector. ****Competitive growth experiments**** _Note: It is important to maintain coverage (ideally 200X) at all stages of the competitive growth experiment. Also, at least 2 biological replicates (independent transductions) are needed for each cell line and/or condition of interest._ 1) Package the pooled library of double gRNA constructs into lentivirus (see above) 2) Titer the library virus in the cell line of interest. The transduction efficiency of certain cell lines can be significantly less than 293T, so it is important to establish the functional titer in your specific cell line. 3) Day -1: On day prior to transduction plate sufficient cells so that by next day there will be at three times the desired minimum coverage by day 1. For example, for 200X coverage of a 50,000 element library 30 mil cells would be needed. Use sufficient number of plates so that cells will be ~50%-60% confluent on day 1. To facilitate growing these large number of cells we use 245 mm x 245 mm Bioassay dishes (Corning). 4) Day 0: Transduce cells with gRNA library virus at MOI of 0.3 or lower to ensure that each cell had 0-1 double gRNA constructs. Because approximately 2/3 of the cells will be killed it is important to start transduction with three times the desired minimal coverage. 5) Day 1: Change media to remove virus 6) Day 2: Passage cells if they are greater than 85% confluent, otherwise leave until Day 3 7) Day 3: Count and passage cells. Freeze down enough cells to account for minimum desired coverage (snap freeze in N₂, store at -80 C). Re-plate enough cells to account for minimum desired coverage. Start puromycin selection at predetermined selection dose. 8) Days 4-28: Keep cells in exponential growth by passaging every 2-4 days as needed for your cell line. At every passage, re-plate enough cells to account for minimum desired coverage and freeze the rest. 9) Extract DNA from cells harvested at 3, 14, 21 and 28-day time points using the Blood and Cell Culture DNA Mini Kit (Qiagen) according to manufacturer protocols. ****NGS library preparation**** 1) Quantitate DNA extracted from step #9 above to ensure there is adequate coverage prior to starting library preparation. The dual gRNA cassette will be amplified and prepared for deep sequencing using two steps of PCR. 2) PCR 1: Perform this as 10 separate 50 µL reactions with 2 µg input genomic DNA per reaction (total 20 µg for each sample) using Kapa Hifi enzyme. PCR primers are as follows: NGS_dual-gRNA_SP_Lib_F: ACACTCTTCCCTACACGACGCTCTCCGATCT TATATATCTGTGGAAAGGACGAAACACCG NGS_dual-gRNA_SP_Lib_R: GACTGGAGTTCAGACGTGTGCTCTCCGATCT CCTATTTTAACTTGCTATTCTAGCTCTA Thermocycling parameters are: 95°C for 30 s, 21-26 cycles of (98°C for 15 s, 55°C for 15 s, 72°C for 45 s), and 72°C for 5 min. Test the numbers of cycles to fall within the linear phase of amplification. 3) Pool amplicons (600 bp) of 10 reactions for each sample then size-select and purify with Agencourt AMPure XP beads (Beckman Coulter) at 0.8 ratio. 4) Further purify with QIAquick PCR Purification (Qiagen). 5) PCR 2: Use 4 separate 50 µL reactions with 5 ng of first-step PCR product per reaction (total 20 ng for each sample), using Next® Multiplex Oligos for Illumina (New England Biosciences) to attach Illumina adaptors and indexes. The thermocycling parameters are: 95°C for 30 s, 7-8 cycles of (98°C for 15 s, 72°C for 45 s), and 72°C for 5 min. 6) Pool amplicons from these 4 reactions then size-select and purify twice using Agencourt AMPure XP beads at 0.8 ratio. 7) Quantify the purified second step PCR library using real-time PCR with Illumina Library Quantification kit (Kapa Biosystems). ****Next-generation sequencing**** Sequencing should be performed at a core facility with experience working with amplicon-based sequencing. To identify both gRNA on each construct 75 bp paired-end reads are required. Given the lack of base diversity in the amplicon it is necessary to spike in ~20% PhiX when sequencing using Illumina HiSeq Rapid run. ****Processing of paired-end reads, fitness estimation, and calculation of interaction scores**** A software pipeline to analyze the raw sequencing data is available here: <http://ideker.ucsd.edu/papers/rsasik2017/> Below is a brief description of how the pipeline processes the raw data. Inputs to the pipeline are FASTQ files for two replicates at Day 3, Day 14, Day 21, and Day 28 time points. 1) Trim FASTQ files of scaffold sequence using cutadapt16, discarding trimmed reads of unexpected lengths <19 or >21 bps. 2) Truncate remaining reads to 19 bases from the appropriate end, reverse-complement reads from reverse strand. 3) Check both reads in a pair for sequence matches against gRNA sequences used in the library, allowing one mismatch anywhere in a read. 4) Aggregate read pairs that matched to a known construct to compute the total counts for each construct in the relevant sample. 5) Fit linear regression to each gRNA to determine fitness, removing data points that fall below minimum threshold of raw sequencing reads. 6) Rank gRNA for each gene by absolute fitness effect and assign weights. 7) Combine the weighted values for the 3 gRNA per gene to determine fitness effect for each gene 8) Determine genetic interactions from measured fitness of double gRNA constructs and expected double knockout fitness

calculated from single gene fitnesses $(f_c = f_p + f_p + \pi_{pp})$ where π_{pp} is the interaction, f_c is the measured double gRNA fitness, and f_p and f_p are the single gene effects.

References

1. Mali, P., Esvelt, K.M. & Church, G.M. Cas9 as a versatile tool for engineering biology. *Nat Methods* 10, 957-63 (2013). 2. Cong, L. et al. Multiplex genome engineering using CRISPR/Cas systems. *Science* 339, 819-23 (2013). 3. Jinek, M. et al. RNA-programmed genome editing in human cells. *Elife* 2, e00471 (2013). 4. Gilbert, L.A. et al. Genome-Scale CRISPR-Mediated Control of Gene Repression and Activation. *Cell* 159, 647-61 (2014). 5. Shalem, O. et al. Genome-scale CRISPR-Cas9 knockout screening in human cells. *Science* 343, 84-7 (2014). 6. 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). 7. Konermann, S. et al. Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex. *Nature* 517, 583-8 (2015). 8. Wong, A.S.L. et al. Multiplexed barcoded CRISPR-Cas9 screening enabled by CombiGEM. *Proceedings of the National Academy of Sciences of the United States of America* 113, 2544-2549 (2016). 9. Bassik, M.C. et al. A systematic mammalian genetic interaction map reveals pathways underlying ricin susceptibility. *Cell* 152, 909-22 (2013). 10. Horn, T. et al. Mapping of signaling networks through synthetic genetic interaction analysis by RNAi. *Nat Methods* 8, 341-6 (2011). 11. Laufer, C., Fischer, B., Billmann, M., Huber, W. & Boutros, M. Mapping genetic interactions in human cancer cells with RNAi and multiparametric phenotyping. *Nat Methods* 10, 427-31 (2013). 12. Wong, A.S. et al. Multiplexed barcoded CRISPR-Cas9 screening enabled by CombiGEM. *Proc Natl Acad Sci U S A* 113, 2544-9 (2016). 13. Sanjana, N.E., Shalem, O. & Zhang, F. Improved vectors and genome-wide libraries for CRISPR screening. *Nat Methods* 11, 783-4 (2014). 14. Doench, J.G. et al. Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. *Nat Biotech* 34, 184-191 (2016). 15. Chari, R., Mali, P., Moosburner, M. & Church, G.M. Unraveling CRISPR-Cas9 genome engineering parameters via a library-on-library approach. *Nat Meth* 12, 823-826 (2015). 16. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.Journal* 17, <http://dx.doi.org/10.14806/ej.17.1.200> (2011).

Figures

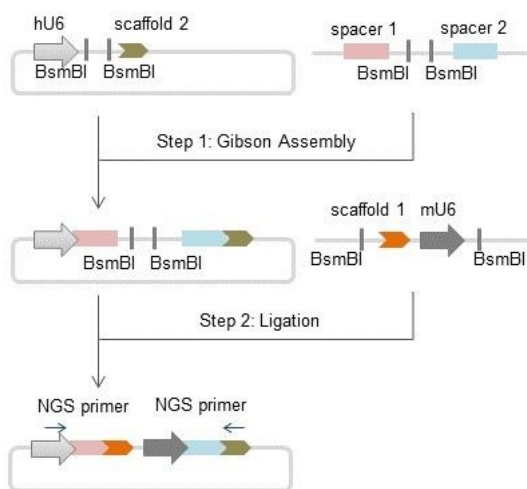


Figure 1

Schematic of cloning strategy for dual-gRNA library synthesis Preparation of the dual-gRNA library involves a two-step cloning process whereby each synthesized oligonucleotide is assembled progressively with promoters and 3' gRNA scaffolds.