Title

Rapid Assembly of Multiplex Natural CRISPR Arrays

ISBN

9781071616567

Authors

Cooper, Robert M

Hasty, Jeff

Publication Date

2021

DOI

10.1007/978-1-0716-1657-4_6

Rapid Assembly of Multiplex Natural CRISPR Arrays

Robert M. Cooper[1,*], Jeff Hasty[1,2,3]

[1]BioCircuits Institute, University of California, San Diego, La Jolla, CA
[2]Molecular Biology Section, Division of Biological Sciences, University of California, San Diego, La Jolla, CA
[3]Department of Bioengineering, University of California, San Diego, La Jolla, CA
[*]For correspondence email rcooper@ucsd.edu

Running head: CRISPR Array Assembly

Rapid Assembly of Multiplex Natural CRISPR Arrays

Abstract

One of the key advantages of CRISPR-Cas systems for biotechnology is that their nucleases can use multiple guide RNAs in the same cell. However, multiplexing with CRISPR-Cas9 and its homologs presents various technical challenges, such as very long synthetic targeting arrays and time-consuming assembly. Recently, other CRISPR associated, single-effector nucleases such as Cas12a have been shown to process their own CRISPR arrays, enabling the use of much more compact natural arrays. Unfortunately, these highly repetitious arrays can be difficult to synthesize commercially or assemble in the lab. Here, we describe a simple method to accurately assemble completely natural, multiplex CRISPR arrays that can be completed in 1-2 days. This should be of great use both in prokaryotes with their own native CRISPR systems, and in eukaryotes when paired with Cas12a or other CRISPR nucleases that also process their own arrays.

Key Words: CRISPR, CRISPR-Cas, CRISPR arrays, multiplexing, Cas12a

Introduction

Natural CRISPR-Cas systems are inherently multiplex. Bacteria and archaea can encode hundreds of targeting spacers in very compact CRISPR arrays ***(1)***. However, scaling up CRISPR for use by humans has been more difficult ***(2)***. The majority of early work has used the single effector nuclease Cas9. Cas9 itself is very simple to port to other organisms, because it requires only a single gene. However, the simplicity of the coding gene comes at the expense of greater sequence length and complexity for the targeting array. Cas9 does not process its own arrays and requires a trans-activating CRISPR RNA (tracrRNA), so to port it to other organisms, scientists usually use synthetic tracrRNA-guide RNA (gRNA) fusions called single guide RNAs (sgRNAs), which are each expressed from an independent transcriptional unit. The resulting

array complexity rapidly becomes a problem when using more than one guide RNA ***(2)***. Performing multiplex targeting with Cas9 often requires many cloning steps and/or long sgRNA arrays that can exceed the length capacity of viral vectors. Another chapter in this book covers one solution for more compact multiplexing using the CRISPR-Cas9 system.

More recently, the CRISPR-Cas12 system was shown to process its own CRISPR array using the same single enzyme cleaves its target ***(3)***. This system allows the best of both worlds for synthetic multiplexing applications – a compact single gene paired with a compact natural CRISPR array ***(4-6)***. Unfortunately, the eponymous palindromic repetition of natural CRISPR arrays makes longer multiplex arrays difficult for commercial providers to synthesize and for individual researchers to assemble ***(2)***. Thus, while Cas12 solves the array length problem of synthetic Cas9 systems, multiplexing with longer natural CRISPR arrays has still required either time-consuming cloning with each spacer added to the array one at a time ***(7-9)***, or sequence modifications to the ends of the spacers ***(10)***.

Here, we describe a technique that can accurately assemble a multiplex natural CRISPR array in just one day ***(11)***. The technique requires no sequence modifications and uses only standard-length DNA oligos. We have used this strategy to assemble multiplex CRISPR arrays of up to 9 spacers and demonstrated them in bacteria, including arrays from both a Type I-F CRISPR system and a Cas12a system ***(11)***.

The key insight of our method is that it assembles only the top strand of the array using ligation, and then later fills in the bottom strand using PCR (Figure 1). During annealing and ligation, the top strand oligos are joined by shorter bottom oligos that only cover the spacer regions. This restricts ligation junctions to the unique spacer regions of the array, while leaving single-stranded gaps that cover the repeat portions of the array. In this way, the method avoids

incorrect annealing, ligation junctions, or spacer order, which could otherwise result from annealing between repeat regions.

Materials

1. DNA oligos for array assembly, as described in Methods. Standard quality desalted oligos in TE buffer or water have worked for us.
2. T4 DNA polynucleotide kinase.
3. T4 DNA ligase with buffer.
4. High-fidelity DNA polymerase (we used Q5 from New England Biolabs).
5. DpnI (if using a plasmid vector).
6. PCR tubes.
7. PCR thermocycler.
8. DNA electrophoresis machine for running gels.
9. PCR purification kit (we used Qiagen).
10. Gel purification kit (we used Qiagen).
11. Depending on your strategy for insertion into the vector, one of the following:
    a. BsaI or another Golden Gate Assembly-compatible restriction enzyme.
    b. Gibson Assembly master mix.
12. Vector template, which can be either a plasmid or linear DNA.
13. Competent cells.

Methods

1. Prepare your vector. One vector compatible with a broad range of hosts that we have had success with is pBAV1k (Addgene #26702) {Bryksin:2010js}. For plasmids, PCR the plasmid with compatible Golden Gate adaptors ***(12)***. If using the restriction enzyme BsaI, append the Golden Gate adaptor sequence 5'-TTTGGTCTCA-3' to the 5' end of each

primer (See Note 1). For the primer adjacent to the beginning of the array, after the Golden Gate adapter add the reverse complement of the first 4 bases of the CRISPR repeat. For the primer adjacent to the end of the array, add the last 4 bases of the final spacer and then the full CRISPR repeat, after the Golden Gate adapter and before the vector sequence (see Note 2, Figures 1, 2, and Table 1). Check your PCR on a gel, and if it looks good, purify it with a PCR purification kit. If the PCR product is significantly different in size from the parent plasmid, you can gel extract the product to separate it from the parent plasmid and reduce background when cloning.

2. Design oligos to use in assembling your CRISPR array (Figures 1, 2, Table 1). For an array of *n* spacers, you will need *n* top oligos and *n* bottom oligos. Bottom oligos should simply be the reverse complement of each spacer, followed by the reverse complement of the last 4 bases of the repeat at their 3' ends (See Note 2). The bottom oligo for the final spacer in the array should also include a Golden Gate adaptor sequence at its 5' end. All top oligos except the first should begin halfway through one spacer, span the repeat, and end halfway through the next spacer. The first top oligo should begin at the first repeat, end halfway through the first spacer, and include the Golden Gate adaptor at its 5' end. Order standard desalted oligos and normalize to 100 µM in elution buffer, TE, or water.
3. Phosphorylate top oligos. Mix 1-2 µl of each top oligo (from 100 µM stock solutions), 1 µl T4 polynucleotide kinase, and T4 ligase buffer to 1x (See Note 3). Incubate at 37 °C for an hour. Alternatively, you could order 5' phosphorylated top oligos.
4. Anneal oligos. Mix 2-6 µl of each bottom oligo, and then combine 1 part phosphorylated top oligos with 2-3 parts bottom oligos in a PCR tube. Heat to 85 °C in a thermocycler, and then slowly cool back to 37 °C at 0.1 °C per second (See Note 4).

5. Ligate oligos. Add 1 µl T4 DNA ligase and fresh T4 DNA ligase buffer to 1x. Incubate at 37 °C for about an hour. Leaving the ligation overnight is fine.
6. Remove unligated oligos. Purify the ligation using a PCR purification column.
7. Fill in the bottom strand and amplify. PCR the ligation using the first top oligo and final bottom oligo as primers. We used Q5 DNA polymerase, annealed at 72 °C (see Note 5), extended for 20 seconds, and ran for 20 cycles.
8. Purify the PCR product. For smaller, easier assemblies, purify the product using a PCR purification kit. For higher accuracy on difficult assemblies, instead run the ligation on a gel (after diluting to avoid overloading the wells), cut out the correct band, and purify the DNA using a gel extraction kit. If in doubt, run a test gel, and use gel extraction if the intended band is not the only clear product.
9. Insert the array into a vector. Combine 4 µl total of the vector and the PCR product at equimolar concentrations, 0.25 µl T4 DNA ligase, 0.25 µl BsaI, and 0.5 µl T4 DNA ligase buffer in a PCR tube. If your vector PCR came from a plasmid, also add 0.25 µl DpnI to cleave the parental plasmid. Incubate for 30-50 cycles of 1 minute each at 37 °C and 24 °C, followed by 10 minutes at 50 °C to inactivate the enzymes. If you prefer to use Gibson Assembly ***(13)*** to insert the array into your vector rather than a Golden Gate strategy, see Note 6.
10. If your vector is linear DNA, PCR amplify the final product.
11. Transform the product into your competent cells using a protocol appropriate for those cells, and grow clonal transformants.
12. Pick several clones, extract their DNA using a protocol appropriate for your cells, and PCR and sequence across the array to verify correct assembly. For a representative

screening PCR of clonal arrays, see Figure 3. In this example, 11 of 16 clones had the correct number of spacers. Sequencing showed all of those 11 were assembled in the correct order. Seven of those were completely correct, and the remainder had small insertions, deletions, or substitutions.

Notes

1. The Golden Gate adaptor sequence 5'-TTT GGTCTC A-3' consists of 3 parts. The first three Ts simply extend the end of the DNA to help the restriction enzyme find its target site, and they could be replaced with any sequence. Here, we used BsaI with target site GGTCTC, but any other Golden Gate-compatible restriction enzyme would work as well. The final A is a spacer required because of the restriction enzyme's offset cutting site.
2. The exact end points of the assembled array are not critically important, so long as they provide unique ligation junctions for insertion into the vector. In the design provided here, the final repeat of the array is included in the vector PCR to reduce the length of the array to be assembled. The bottom oligo for the final spacer extends 4 bases into the repeat at its 3' end to provide a 20-base annealing sequence for the primer in the PCR amplification step. Our spacers were 32 base pairs long, and only half of each spacer is included in the top oligo, so we added 4 bases to the bottom oligo to reach an annealing length of 20 base pairs (see Figures 1, 2). If your spacers are longer or shorter, you should adjust the extension of the bottom oligo into the repeat to ensure a 20 base annealing region for PCR. This is only important for the final spacer in the array, but we suggest ordering all bottom oligos with the same design to make them compatible with potential alternate array designs you may wish to assemble.

3. T4 polynucleotide kinase buffer generally omits ATP to allow users to supply their own radiolabeled version. T4 ligase buffer works as well and does not require additional ATP. Without ATP, the kinase will not work.
4. If your thermocycler cannot be programmed for a slow cooling step, you could heat a volume of water to near boiling, place the PCR tube containing the oligos in it, place it in a 37 °C water bath, and let it slowly come to equilibrium.
5. A high annealing temperature is critical for accurate amplification in this step. You can check the recommended annealing temperature for your primers when using Q5 at https://tmcalculator.neb.com/. If using another DNA polymerase, check the maximum allowed annealing temperature for your primers. Note also that using too many PCR cycles can make the PCR product less clean.
6. We have also successfully used Gibson Assembly to insert assembled arrays into their vectors. We find Golden Gate to be more accurate than Gibson Assembly in general, but both can work. The Gibson variation uses the same top strand-only ligation strategy to assemble the actual array; it just uses a different method to insert the array into a final vector. To use the Gibson method, you will need to prepare your vector differently in Methods Step 1, slightly change your oligo designs in Methods Step 2, and use a different vector insertion method in Methods Step 9.
    a. In Methods Step 1, the forward primer for the vector (at the end of the CRISPR array) should begin just after the terminal CRISPR repeat in your final design. The reverse primer for the vector (at the beginning of the array) will begin just before the initial repeat. Depending on the length of the repeat units in your array, you can extend the primers slightly into the terminal repeats to ensure a 20-base

overlap with the assembled array from Methods Step 8, for the final Gibson Assembly (see also below). Just be sure not to extend these overlaps so far into the repeats that the vector primers would anneal to each other.

b. In Methods Step 2, you will now need *n*+1 top oligos. The top oligo at the beginning of the array should begin 20 bases into the adjacent vector sequence, span the initial repeat, and end halfway through the first spacer. The top oligo at the end of the array should begin halfway through the last spacer, span the terminal repeat, and extend 20 bases into the adjacent vector sequence. The final bottom oligo should not include a Golden Gate adaptor sequence. If desired, you can reduce the top oligo overlaps with the vector sequence to avoid overly long oligos, and instead place the overlaps on the vector primers as described above.

c. In Methods Step 9, use Gibson Assembly to insert the assembled array into your vector. Combine 2 µl total of vector and array DNA at equimolar final concentrations in a PCR tube. Place in a thermocycler block preheated to 50°C and add 2 µl of 2x Gibson Assembly master mix. Incubate at 50°C for 1 hour.

References

1. K.S. Makarova, Y.I. Wolf, O.S. Alkhnbashi, et al. (2015) An updated evolutionary classification of CRISPR–Cas systems, Nature Reviews Microbiology. 13, 722–736.
2. B. Adiego-Pérez, P. Randazzo, J.M. Daran, et al. (2019) Multiplex genome editing of microorganisms using CRISPR-Cas, FEMS Microbiology Letters. 366, 2307.
3. I. Fonfara, H. Richter, M. Bratovič, et al. (2016) The CRISPR-associated DNA-cleaving enzyme Cpf1 also processes precursor CRISPR RNA, Nature. 532, 517–521.
4. Y.E. Tak, B.P. Kleinstiver, J.K. Nuñez, et al. (2017) Inducible and multiplex gene regulation using CRISPR–Cpf1-based transcription factors, Nature Methods. 14, 1163–1166.
5. B. Zetsche, M. Heidenreich, P. Mohanraju, et al. (2017) Multiplex gene editing by CRISPR–Cpf1 using a single crRNA array, Nature Biotechnology. 35, 31–34.
6. C.C. Campa, N.R. Weisbach, A.J. Santinha, et al. (2019) Multiplexed genome engineering by Cas12a and CRISPR arrays encoded on single transcripts, Nature Methods. 16, 887–893.

7. A.A. Gomaa, H.E. Klumpe, M.L. Luo, et al. (2014) Programmable Removal of Bacterial Strains by Use of Genome-Targeting CRISPR-Cas Systems, mBio. 5, e00928–13–e00928–13.
8. M.L. Luo, A.S. Mullis, R.T. Leenay, et al. (2015) Repurposing endogenous type I CRISPR-Cas systems for programmable gene repression, Nucleic Acids Research. 43, 674–681.
9. B.F. Cress, Ö.D. Toparlak, S. Guleria, et al. (2015) CRISPathBrick: Modular Combinatorial Assembly of Type II-A CRISPR Arrays for dCas9-Mediated Multiplex Transcriptional Repression in E. coli, ACS Synthetic Biology. 4, 987–1000.
10. C. Liao, F. Ttofali, R.A. Slotkowski, et al. (2019) Modular one-pot assembly of CRISPR arrays enables library generation and reveals factors influencing crRNA biogenesis, Nature Communications. 10, 2948.
11. R.M. Cooper and J. Hasty (2020) One-Day Construction of Multiplex Arrays to Harness Natural CRISPR-Cas Systems, ACS Synthetic Biology. 9, 1129–1137.
12. C. Engler, R. Kandzia, and S. Marillonnet (2008) A One Pot, One Step, Precision Cloning Method with High Throughput Capability, PLoS ONE. 3, e3647.
13. D.G. Gibson (2011) Enzymatic assembly of overlapping DNA fragments, Methods in enzymology. 498, 349–361.