Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

CleaveSeq: Scalable characterization of ribozyme-based RNA biosensors

Brent Townshend (bst@stanford.edu)
Stanford University https://orcid.org/0000-0003-1927-6341

Method Article

Keywords: aptamer, biosensor, ribozyme, RNA

DOI: https://doi.org/10.21203/rs.3.pex-1346/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Abstract

Biosensors are key components in engineered biological systems, providing a means of measuring and acting upon the large biochemical space in living cells. CleaveSeq, described here, allows high-throughput characterization aptamer-coupled ribozyme libraries.

Introduction

DRIVER and CleaveSeq are methods for generation and characterization respectively of biosensors based on a self-cleaving ribozyme coupled to an RNA aptamer[1]. This protocol describes the steps of CleaveSeq that can be used to characterize a library of biosensors under multiple conditions. The number of distinct sequences that can be characterized in parallel are limited only by the number of NGS reads dedicated to this – typically thousands of sequences can be characterized using a single Illumina run. Each different condition, such as a different chemical addition or concentration, requires a separate tube for the reactions that form the steps of CleaveSeq, with them all combined after barcoding into a single Illumina run. Although the protocol below is written as a manual set of steps, the bulk of the procedure is amenable to operation on a liquid-handling robot.

Reagents

- Chemicals to be tested
- T7 RNA Polymerase@50U/µl (NEB M0251)
- RNApol Buffer@10x (NEB M0251)
- rNTP@25mM/NTP (NEB N0466)
- Omniscript Reverse Transcription Kit (Qiagen 205113)
- T4 DNA Ligase Buffer@10x (NEB M0202)
- T4 DNA Ligase@400U/µl (NEB M0202)
- Thermo Pol Buffer@10x (NEB B9004)
- Uracil-DNA Glycosylase (UDG)@5U/µl (NEB M0280)
- Endonuclease IV@10U/µl (NEB M0304)
- MgCl2@25mM (NEB B9021)
- Salmon Sperm DNA diluted to 100µg/ml (Life Technologies AM9680)
- Hot Start Taq (NEB M0495)

· TE8 (IDT 11-05-01-09)

· Kapa HiFi PCR Kit (Roche KK2101)

· NEBNext Dual Unique Index Barcodes (NEB E6440, E6442, E6444, or E6446)

· Kapa Library Quantification Kit for Illumina (Roche KK4855, KK4824, KK4835, KK4854, or KK4873; depending on real-time PCR cycler used)

· PhiX Control V3 (Illumina FC-110-3001)

· Oligonucleotides

o Library (see Procedure)

o T7Promoter (IDT; AATTTAATACGACTCACTATAGG; PAGE-purified)

o Z-Stop (IDT; /5Phos/G ACAGC GTTTTGTTTTGT CCC AC GCUGT C ACTGGA /iSpC3/ U TTTCTTTTT GCTGT TTC GTCC; PAGE-purified)

o Ref_ZStop_W_R6R30m_X (IDT; GAC AGC GTT TTG TTT TGT CCC ACG CUG TCA CTG GA /iSp18/ UTT TCT TTT TGC TGT TTC GTC CTC GGC ATA GCA TCG CCT AAG TCA TTG ACA CGG ACT CAT CAG ACC GGA ATG ACA TCC AGT GAC AGC TTT GTT TGT TTC CC)

o Ref_ZStop_W_R6R30n_X_cut (IDT; TTT CTT TTT GCT GTT TCG TCC TCG GCA TAG CTC GGC ACA CTT CAT TGA CAC GGA CTC ATC AGA CCG GAA CGT AAT CCA GTG ACA GCT TTG TTT GTT TCC C)

o Ref_ZStop_clvd_R6R30o_X (IDT; /5Phos/GAC AGC GTT TTG TTT TGT CCC ACG CUG TCA CTG GA /iSp18/ UTT TCT TTT TGC TGT TTC GTC CTC GGC ATA GCA AGG ATT CCA TCA TTG ACA CGG ACT CAT CAG ACC GGA AGC ATA TCC AGT)

o Ref_ZStop_clvd_R6R30p_X_lig_cut (IDT; TTT CTT TTT GCT GTT TCG TCC TCG GCA TAG CAG TGA ACC TCT CAT TGA CAC GGA CTC ATC AGA CCG GAA TCG AAT CCA GTG ACA GCG TTT TGT TTT GTC CCA CGC)

o WFU (IDT; ACAC TCTTTCCCTACACGAC GCTCTTCCGATCT GGG AAACAAACAAA GCTG/3SpC3/)

o ZFC (IDT; ACAC TCTTTCCCTACACGAC GCTCTTCCGATCT GCGT GGG ACAAAACAAAAC GC/3SpC3/)

o XRC (IDT; GACTGGAGTTCAGACGTGT GCTCTTCCGATCT TTT CTTTTT GCTGT TTC GTCC/3SpC3/)

o RD1 (IDT; ACAC TCTTTCCCTACACGAC GCTCTTCCGATCT)

Equipment

· PCR Tubes (USA Scientific 1402-4700)

· DNA Lo-bind tubes (Eppendorf 22431021)

· Thermocycler

· Illumina Sequencer

· Real-time PCR Cycler

Procedure

All steps up to and including reverse transcription should be carried out in an RNase-free environment using RNase-free reagents, tubes, and barrier pipette-tips.

Recipes

Reference Mix

1. Mix the following ingredients to form 1000 µl of the reference mix@20nM

a. 992µl TE8

b. 2µl Ref_ZStop_W_R6R30m_X oligo@10µM

c. 2µl Ref_ZStop_W_R6R30n_X_cut oligo@10µM

d. 2µl Ref_ZStop_clvd_R6R30o_X oligo@10µM

e. 2µl Ref_ZStop_clvd_R6R30p_X_lig_cut oligo@10µM

2. Dilute the reference mix to 50pM by mixing 2µl@20nM and 798µl of TE8 in a DNA Lo-Bind Eppendorf tube.

PCR1 Oligo Mix

1. Mix the following ingredients to form 250µl of the PCR1 oligo mixture @4x

a. 225µl TE8

b. 5µl WFU oligo@100µM

c. 5µl ZFC oligo@100µM

d. 5µl XRC oligo@100µM

e. 10µl RD1 oligo@100µM

Library Preparation

1. Prepare library of transcription templates based on the sTRSV ribozyme with chosen loop sequences and preceded by the T7 RNA polymerase's promoter sequence.

2. Synthesize the library as a pool of oligonucleotides in the reverse-complement direction.

3. For example, for the native ribozyme, sTRSV, in this context, use the sequence: < TTTTTATTTTTCTTTTT GCTGT TTC GTCC TCAC GGAC TCATCAG ACCGGA AAGCACA TCCGGT G ACAGC TTTGTTTGTTT CCC TATAGTGAGTCGTATTAAATT >. Desired sequences, which may contain aptamers, replace the loop sequences, TCAC and AAGCACA, above (in reverse-complement orientation).

4. Anneal the library to an equimolar solution of the T7 promoter in the forward direction to allow runoff transcription.

5. Dilute the library to a concentration of 10nM in TE8.

Transcription

1. Prepare enough transcription master mix@2.5x for N reactions, by mixing in the following order:

a. water to bring the total volume to 8N µl

b. 2.0N µl of 10x RNA Pol Buffer

c. 1.8N µl of rNTP@100mM

d. 0.4N µl of DTT@500mM

e. 2.0N µl of T7 RNA Polymerase@50U/µl

2. Add the following ingredients in the listed order in one PCR tube per reaction:

a. water to bring the total volume to 20µl

b. 2 µl of the above library

c. up to 11.8 µl of target chemical(s) (if chemicals are in DMSO, limit to 2µl to keep final DMSO concentration <= 10%)

d. 8µl of 2.5x transcription master mix

3. Vortex briefly and spin down mixtures.

4. Incubate at 37ºC for 30 minutes

5. Spin down products.

Reverse Transcription

1. While transcription is running, prepare 4µl/reaction of RT master mix @2x concentration, by mixing in the following order:

a. 0.72N µl water to bring the total volume to 4N µl

b. 0.8N µl of Omniscript buffer@10x

c. 0.8N µl of Omniscript dNTPs@5mM

d. 1.28N µl of MgCl2@25mM

e. 0.4N µl of Omniscript enzyme@4U/µl

2. Remove 2µl of each transcription reaction to clean PCR tubes.

3. Add 2µl of stop oligonucleotide @4µM to each tube.

4. Vortex briefly and spin down mixture.

5. Add 4µl of RT master mix to each reaction

6. Vortex briefly and spin down mixture.

7. Incubate at 50ºC for 20 minutes.

8. Heat inactivate enzyme and refold cDNA at 95ºC for 2 minutes.

9. Spin down products.

Ligation

1. While reverse transcription is running, prepare 4µl/reaction of ligation master mix@5x concentration, by mixing in the following order:

a. 1.0N µl water

b. 2.0N µl T4 DNA Ligase Buffer@10x

c. 1.0N µl T4 DNA Ligase@400U/µl

2. Add 8µl of water to the same tubes use for the reverse transcription.

3. Add 4µl of the ligation master mix to each reaction.

4. Vortex briefly and spin down mixture.

5. Incubate at 37ºC for 15 minutes

6. Heat inactivate enzyme at 65ºC for 10 minutes.

7. Spin down products.

UDG/Exo digestion

1. While ligation is running, prepare 8µl/reaction of digestion master mix@1.25x, by mixing in the following order:

a. 6.4N µl of water

b. 1.0N µl of Thermo Pol Buffer@10x

c. 0.4N µl of Reference Mix@50pM

d. 0.1N µl of UDG@5U/µl

e. 0.1N µl of Endoncuclease IV@10U/µl

2. Dilute ligation products 25x by adding 6µl of each reaction to a new PCR tube containing 144µl of TE8.

3. Vortex briefly and spin down mixtures.

4. Add 2µl of each reaction to 8µl of digestion master mix@1.25x in a new PCR tube.

5. Vortex briefly and spin down mixtures.

6. Incubate at 37ºC for 15 minutes.

7. Heat inactivate enzymes at 85ºC for 20 minutes.

8. Spin down products.

NGS Adapter Addition

1. While digestion is running, prepare 4µl/reaction of PCR1 master mix@3.5x, by mixing in the following order:

a. 0.37N µl of water

b. 0.08N µl of Thermo Pol Buffer@10x

c. 0.08N µl of MgCl2@25mM

d. 0.08N µl of Salmon Sperm DNA@100µg/ml

e. 0.04N µl of PCR1 oligonucleotide mixture@10x

f. 0.02N µl of Hot-Start Taq@200x

2. Add 4µl of PCR1 master mix to each digestion reaction in the same PCR tube.

3. Vortex briefly and spin down mixtures.

4. Run the following PCR cycling: (PCR1)

a. Activation for 30s at 95ºC

b. Repeat for 5 cycles:

    i. 30s at 95ºC

    ii. 30s at 57ºC

    iii. 30s at 68ºC

c. 60s at 68ºC

d. Hold at 4ºC

5. Dilute each tube 10x in place by adding 126µl of water, vortex briefly and spin down.

6. Spin down products.

NGS Index Addition

1. While PCR1 is running, prepare 16.7µl/reaction of PCR2 master mix@1.5x, by mixing in the following order:

a. 10.4N µl of water

b. 5.0N µl of Kapa HiFi Buffer@5x

c. 0.75N µl of Kapa dNTP's@10mM

d. 0.5N µl of Kapa HiFi Polymerase@50x

2. Also, prepare a 6.25x dilution of the NEBNext Unique Dual Index Primer Pairs by diluting in TE8 and aliquoting 6.25µl/well to new plates for future use.

3. In clean PCR tubes, add the following:

a. 16.7µl of PCR2 master mix@1.5x

b. 6.25µl of a unique dual index primer pair@0.8µM each (this is a 6.25x dilution of the stock dual index primers); each sample should use a distinct primer pair.

c. 2µl of the diluted product from PCR1

4. Vortex briefly and spin down mixtures.

5. Run the following PCR cycling: (PCR2)

a. Activation for 180s at 95ºC

b. Repeat for 14 cycles:

i. 30s at 98ºC

ii. 30s at 64ºC

iii. 30s at 72ºC

c. 60s at 72ºC

d. Hold at 4ºC

6. Perform a 1.8 Ampure cleanup of products using manufacturer's instructions, elute products in 30µl of TE8.

7. Quantify products using Kapa Quantification Kit using manufacturer's instructions

a. Expect each product to be 2-20nM

Sequencing Library Mixdown

1. Mix products above such the number of moles added to the mixture correspond to the relative number of NGS reads that are desired for each condition. For equal coverage the mixture should include the same number of moles of each component. A minimum of 50 usable reads per library entry is required and >100 is preferable. Note that after removal of low-quality reads, single-base errors, references, amplicons, and other undesired sequences, approximately half of the total reads are typically usable.

2. Compute the concentration of the mixture based on the contribution of each component.

3. Dilute the mixture with TE8 to the required concentration for the sequencing platform of choice. For example, for iSeq, dilute to 100pM, for other Illumina platforms dilute to 4nM.

4. Addition of ~10% PhiX to the library is recommended to reduce issues related to amplicon sequencing on Illumina platforms.

Sequencing

Sequencing should be performed using paired-end reads long enough to read the entirety of the longest library entry including prefixes, suffixes and an overlap of at least 10nt for the two reads. Typically, 2x75 is adequate for short loop sequences (<40 nt in the two loops combined), or 2x150 for longer libraries. Both indexes should also be read using 8 cycles each.

Data Analysis

The sequencing data should be demultiplexed into separate fastq files for each of the dual unique indexes by the sequencing platform such as Illumina basespace. Following this, the following steps should be performed:

1. Paired reads can be assembled using PEAR[2]

2. Joined fastq files can be preprocessed to form a text file of total reads of each unique sequence in each file. This can be done using shell scripts (such as fastq2counts.sh from http://github.com/btownshend/NGSDatabase) or other tools.

3. For each sequence in the library, there will be reads of that sequence prefixed by one of two sequences. Reads corresponding to uncleaved RNA will be prefixed the same as they were in the input

library: <GGGGAAACAAACAAA>. Reads corresponding to cleaved RNA will be prefixed with <GGGGACAAAACAAAAC>. For each sequence, find the counts, Nc and Nu, of cleaved and uncleaved reads.

4. Calculate cleavage of each sequence as Nc/(Nc+Nu).

5. Compare cleavage of each sequence under each condition tested. For example, if one condition included no added chemicals and another had a chemical addition, then the ratio of cleavages in those two conditions, which mapped to different fastq files from the demultiplexing, will give the fold change due to addition of the chemical.

Troubleshooting

· Low final concentration may result from failure at any step. Use of negative and positive control throughout the process are recommended. For example, at least one sample run in parallel should be a fixed sequence with known cleavage properties and a no-template control should also be included. Additional no-template controls can be added before PCR1 and PCR2

· If final concentration is slightly lower than the required sequencing concentration, products can be concentrated by a further cleanup step (and subsequent re-quantification) or by evaporation.

· References added during the digestion steps should result in 1-10% of the final reads. The presence or absence of these references in the NGS reads can be used to diagnose problems in the protocol.

Time Taken

A CleaveSeq run, starting from a library of ribozyme-based sensors in the form of the DNA templates, through generation of a library that can be submitted for next-generation sequencing, takes approximately 8 hours. Sequencing time varies with the platform, but generally takes less than one day and analysis of the resulting sequence data approximately 2 hours.

Anticipated Results

See [1] for example results.

References

1. Townshend, B., Xiang, J., Manzanarez, G., Hayden, E. & Smolke, C. *A multiplexed, automated evolution pipeline enables scalable discovery and characterization of biosensors*. http://biorxiv.org/lookup/doi/10.1101/2020.05.29.117960 (2020) doi:10.1101/2020.05.29.117960.

2. Zhang, J., Kobert, K., Flouri, T. & Stamatakis, A. PEAR: a fast and accurate Illumina Paired-End reAd mergeR. *Bioinformatics (Oxford, England)* **30**, 614–20 (2014).