Title

Library construction and imaging for high-throughput image-based screening using multiplexed FISH

Authors

George Emanuel, Jeffrey R. Moffitt, and Xiaowei Zhuang

Laboratory group: Zhuang lab

Abstract

Image-based, high-throughput screening of pooled libraries of genetic perturbations will greatly advance our understanding of biological systems and facilitate many biotechnology applications. Recently, we introduced a method for high-throughput, image-based screening in which the genotypes of individual cells are determined optically using multiplexed FISH to identify expressed RNA barcodes unique to each genotype. Here we describe protocols to construct a library of expression plasmids that express these RNA barcodes and the associated genetic variants and to perform the multiplexed FISH required to identify these barcodes. This protocol is suitable for screening a wide variety of genetic variants in *E. coli* cells and can be adapted to image other cell types.

Introduction

This high-throughput image-based screening methods utilizes optical microscopy to measure both (1) the phenotype produced by each member of libraries of genetic variants and (2) a unique RNA barcode expressed within each cell associated with a given genotype. The method is compatible with a wide range of phenotypes and many different methods for constructing libraries of genetic variants; thus, here we describe only the protocols associated with (i) constructing a library of RNA barcodes, (ii) fusing this library with an existing genetic variant library, (iii) using next-generation sequencing to determine which barcode is associated with each genetic variant, and (iv) using a modified version of multiplex error-robust fluorescence in situ hybridization (MERFISH)[1,2] to optically determine the barcode associated with each cell after the phenotype of that cell has been imaged. We have utilized this method to screen libraries containing tens of thousands of unique genetic variants in tens of millions of *E. coli* cells.

Reagents

Barcode construction and fusion with genetic variant libraries

- Oligonucleotides for barcode assembly (Integrated DNA Technologies; Sequences included in Table 1)
- pZ Series expression plasmids[3] (ExpresSys)
- Plasmid library of genetic variants
- Several sets of PCR primers (the sequences will depend on the expression plasmids used)

- o *Barcode-amplification primers* designed to amplify the DNA containing the barcode sequence and integrate it into the desired expression plasmid via isothermal assembly[4].
- o *Barcode-genetic-variant-fusion primers* designed to extract the expression region for the barcodes from a plasmid barcode library and the expression region from a genetic variant and then create a library of plasmids that fuse these expression cassettes via isothermal assembly. In addition, these primers will introduce a 20-nt stretch of random nucleotides that will serve as a unique molecular identifier (UMI).
- o *Illumina-sequencing-adaptor primers* designed to extract the barcode sequence, genetic variant sequence, and UMI from each plasmid and add the appropriate adaptors to permit Illumina sequencing.
- Nuclease free water
- UltraPure agarose (ThermoFisher, 16500500)
- T4 ligase (NEB, M0202S)
- T4 polynucleotide kinase [PNK] (NEB, M0201S)
- Phusion polymerase (NEB, M0531S0)
- EvaGreen (Biotium, 3100)
- NEBuilder HiFi DNA Assembly Master Mix (NEB, E2621L)
- Turbo electrocompetent *E. coli* (NEB, C2986K)
- Super optimal broth with catabolite repression [SOC] (NEB, B9020S)
- Zymoclean Gel DNA Recovery Kit (Zymo, D4002)
- DNA Clean and Concentrator (Zymo, D4003)
- Zyppy plasmid miniprep kit (Zymo, D4019)
- Electroporation cuvettes (ThermoFisher, P41050)
- Luria-Bertani broth [LB] (Teknova, L8000)
- Carbenicillin (ThermoFisher, 10177-012)
- Glycerol (Teknova, G1796)
- Qubit dsDNA HS assay kit (ThermoFisher, Q32851)
- MiSeq reagent kit v2 (Illumina, MS-102-2003)

Barcode imaging

- Imaging coverslips (Bioptechs, 0420-0323-2)
- LB (Teknova, L8000)
- Carbenicillin (ThermoFisher, 10177-012)
- Chloramphenicol (Sigma, C0378)
- Phosphate-buffered saline [PBS] (ThermoFisher, AM9625)
- Ethanol
- Methanol
- Acetone
- Nuclease free water
- Polyethylenimine [PEI] (Sigma-Aldrich, P3143-500ML)
- Tris(2-carboxyethyl)phosphine [TCEP] (Sigma-Aldrich, 646547-10X1ML)
- Dextran sulfate (EMD Millipore, 3730-100ML)
- Ethylene carbonate (Sigma-Aldrich, E26258-500G)

- Murine RNase inhibitor (NEB, M0314L)
- DNA oligonucleotides labeled with either ATTO565, Cy5, or Alexa750 via a disulfide bond (Bio-Synthesis Inc; Sequences in Table 2)
- Glucose (Sigma-Aldrich, G8270)
- Trolox (Sigma-Aldrich, 238813)
- Glucose oxidase (Sigma-Aldrich, G2133)
- Catalase (Sigma-Aldrich, C100-500mg)

Equipment

Barcode construction and fusion with genetic variant libraries

- Standard molecular biology lab equipment
- Bio-Rad CFX96
- Amaxa Nucleofector II
- Qubit 3.0 Fluorometer
- Illumina MiSeq

Barcode imaging

- Computer-controlled fluidic system (Detailed protocols for construction are provided in Ref. [5])
- Epi-fluorescence microscope (custom built or commercial) with the following components
  - Automatic focus lock
  - Motorized stage
  - 100-500 mW illumination sources at 750 nm, 650 nm, 560 nm
  - Brightfield lamp
  - High numerical aperture objective, e.g. 1.4 NA 60X Objective
  - High performance camera, e.g. EMCCD or sCMOS
  - Optional: Optics to flatten the illumination profile
- Open-source software for integrated microscope and fluidic system control (github.com/ZhuangLab/storm-control)
- Open-source software for the analysis of phenotype and barcode images (http://zhuang.harvard.edu/merfish.html)

Procedure

Barcode construction and fusion with genetic variant libraries

1. Construction of the barcode library
   1.1. Mix all of the barcode construction oligonucleotides listed in Table 1 in water to a final concentration of 100 nM each.
   1.2. Mix 16 μL of the above oligonucleotide mixture, 2 μL T4 ligase buffer, 2 μL PNK in a PCR tube.
   1.3. Incubate this mixture at 37 °C for 30 minutes to phosphorylate any oligonucleotides that may have lost the phosphorylation provided during oligonucleotide synthesis. Incubate at 65 °C for 20 minutes to heat inactivate the PNK.

1.4.   Add 1 μL T4 ligase (NEB, M0202S) to this mixture and incubate this reaction for 1 hour at room temperature to ligate the oligonucleotides forming the DNA template for the RNA barcodes.

1.5.   To amplify this library of DNA templates, dilute the ligation product 10 fold into water and mix 5 μL of this dilution with 25 μL Phusion 2X master mix, 2.5 μL 20X EvaGreen, the barcode amplification primers to a final concentration each of 0.5 μM, and water to a total volume of 50 μL.

1.6.   PCR amplify this reaction via limited cycle PCR.

1.7.   Gel purify this PCR product using the Zymoclean DNA recovery kit and the manufacturer's instructions. Elute the product into 20 μL water.

1.8.   To create the desired backbone vector for barcode expression, PCR amplify 50 μL of the specific backbone vector and gel purify (as in 1.5-1.7). Elute in 20 μL water.

1.9.   Mix 5 μL of the gel purified barcode ligation product, 5 μL of the gel purified barcode expression backbone, and 10 μL HiFi DNA assembly master mix. Incubate for 1 hour at 50 °C to fuse the barcode DNA template with the plasmid backbone, purify using the Zymo DNA clean and concentrator kit according to the manufacturer's instructions, and elute the plasmid library with 6 μL water.

1.10.   To transform this plasmid library into *E. coli*, mix the assembled plasmids with 10 μL of electrocompetent *E. coli* on ice, transfer into an electroporation cuvette on ice, and electroporate using an Amaxa Nucleofector II set to bacteria program 3.

1.11.   Immediately add 1 mL SOC, transfer into a 14mL culture tube, and incubate at 37 °C on a shaker for one hour.

1.12.   Dilute this culture into 50 mL of LB supplemented with 0.1 mg/mL carbenicillin in a sterile, beveled 250 mL flask. Place on the shaker at 37 °C overnight to allow the cells which have taken up a plasmid to grow.

1.13.   Miniprep the overnight culture using the Zyppy plasmid miniprep kit according to the manufacturer's instructions. Elute the final barcode plasmid library into 30 μL Zymo elution buffer and store at -20 °C until needed.

2.   Merging the barcode expression library with an existing genetic variant library

    2.1.   Amplify and gel purify the corresponding halves of the barcode plasmid library and an existing genetic variant library by limited-cycle PCR using the appropriate barcode-genetic-variant-fusion primers (as in 1.5 – 1.7).

    2.2.   Mix 5 μL of each of the purified plasmid library fragments with 10 μL HiFi DNA assembly master mix. Incubate for 1 hour at 50 °C to assemble these PCR fragments, and purify the resulting plasmids using the Zymo DNA clean and concentrator kit according to the manufacturer's instructions. Elute the plasmid library into 6 μL of water.

    2.3.   Transform this library into electrocompetent *E. coli* and then re-isolate this plasmid library via plasmid extraction as described in 1.10 – 1.12.

    2.4.   To ensure that each *E. coli* contains at most one plasmid, dilute this plasmid library to a defined concentration of 100 pg/μL in water and re-electroporate into *E. coli* (as in 1.10 – 1.12).

    2.5.   Measure the OD600 of the overnight culture and dilute this culture to a concentration of ~1000 cells/μL using the conversion of 800 million cells per 1 unit of OD600.

2.6. Inoculate 2mL of LB containing 0.1 mg/mL carbenicillin with a volume of culture containing the number of cells corresponding to the desired library size. (We typically keep a number of cells that is at most 10% of the expected number of unique barcodes in the original barcode library). Incubate at 37 °C overnight on a shaker to grow this culture to saturation.

2.7. Miniprep 600 µL of the overnight culture (as in 1.13) and elute into 30 µL Zymo elution buffer. Store this plasmid library at -20 °C. It will be used as a source for material for next-generation sequencing.

2.8. To archive the *E. coli* containing this barcoded genetic-variant library for later high-throughput image-based screening, mix the remaining overnight culture with an equal volume of 50% glycerol. Separate this mixture into 100 µL single-use aliquots and store at -80 °C.

3. Sequencing the merged genetic variant and barcode expression library to determine which barcode was associated with each genetic variant
   3.1. Dilute the barcoded genetic variant plasmid library (prepared in 2.7) 10-fold in water. To extract the desired regions of this plasmid, where each region contains the UMI, amplify 5 µL of this dilution via limited cycle PCR using the Illumina-sequencing-adaptor primers and gel purify this product (as in 1.5-1.7). Elute the product in 200 µL of water. To address complications associated with sequencing of libraries with identical regions within many different molecules, these primers should contain short regions of random nucleotides of variable length. These variable length regions will introduce apparent variability into these common regions during sequencing.

   3.2. Because the desired region to be sequenced is typically larger than the read length associated with the sequencing platform used, repeat 3.1 using different primer sets that allow amplification of different portions of the desired region of these plasmids. It is crucial that each amplification product contain the UMI region so that these different fragments can be associated with the same plasmid during analysis.

   3.3. To add the adapters necessary for Illumina sequencing, amplify and gel purify 5 µL of this product (as in 1.5-1.7). To limit the degree of amplification, use these primers at a reduced concentration of 0.05 µM. Elute the product into 10 µL water.

   3.4. Quantify the concentration of the elution using the Qubit fluorimeter and dilute to 4 nM in water.

   3.5. Sequence these molecules following the MiSeq protocol from Illumina.

   3.6. Construct the barcode to genetic variant look-up table by grouping sequencing reads by UMI and assigning to each the most frequently occurring barcode and gene variant seen for that UMI.

Barcode imaging

1. Preparing the library for barcode imaging
   1.1. Thaw a 100 µL aliquot of the barcoded genetic-variant library (prepared in 2.8) to room temperature and dilute into 2 mL LB supplemented with 0.1 mg/mL carbenicillin. Incubate on the shaker at 37 °C for 3 hours to allow these cells to grow.

1.2. Optional: If the barcode library is under the control of a plasmid utilizing a relaxed origin, e.g. an origin capable of dividing even when translation is inhibited, barcode expression can be enhanced by inhibiting protein translation. Add 20 μL of chloramphenicol to the culture to a final concentration of 34 μg/mL to inhibit growth. Incubate at 37 °C for an additional 3 hours.

1.3. In parallel, coat an imaging coverslip so that *E. coli* will stick to it by placing the coverslip in a petri dish and covering with 1% PEI in water for 30 minutes. Wash the coverslip once in PBS. Dry the coverslip by removing the PBS.

1.4. To adhere *E. coli* to the coverslip, dilute the above culture 10-fold into PBS and pour onto the coated coverslip. To force *E. coli* cells to lie flat on the surface of the coverslip, spin the petri dish in a centrifuge at 100g for 5 minutes.

1.5. Wash the coverslip once with PBS.

2. Prepare the hybridization and imaging buffers for measurement of the barcode RNAs
   2.1. Prepare aliquots of water, PBS, 2X SSC, and fixation buffer[6] (80% v/v methanol and 20% v/v acetone).
   2.2. Prepare an imaging buffer[7] which will reduce the rate of fluorophore photobleaching during barcode imaging. This buffer is comprised of 50 mM TrisHCl pH 8, 10% w/v glucose, 2 mM Trolox, 0.5 mg/mL glucose oxidase, and 40 μg/mL catalase in 2X SSC.
   2.3. Prepare a cleavage buffer to cleave the disulfide bond linking fluorophores to readout probes, allowing these dyes to be rapidly washed away. This buffer is comprised of 50 mM TCEP in 2X SSC.
   2.4. Prepare a master hybridization mix consisting of 5% w/v dextran sulfate, 5% w/v ethylene carbonate, 0.05% w/v yeast tRNA, and 0.1% v/v Murine RNase inhibitor in 2X SSC.
   2.5. Aliquot 1 mL of this hybridization master mix into 15 mL falcon tubes, making one tube for each round of hybridization. Add to each tube the appropriate set of readout probes (as described in Table 2) to a final concentration each of 10 nM. Each readout probe will determine the presence of a "1" or a "0" in a given bit, and each round can probe the value of multiple bits simultaneously by using multiple readout probes each conjugated to a spectrally distinct fluorescent dye.
   2.6. Load all buffers and hybridization mixes into the fluidics system.

3. Image the phenotypes and the barcodes
   3.1. Assemble the coverslip into the flow chamber following the manufacturer's instruction.
   3.2. Flow PBS into the chamber and place it on the microscope.
   3.3. Proceed with the appropriate high-resolution imaging protocol for the desired phenotype measurement.
   3.4. After phenotype imaging, fix the cells by washing the flow chamber with water to remove salts (that can precipitate upon contact with methanol and acetone) and then flowing on the fixation buffer. Incubate the sample at room temperature for 30 minutes in this buffer. Wash again with water to remove residual fixation buffer.
   3.5. Flow on the hybridization buffer containing the readout probes assigned to the first hybridization round.

3.6.   Incubate at room temperature for 30 minutes to allow the readout probes to hybridize to the complementary barcode RNAs.

3.7.   Flow the imaging buffer onto the sample.

3.8.   Collect an image of each field-of-view with the illumination channels appropriate to the readout probes utilized. For example, for the readout probes listed in Table 2, collect an image of each field-of-view when illuminated with 750-nm, 650-nm, and 560-nm light to image the readout probes conjugated to Alexa750, Cy5, and ATTO565, respectively, in addition to an image with bright-field illumination for image alignment.

3.9.   Collect a brightfield image of this field-of-view which will be used during analysis to correct any small offset of the stage observed between imaging rounds

3.10.   Repeat 3.8-3.9 for as many fields of view as desired.

3.11.   Flow the cleavage buffer onto the sample to extinguish fluorescence from the sample. Incubate for 15 minutes.

3.12.   Flow 2XSSC onto the sample to remove residual cleavage buffer.

3.13.   Repeat 3.5-3.12 for each round of hybridization. For example, readout of a 21-bit barcode will require 42 distinct readout probes which, with 3 color imaging, can be measured in 14 rounds of hybridization.

3.14.   Recommended: The use of integrated microscope and fluid handling software can automate the repetitive process of imaging, fluid handling and flow, and staining of the sample, substantially simplifying this protocol.

4.   Identify the barcode expressed within each cell

4.1.   Correct inhomogeneities in the illumination using a flat-field correction. This correction can be determined by averaging the fluorescence signal observed across many fields of view.

4.2.   Align images from each round of imaging by maximizing the cross-correlation of the bright-field images.

4.3.   Segment cells by determining closed boundaries after applying the Canny edge detection algorithm.

4.4.   For each cell, calculate the mean fluorescence intensity across all constituent pixels in each image.

4.5.   To account for differences in dye brightness and staining between hybridization rounds, normalize the intensity observed for each color channel in each round by the mean intensity observed for that color channel and round across all fields of view.

4.6.   For each bit, determine the ratio of the fluorescence signal observed in each cell for the readout probe associated with a "0" signal to that associated with a "1" signal. This "0"-to-"1" ratio will be used to determine if the barcode associated with each cell should be assigned a "1" or "0" in a given bit.

4.7.   To determine the best bit-calling-threshold for assigning a bit value from the "0"-to-"1" ratio, select ~100 barcodes known to be present in the sample via next-generation sequencing. Determine the frequency with which these barcodes are detected as a function of this bit-calling-threshold for each bit. Select the set of thresholds that maximize this number.

4.8.   For each cell, determine the value of the measured barcode using these thresholds and the measured "0"-to-"1" ratios.

4.9. Discard cells with measured barcodes that were revealed by next-generation sequencing to have been assigned to multiple genetic variants

4.10. Recommended: Some cells will appear dim in both the "1" and "0" channels associated with a given bit, potentially leading to incorrect calling of this bit. Thus, to increase the accuracy, cells with a "1" or "0" brightness below a given threshold can be discarded from subsequent analysis.

4.11. Group cells together that are expressing the same barcode to determine the average properties of the measured phenotype.

References

1. Chen, K. H., Boettiger, A. N., Moffitt, J. R., Wang, S. & Zhuang, X. Spatially resolved, highly multiplexed RNA profiling in single cells. *Science (80-.).* **348,** aaa6090-aaa6090 (2015).

2. Moffitt, J. R. *et al.* High-throughput single-cell gene-expression profiling with multiplexed error-robust fluorescence in situ hybridization. *Proc. Natl. Acad. Sci.* **113,** 11046–11051 (2016).

3. Lutz, R. & Bujard, H. Independent and tight regulation of transcriptional units in Escherichia coli via the LacR/O, the TetR/O and AraC/I1-I2 regulatory elements. *Nucleic Acids Res.* **25,** 1203–10 (1997).

4. Gibson, D. G. *et al.* Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* **6,** 343–345 (2009).

5. Moffitt, J. R. & Zhuang, X. RNA Imaging with Multiplexed Error-Robust Fluorescence In Situ Hybridization (MERFISH). in 1–49 (2016). doi:10.1016/bs.mie.2016.03.020

6. Shaffer, S. M., Wu, M.-T., Levesque, M. J. & Raj, A. Turbo FISH: A method for rapid single molecule RNA FISH. *PLoS One* **8,** e75120 (2013).

7. Rasnik, I., McKinney, S. A. & Ha, T. Nonblinking and long-lasting single-molecule fluorescence imaging. *Nat. Methods* **3,** 891–893 (2006).

Associated publications

George Emanuel, Jeffrey R. Moffitt, and Xiaowei Zhuang, High-throughput, image-based screening of pooled genetic variant libraries

Software for sequence and image analysis can be found at github.com/ZhuangLab.

Author information

Affiliations

Howard Hughes Medical Institute

George Emanuel, Jeffrey R. Moffitt, Xiaowei Zhuang

Graduate Program in Biophysics, Harvard University, Cambridge, MA 02138, USA

George Emanuel

Department of Chemistry and Chemical Biology, Harvard University, Cambridge, MA 02138, USA

George Emanuel, Jeffrey R. Moffitt, Xiaowei Zhuang

Department of Physics, Harvard University, Cambridge, MA 02138, USA

George Emanuel, Jeffrey R. Moffitt, Xiaowei Zhuang

Correspondence should be addressed to Xiaowei Zhuang (zhuang@chemistry.harvard.edu) and Jeffrey R. Moffitt (lmoffitt@mcb.harvard.edu)