**Inverted GFP expressed lentiviral barcoding plasmid**

The map and protocols can be found in benchling <https://benchling.com/s/seq-F1D5aW7t9lBn3q8oywBg>

Barcodes are cloned using the NdeI and/or BamHI sites.

Barcode structure (in the sense of the lentiviral construct and plasmid): NNNN**CT**NNNN**AC**NNNN**TC**NNNN**GT**NNNN**TG**NNNN**CA**NNNN

**Barcoded Library Retransformation Protocol**

(Weinreb, Rodriguez-Fraticelli et al. *bioRxiv* 2018)

Day 0/1:

1) Dilute 6 liters of LB-Agar (32 grams per liter of distilled water). Autoclave.

2) Prepare 20 large LB-Agar Ampicillin plates 25x25 by adding 250 ml of Agar per plate.

3) Let the plates solidify/dry for 24-48h at room temperature with the lid semi open.

4) Chill some pipette tips and tubes by placing them in the cold room overnight.

Day 2:

1) Make sure the plates are ready by spreading 50 ul of SOC on one of them. The SOC should be “absorbed” within a minute.

In the cold room:

2) Dilute library DNA (from midi/maxi) with water at 20 ng/ul (make at least 20 ul)

3) Chill 10 eppendorf tubes and 10 electroporation cuvettes (0.1 cm width, Thermo Fisher) on ice.

4) Add 20 ng of library plasmid DNA in each eppendorf tube.

5) Thaw 250 ul of Stbl4 ElectroMAX competent cells on ice.

6) Mix competent cells by gentle tapping and immediately pipette 20-25 ul of cells per eppendorf, directly on top of the aliquoted DNA.

7) Gently pipette the cell/DNA mixture up and down and then transfer each mix into a separate electroporation cuvette. Take the cuvettes on ice to the electroporator.

At the electroporator:

8) Take one cuvette out from the ice and make sure cuvette electrodes are dry and do now form condensation.

9) Electroporate the cuvette using BioRad GenePulser at 1200 V, 25 µF, 200 Ω and then immediately add 1 ml of warm SOC medium to it, resuspend the mix slowly up and down and transfer into a 10 ml Falcon snap-cap tube and leave at room temp.

10) Repeat for each of the remaining 9 cuvettes.

11) Incubate the tubes at 37ºC for 1h. Agitation is recommended (250 rpm).

12) Spread the content of each tube into 2 Large LB-Agar plates. Incubate for 24h at 30-32ºC.

Day 3:

1) Make sure that small, even, round colonies have formed. Using a ruler, count the colonies in ten “randomly" selected 1 cm x 1 cm squares. The total number of those ten squares summed should be close to 800 colonies. Do not let the plates overgrow or overdry.

2) While you count, prepare and prewarm some LB supplemented with Ampicillin at 37ºC.

3) Add 20 ml of pre-warmed LB supplemented with Ampicillin per plate, incubate for 10 minutes at room temp, and then scrape off the colonies with a bacterial spreader. Resuspend colonies in the plate and transfer to a 3L culture flat erlenmeyer with 100 ml of LB-Amp.

4) Repeat this process 3 times per plate to ensure that most colonies are resuspended and rinsed off from the plate.

5) The final culture will contain ~1.5 L of LB-Amp and should look like a well-grown Maxiprep. Mix well by agitation by hand.

6) Transfer 500 mL of this culture into 3 different flasks and incubate at 37ºC for 2-3h to allow bacteria to recover. Do not let O.D. (600 nm) be larger than 3. At this point, library can be glycerol-stocked in 10 ml Falcon tube aliquots (5 ml Culture + 5 ml Glycerol) and frozen at -20.

7) Maxiprep the entire 1.5L of bacterial culture with Maxi kit (Macheray Nagel). Yield should be in the range of 0.5-1.5 mg of DNA for the entire culture.

8) For library regrowth from glycerol stocks, add an entire 10 ml Falcon tube aliquot per 1L LB-Amp culture and grow for 6h-8h during the day (never allow O.D. to be larger than 2-3).

**Lentivirus production and barcode labeling (modified from** [**http://www.bu.edu/dbin/stemcells/files/Protocol%201-%20Lentivirus%20Packaging%20by%20293T%20Transfection.pdf**](http://www.bu.edu/dbin/stemcells/files/Protocol%201-%20Lentivirus%20Packaging%20by%20293T%20Transfection.pdf)**)**

Barcoded GFP plasmid and third generation lentivirus components were transfected into HEK293T Lenti-X cells using the Trans-IT 293 kit. Lentivirus was harvested every 12 hours for 72 hours and concentrated into 100 fold using ultra-centrifugation at 80,000g. HEK293T Lenti-X cells were grown in DMEM with 10% fetal bovine serum (FBS) and 1% PenStrep. Viral titers are estimated from transducing 10E6 293T cells with 0.1, 0.5, 1, 2 and 5 ul of concentrated virus and measuring %GFP after 3 days. Hematopoietic stem and progenitor cells (HPCs) are transduced using spin infection (800g for 90 minutes at 30ºC) with 20-50 MOI and incubated for 8h.

***InDrop RNA barcode amplification:***

For single cell RNAseq, use a fraction (typically 1/4th) of the in vitro transcribed RNA (IVT) for GFP-specific RT with the following primer (Tm 60):

TGAGCAAAGACCCCAACGAG

Use PrimeScript as the RT kit of choice and purify with SPRI beads into 10 ul water.

Then amplify the eluted cDNA with the following primers for 6-7 cycles (adjust using qPCR):

Fw (R1-GFP primer): TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGNNNNCGTTGCTAGGAGAGACCATATG

- purple: R1 adapter for v3 inDrops

- blue: diversity generator for sequencing cluster detection

- green: barcode amplification Fw sequence

Rv (InDrop R2 primer): CAAGCAGAAGACGGCATACGAGATGGGTGTCGGGTGCAG

Then index the resulting product with the indexed R1 primer for 6-7 cycles (adjust using qPCR):

Fw (R1-index primer):AATGATACGGCGACCACCGAGATCTACAC(libraryID)TCGTCGGCAGCGTC

Rv (InDrop R2 primer): CAAGCAGAAGACGGCATACGAGATGGGTGTCGGGTGCAG

When sequencing, we typically just need about 1/10th of a lane for every 10,000 cell transcriptomes. We use the same InDrop indexes (R1) for both whole transcriptome and barcode transcripts, and mix them into the same sequencing lane. Then sequence according to the usual sequencing guidelines for InDrop v3 libraries (https://projects.iq.harvard.edu/files/singlecellcore/files/sequencing\_indrops\_libraries\_06\_20\_2018.pdf).

***Genomic amplification:***

For amplification, from 10-100 ng of pure genomic DNA (100% labeled), run 10 PCR cycles (+3 cycles if starting from 10 ng DNA), using the following primers:

ctgagcaaagaccccaacgagaa

gaaggcacaggtcgacaccagt

With the HiFi Kapa kit, do 98ºC 2 min, 98ºC 10 sec, 55ºC 25 sec, 72ºC 30 sec, 72ºC 5 min, 4ºC.

Purify the entire reaction yield in 25 ul dH2O and use 1/2 of the eluate as a template for a second PCR, same parameters, except anneal at 58ºC, same 10 cycles:

ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNNNGAGTAACCGTTGCTAGGAGAGACCATAT

GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTcacaggtcgacaccagtctcatt

Purify the PCR product in to 25 ul and use up to 10 ng of the eluate (typically about 1 ul) as a template for a third PCR to add Illumina sequencing adapters. Usually 8 cycles should be enough.

With the HiFi Kapa kit, do 98ºC 2 min, 98ºC 10 sec, 65-60ºC20 sec, 72ºC 30 sec, 72ºC 5 min, 4ºC.

(P5): AATGATACGGCGACCACCGAGATCTACAC(index5)ACACTCTTTCCCTACACGACGCTCTTCCGATCT

(P7): CAAGCAGAAGACGGCATACGAGAT(index7)GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT

***Library diversity testing:***

The same protocol should be followed for library diversity testing, but one can skip the first primer step and proceed directly to the second PCR step. PCRs should be tested by qPCR to avoid library amplification bias and saturation. PCR reactions should always be stopped at the beginning of the exponential phase, about 3 cycles after it begins. A mock reaction can be run with SYBR to find the exact number of cycles, and then the adequate number of cycles is chosen for the sample that will be sequenced.

We typically do 2 replicate reactions and count only the sequences are observed in both. The sequencing depth should be at least 5 million reads per replicate. Once sequencing fastqs are demultiplexed, we run a python script, available at <https://www.github.com/Rodriguez-Fraticelli/LARRY_whitelist_v2>. For usage of the script, see the github link.