Lessons from the Single Cell Core

Sarah Boswell
Senior Advisor

http://iccb.med.harvard.edu/single-cell-core
sarah_boswell@hms.harvard.edu

Armenise 561
Overview

- Working with the Single Cell Core (SCC)
- Experimental Design
- Sample Preparation
- Library Preparation & Sequencing
- Cost Calculations
- Basic Data QC
Credits

SCC inDrops Technical Expert

鹡鸰 Alex Ratner 鹡鸰

Klein Lab

Allon Klein
Rapo Zilionis

ICCB-Longwood

Caroline Shamu

Harvard Medical School
Tools and Technology Grant
Systems Biology Department

Harvard Chan Bioinformatics Core
Shannan Ho Sui
Rory Kirchner
SCC: Klein Lab inDrops Method

- Lysis and reverse transcription occurs in the beads
- Samples are frozen after RT as RNA:DNA hybrid in gel
- Library prep is based on CEL-Seq method

Working with the SCC

weeks to months of optimization
Example Project:
- 12 samples – 1500 cells per library
- 3 days of InDrop encapsulation (4 samples / day)

Estimated sample cost: $5,500
Sequencing cost: $5000

Planning a well thought out experiment is the most important part of a single cell project.
SCC Consult: Experimental Design

- What is the goal of the experiment?
  - Mutant vs control
  - Time course
  - Sorted vs unsorted

- Can you get your answer with bulk sequencing?
  - Sorted vs unsorted
Experimental Design: Numbers

- Population of interest
  - 1% or 10% of cells

- Number of cells available vs needed
  - Minimum ~10-25,000 (to collect 3000)

- Cells at > 90% viability after 1h on ice
  - Dead cells add to background noise
  - Dying cells show up as mitochondrial high and are excluded from analysis.

Need 10-30 cells with unique signature
Optimize dissociation protocol for high viability
  - Shorten time, lower temperature, inhibitors in buffer

Clean-up cells / keep in suspension with Optiprep
  - 15% Optiprep is a good starting point

Enrich for population of interest
  - FACS sort only if absolutely necessary

Sample QC
  - qPCR for housekeeping genes
Mouse dissection in the cold room
Shorten perfusion times
TrypLE, StemPro Accutase, Liberase
Add ROCK / apoptosis inhibitors to mix
Shorten dissociation time
Clean up dead cells with OptiPrep
SCC Consult: OptiPrep

- Density gradient media
- Non-toxic
- Metabolically inert
- Low viscosity and osmolarity

- Cells mixed with OptiPrep to keep in single cell suspension while loading for inDrops.

- 15% OptiPrep is standard, though can vary with cell type.
**SCC Consult: Enrichment FACS**

- Look at unsorted population
- Mechanical dissection of region of interest
- Magnetic-activated cell sorting (MACS)
- Cell strainers (20, 40, 70, 100 μM)
- OptiPrep gradient
SCC Consult: FACS

- Is your cell type durable or fragile?
- Can you sort on a broad marker?
- How quickly can your sort be done?
- What kinds of cell numbers do you get after sorting?
- Prep more samples than you need on day of run!
SCC Consult: Quality Controls

- > 90% (95% preferred) Trypan Blue excluded before and after ½-1h on ice
- Remove clumps by passing cells through mesh strainer: 40 or 20 uM
- Confirm cell concentration – don’t trust FACS
- Check for free floating RNA in solution released from dead cells
Working with the Single Cell Core

weeks to months of optimization
Option 1:

- User bring over dissociated samples on ice
- Can stagger arrival of samples for better viability

Option 2:

- User can bring over samples and perform dissociation in our space. TC room and standard lab equipment available.

**NOTE:** user must bring SSIII and RNaseOUT for samples

- 1 tube SSIII / 7,000 cells
- 1 tube RNaseOUT / 15,000 cells
SCC inDrops: day of your run

- Minimum samples concentration 25,000 cells in 150ul
- 10,000 cells if delivered in Optiprep (~200ul)

- matching the speed of bead injection with the speed of droplet generation it is possible to set conditions in which nearly every droplet would be loaded with one hydrogel bead
Cells encapsulated and RT occurs within droplet
  - Cells immediately encounter lysis buffer
  - Emulsion then divided into number of desired libraries
  - Droplets broken and cDNA stored at -80°C.

Library prep performed every 2 weeks
inDrops Library Prep

- Second Strand Synthesis to make full length dsDNA
- In vitro transcription (IVT) back to RNA off T7 promoter from primer
- Fragmentation RNA
- RT with random hexamer primer containing adaptor
- PCR off adaptors to add index and illumina adaptors

Provided by Klein Lab
inDrops Library Prep

- Library – Qubit concentration and size range
- Representative BioAnalyzer trace of final library

Provided by Klein Lab
Sequencing your inDrops Library

- Biopolymers Facility (NRB)
  https://genome.med.harvard.edu/
- Dana Farber Molecular Biology Core (Fenway)
  https://mbcf.dfci.harvard.edu/
- Bauer Core (Cambridge)
  http://bauercore.fas.harvard.edu/
- Tufts Genomics (Boston)
  http://tucf-genomics.tufts.edu/

Have core qPCR library pool!
Sequencing V3 inDrops Library

1. 61 bp Read 1 transcript
2. 8 bp Index Read 1 (i7) single cell barcode
3. 8 bp Index Read 2 (i5) library index
4. 14 bp Read 2 barcode/UMI
Cost of a Single Cell Experiment: Example Project

- 12 samples, collecting 3000 cells / sample

<table>
<thead>
<tr>
<th>Line Item</th>
<th>Price/Unit</th>
<th>Units</th>
<th>Price</th>
</tr>
</thead>
<tbody>
<tr>
<td>Set up fee per project</td>
<td>$520</td>
<td>1</td>
<td>$520</td>
</tr>
<tr>
<td>Per run fee</td>
<td>$25.21</td>
<td>3</td>
<td>$76</td>
</tr>
<tr>
<td>Per sample fee</td>
<td>$6.20</td>
<td>12</td>
<td>$74</td>
</tr>
<tr>
<td>Per 1000 cells</td>
<td>$11.10</td>
<td>36</td>
<td>$400</td>
</tr>
<tr>
<td>Full day effort</td>
<td>$273.44</td>
<td>3</td>
<td>$820</td>
</tr>
<tr>
<td>Library preparation</td>
<td>$128</td>
<td>12</td>
<td>$1,537</td>
</tr>
</tbody>
</table>

Total InDrop Fee: $3,427

- Triplicates over 3 days
- Make libraries of 1,500 cells
- User brings 5 tubes SSIII; 3 tubes RNaseOUT
## Cost of a Single Cell Experiment: Example Project

<table>
<thead>
<tr>
<th>Samples</th>
<th>Cell Libraries</th>
<th>Total Cells</th>
<th>Nextseq Runs</th>
<th>Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>12</td>
<td>1,500</td>
<td>18,000</td>
<td>3</td>
<td>$5,000</td>
</tr>
<tr>
<td>3,000 cell libraries</td>
<td>36,000 cells total</td>
<td>6 Nextseq runs ($10,000)</td>
<td>6 Nextseq runs ($10,000)</td>
<td>6 Nextseq runs ($10,000)</td>
</tr>
</tbody>
</table>

### Sequencing
- ~$1,600 / NextSeq
- QC adds one $400 fee
- Expect 400-500 Million reads / NextSeq
- Estimates here for ~75,000 reads per cell.
Cost of a Single Cell Experiment: Example Project

12 samples, collecting 3000 cells / sample

- Enzyme cost ~$2,000
- inDrop cost with library prep ~$3,500
- Sequencing cost $5,000-$10,000

- Minimum cost is $10,500
Single Cell: Data Analysis

- Harvard Chan Bioinformatics Core (HSPH)
  - inDrops pipeline available on Orchestra
    - http://bcbio-nextgen.readthedocs.io/
  - Free consultations / Fee for service
  - Partner with the SCC / Klein Lab

- Kharchenko Lab (HMS)
  - Software available
  - Web-interface package is in the works
high level analysis overview

- inDrops pipeline available on Orchestra
- Reports are generated similar to the images that follow for QC
Single Cell: HSPH Pipeline

K562 cells
- Ideal data
- Reads / cell barcode

Zebrafish Blood
- Messier distribution

Failed Run
- Free floating RNA contamination

- Cut off usually remove any cell with <10,000/20,000 barcodes per cell
- Exact threshold may depend on sample
- Around 5000 UMI disambiguated reads / cell in this example
- 1000 genes detected / cell
- Cell type dependent – lots of other cell types have more
> Around 1000 genes detected / cell
> Cell type dependent – lots of other cell types have more than here
These samples could have been sequenced deeper.
If at saturation plot would level off toward the right.
Cells with mostly mitochondrial reads

- Around 5% of cells have high mitochondrial markers
- Directly correlates to health of cells

Provided by HSPH
Cells with low complexity

- Fewer genes detected than expected from total cell counts
- Sometimes this is biology / sometimes cells should be excluded
Filtering and Correction

- Remove cells with high mitochondrial RNA
- Remove cells with abnormally low high or low genes detected
- Correct mitochondrial RNA percentage
- Correct total UMI
- Correct genes detected
- Correct/filter low complexity
Clusters of T and B cells
Working with the Single Cell Core

1. Submit Application
2. Consultation
3. Users Optimize Experiment
4. Book Experiment
5. InDrop Run
6. Library Preparation
7. Library Delivered to User
8. Sequencing at Core of Choice
9. Data Analysis
SCC Consult: Key Points

- What is your scientific question?
- Samples >90% viable for ½-1h on ice?
- 4-7 samples can be run in one day.
- How will you analyze your data?

Batch effects – biological, inDrops run, library prep
SCC: Final Thoughts

- Practice your protocol!
- Do not make your inDrops run day the first day you run through the whole protocol

Be sure sequencing core understands the specific sequencing parameters for inDrops

qPCR

Precise quantitation is key to good sequencing!