The importance of single cell sequencing

“The single cell ‘omics revolution is firmly underway. Nearly every expression study worth doing will be worth doing at single cell level…” Ewan Birney, EMBL
Single-cell analysis – a brief history

- Single cell qPCR of neurons, Eberwine, et al.
- First single-cell mRNA sequenced, Tang, et al.
- DNA-Seq of single human cancer cell, Navin, et al.
- Single cell WGS of Neurons, Evrony, et al.
- Organ lineage tracing with SC RNA-Seq, Truettlein, et al.
- SC ATAC-Seq of human cell lines, Cusanovich, et al.

1992
- SC whole transcriptome microarrays

2003
- STRT-Seq 5' end Seq
- Islam, et al.

2009
- CEL-Seq
- Remove PCR bias with IVT
- Hashimshony, et al.

2010
- SMART-Seq
- Full length pre-amp
- Ramskold, et al.

2011
- SMART-Seq2
- Increased sensitivity
- Picelli, et al.

2012
- MALBAC
- Zong, et al.

2013
- Incorporation of UMI’s to methods
- Jaitin, et al.
- Islam, et al.

2014
- SC RNA-Seq of immune cells, Shalek, et al.

2015
- First single-cell exome sequenced, Xu, et al.
- First single-cell T-cell epigenome with HiC, Nagano, et al.
Single-Cell Sequencing

Pubmed: "Single-cell Sequencing"
Why Perform Single Cell Analysis?

1. Assess cell-to-cell heterogeneity

2. Map cell trajectories

3. Dissect transcriptional mechanics

4. Infer gene regulatory networks
Assessing Heterogeneity in Cancer

Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq

Tirosh et al., 2016, Science
Illumina® | Bio-Rad®
Single Cell Transcriptome Sequencing Solution
The Illumina | Bio-Rad
Single Cell Sequencing Solution

Isolate and barcode single cells

Prepare your library

Sequence efficiently

Analyze and store data with BaseSpace Sequence Hub

ddSEQ Single-Cell Isolator

NextSeq, HiSeq, NovaSeq

SureCell WTA 3’ Library Prep Kit

Single Cell RNA Seq v1.0.0

High-throughput single-cell sequencing workflow. Integrating the Bio-Rad industry-leading Droplet Digital partitioning technology with leading Illumina next-generation sequencing (NGS) technologies, Bio-Rad and Illumina will launch an isolation-to-analysis commercial solution that will enable high-throughput sequencing of thousands of individual cells quickly and cost effectively.

For Research Use Only. Not for use in diagnostic procedures.
Single Cell Isolation Principle

ddSEQ™ Single-Cell Isolator with Droplet Digital™ Technology

Cell Isolation Reagents

(including barcodes)

Single-Cell Suspension

Isolate thousands of cells in minutes

For Research Use Only. Not for use in diagnostic procedures.
Single Cell Isolation Principle

ddSEQ™ Single-Cell Isolator with Droplet Digital™ Technology

Single-Use Microfluidic Cartridges

Cell Isolation Reagents *(including barcodes)*

Single-Cell Suspension

Isolate thousands of cells in minutes

Oil

For Research Use Only. Not for use in diagnostic procedures.
Experimental Design for the ddSEQ™ System

1 Cartridge:
- 4 wells with samples
- Load 4.5ul of cell suspension per well
- 2,500 cells/ul
- 11,250 cells loaded per well
- Collect 300-375 cell+bead droplets per sample

Goal = 1 cell & 1 bead

Cell+Bead capture rate 3-4%
Cell Preparation Methods

Solid tissue dissociation

Tissue culture

Trypsinize, Wash & Filter (35um strainer)

Clump-free Single Cell suspension

FACS

S3e™ Cell Sorter
Quality Assessment and Cell Counting

- Cells must be free of clumps to prevent multiple cells per droplet
- Suspension concentration must be accurately determined
- Free of debris from dead or damaged cells
  - Trypan blue staining

ZOE™ Cell Imager
TC20™ Automated Cell Counter
SureCell™ WTA 3’ Library Preparation

- Single cell encapsulation with the ddSEQ™ Single-Cell Isolator

Microspheres ▶️ Cells

Lysis and barcoding

Read1  Cell Barcode  UMI

TTTTT  TTTTT  TTTTT  TTTTT
SureCell™ WTA 3’ Library Preparation

- Single cell encapsulation with the ddSEQ™ Single-Cell Isolator
- Sensitive assay chemistry without pre-amplification
- First strand synthesis in droplets
- Modified Nextera® Library prep

For Research Use Only. Not for use in diagnostic procedures.
Sequencing of SureCell™ WTA 3’ Libraries
Sequencing Conditions for SureCell 3’ Libraries

1. Custom Rd 1 Seq Primer Binding Site
   Read 1 (68 cycles):
   - Bead/Cell Barcode
   - Unique Molecular Identifier

2. Paired End Turnaround

3. Rd 2 Seq Primer Binding Site
   Read 2 (75 cycles):
   - cDNA insert
   - Reads will map to the sense strand

68 x 75 Cycle PE Run with Single Index

Read 1: Bead/Cell Barcode:
- Identifies original cell per droplet

UMI:
- Identifies original mRNA transcript

Read 2: cDNA insert sequence that maps to sense strand

For Research Use Only. Not for use in diagnostic procedures.
How much sequencing will I need?

- How many genes do I want to detect per cell?
- Am I looking for potentially rare cell types?
- Saturating the number of detected genes requires more reads

Range depends on cell type and expression levels and also the biological question you’re asking.
Sequencing Power for Every Scale

Single Cell RNA-Seq

- **MiniSeq™**
  - 7.5 Gb | 25M
  - 2x150

- **MiSeq®**
  - 15 Gb | 25M
  - 2x300

- **NextSeq®**
  - 120 Gb | 400M
  - 2x150

- **HiSeq® 2500**
  - 1000 Gb | 4B
  - 2x125

- **HiSeq® 3000**
  - 750 Gb | 2.5B
  - 2x150

- **HiSeq® 4000**
  - 1500 Gb | 5B
  - 2x150

- **HiSeq X® Five**
  - 9000 Gb | 30B
  - 2x150

- **HiSeq X Ten**
  - 18000 Gb | 60B
  - 2x150

- **NovaSeq 6000 S2**
  - 2000 Gb | 6.7B
  - 2x150

For Research Use Only. Not for use in diagnostic procedures.
Sequencing cost per cell

2017 USD List

# of reads per cell

# of genes detected per cell (APPROXIMATE)

$0.18 $0.35 $0.69 $1.74 $3.48 $6.96

NextSeq 500

For Research Use Only. Not for use in diagnostic procedures.
Sequencing cost per cell

2017 USD List

<table>
<thead>
<tr>
<th># of reads per cell</th>
<th># of genes detected per cell (APPROXIMATE)</th>
<th>Cost per cell</th>
</tr>
</thead>
<tbody>
<tr>
<td>25,000</td>
<td>2,000</td>
<td>$0.11</td>
</tr>
<tr>
<td>50,000</td>
<td>4,000</td>
<td>$0.23</td>
</tr>
<tr>
<td>100,000</td>
<td>8,000</td>
<td>$0.44</td>
</tr>
<tr>
<td>250,000</td>
<td>20,000</td>
<td>$1.12</td>
</tr>
<tr>
<td>500,000</td>
<td>50,000</td>
<td>$2.25</td>
</tr>
<tr>
<td>1,000,000</td>
<td>100,000</td>
<td>$4.50</td>
</tr>
</tbody>
</table>

HiSeq 3000

For Research Use Only. Not for use in diagnostic procedures.
Sequencing cost per cell

2017 USD List

# of genes detected per cell (APPROXIMATE)

# of reads per cell

- $0.10 for 25,000 reads
- $0.20 for 50,000 reads
- $0.41 for 100,000 reads
- $1.02 for 250,000 reads
- $2.05 for 500,000 reads
- $4.10 for 1,000,000 reads

For Research Use Only. Not for use in diagnostic procedures.
What about data analysis?
BaseSpace® Single-Cell RNA App

- **Simple analysis setup** for samples across multiple sequencing runs
  - Up to 96 samples per analysis

- **Easily choose analysis parameters**
  - Reference genome
  - ERCC spike-ins
  - Subsampling for QC

- **Rapid** alignment, cell and gene counting, and filtering

For Research Use Only. Not for use in diagnostic procedures.
BaseSpace® Single-Cell RNA App

- **Per-sample reports** available in-browser and as PDF
  - Alignment quality
  - Coverage information
  - Abundant sequences

- Perform global PCA and tSNE clustering

- Cell-cycle heatmap

- All output files available for download, or as input into downstream applications
  - Includes cell-gene expression table
tSNE analysis identifies a sub-population in a heterogeneous cell mixture mouse/human

Total cells: 602

hg19 RPL13

mm10 Rpl13
PCA clustering of 1:1 mixture of 1,400 mouse and human cells detects distinct populations

HEK293 / NIH3T3 (human/mouse)

hg19 RPL13

mm10 Rpl13
Two-species cell mixture (HEK293/NIH3T3) demonstrates low crosstalk and high purity

Cells: 1384
Duplets: 5.8%
Purity: 99.1%

For Research Use Only. Not for use in diagnostic procedures.
Sub-populations identification in PBMCs

- **NK cells (11.4%)**
- **CD16+ Monocytes (5.4%)**
- **CD14+ Monocytes (17.6%)**
- **Dendritic cells (1.2%)**
- **Platelets (0.8%)**
- **B-Cells (9.0%)**
- **Memory T-Cells (12.7%)**
- **Cytotoxic T-Cells (13.4%)**
- **Naïve CD4 & Naïve CD8 T-Cells (15.2%)**

**t-SNE clustering**
- 6,000 cells
- 4 cartridges
- ~20m of hands on time for barcoding
- Major subclasses of peripheral blood mononuclear cells (PBMCs) identified down to <1%
- 60,000 reads per cell

For Research Use Only. Not for use in diagnostic procedures.
Further Analysis with 3rd party software

- Import cell-gene expression tables from BaseSpace
- Discover and characterize populations with gene sets
- Extended heatmapping capabilities
System performance on reference cells
Sensitivity & Reproducibility Across Cell Lines

HEK293 Genes vs. Reads Per Cell

NIH3T3 Genes vs. Reads Per Cell

Human Genes Detected

Mouse Genes Detected

Gene Count Reproducibility

R2: 0.975
NumGenes: 10,011

Genes Detected Across Different Cell Sizes

A20 3T3 HEK BJ

For Research Use Only. Not for use in diagnostic procedures.

Illumina internal data on file
Correlation between ddPCR and ddSEQ

Good correlation between bulk and aggregate single cell data

Data courtesy of
Drs. David Schaffer & Maroof Adil
University of California, Berkeley
Department of Chemical & Biomolecular Engineering

For Research Use Only. Not for use in diagnostic procedures.
ddSEQ SCI & SureCell Kit Future Scalability

More cells, same robust assay

More sensitivity through target enrichment

More assays for single cell characterization

For Research Use Only. Not for use in diagnostic procedures.
# Illumina | Bio-Rad Single Cell Sequencing Solution

<table>
<thead>
<tr>
<th>Isolate and barcode single cells</th>
<th>Prepare your library</th>
<th>Sequence efficiently</th>
<th>Analyze and store data with BaseSpace Sequence Hub</th>
</tr>
</thead>
</table>

1. Isolation and barcoding of several thousand single cells in minutes
2. Affordable platform
3. Scalable to profile increasing number of cells in a cost-effective manner
4. Proven Illumina library prep with streamlined workflow
5. Single-Cell RNA-Seq BaseSpace® App available

For Research Use Only. Not for use in diagnostic procedures.
A solution to meet your needs

- **Sensitive**: 5-7K genes per cell
- **Accurate**: No pre-amp, correlates with orthogonal method (ddPCR)
- **Scalable**: Kit formats to support $10^2$-$10^4$ cells per day
- **Simple**: Nextera. Easy stopping points, sequence the second day
- **Reproducible**: High correlation experiment to experiment
- **Integrated**: Built-in setup and analysis. End-to-end solution
Thank you