Single Cell Genomics – CCR new initiative & practical tips

Liz Conner, CCR Genomics Core

Michael Kelly, Single Cell Analysis Facility

February 15, 2018

CCR Genomics Core

- **20** years in Bldg 37
- **1035** Registered iLab members
- **301** Principal Investigators/Lab groups
- **10** ICs (NIAID, NHLBI, NIDDK, NEI, NIAMS, NIMH, NINDS, NIAAA, NIA, NICHD)
- **340** Requests per month
- **> 50,000** Samples processed/year across all platforms
- **4** Staff members: Liz Conner, Steve Shema, Qin Wei, Val Bliskovsky
- **>6** Platforms

Bldg. 37, Room 2135

KEY TECHNOLOGIES

Sanger Sequencing

Ncounter Technology

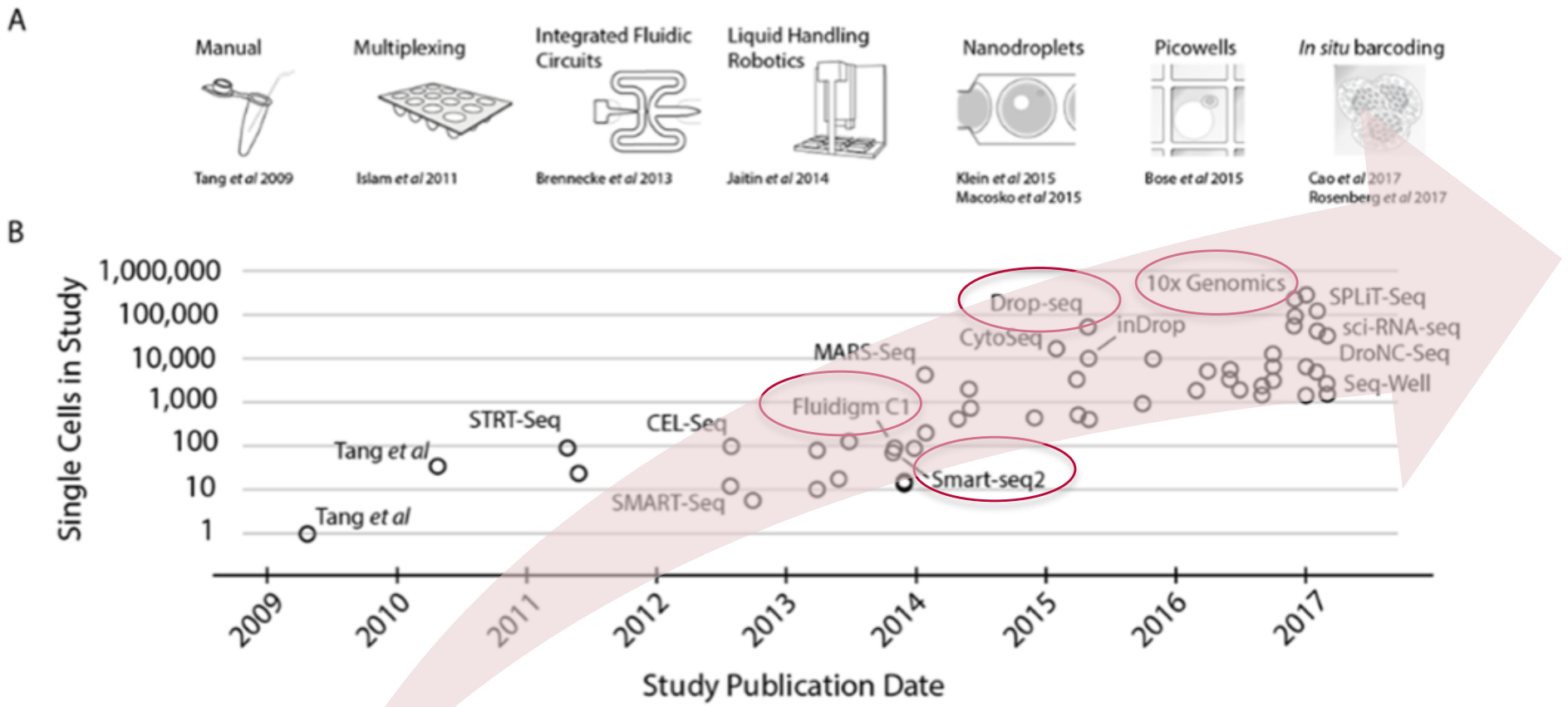
Next-Generation Seq

BioRad QX200 ddPCR

Fluidigm C1 Single Cell

Other

Scaling of scRNA-seq experiments

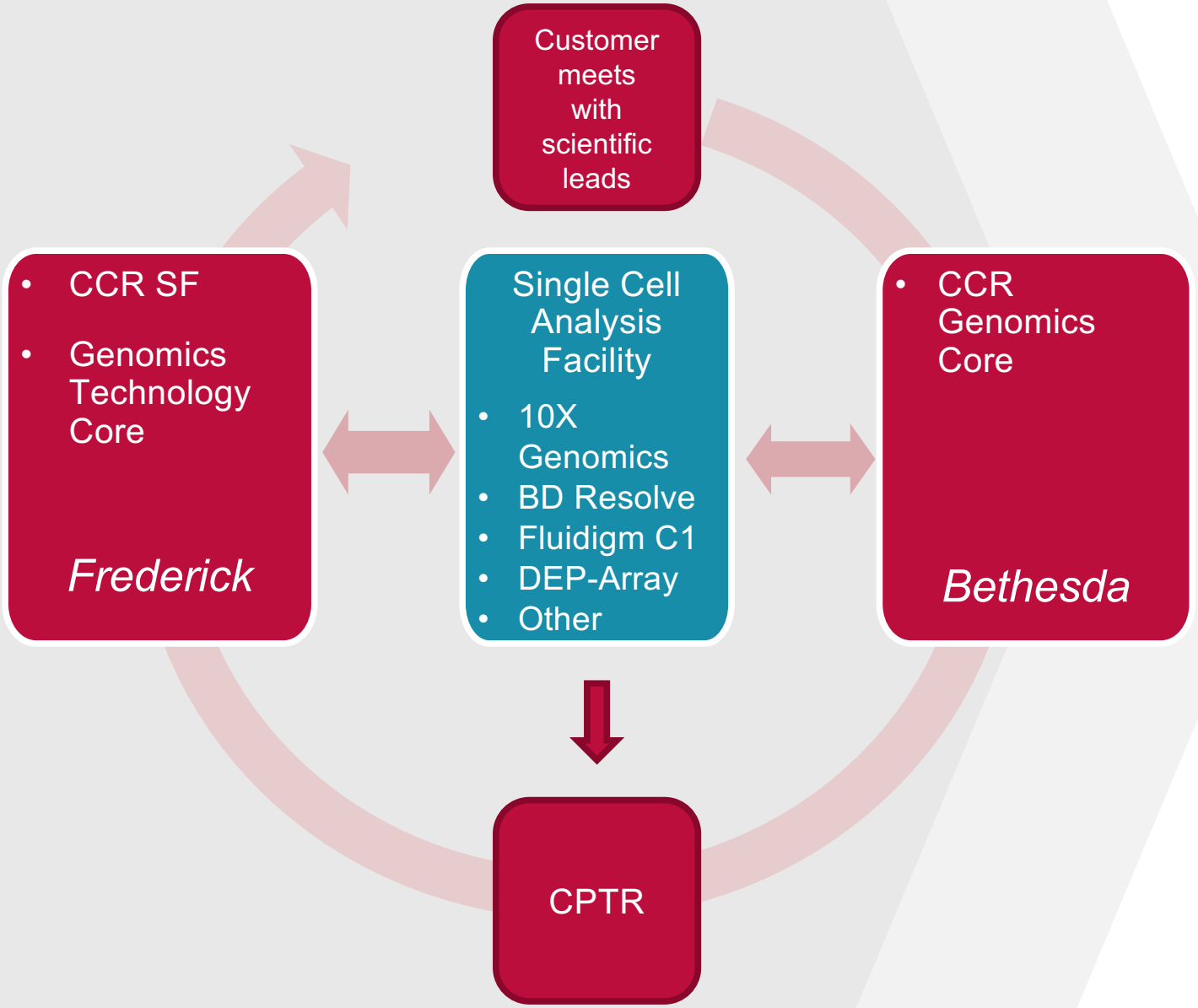


Title:	Exponential scaling of single-cell RNA-seq in the last decade
Authors:	Svensson, Valentine; Vento-Tormo, Roser; Teichmann, Sarah A
Publication:	eprint arXiv:1704.01379
Publication Date:	04/2017

CCR New Initiative

The Single Cell Analysis Facility (SCAF) will implement state-of-the-art single-cell technologies and provide high-throughput isolation of rare and single cells for downstream library preparation and sequencing of RNA and DNA. The Facility will also serve as a central hub for SC technology assessment, data analysis, and focal point for collaboration and idea exchange with CCR Laboratories and Centers of Excellence, NCI Cores and Facilities, as wells as other research groups at NIH focused on cutting-edge SC research.

Single Cell Analysis Facility (SCAF)



The players……

NIH NATIONAL CANCER INSTITUTE
NCIatFrederick Accessioning

Search... Go Sign In

SERVICES ABOUT HELP + NEW REQUEST

Home » Services » Service Area

Single Cell Analysis Facility (CCR) - Center for Cancer Research (CCR) Dedicated Services

Overview FAQ Billing Catalog

Begin Request

Overview

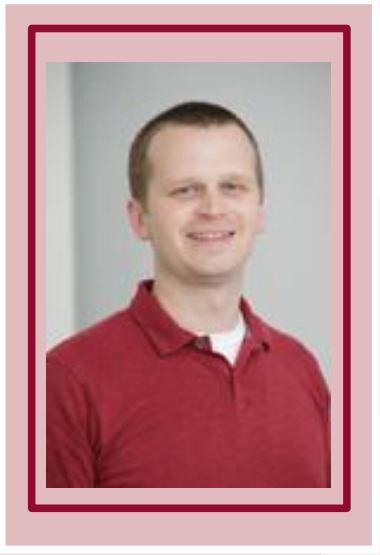
CCR SCAF is a dedicated Single Cell Analysis Facility with state of the art technologies for single cell research. The rapid advancement in single cell technologies in recent years has provided cancer researchers powerful new means to study cancer including tumor heterogeneity, tumor evolution, drug resistance, etc. Our current technologies include single cell isolation, single cell RNAseq, single cell DNA seq, and single cell protein expression analysis.

Capabilities

1. BD Genomics Rhapsody Single Cell RNA/protein sequencing
2. 10xGenomics Chromium Single Cell RNAseq
3. DEPArray Single Cell Isolation

Contacts

MANAGER
Xiaolin Wu
forestwu@mail.nih.gov
301-846-7677
ATRF, Rm C3016
8560 Progress Drive
Frederick, MD 21701



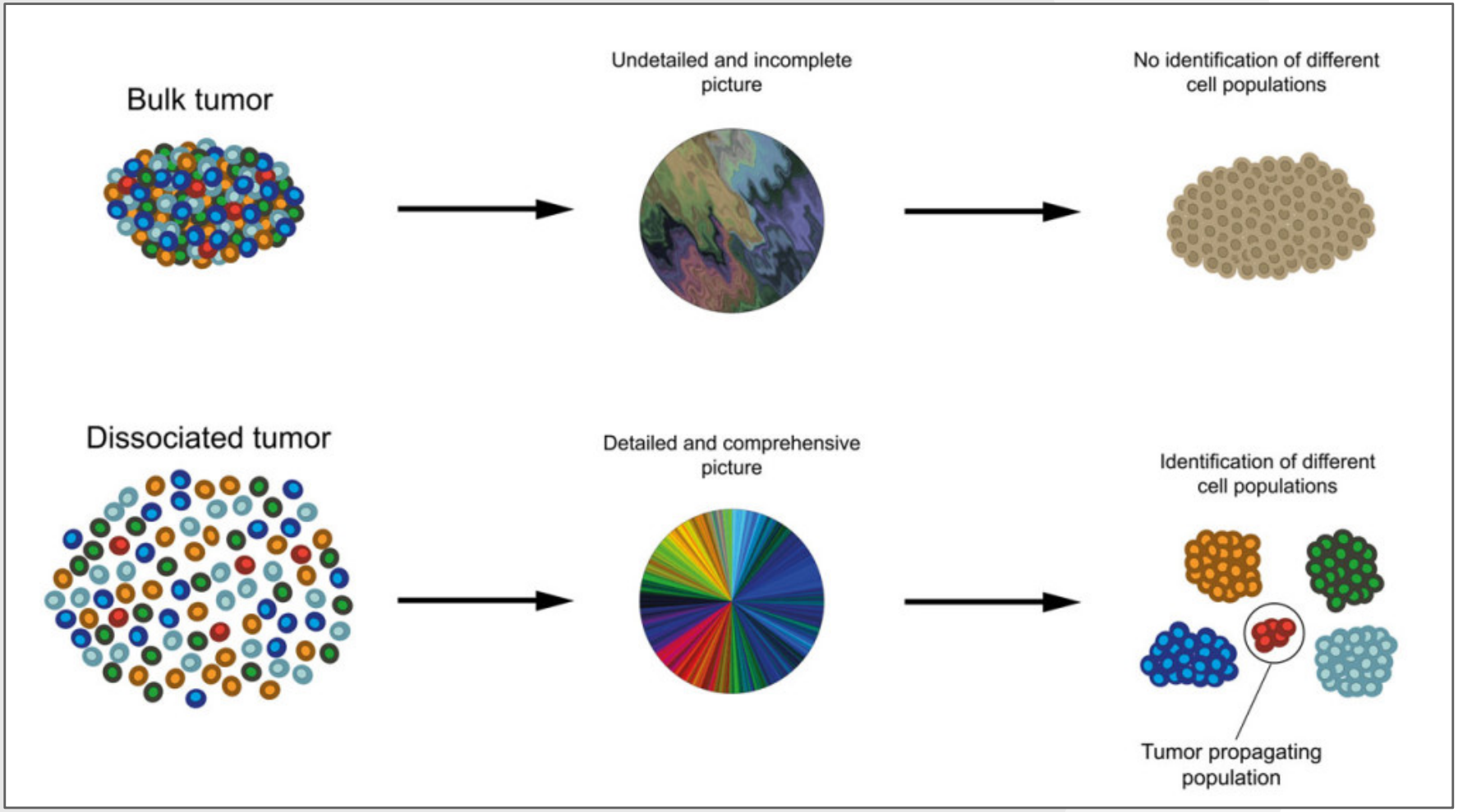
Michael C. Kelly, Ph.D.
Bldg. 37, Room 1042

Outline:

- **General Background & Considerations**
- What Support Will SCAF Provide?
- Platform / Method Selection
- What's on the horizon?

General Background & Considerations

Why Single Cell?



http://fightdipg.org/research-projects/

General Background & Considerations

Is my biological question a “single cell problem” and other things to think about?

https://btep.ccr.cancer.gov/november-2017-single-cell-rna-seq-mind-read-starting-adventure/

Or Google: “BTEP TOTM Single Cell 2017”

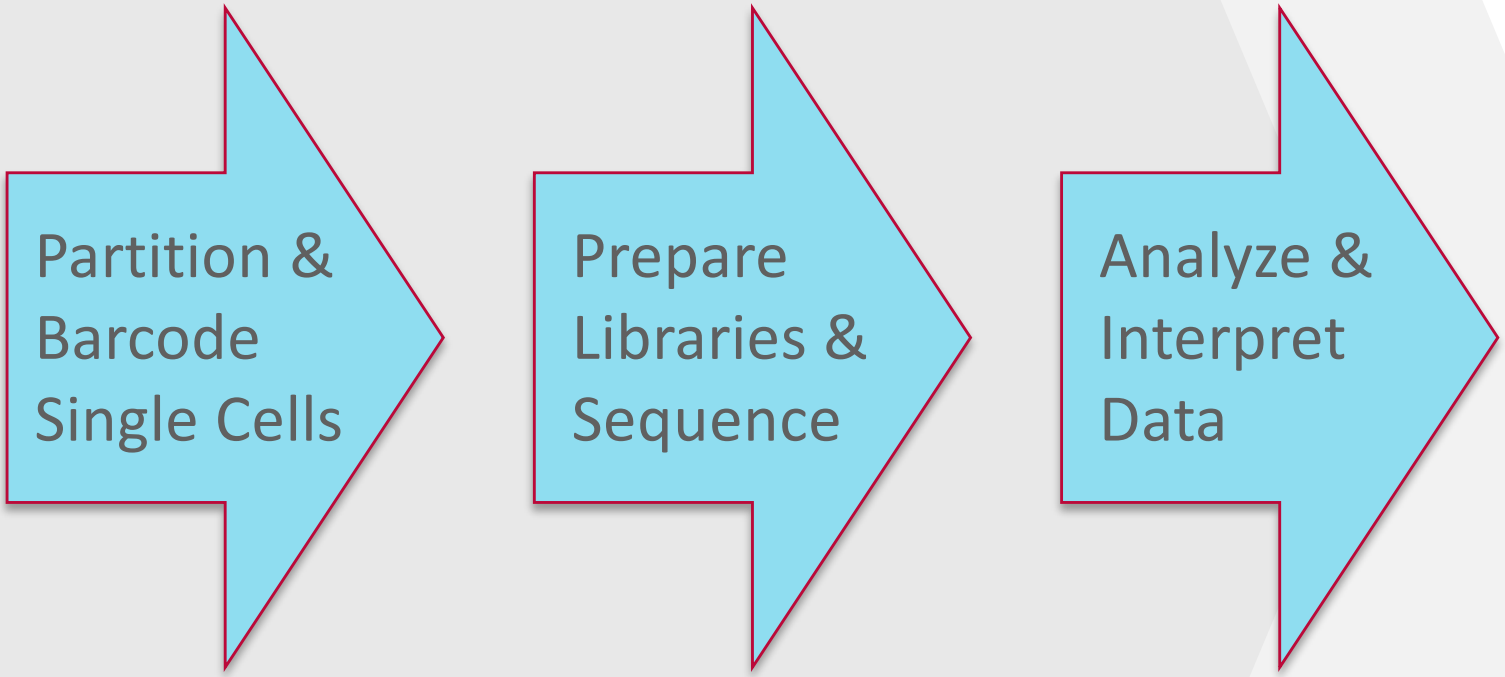
General Background & Considerations

Things to Think About Before Getting Started -
The Short List (TL;DR version)

- Have a clear idea of why you want to use “single cell” and understand the limitations
- Spend time on optimizing the input cell prep. Healthy viable single cells give the best data and fewest headaches for analysis
- Anticipate considerable ***non-standard*** bioinformatics needs

General Background & Considerations

How does it work?

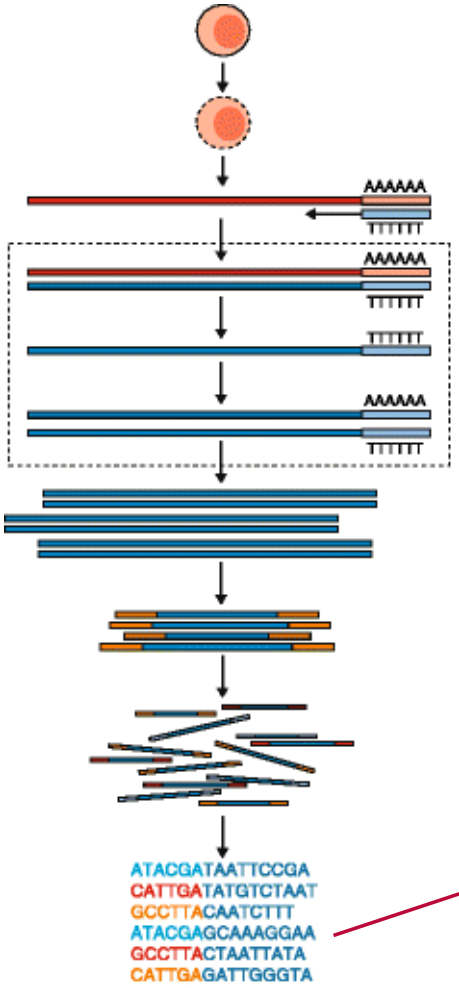


- Rare populations?
- Difficult to dissociate?
- Cell viability?

- Very low input molecular biology
- Currently limited to certain assays

- Analysis can be more challenging
- Often requires informatics and subject matter expertise

Generalized Workflow for Single Cell RNA-Seq



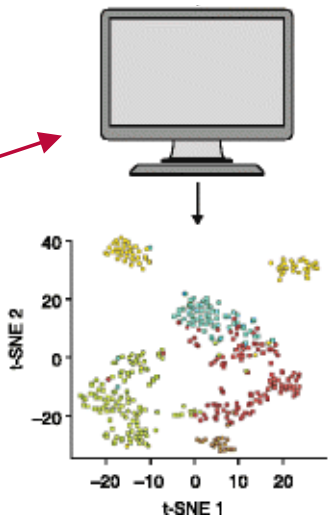
1. Isolate single cells from a tissue sample (including micro-dissection and manipulation, flow cytometric cell-sorting, microfluidic platforms, and droplet-based methods)
2. Single cell lysis in a way that preserves cellular mRNA
3. mRNA molecule capture using poly[T] sequence primers that bind to mRNA poly[A] tails
4. Convert poly[T]-primed mRNA into cDNA using reverse transcription
5. cDNA amplification (usually by PCR or by *in vitro* transcription)
6. cDNA sequencing library preparation (insert 'index' nucleotide barcodes to identify each library)
7. Pool cDNA sequencing libraries

Sequence libraries (via Next Generation Sequencing)

1 – Partition & Barcode Single Cells

2 – Sequence & Quantitate

3 - Analyze & Interpret



8. Use bioinformatic methods to perform quality control and to assess technical variability in the scRNA-seq data

9. Use bioinformatic and/or computational methods to interpret robust data biologically

Haque et al 2017

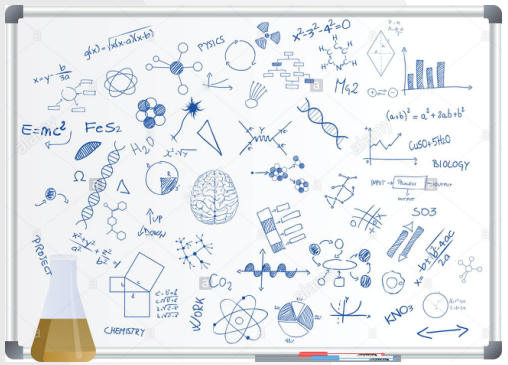
Outline:

- General Background & Considerations
- **What Support Will SCAF Provide?**
- Platform / Method Selection
- What's on the horizon?

What Support Will SCAF Provide?

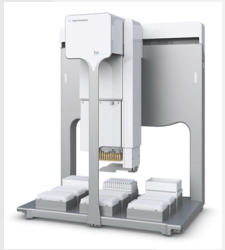
Project consultation:

- Early-phase discussion of options and considerations
- Experimental design discussion
- Advice on platform selection
- Bioinformatics-needs assessment
- Cost estimates & rough timeline estimate



What Support Will SCAF Provide?



Equipment access and support:

- Easy, local access to cutting-edge technology
- Expert support for cell prep, capture, library preparation, sequencing, and primary informatics
- Walk-up access to certain equipment (after training)
- New platforms and methods will continue to be evaluated and validated – ***opportunities for collaboration***

What Support Will SCAF Provide?

Training, education, and community support

- Observation and hands-on training welcomed / encouraged
- One-on-one discussions and demonstrations for projects
- Supporting NIH single cell SIG & users group
- Future workshops and training sessions (e.g. sample handling, bioinformatics, etc.) either through BTEP or other mechanism

BTEP

https://nih-irp-singlecell.github.io

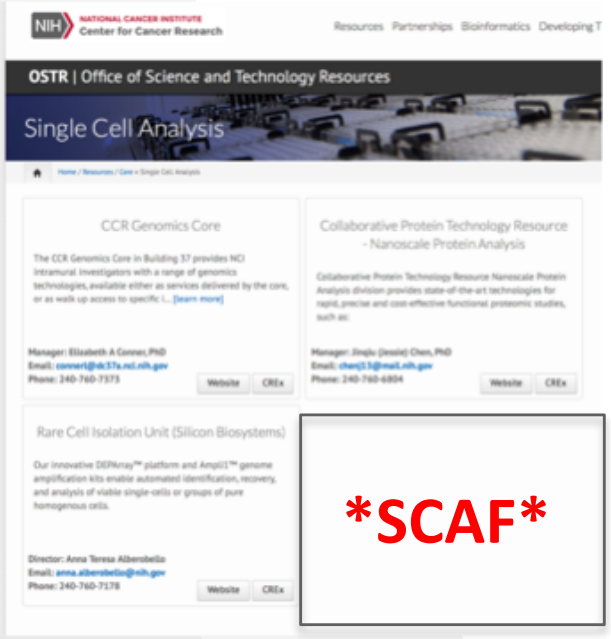
What Support Will SCAF Provide?

How to Contact for Initial Project Inquiry?

- Email me: michael.kelly3@nih.gov

- OSTR website *(will be updated shortly)*
  - https://ostr.cancer.gov/resources/fnl-cores
  - https://ostr.cancer.gov/resources/core
- Future presence on CREx / TheScientist.com

- Equipment access for walk-up use will likely be scheduled using Agilent CrossLab web portal (used by CCR Genomics Core)



What Support Will SCAF Provide?

Formal Project Initiation (Workflow)

- OSTR website -> Frederick Nat Lab NCI Accessioning System (NAS): https://ncifrederick.cancer.gov/Services/Accessioning/Services/LabServices/Area/39
  - Setup Profile & Login
  - New Request -> Single Cell Analysis Facility
  - Enter Brief Project Description
  - Submission & Follow-up on Project Details
  - Cost Estimation, Informatics Needs Assessment, Timeline
  - PI Funds & Subsidy Approval -> Project Initiation

Outline:

- General Background & Considerations
- What Support Will SCAF Provide?
- **Platform / Method Selection**
- What's on the horizon?

Platform / Methods Selection

10X Genomics Chromium

(Droplet-based, high-throughput, lower sensitivity for broad survey)

Now: 3’ gene end-counting whole transcriptome RNA-Seq
~Now: 5’ and TCR / BCR profiling
Upcoming: AbSeq/CITE-Seq, ATAC-Seq, CNV

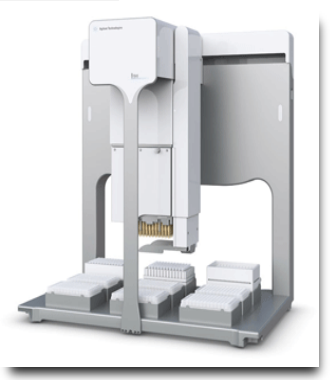
BD Rhapsody

(Microwells, high-throughput, targeted assays with good quantitative sensitivity)

~Now: Targetted panel gene detection (cancer immun)
~Now: Custom assay panels ordered through BD
Upcoming: AbSeq, and Other feature barcoding?

96-well, 384-well Plate Format (of Fluidigm C1) *

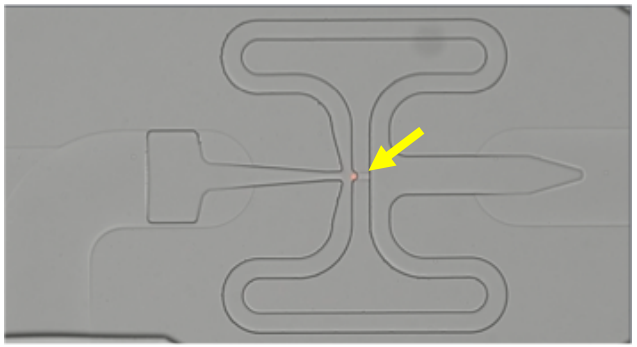
(Lower throughput, but flexible chemistries (good “sandbox” platform, full-length isoform-level RNA-Seq)

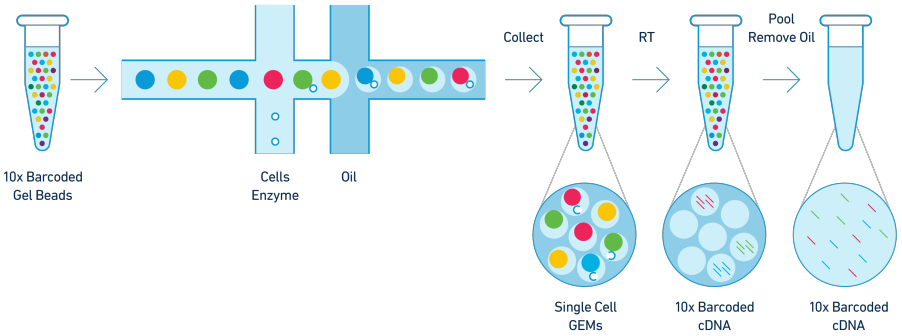


~Now / Upcoming: Smart-Seq2 or Clontech SMARTer v4
Future: Combinatorial indexing, ATAC-Seq, Other Features?

** Collaborative effort with Liz & Val to establish methods*

Overview of common single cell RNA-seq methods





Method	$ system	$ per cells	No. cells	Doublets	Transcript type	UMIs	Capture Efficiency
DROP-seq	$50000	$0.65	up to 50000	0.36-11.3%	3' mRNA	Yes	~2%
Fluidigm C1	$150,000	$1.5-10	96, 800 (10k?)	10-23%	mRNA	No	~10%
10X Genomics	$125,000	$0.20-1.00	1000-6000	1-5%	3' mRNA	Yes	65%
Wafergen	$200,000	$1.5-2.5	~1800	1-5%?	3' mRNA	Yes	?

Modified from core-genomics.blogspot.com

Single Cell-Per-Well Protocols

Individual samples remain partitioned until library prep indexing

-Fluidigm C1

-FACs-based protocols

Microwell Protocols

Although using individual wells, more like droplet methods – barcoding at RT step

-BD Rhapsody

-iCell8 (Wafergen/Clontech/Takara)

-SeqWell

Droplet-Based Massively Parallel Protocols

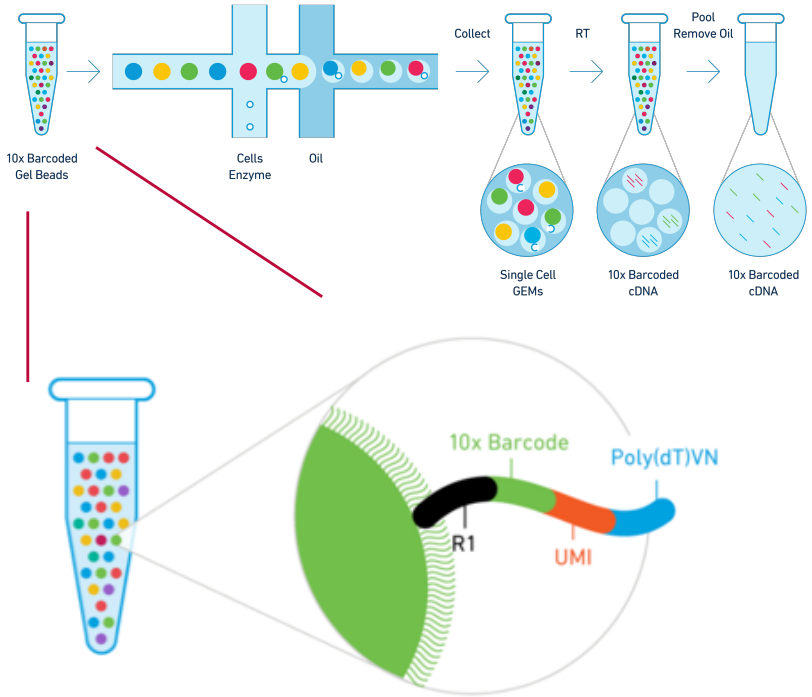
Cells and barcodes partitioned in droplets; indexing occurs at RT

-DropSeq / InDrop (1CellBio)

-10X Genomics Chromium

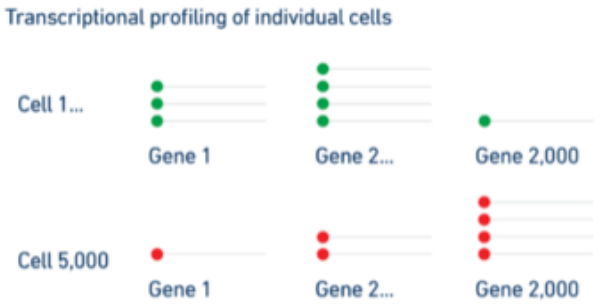
-BioRad/Illumina SureCell on ddSEQ

Droplet-based barcoding allows high-throughput scRNA-Seq gene counting

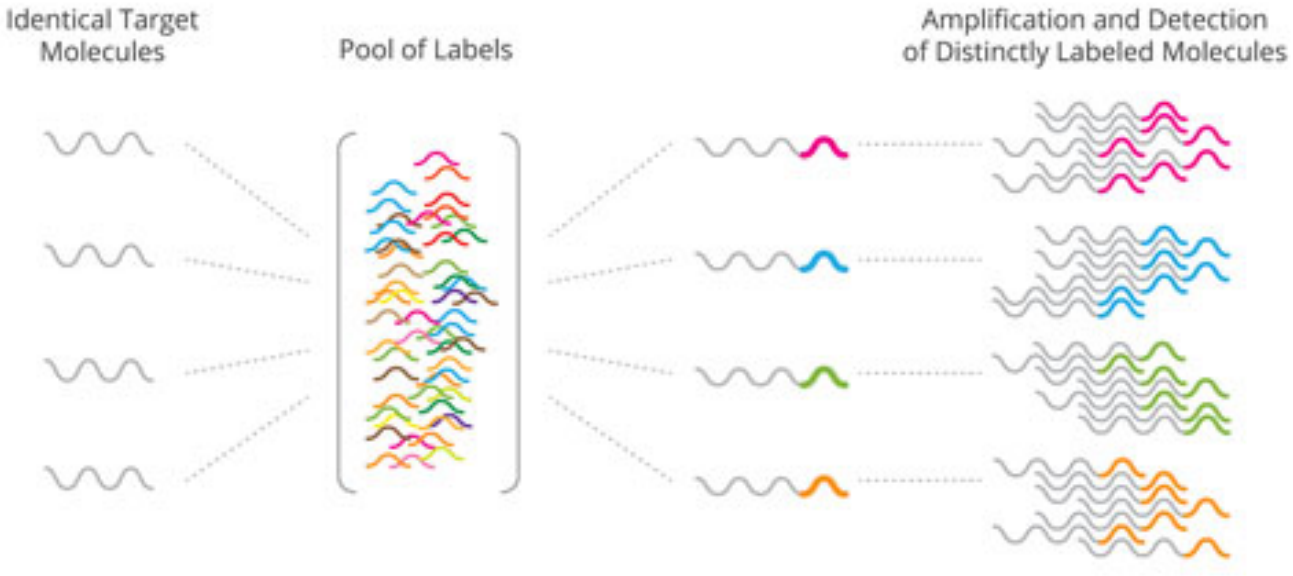


From 10X Genomics Promotional Material

- Barcode added during reverse transcription
- 3’ (or 5’ end) of transcript is selectively enriched by PCR
- Interestingly, these methods give inherent strand information
- Originally lower sensitivity than single cell-per-well protocols, but now approaching similar levels



Unique Molecular Indices Help Reduce Undesirable Biases From PCR Amplification

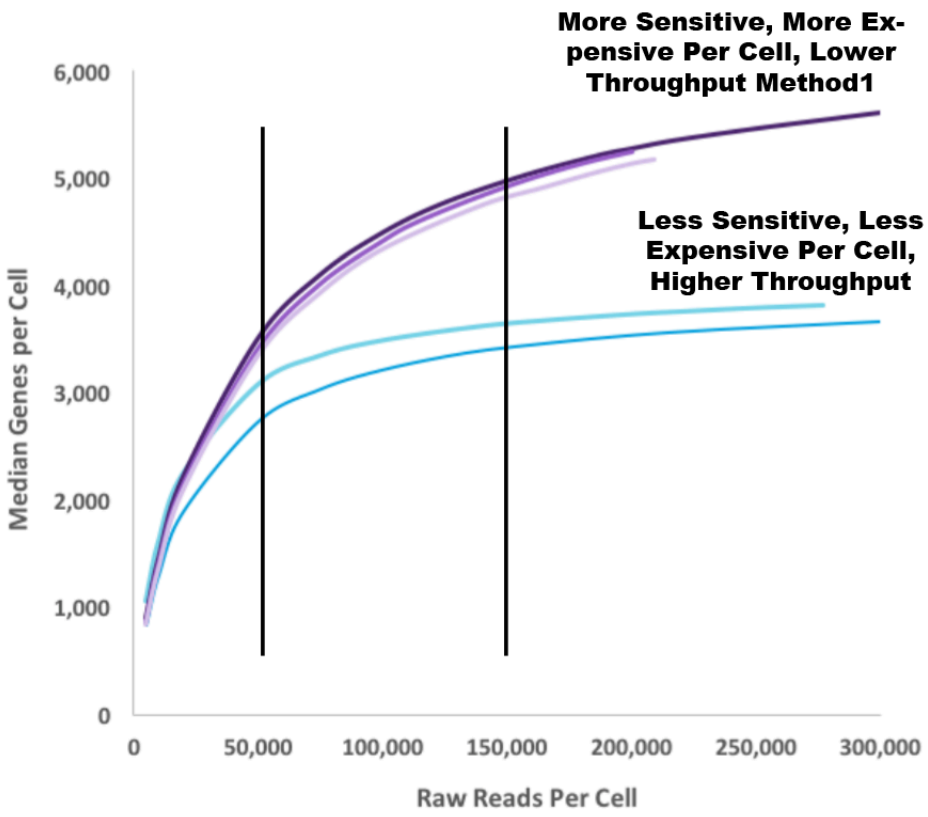


From: http://www.genengnews.com/gen-articles/molecular-indexing-with-precise-assays/5607

Note: Unique molecular identifiers are currently only possible with 5' or 3' end only methods on Illumina sequencers

Islam et al 2014 Nature Methods PMID: 24363023

Which is better – more cells or greater depth?



Modified from 10X Genomics material

- More information can be gained by sequencing to greater depth – especially using sensitive methods
  - More genes detected; fewer "drop-outs"
  - Better isoform discrimination (when full-length libraries sequenced)
- More independent observation (more cells) is better for cell identity classification – averages out noise
- Classic scientific non-answer: it depends on what you are looking for
  - Broad survey of cell types or dynamics processes best modeled by higher-throughput data
  - Investigation of presumably low-expressed (or specific isoforms) requires greater depth

Platform / Methods Selection

Cost Estimates

10X Chromium (3' Single Cell Gene Expression)
~$5000 for 2 capture lanes *with sequencing*
Up to 10,000 cells per capture lane

FACs-Based SMARTer / C1 (full-length scRNA-Seq)
-Still working on the protocol – check in with us soon
-$2k-$7k for 96-cells for Clontech
-We are working on bringing this cost down significantly (~$1k for 96-cells)
-C1 is about $2.5k for up to 96 cells
-All costs here *do not include sequencing*

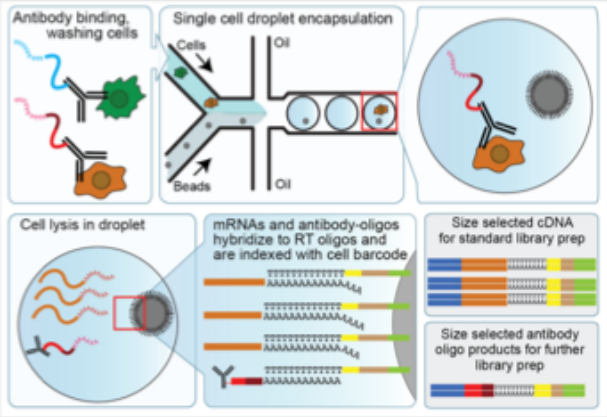
BD Rhapsody - TBD

Note 1: These are very rough estimates for early planning stages. More accurate cost estimates will be provide via "NAS" portal.
Note 2: Above estimates do not include any subsidies

Outline:

- General Background & Considerations
- What Support Will SCAF Provide?
- Platform / Method Selection
- **What's on the horizon?**

Feature Barcoding

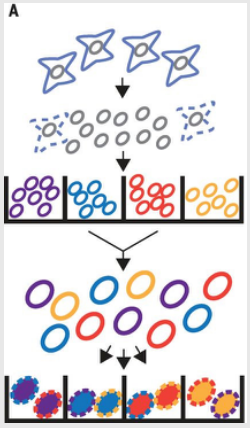


Protein, Spatial, Lineage, etc.

Beyond 3' End Counting

- Established / Reliable Smart-Seq Workflow
- TCR / BCR
- Targeted Enrichment
- Long-read sequencing?

Higher Throughput



Combinatorial Indexing

Epigenomics

ATAC-Seq
DNase HS
ChIP?

Questions?

We are always happy to talk with you about how we might be able to support your work.

Email: michael.kelly3@nih.gov
Location: Bld 37 / Rm 1042