Genomic Facility Overview and practical aspects of single-cell experiments

Andrea Rinaldi, PhD, PD
Head of Genomic Facility at BIOS+
Bellinzona (CH)
www.ior.usi.ch

Genomic Facility Overview


A. Rinaldi, PhD PD

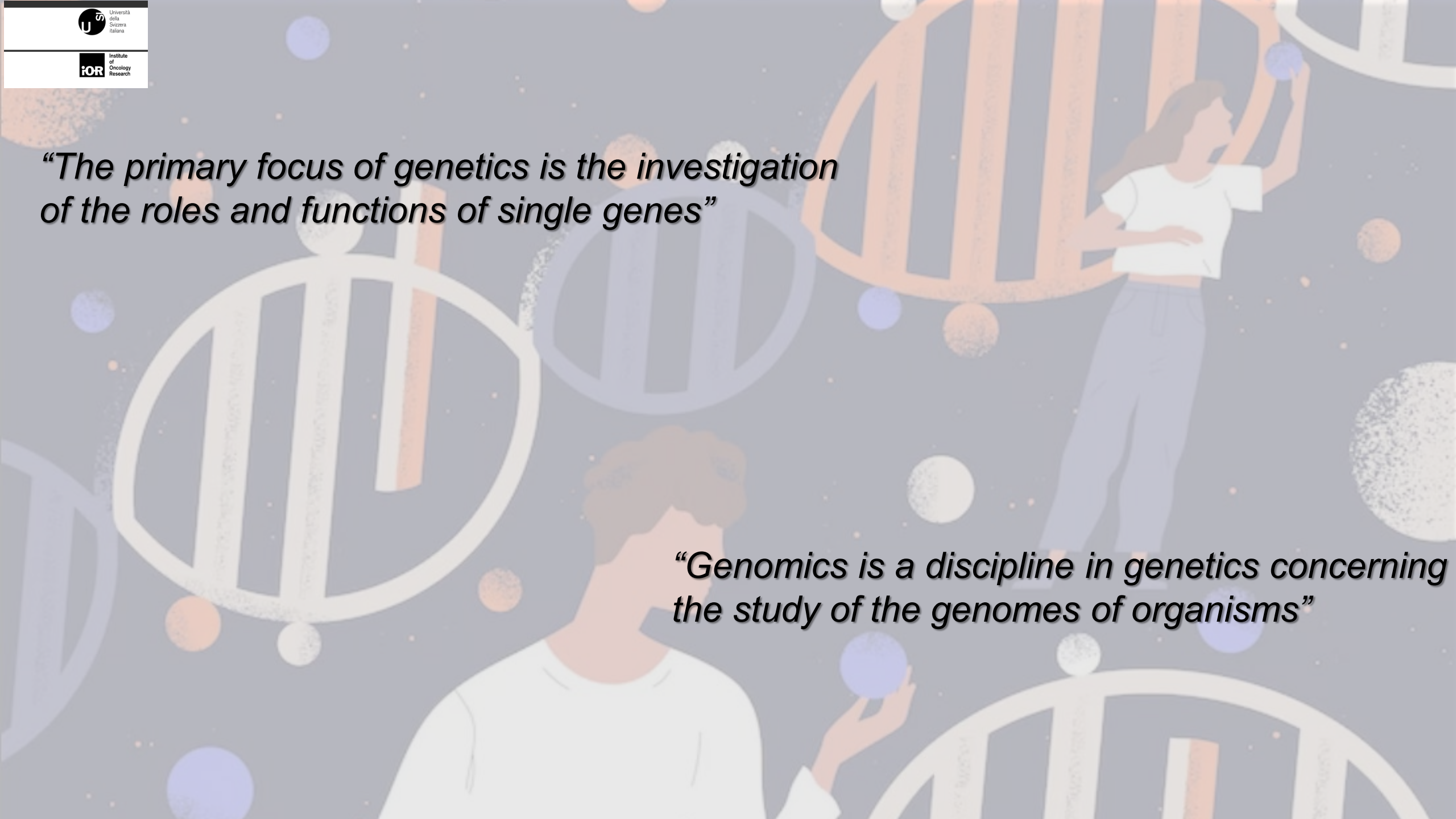

A. Cerana


G. Galimberti

The IOR Genomics Facility is fully equipped to perform genomic studies, including next-generation sequencing (NGS) for DNA and RNA analyses

The services include sample preparation, data generation, and, on request, delivery of partially preprocessed or analyzed sequencing data. Currently, we support the library construction for:

- RNA-Sequencing (polyA and Ribodepleted libraries)
- Whole Exome Sequencing (WES)
- Whole Genome Sequencing (WGS)
- small RNA-Sequencing
- ChIP-Sequencing, ATACseq, Cut&Run seq
- Methyl-Sequencing
- Amplicon sequencing
- Spatial transcriptomics
- scRNA (smartseq, 10xGenomics, BD Rhapsody, ParseBiosciences)
- scDNA (MissionBio – Tapestri)
- Cell surface Protein Labeling for Single Cell RNA
- 16s Microbiome
- Long Reads sequencing



“The primary focus of genetics is the investigation of the roles and functions of single genes”

“Genomics is a discipline in genetics concerning the study of the genomes of organisms”

Next Generation Sequencing (NGS)

Sequencing is a method for determining the exact order of nucleotides in a given DNA or RNA sequence

Next-generation sequencing (NGS) platforms perform massively parallel sequencing, during which millions of DNA fragments are sequenced in unison.

- Rapid (sequence an entire genome in less than one day)
- Low cost in comparison to traditional techniques

⇦ 1953 - Discovery of DNA structure by Watson and Crick

⇦ 1977 - Publication of Sanger sequencing method

⇦ 1983 - PCR development

⇦ 1986 - Launched the 1st automated sequencer: Applied Biosystems Prism AB370A

⇦ 1998 - Sequencing of C. Elegans Genome (100 million bp)

⇦ 2001 - Sequencing of Human genome (3.2 billion bp)

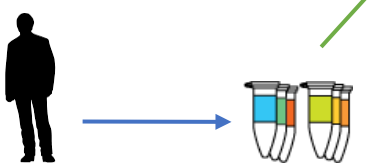
⇦ 2005 - Launched the Roche 454-NGS System

⇦ 2006 - 1st Solexa NG Sequencer: genome analyzer Illumina

⇦ 2011 - PacBio released the first “single molecule real-time” sequencer

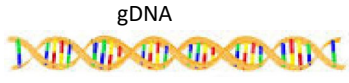
⇦ 2015 - 1st Oxford Nanopore Technologies sequencer: the pocket-sized minION

NGS: preparing the library



Extract nucleic acids from

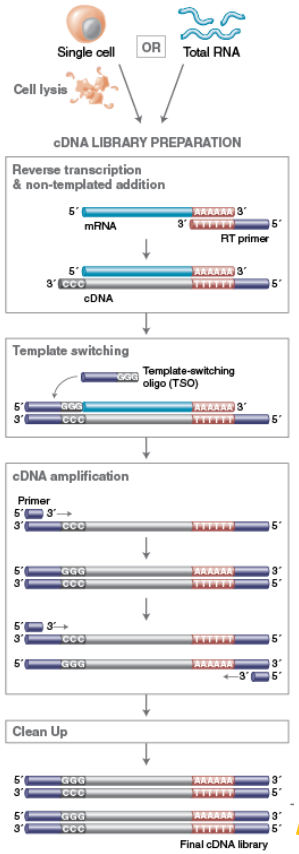
blood, tissue, saliva

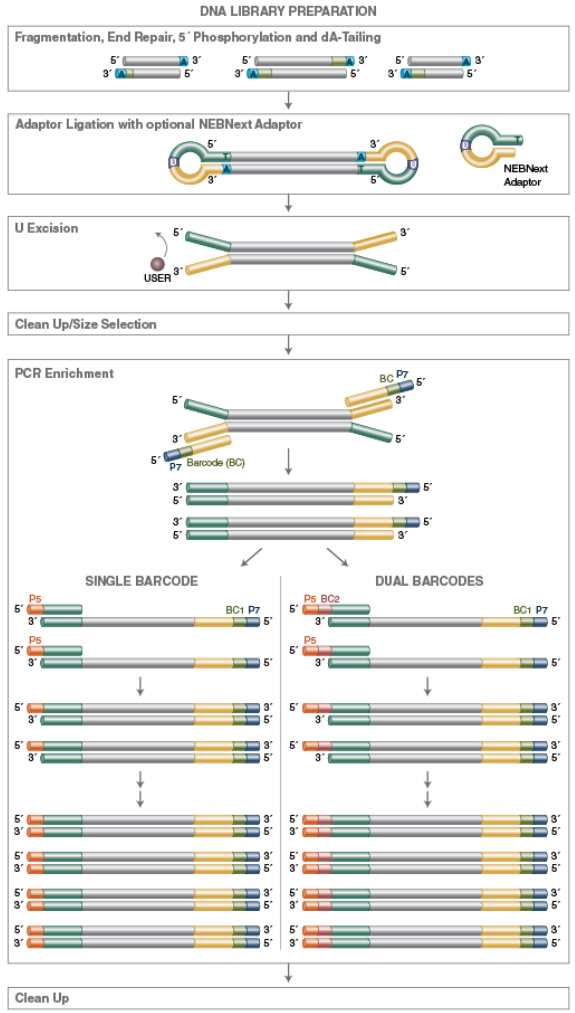


gDNA

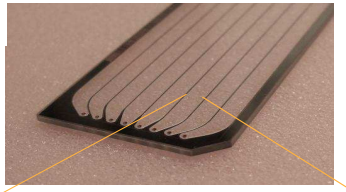
Shear dsDNA into fragments

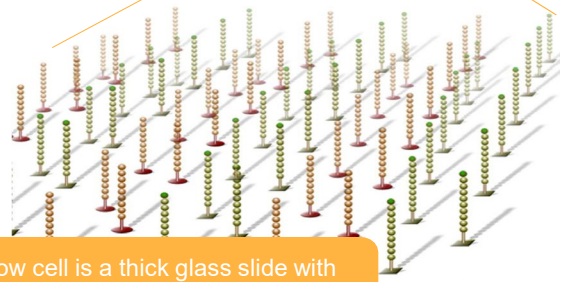
totalRNA





What is a Flow Cell?

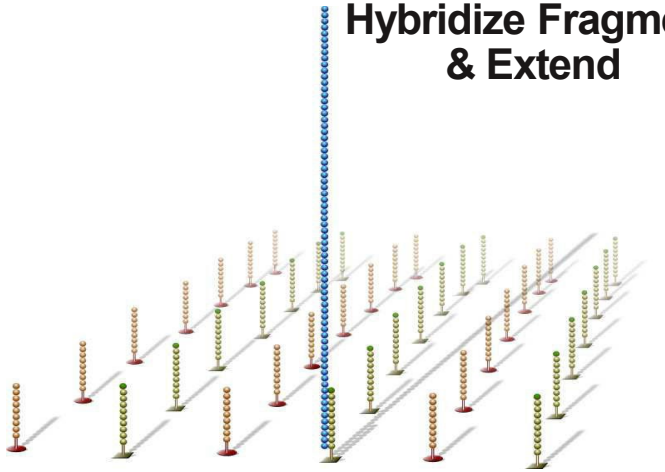




A flow cell is a thick glass slide with channels or lanes

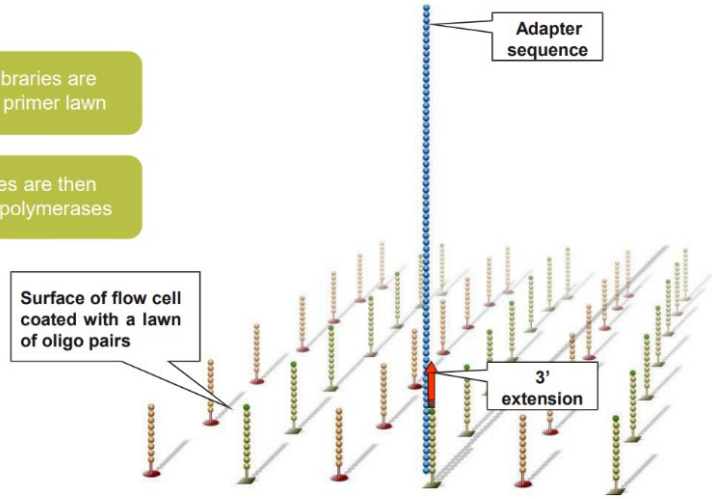
Each lane is randomly coated with a lawn of oligos that are complementary to library adapters

Hybridize Fragment & Extend



Single DNA libraries are hybridized to primer lawn

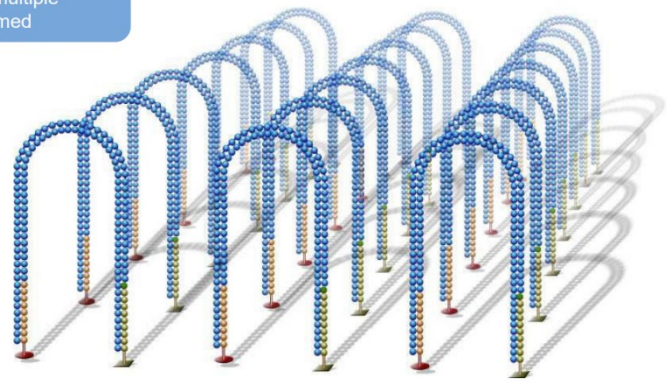
Bound libraries are then extended by polymerases



Single-stranded molecule flips over and forms a bridge by hybridizing to adjacent, complementary primer

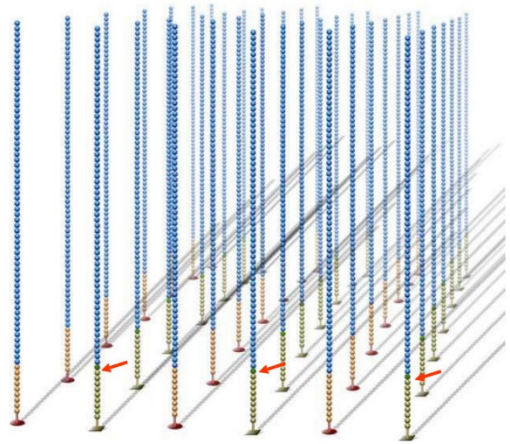
Hybridized primer is extended by polymerases

Bridge amplification cycle is repeated until multiple bridges are formed

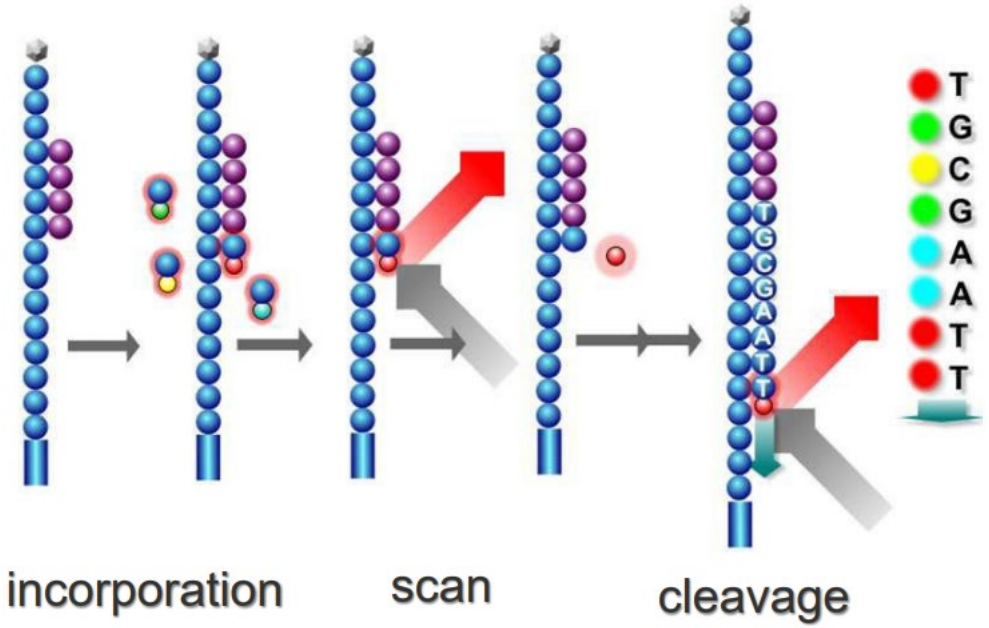


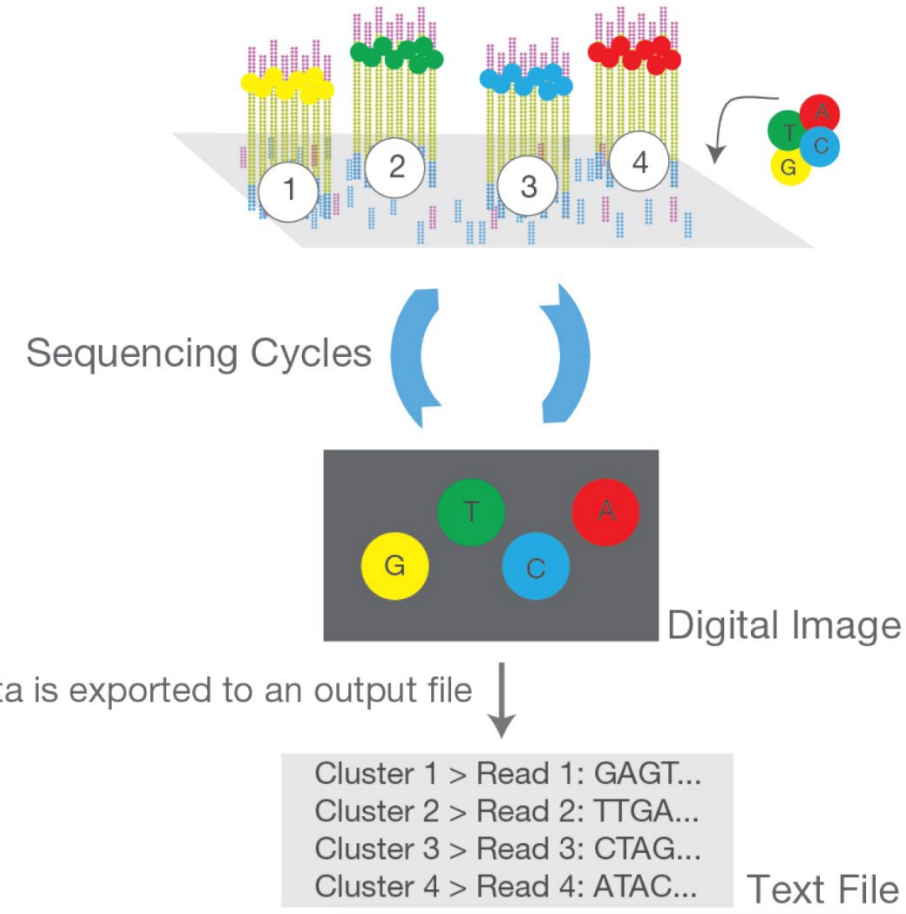
Bridge Amplification

dsDNA bridges are denatured



Linearization





illumina® Sequencing by Synthesis (SBS)

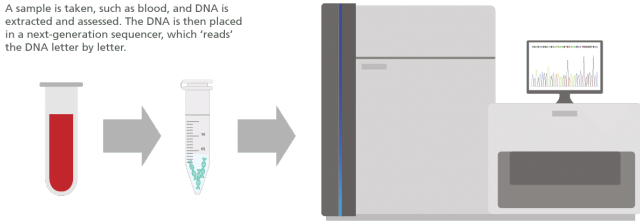
Sequencing reagents, including fluorescent nucleotides, are added and the first base is incorporated. The flow cell is imaged and the emission from each cluster is recorded. The wavelength and intensity of emission are used to identify the base. The cycle is repeated “n” times to create a read of the length of ”n” bases

Università della Svizzera italiana

Institute of Oncology Research

How the genome is interpreted: 1. Sequencing

A sample is taken, such as blood, and DNA is extracted and assessed. The DNA is then placed in a next-generation sequencer, which 'reads' the DNA letter by letter.



How the genome is interpreted: 2. Bioinformatics

The vast quantity of data generated by the sequencer is checked and filtered by bioinformaticians using computers and software tools.

The new DNA sequence is **quality assured** to ensure the bases are correctly identified.

The genome sequence is 'reassembled'. This is called **sequence alignment**.

The newly sequenced genome is compared against a reference genome to identify variants. This is known as **variant calling.**

This produces a file known as a **variant call file**. This details all the variants that have been discovered in the newly sequenced genome.

How the genome is interpreted: 3. Analysis

A team of scientists and clinicians work together to review and interrogate the variant call file to determine which information is relevant to report.

They will use a variety of tools and resources, depending on the clinical question:

Genetic databases

Research and literature

Statistics and modeling

Health information (**phenotypic data**) about the patient

How the genome is interpreted: 4. Results

A report detailing the findings of the genomic test is then produced. It will be sent to the patient's clinical team and used to inform their care.

Here are some of the possible outcomes:

Make a **diagnosis**

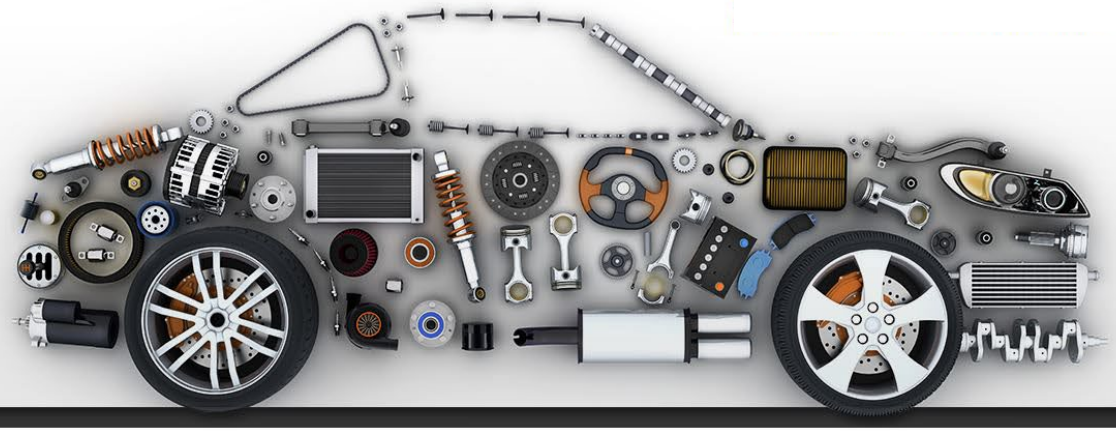
Identify possible **treatments**

Referral to relevant **clinical trials**

No result (though new evidence may emerge at a later date to support a diagnosis)

Practical aspects of single-cell experiments

Sometimes the sum of the PARTS is **GREATER** than the WHOLE



Credits: Weizmann institute

Single-cell Sequencing: Reviewing the technology landscape

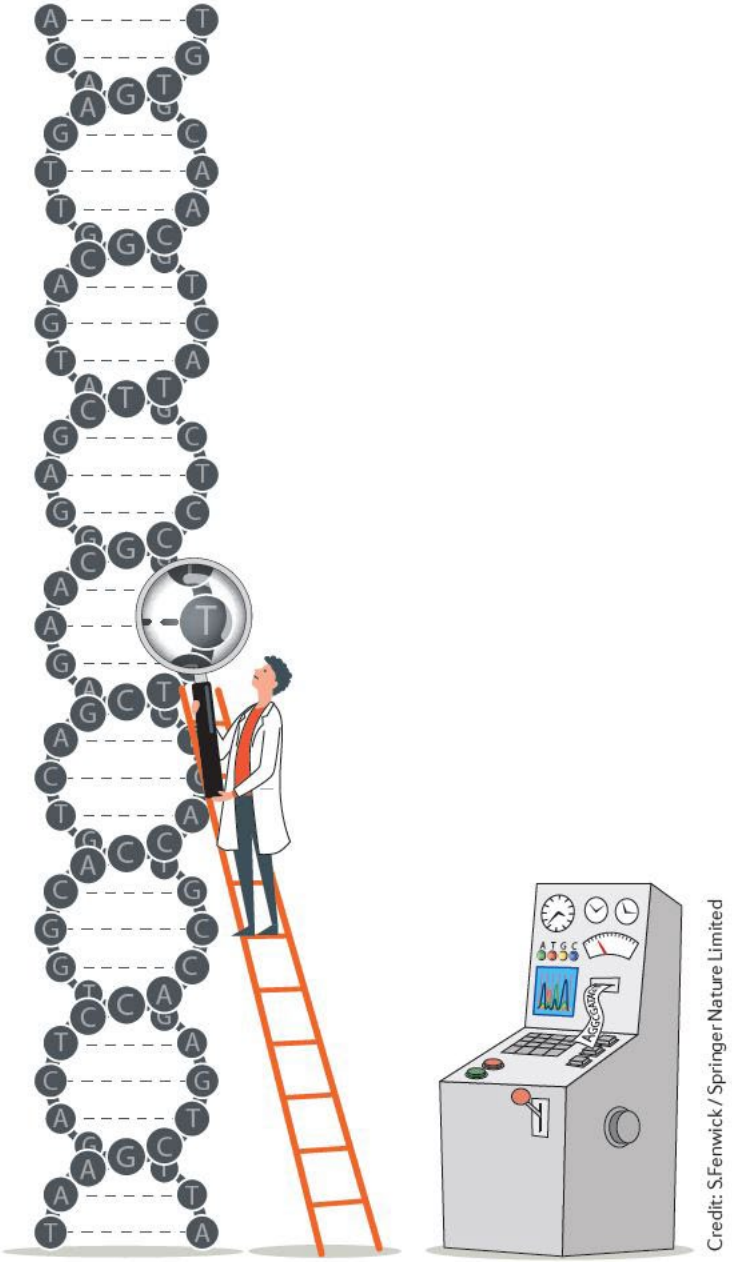
1) Cells in wells – active (Flow sorting, CellenOne, Fluidigm C1, SmartSeq)
- Screen for and retrieve single cells of interest
- Enrich for rare cells with decided properties
- Control the cellular microenvironment
- Monitor and control cell-cell interactions
- Precise/extensive manipulation of single cells

2) **Droplets** (Drop-seq, 10x Genomics)
- Introduce distinct 'packets' of reagents to single cells (e.g. barcodes)
- Perform amplification on individual cells
- Sort large population of single cells

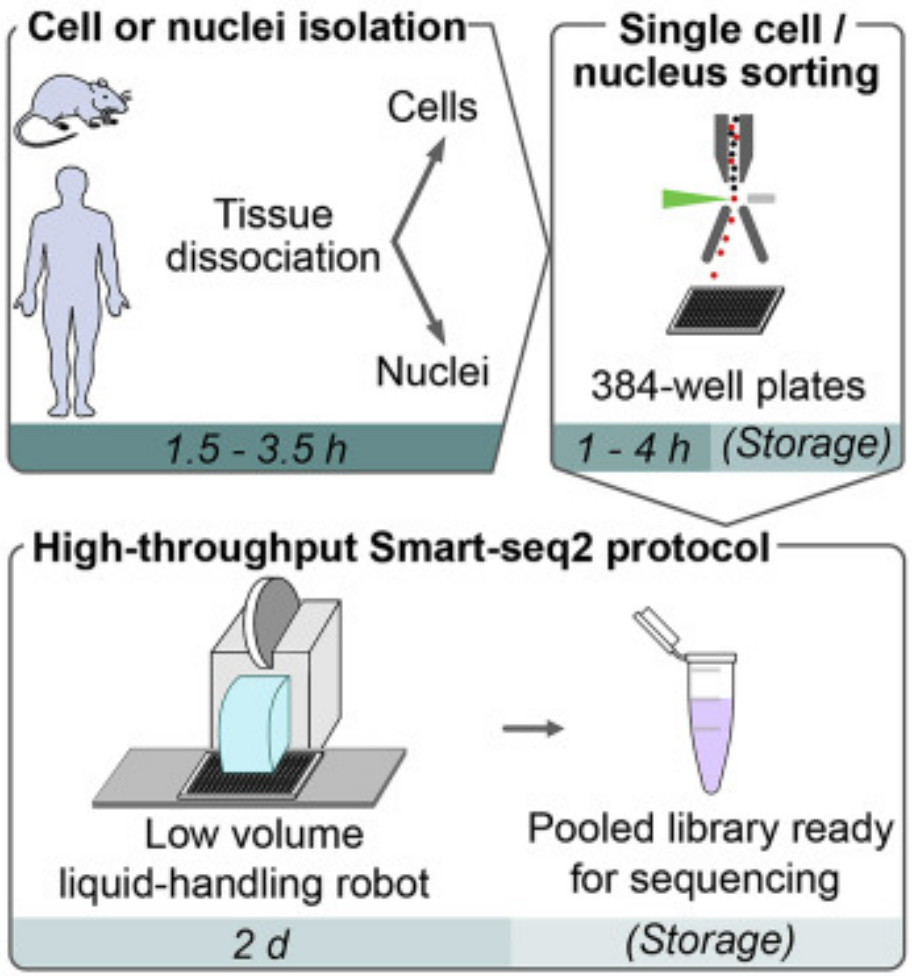
3) **Combinatorial indexing** (SCI-seq, SPLiT-seq, Parse)
- Economic use of reagents for cell separation
- Efficiency of handling a larger population than Drop-seq
- Maintain complexities of the population without bias from droplet or well

4) **Cells in wells - passive** (BD Rhapsody)
- large population of single cells – as droplets systems
- Efficiency of handling delicate population than Drop-seq (e.g. neutrophils)
- Maintain complexities of the population without bias from droplet or well

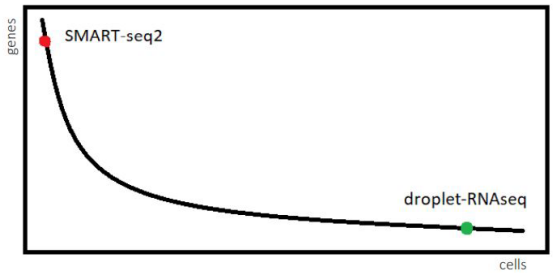
5) **Spatial Transcriptomics** (10xGenomics, …..)
- Bidimensional assessment of gene transcription



SmartSeq scRNA pipeline



More cells or more genes?

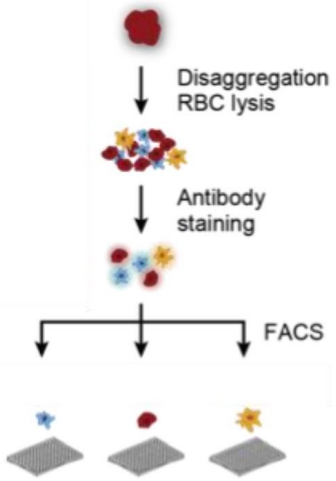


SMART-seq2
- 100 cells
- 1M reads per cell

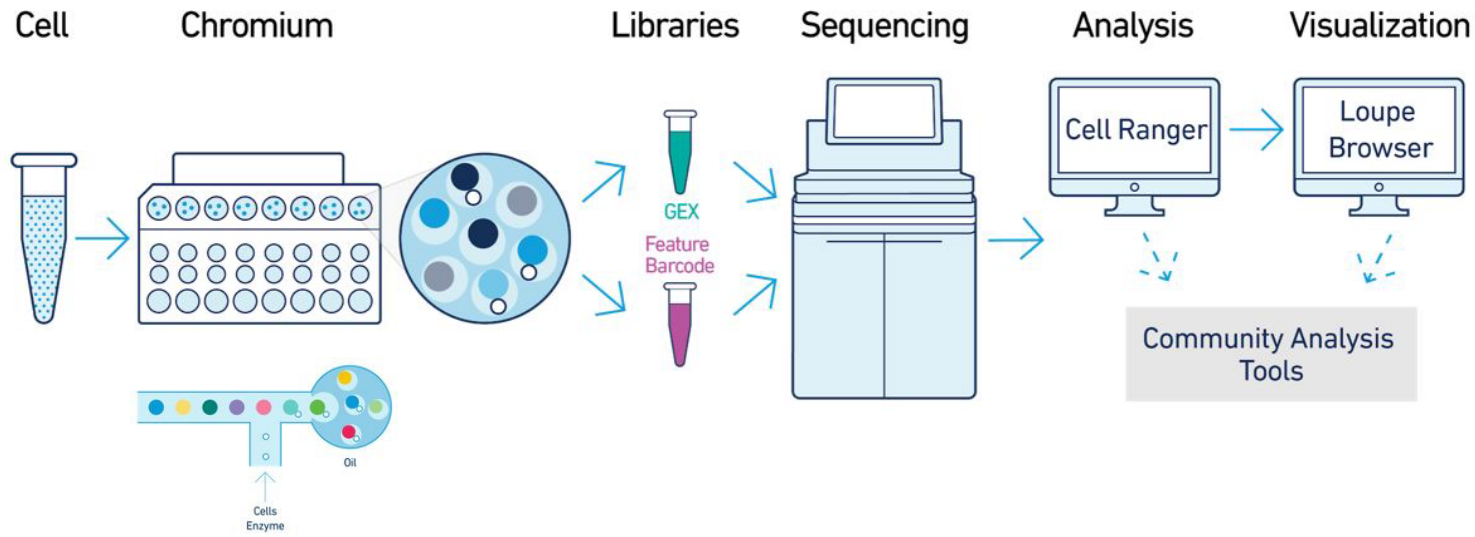
Droplet-RNAseq
- 10’000 cells
- 50k reads/cell

- You can modulate your experiment
- You can run all cells type
- Full length libraries

- You need a Sorter Facility
- Per-cell costs are considerable
- Mid-Low throughput



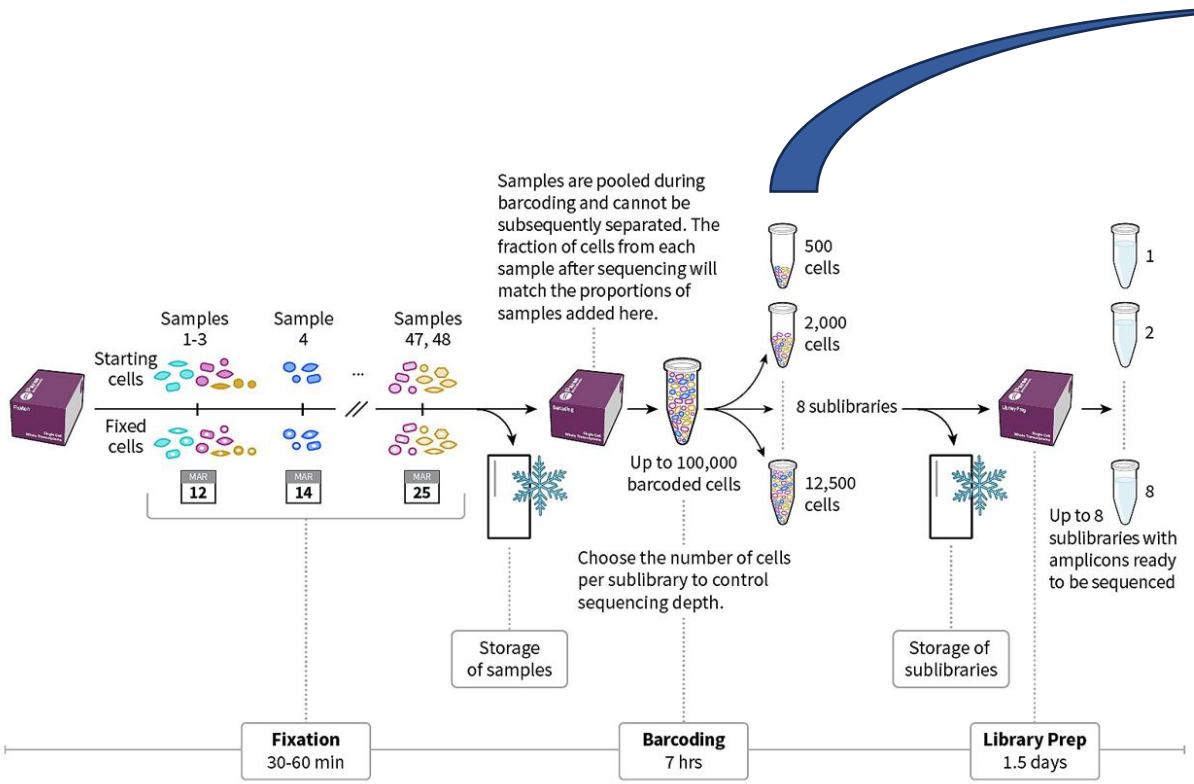
10XGenomics scRNA pipeline

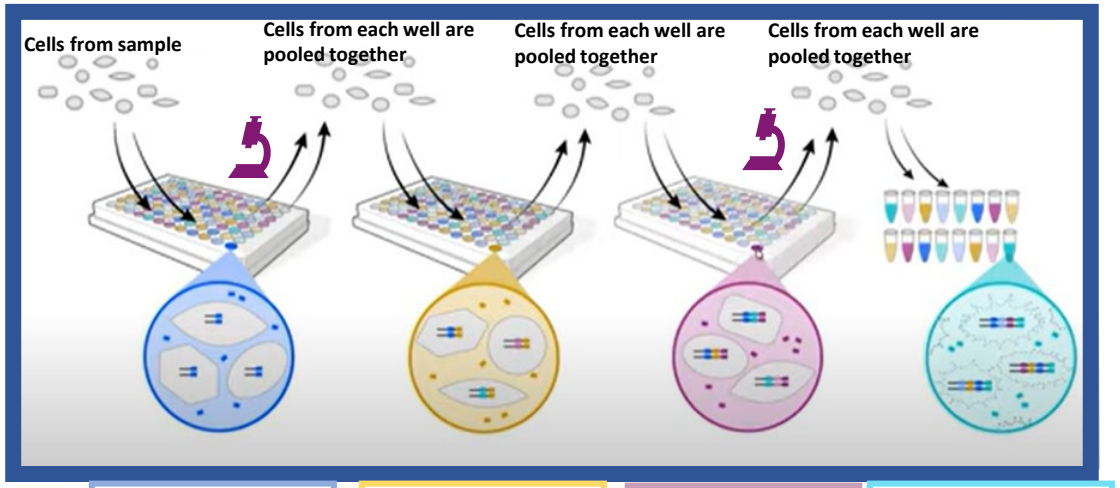


- RNAseq; TCR or BCR; Cell Surface sequencing; CRISPR screening; ATACseq
- Standardized instrumentation and reagents
- Easy to use, High-throughput scaling (up to 16 samples can be processed simultaneously)

- NOT all the cells are suitable for 10XGenomics

Parse scRNA pipeline (split-pool barcoding)





cDNA is generated with an in-cell RT reaction using well-specific barcoded primers

Cells are then distributed into 96 wells, in-cell ligation reaction append a second well-specific barcode to the cDNA

The third-round barcode is appended with another round of in-cell ligation

Cells spit into 8 distinct populations (sublibraries). The cells are then lysed, and the barcode is isolated. A fourth sublibrary-specific barcode is introduced by PCR to each cDNA molecule.

- Time flexibility – single experiment for samples collected on different dates (up to 6 months storage)
- No instrument is required for the experiment.
- Up to 48 samples / 100k cells in total – kit has to be used at once
- No 3'/5' bias – random hexamers method
- Works with any species, any sizes of cells/nuclei
- The lab work is arduous and burdensome

Parse

WTK Barcoding Minimum Retention Rate (MRR):
20%– 35%

WT Mini:	50K Fixed Cell Input	X	**20% MRR**	=10K to 60k Barcoded Cell Output
WT:	350K Fixed Cell Input	X	**28% MRR**	=100K Barcoded Cell Output
WT Mega:	2.85M Fixed Cell Input	X	**35% MRR**	=1M Barcoded Cell Output
WT Penta:	14.4M Fixed Cell Input	X	**15% MRR**	=2.2M Barcoded Cell Output

Evercode™ WT Mini
Up to 10K cells, 1-12 samples

Evercode™ WT
10K-100K cells, 1-48 samples

Evercode™ WT Mega
100K-1M cells, 1-96 samples

Evercode™ WT Penta
1M-5M cells, 1-384 samples

Parse

BD Rhapsody scRNA pipeline

Load cells and beads

Pair ONE cell with ONE barcoded bead in microwell

Lyse cells

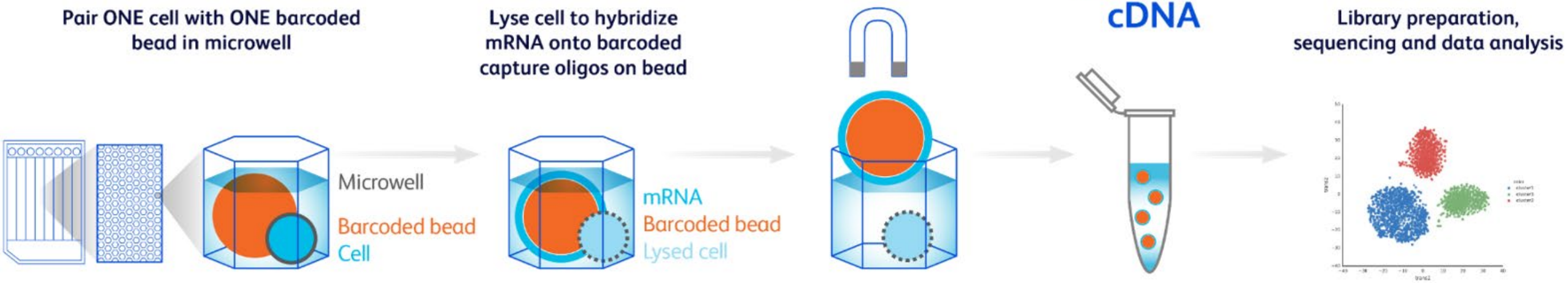
Lyse cell to hybridize mRNA onto barcoded capture oligos on bead

Retrieve beads

Synthesize cDNA

Analyze data

Library preparation, sequencing and data analysis

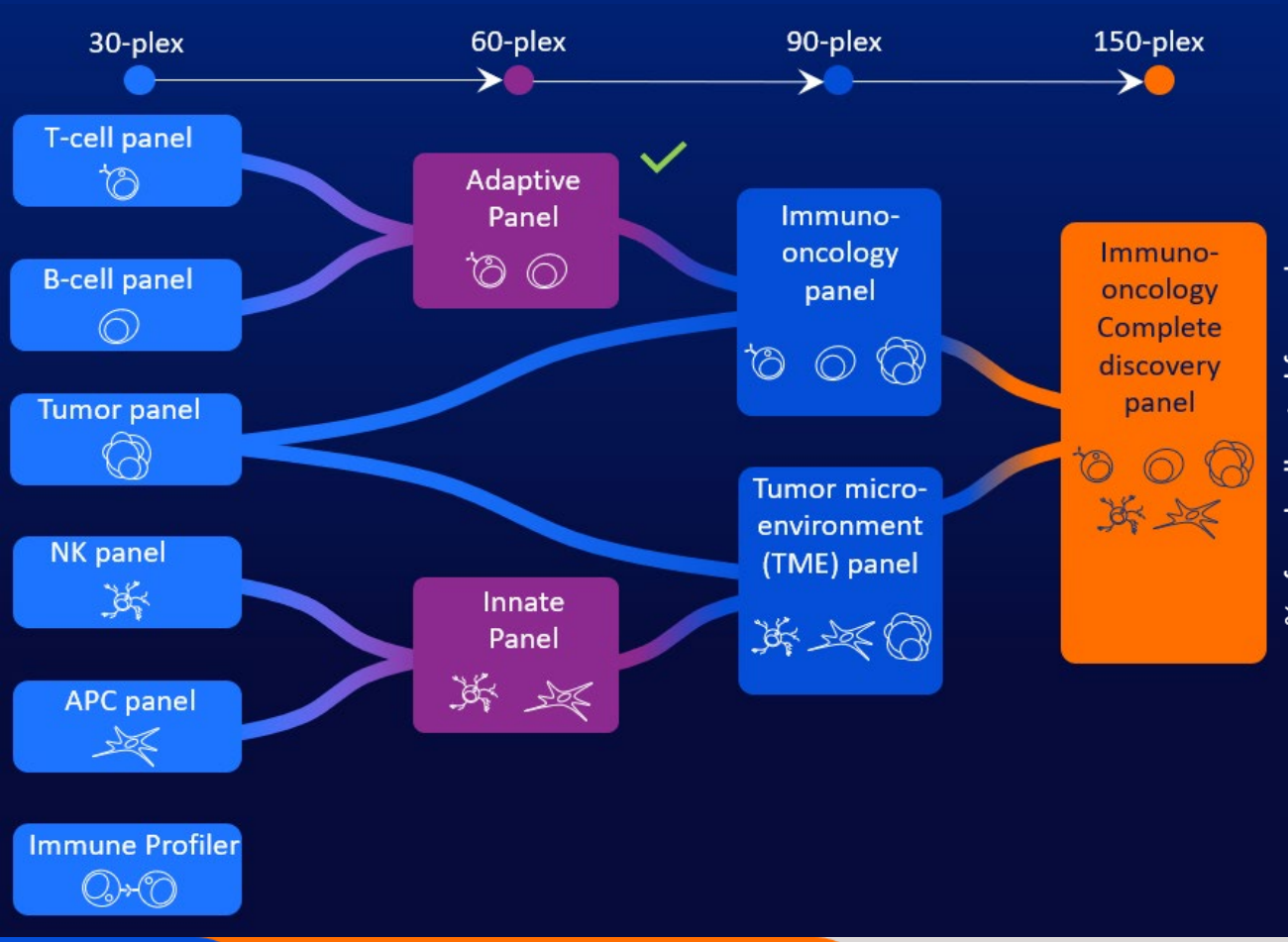


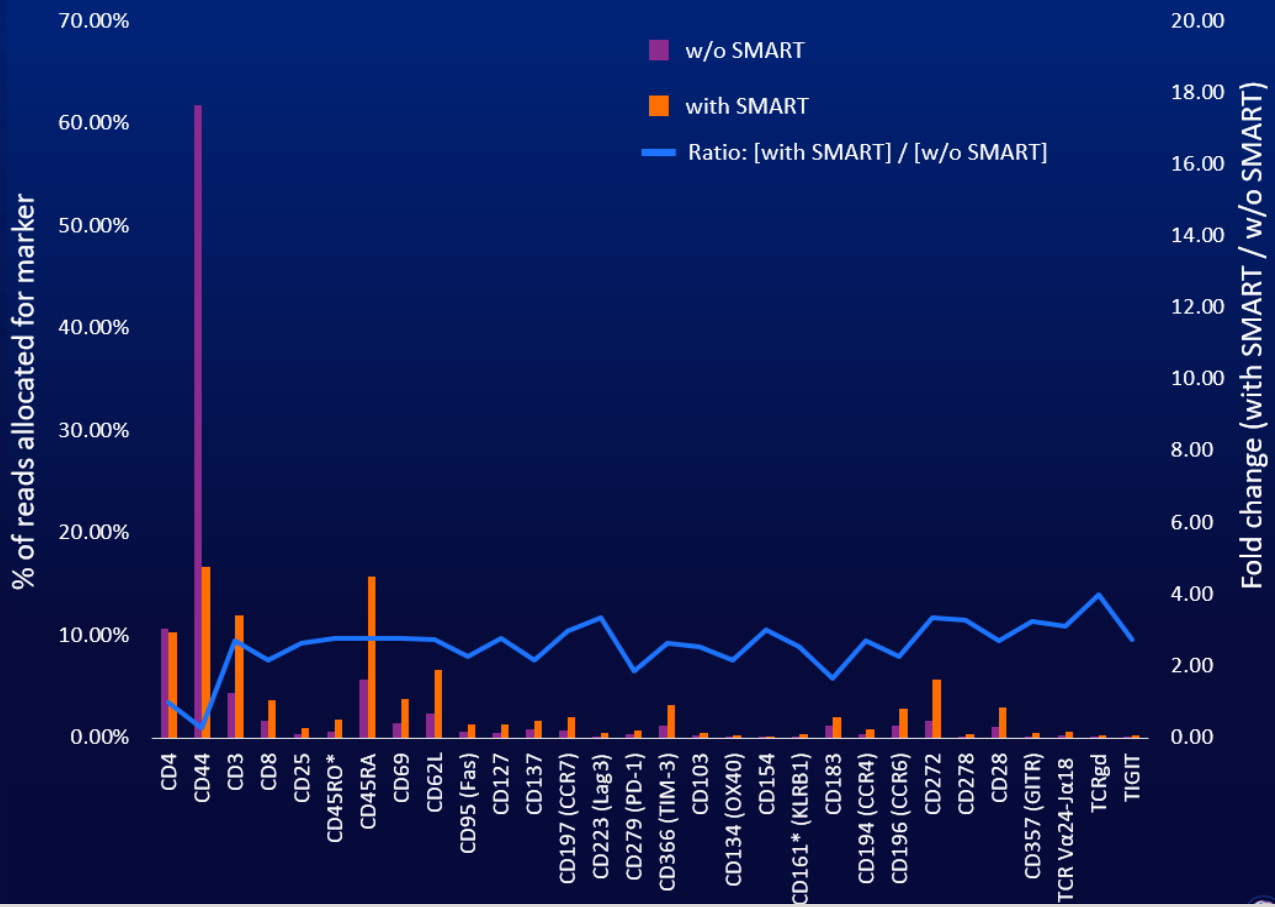
- 100k cells /lane (800k / cartridge)
- RNAseq; TCR or BCR; Cell Surface sequencing
- Works with any species, any size of cells

CITE-Seq data quality

Cellular Indexing of Transcriptomes and Epitopes by Sequencing

Example: T-cell panel: Detect 30 critical T-cell markers in single cell samples





BD

GeneMind
真迈生物

MobiNova-100

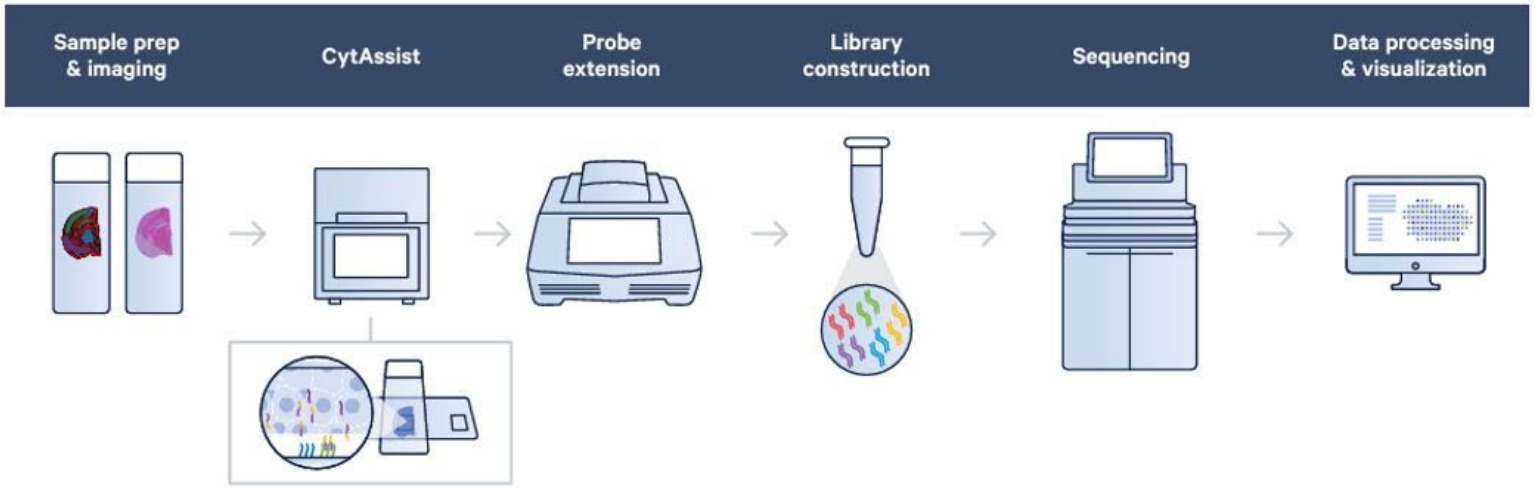
Single Cell System

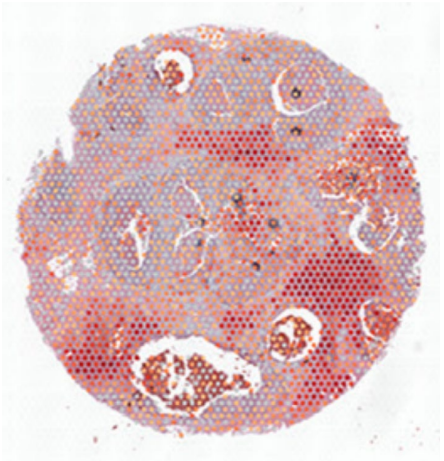
Single-cell Library Preparation Kit
3' Transcriptome, 5' Transcriptome
Microbial transcriptome, ChIP-seq

MobiNova®-D1

Tissue Dissociation Instrument

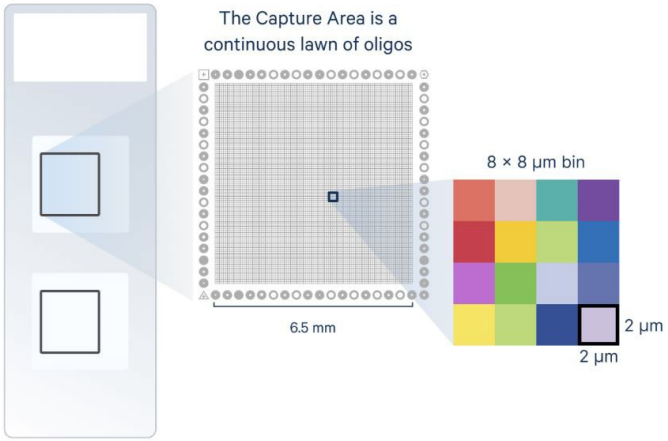
10XGenomics Spatial Transcriptomics pipeline – VISIUM HD



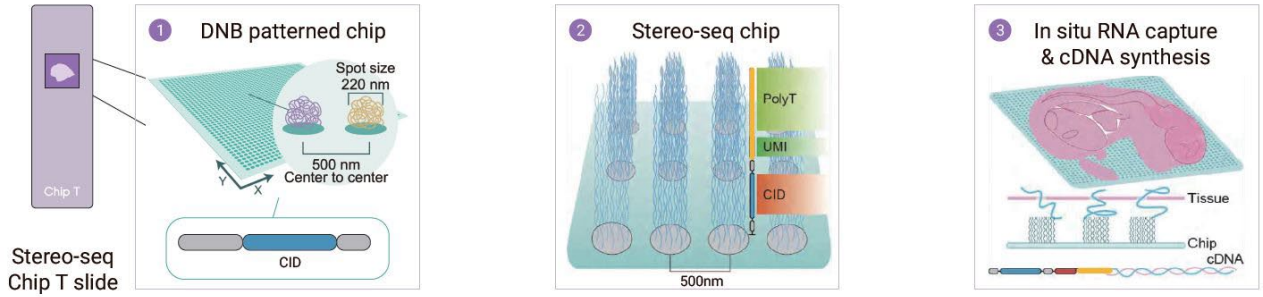


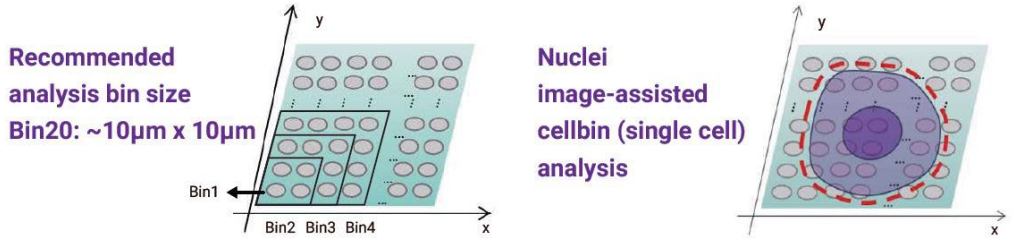
- The data is output an 2um resolution
- Capture areas are 6.5x6.5 mm

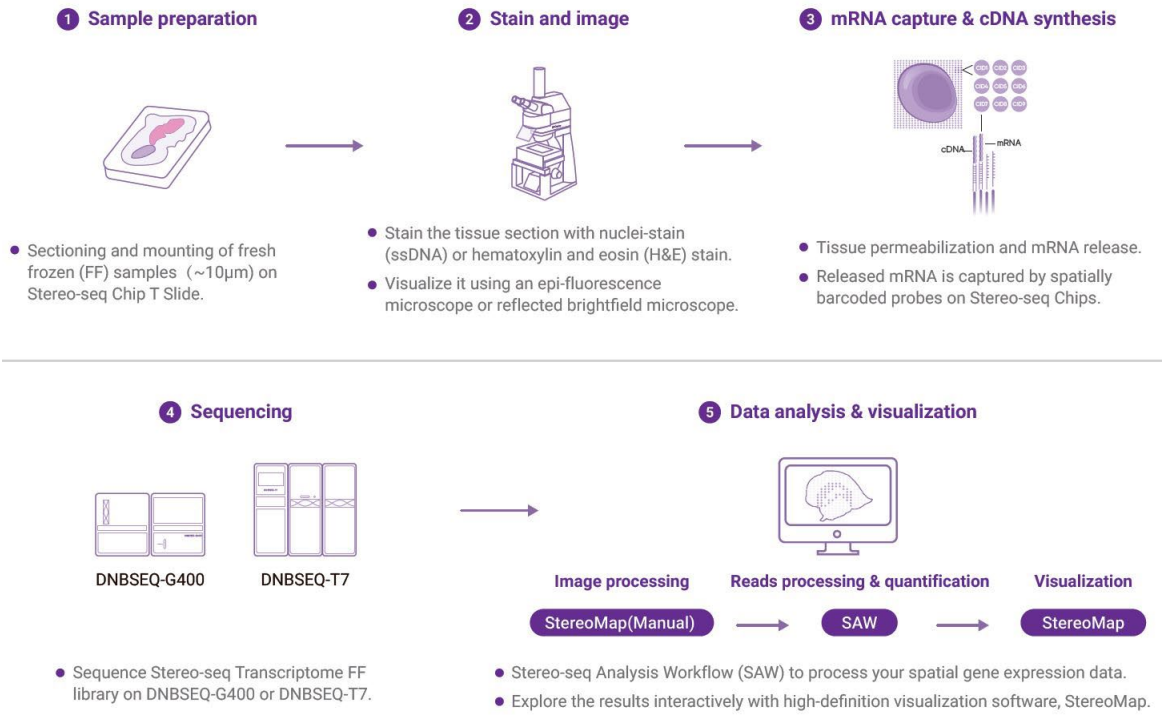
- Coordination with Pathologists and Imaging Facility



STOmics Spatial Transcriptomics pipeline – Stereoseq







- The data is output an 500nm resolution
- Capture areas are 50x50 mm or 100x100mm

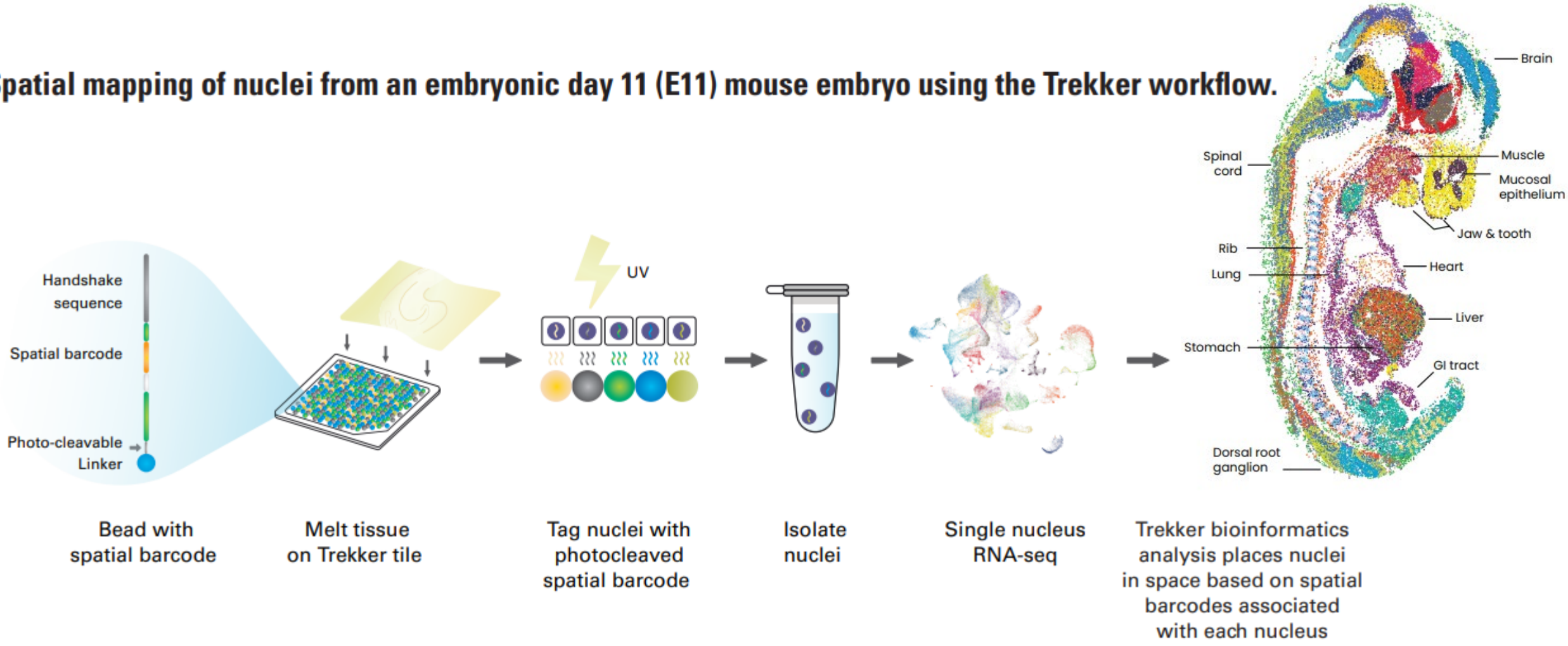
The smallest objects that the unaided human eye can see are about 0.1 mm long. Some cells are visible to the unaided eye. Smaller cells are easily visible under a light microscope. The power of a light microscope is limited by the wavelength of visible light, which is about 500 nm. To see anything smaller than 500 nm, you will need an electron microscope

STOmics

Trekker™ Single-Cell Spatial Mapping Kit

TaKaRa

Spatial mapping of nuclei from an embryonic day 11 (E11) mouse embryo using the Trekker workflow.



Bead with spatial barcode

Melt tissue on Trekker tile

Tag nuclei with photocleaved spatial barcode

Isolate nuclei

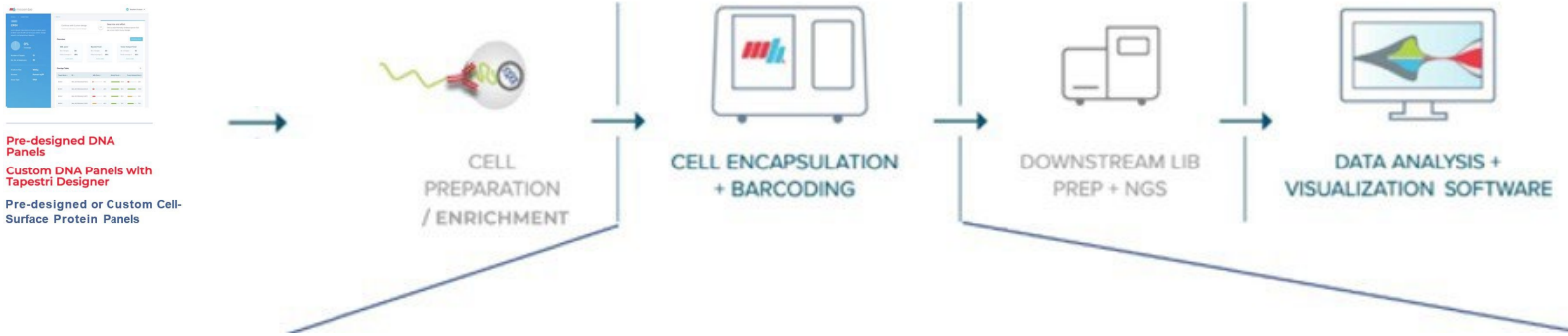
Single nucleus RNA-seq

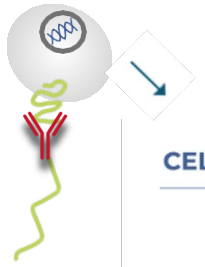
Trekker bioinformatics analysis places nuclei in space based on spatial barcodes associated with each nucleus

Tapestri – Mission bio scDNA

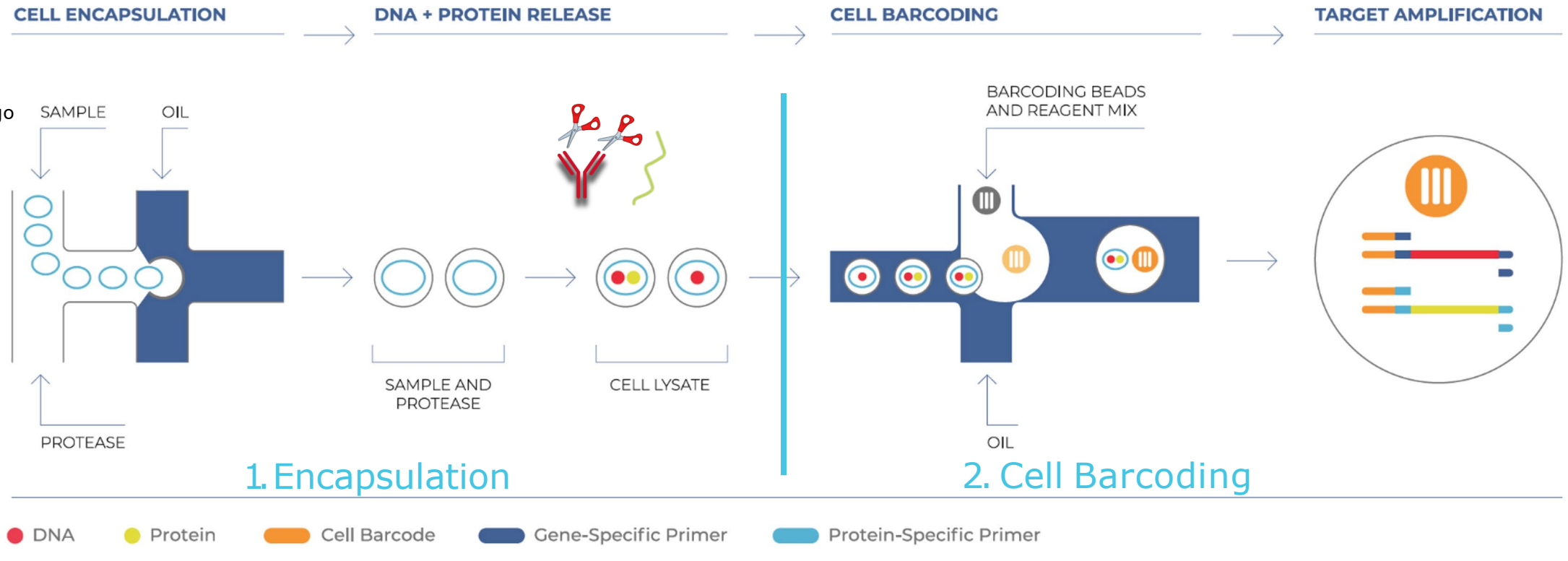
- Targeted single-cell DNA and protein profiling
- SNV/indel detection across 100s of targeted loci
- Gene-level and chromosome-level CNV detection
- Single-cell DNA throughput 5’000 cells – 10’000 cells
- Catalog panels or customizable content
- High sensitivity for rare clones – down to 0.1
- Compatible with any tissue type

mission bio





Antibody-Oligo conjugate for cell-surface protein detection



mission bio

BioSkryb GENOMICS

Whole genome or exome and transcriptome sequencing from a single cell

Uses a single cell for the construction of a whole-genome and full-length mRNA transcriptome library. Workflow available for hybrid capture compatibility.

Provides industry-leading genomic coverage and resolution

Leverages a novel patented technology, primary template-directed amplification (PTA), to dramatically increase genomic capture and coverage to 97%.[1,2]

Superior transcriptome capture and coverage

Increases gene body coverage, representation across transcript sizes, and variant calling versus droplet-based RNA sequencing methods.[3,4]

Università della Svizzera italiana

Institute of Oncology Research

Thank You!

andrea.rinaldi@ior.usi.ch