PRODUCED BY:

Front Line Genomics

THE SINGLE-CELL AND SPATIAL BUYER'S GUIDE

SPONSORED BY

Miltenyi Biotec

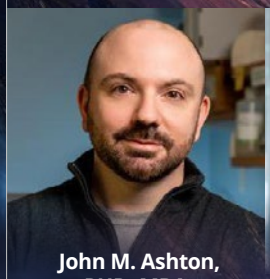
S2 genomics

sapio sciences

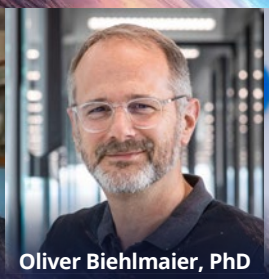
TaKaRa

TECAN

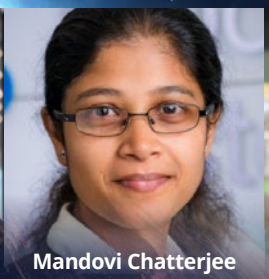
CONTRIBUTORS



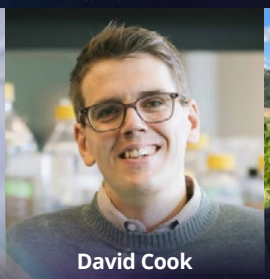
John M. Ashton, PHD, MBA
Associate Professor, Department of Biomedical Genetics, Director, Genomics Research Center
Wilmot Cancer Institute, University of Rochester



Oliver Biehlmaier, PhD
Head of the Imaging Core Facility at the Biozentrum
University of Basel



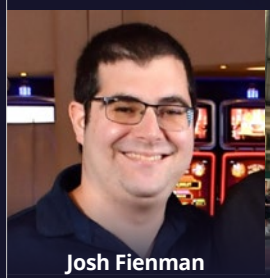
Mandovi Chatterjee
Director, Single-cell Core
Harvard Medical School



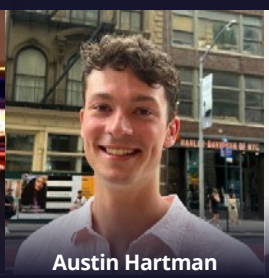
David Cook
Scientist, **Ottawa Hospital Research Institute**
Assistant Professor, Department of Cellular and Molecular Medicine
University of Ottawa

Andrea Corsinotti
Single-cell Multi-omics Facility Manager, Centre for Regenerative Medicine, Institute for Regeneration and Repair
University of Edinburgh

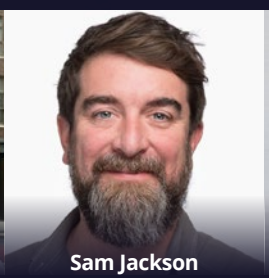
Jonathan Coxhead
Senior Experimental Officer
Newcastle University



Josh Fienman
Scientist, Genomics (NGS Technology Center)
Pfizer

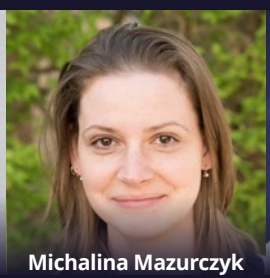


Austin Hartman
PhD Candidate
Stanford University

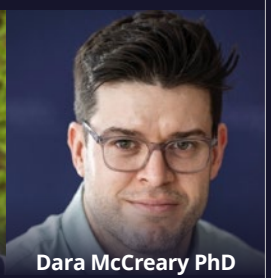


Sam Jackson
Tools and Technology Platform Manager
UKDRI

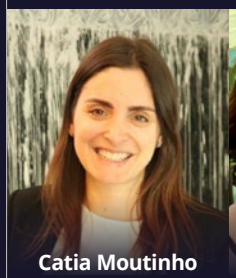
Jan-Phillipp Mallm
Head, Single-cell Open Lab
DKFZ German Cancer Research Center



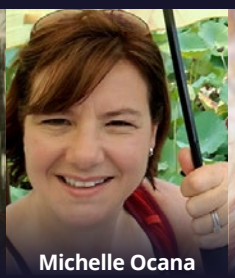
Michalina Mazurczyk
Mass Cytometry Facility Manager, Medical Research Council Weatherall Institute of Molecular Medicine, Radcliffe Department of Medicine
University of Oxford



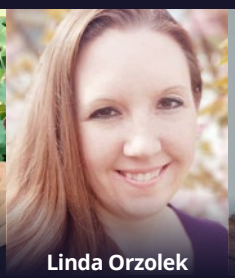
Dara McCreary PhD
Scientific Business Analyst
Sapio Sciences



Catia Moutinho
Founder & Scientific Advisor
The Single-Cell World

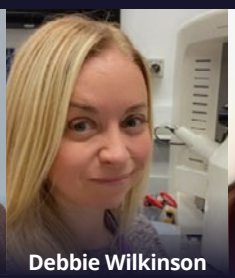


Michelle Ocana
Managing Director, Neurobiology Imaging Facility
Harvard Medical School

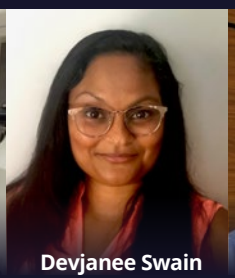


Linda Orzolek
Director, Single Cell & Transcriptomics Core, **John Hopkins University** (at time of contribution)
Single cell & Transcriptomics Senior Product Manager
Psomagen (Now)

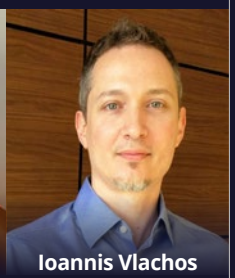
Carolina Oses Sepulveda
Researcher and Lab Manager
Spatial Proteomics Unit in SciLifeLab (KTH Royal Institute of Technology)



Debbie Wilkinson
Co-Manager and Senior Microscopy Application Specialist, Microscopy and Histology Core Facility
University of Aberdeen



Devjanee Swain Lenz
Director, Sequencing and Genome Technologies
Duke Center for Genomic and Computational Biology



Ioannis Vlachos
Associate Professor, **Harvard Medical School**
Director of the Spatial Transcriptomic Technologies Unit
Beth Israel Deaconess Medical Center

FOREWORD

Matt Higgs

Science Writer
Front Line Genomics

WHILE SINGLE-CELL AND SPATIAL WORKFLOWS DIFFER GREATLY IN PROCEDURE AND NUANCES, THE OVER-ARCHING REASON FOR CHOOSING ONE OF THESE METHODS REMAINS CONSISTENT – TO PROFILE A TISSUE WITH GREATER SPECIFICITY THAN A BULK METHOD.

But for your project, should you use single-cell, or spatial, or both? Which analytes should you measure? Which platform and workflow would be best for you? There are many convoluted questions for those wanting to use these technologies. Answering these questions is the rationale behind the production of this Single-cell & Spatial Buyer's Guide. In an exciting technological market with a variety of options, a resource that provides a dose of clarity is much needed. I believe this report is that dose of clarity.

Within these pages, you will find separate guides for picking the best workflow for a single-cell and for a spatial biology experiment. The report covers the end-to-end decision making, from sample prep, to instrument choice through to analysis options and microscopy advice. Additionally, it provides direct comparisons between the latest commercial offerings for single-cell and spatial. Finally, it looks at the broader points that may impact your purchasing choice, such as the ESG policies of the companies operating in the single-cell and spatial market.

A brief disclaimer on the contents of this Buyer's Guide. As I have said, you will find specifications for, and detailed descriptions of, many products in the single-cell and spatial market. This information was gathered to the best of our ability, using the information that is made publicly available on company websites, information from academic articles and from conversations with the companies themselves. We have strived to ensure accuracy in all the information we have provided but we do not claim that the information is without error, nor can this information reflect any developments that occur after being written.

Another exciting angle of this Buyer's Guide is that it was forged by interviewing a series of experts in the field. Through this approach, we have gained a set of unique insights and guidance, which have shaped this resource. Excerpts from our discussions with these experts, and discussions between the experts themselves, are found throughout the chapters. Within these interviews and panel discussion, you will find advice on how to get the most out of your single-cell and spatial workflow, hard fought wisdom gained from working with these technologies, as well as perspectives and views on current topics in single-cell and spatial.

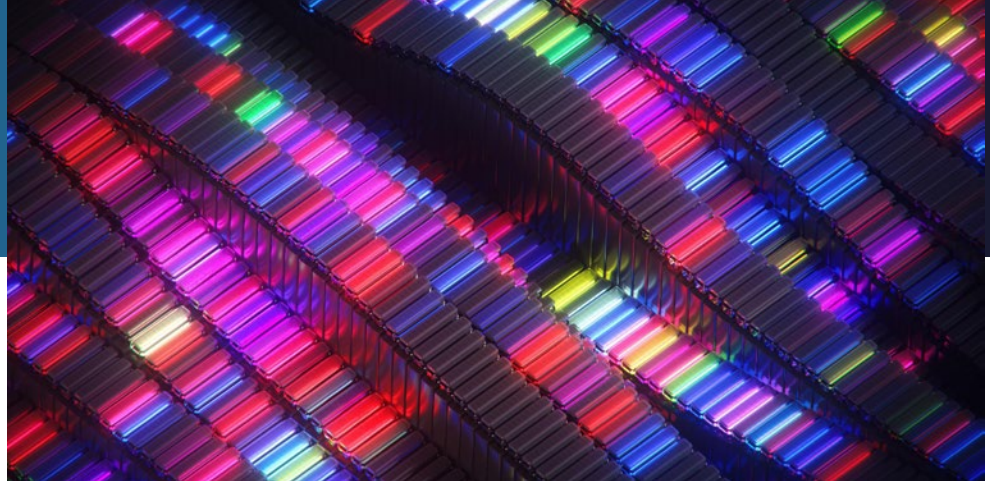
We would like to take this opportunity to thank all of those contributors for their time and insights when writing this Buyer's Guide.

We would also like to thank the sponsors of this report: Miltenyi Biotec, Sapio Sciences, Takara Bio, Tecan and S2 Genomics.

We hope this Buyer's Guide helps you during your decision making process.

Thank you for reading.

CONTENTS



CELL PREP POSSIBILITIES : SAMPLE PREP AND CELL ISOLATION METHODS FOR SINGLE-CELL

THIS CHAPTER INTRODUCES THE BASIC SINGLE-CELL WORKFLOW, AS WELL AS BEST PRACTICES AND THE LATEST COMMERCIAL METHODS IN SAMPLE PREPARATION AND CELL ISOLATION, READY FOR SINGLE-CELL SEQUENCING.

Single-Cell Sequencing

In some ways, single-cell sequencing can be seen as the natural progression from bulk tissue sequencing. Biological processes and diseases are complex and heterogenous, and the desire to understand this complexity arose at the same time that single-cell sequencing took off. Bulk sequencing still has its uses, but it can only take you so far. Single-cell sequencing was the tool, with the capacity to capture nuance, necessary for the job.

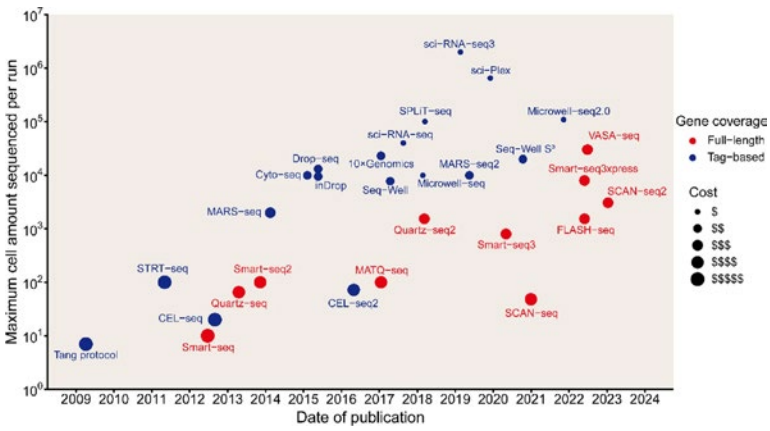
Primarily used for RNA sequencing as opposed to DNA sequencing, single-cell methods deconvolute bulk tissues into individual cells that can then be separately sequenced. Although this results in only a small amount of genetic material per cell, you gain an appreciation of each unit of the tissue, as opposed to an aggregate reading across it. Hence, single-cell RNA sequencing allows researchers to understand many aspects of a disease from the level of each individual cell, which is the level at which diseases tend to act. It allows us to ask how gene expression differs between healthy and diseased cell types and what might this mean for the disease profile and progression[1].

Single-cell DNA sequencing is also valuable. While healthy cells in a tissue will have the same genome, single-cell DNA sequencing can identify somatic or germline mutations in specific cellular populations, which can help with investigations in cancer, ageing and neurodegeneration. The issue with this methodology arises from the fact that there are very small amounts of genomic material in a single cell. Whole genome DNA methods, such as multiple displacement amplifications (MDA)[2], multiple annealing and looping-based amplification cycles (MALBAC)[3] and degenerate oligonucleotide primed PCR (DOP-PCR)[4], allow this analysis at genome-scale by amplifying low abundancies of DNA.

There are two main types of single-cell transcriptomic methods, based on whether they are sequencing mRNA through a primed tag or whether they use full-length transcriptomic methods. Both methods have seen the same overall advancement - an increase in cell throughput, alongside the subsequent decrease in cost per cell to run single-cell methodologies (Figure 1.1).

FIGURE 1.1. DEVELOPMENT OF SINGLE-CELL RNA SEQUENCING (SCRNA-SEQ) TECHNOLOGY.

Timeline and throughput of various scRNA-seq methods. This scatterplot depicts the published date and throughput of sequencing for each technology. The colour indicates the different gene coverage. Size indicates the cost per sequenced cell of scRNA-seq methods. Source: Huang, et al. [12]



This process has advanced, from the early days of manually isolating individual cells, to the first fluidic circuits that allowed the processing of 100s of cells[5,6], and eventually to droplet methods (inDrop[7] and Drop-seq[8]), allowing 10s of 1000s of cells to be processed. Finally, combinatorial indexing methods (sci-RNA-seq[9,10] and SPLiT-seq[11]) have brought us to the current era, in which million-cell experiments are now a viable reality for most researchers.

Single-cell sequencing still relies on NGS sequencers, and single-cell research has benefited from the advancement in sequencing technology capabilities (see our Sequencing Buyer's Guide for a full coverage of this).

The Single-Cell Workflow

A very simplistic single-cell sequencing workflow can be seen in Figure 1.2. Once tissue is procured, a single-cell suspension needs to be generated through gently breaking down the tissue. Individual cells need to then be isolated in well-plates or in contained reaction vesicles. Once individual cells are isolated, these cells are lysed so that the RNA is captured separately for each cell and then RNA is converted to cDNA to undergo standard NGS library prep, sequencing and analysis.

FIGURE 1.2. BASIC WORKFLOW OF SINGLE-CELL SEQUENCING.
Source: Pan, et al. [13]

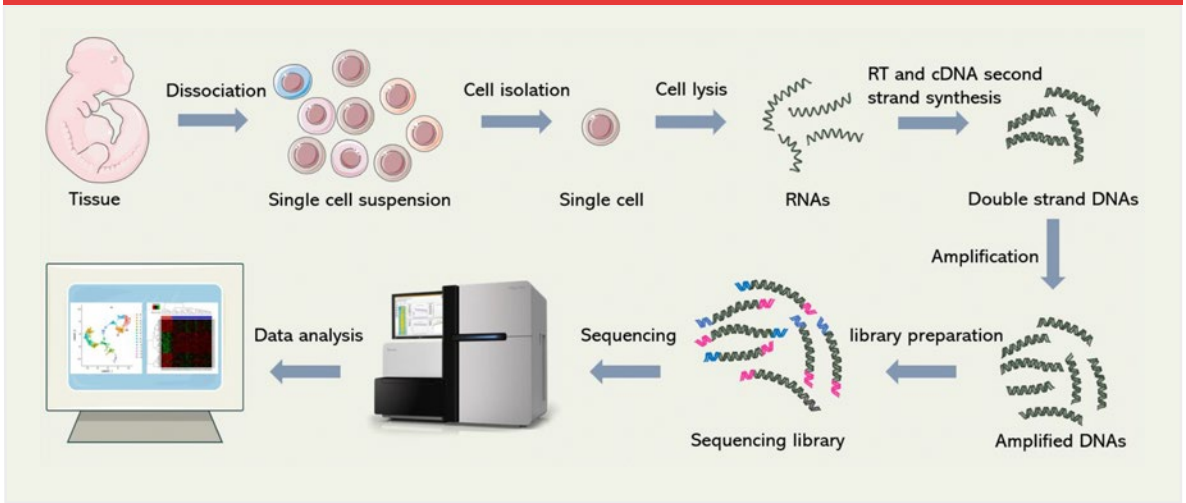


Figure 1.3 highlights each of these different steps of the workflow alongside a list of the methods and some key considerations for each step. Chapter 2 of this Buyer's Guide will go into depth on the single-cell platforms, and Chapter 8 will cover computational analysis. This chapter will cover the methods for tissue dissociation and cell enrichment.

Efficient sample prep is known for being crucial to performing an effective single-cell study. The common phrase in the field is – 'crap in, crap out'. We direct readers to an excellent review paper from 2023 that covers a multitude of sample preparation techniques with advice and guidance[14]. The remainder of this chapter will review the popular methods for the preparation of individual cells and review the commercial, automated instruments that could assist you to standardise the process.

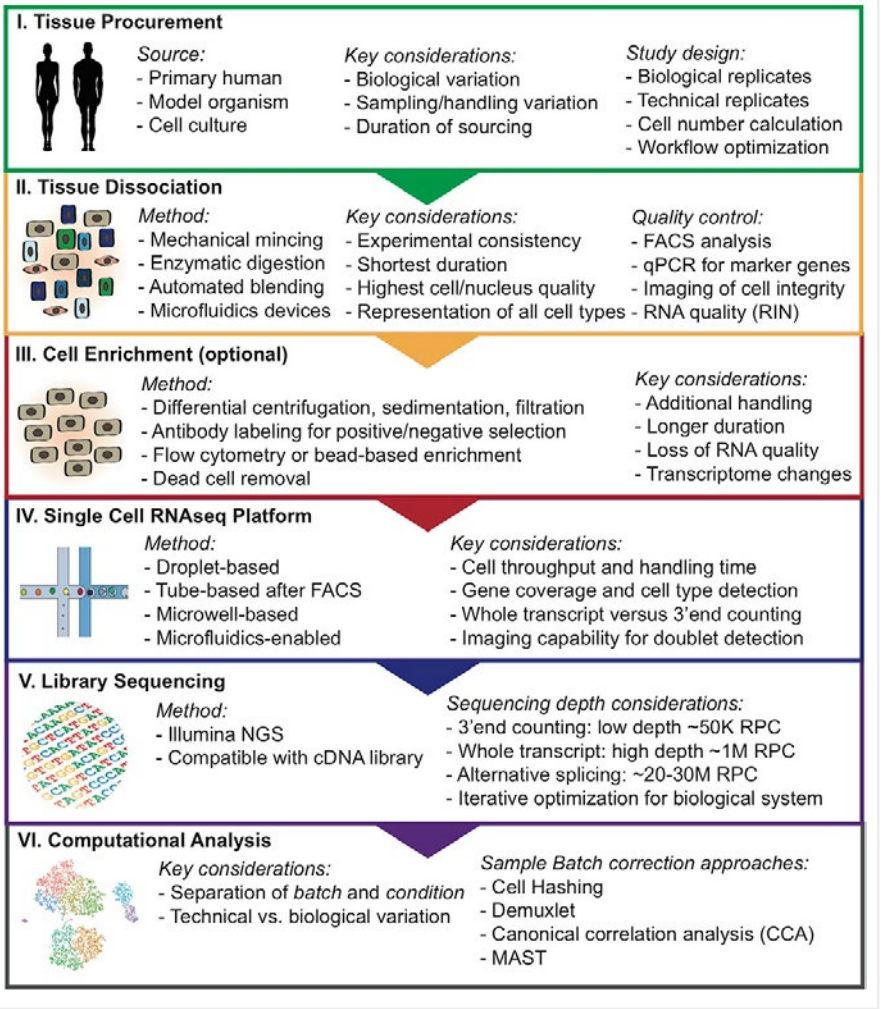
Tissue Dissociation

Once the tissue samples are procured, the principal step in preparing the sample for a single-cell workflow is the tissue dissociation step. The goal is to convert a tissue sample into a suspension of single cells. This process is arguably the greatest source of unwanted technical variation and batch effects. A typical protocol for tissue dissociation involves (1) tissue dissection, (2) mechanical mincing, and (3) enzymatic breakdown.

Since tissue dissociation is such a critical step for ensuring experimental consistency, standardisation is essential. While there is little to replace the standardisation the comes from years of practiced wet lab work, commercial semi-automated platforms go a significant way to allow more reproducible, time-saving and efficient tissue dissection and single-cell preparation.

FIGURE 1.3. STEPS OF THE SINGLE-CELL WORKFLOW WITH METHODS AND KEY CONSIDERATIONS.

Source: Nguyen, et al. [15]



The benefits of automated tissue dissociation for single-cell sequencing can be seen as providing:

1. **Consistency:** they allow for consistent processing of samples, reducing the variability between samples, and improving data accuracy.
2. **Speed:** the dissociation process saves a lot of time, allowing researchers to process more samples in less time. This reduces labour costs and increases the number of samples that can be analysed.
3. **Quality:** It provides more efficient dissociation of tissue, leading to higher yields of viable single cells. It also minimizes the risk of sample contamination.
4. **User-friendly:** they are easy to use and require minimal training. They can be programmed to process various tissue types and sample sizes.
5. **Cost-effective:** although automatic dissociation can be expensive, they can save money in the long term by reducing the need for multiple technicians and increasing the number of samples processed.

Below you will find a selection of the commercially-available dissociators on the marker. When choosing a tissue dissociator, it is important to consider the following factors:

1. The tissue type. Check if there is already a program for it.
2. The throughput requirements for your application.
3. The cost (of the device and all specific materials, if required).
4. The maintenance and service requirements.
5. Other customers' recommendations.

GENTLEMACS™ DISSOCIATOR - MILTENYI BIOTEC

The gentleMACS™ Dissociator is a benchtop instrument for semi-automatic or automatic tissue dissociation. The original dissociator is semi-automatic and can process 1 or 2 samples in parallel. The gentleMACS™ Octo Dissociator with Heaters is the larger, fully automated, counterpart with room to process 8 samples, and can also use heaters. By using one of the tissue-specific MACS Tissue Dissociation Kits with pre-defined enzyme mixes, and pre-defined programs for the gentleMACS Dissociators, there is no need to test multiple dissociation protocols to generate single cell suspensions with high viability.

Both dissociators use dedicated gentleMACS tubes (processing 20 mg – 4,000 mg of tissue per tube). The C tubes are for generating single-cell suspensions or single-nuclei suspensions, while the M tubes are for further homogenization into subcellular material. The recently released gentleMACS Perfusers replaces tedious and complicated manual perfusion methods with automated ex vivo perfusion of liver to get reproducible results isolating viable hepatocytes. Over 40 preset programs exist for various human and mouse tissues plus user-defined programs can be created, saved, and shared. Click here for the user manuals for the original and Octo Dissociators.

PYTHON® TISSUE DISSOCIATION SYSTEM - SINGLERON

The PythoN® system integrates heating, mechanical and enzymatic dissociation in one 15 minute workflow. 8 samples can be processed in parallel with compatibility for 200+ tissue types. It also works with tissue weight from as little as 10mg to 4000mg. Up to 50 custom dissociation programs can be stored.

The system uses disposable Dissociation Units and Singleron's sCelLive® Tissue Dissociation Mix. Demo data shows this combination produces >85% cell viability across a range of tissues.

The PythoN Junior™ is a compact and flexible version of the original. It works on the same principals and process with a 15 minute standard run time but with fewer samples (1-2) per run. This makes it ideal for clinical settings, since suspensions can still be generated from small amounts (~10mg).

SINGULATOR™ - S2 GENOMICS

The Singulator™ Platform is a fully automated single cell and single nuclei isolation solution that is comprised of three components: Singulator instruments, single-use cartridge consumables, and cell or nuclei isolation kits. The reproducibility and precision of the Singulator platform removes a major bottleneck to conducting single cell genomics research, enabling more scientists to process more samples for single cell genomic applications.

The Singulator Platform combines manual dissociation with enzymatic dissociation to isolate cells and manual extraction with chemical dissociation to isolate nuclei. Cells are dissociated from fresh tissue with a 20-60 minute run time achieving viability up to 90%. Nuclei are isolated from fresh, and flash frozen tissue with a 6-10 minute run time from as little as 2 mg of sample input. Finally, the Singulator 200+ automates deparaffinization and dissociation of nuclei from FFPE sections in as little as 40 minutes, enabling snRNA-Seq for FFPE samples.

VIA EXTRACTOR™ - CYTIVA LIFE SCIENCES

The VIA Extractor™ works via single-use sample pouches that allow 3 samples to be processed in parallel. Run times can be as low as 10 minutes but are adjustable and pouches allow up to 1g of tissue. Viability scores tend to score highly (80%+) and yields for difficult tissues are high due to the VIA Freeze™ temperature control function, which can coordinate speed, temperature and time to maximize cell viability.

TISSUEGRINDER - FAST FORWARD DISCOVERIES

The FFX TissueGrinder is an enzyme-free tissue dissociator for single-cell applications. This compact benchtop device has four grinding slots for mechanical dissociation and works in under 5 minutes to generate a suspension from tissue.

Standard labware (Falcon Tubes) are fitted with the FFX grinder and strainer, which apply a combination of shearing and cutting to release viable cells into suspension. The pattern of mechanical processing is controlled by the instrument.

DSC-400/DSC-800 - RWD LIFE SCIENCE

The DSC-400 and DSC-800 are single cell suspension dissociators that use mechanical and enzymatic digestions. Runs typically take between 15 and 30 minutes. The 400 version has 4 independent working channels and the 800 has 8. Both systems can process samples from 20 mg to 4000 mg. Due to patent infringement, the tissue processing tubes of RWD are not available in the USA market.

Reliable nuclei extraction and enrichment has never been easier

Single-nuclei RNA sequencing (snRNA-seq) is a great way to take advantage of frozen tissue or cell types that are not easily dissociated, but isolating high-quality single-nuclei suspensions. That's not always a simple task — until now.

The Miltenyi Biotec end-to-end solution streamlines the process, making it easy to extract intact nuclei from virtually any tissue, followed by enrichment of nuclei suspensions to remove debris. First, automate extraction of nuclei from up to eight fresh or frozen samples in under 10 minutes with the gentleMACS™ Octo Dissociator in conjunction with gentleMACS C Tubes, gentleMACS Octo Coolers, Nuclei Extraction Buffer, and MACS® SmartStrainers. Some tissues, like brain or liver, yield nuclei suspensions with significant amounts of debris, which will diminish snRNA-Seq data quality. The Anti-Nucleus MicroBeads were designed for such challenging sample types and will increase the purity of nuclei through MACS Technology in less than 30 minutes. Visit our website to learn more about our streamlined workflow that reproducibly delivers debris-free nuclei suspensions from fresh or frozen tissue in just 40 minutes.

LEARN MORE

For further information on nuclei extraction and enrichment, visit

▸ **miltenyibiotec.com/nuclei-enrichment**

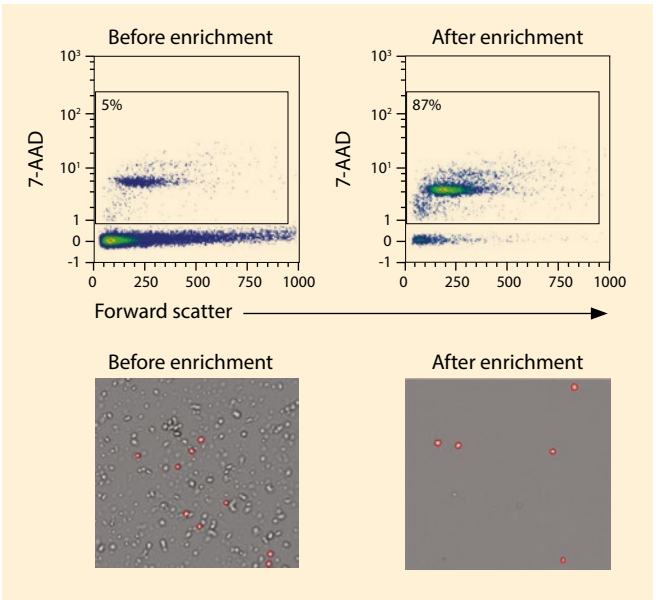


Figure 1: Enrichment of nuclei suspensions from adult mouse brain. Top: Nuclei were stained with 7-AAD and analyzed by flow cytometry (MACSQuant® Analyzer, Miltenyi Biotec) before and after enrichment with Anti-Nucleus MicroBeads. Nuclei increased from 5% to 87% after enrichment. **Bottom:** Nuclei were stained with AO/PI and imaged (CellDrop™ Cell Counter, DeNovix). PI-positive nuclei are circled in red, and all other objects are debris. Visible debris is nearly eliminated after enrichment.

▸ **miltenyibiotec.com**

Miltenyi Biotec, Inc | 6125 Cornerstone Court East | San Diego, CA 92121, USA | Phone 800 FOR MACS | macsus@miltenyi.com | **www.miltenyibiotec.com**

Miltenyi Biotec provides products and services worldwide. Visit **www.miltenyibiotec.com/local** to find your nearest Miltenyi Biotec contact.

Miltenyi Biotec

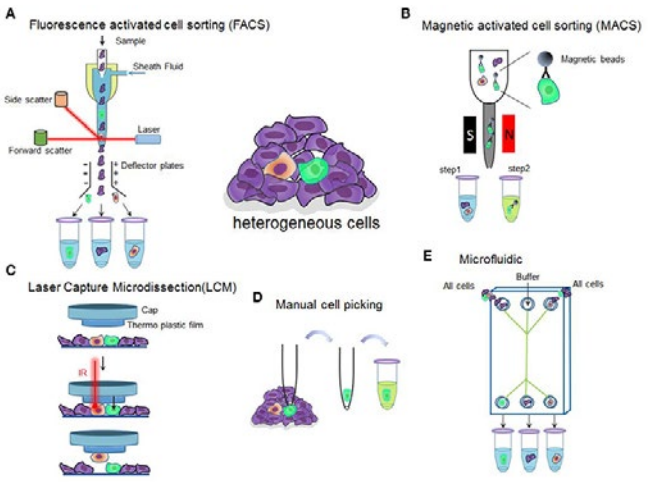
Cell Sorting and Filtering

Filtering liquids or tissue dissociations to remove debris and dead cells is an important step in single-cell workflows. An associated optional step is single cell enrichment or cell sorting. In this step, a single cell suspension is filtered for debris and enriched for cell types of interest. This puts rare cell types to the forefront. There are several methods to accomplish this.

Filtering is a basic method of removing debris and clumps while retaining whole cells, and strainers are a good tool for the job. Commercial strainers include the pluriSelect® pluriStrainer® and the MACS® SmartStrainers. Different mesh sizes tend to be used, starting with large sizes (100 µm) and reducing as needed (e.g. to 70 µm and 40 µm). Every round of straining will result in unwanted cell loss, so a balance needs to be struck

FIGURE 1.4. DIFFERENT METHODS OF CELL ISOLATION AND SORTING

This includes FACS (A), MACS (B), LCM (C), Manual (D) and Microfluidic (E) strategies. Source: Hu, et al. [16]



When it comes to actually sorting those whole cells, the two most common high-throughput methods to achieve this, and enrich for rare cell types, are flow cytometry or fluorescence-activated cell sorting (FACS) and Magnetically Activated Cell Sorting (MACS) (see Figure 1.4). Manual sorting and Laser Capture Microdissection (LCM) can also be an option, but at a much lower throughput and require larger amounts of manual work.

FACS works through fluorescently labelling cells via antibodies. The cells undergo flow cytometry but pass through a laser that allows cells to be sorted based on the colour emitted. This helps sort cells by proteins expressed on their exterior. Advantages of this approach include the extremely high throughput and the much higher cell purity levels that are achieved. However, FACS requires a large quantity of cells in suspension to sort, and the rapid flow can damage cells, impacting their viability downstream.

If the goal is to sort out one or two cell populations (for example T and myeloid cells), a benchtop model with two sort streams and 4-5 colour capabilities is sufficient. If, on the other hand, there is a need for multiple cell populations (for example, naïve, memory and effector subsets of T or B-cells), then a larger floor model with multiple sort streams and multiple fluorescent channels will be necessary (see below for a collection of options).

MACS, on the other hand, is simpler and more affordable and provides more opportunity to scale. It works using supermagnetic nanoparticles to tag target cells and then capture them through magnetically attaching to a column, while the remaining cells pass through. The downsides are that the throughput and resulting purity are lower than FACS. Examples of MACS include MojoSort™ from BioLegend (see below), DynaCellect™ from Thermo Fisher Scientific and the MACS cell separators from Miltenyi Biotec. The latter works using MACS® Microbeads and includes a variety of column types. The automated instrument – autoMACS® NEO Separator (pictured) – automates the process, with capacity for 6 samples, for a high purity, reproducible and gentle cell isolation.

BioLegend's MojoSort™ magnetic bead-based system is designed to reliably isolate desired populations of cells with high purity and yield. Each reagent is optimized for a sample for quick and easy isolation. The beads are small at 130 nm, and isolated cells have been tested in downstream in vivo and in vitro applications to demonstrate cell function, migration, and proliferation. MojoSort™ reagents are compatible with other magnets and column separation systems and have beads and kits designed to separate a wide variety of cell types from both mouse and human samples. They also offer MojoSort™ Human and Mouse Dead Cell Removal Kits, which does not require high calcium buffer exchange, resulting in higher live cell yield compared to Annexin-based kits.

Cell sorting is complicated to set up if high-throughput is required. The popular single-cell provider, 10x Genomics, has a set of best practices for sorting cells prior to their assays, which can be found here. While MACS sorting is pioneered and led by Miltenyi Biotec, FACS sorting has a wide-variety of commercial options available for cell sorting in advance of single-cell sequencing, that we briefly review below.

FACSDISCOVER™ S8 – BD BIOSCIENCES

BD Biosciences has a selection of FACS based Cell Sorters with different specs for different applications. The FACSDiscover™ S8 (pictured) is the latest cell sorter with an impressive 86 detectors, 5 lasers and 3 different nozzle sizes. With BD CellView™ Image Technology and BD SpectralFX™ Technology, this spectral flow cytometer is combined with real-time spatial and morphological insights. A 96 well plate can be filled in less than 80 seconds with 80% cell viability.

WOLF® & VERLO - NANOCELLECT

The WOLF® G2 Cell Sorter uses up to 2 lasers and 9 colours for sorting. Cells are sorted into well plates or conical tubes and the back-to-back sorting speed is 200 events per second, with a sample pressure of <2psi for extremely gentle sorting. Sample volumes are a minimum of 150µl.

The VERLO Image Guided Sorter is the new instrument currently available for early access. This sorter uses image analysis alongside gentle microfluidic cell sorting. It captures actual images of each cells alongside morphology and marker localization.

PALA™ – BIO-TECHNE

The Pala™ benchtop Cell Sorter from Bio-Techne is a lightweight and portable instrument. It uses two lasers and up to 11 fluorescent channels with forward and side scatter. It can efficiently sort cells into well plates, at 1 minutes per plate. It also sorts at a low pressure of < 2 psi for gentle sorting. Input volume is between 100 µl – 600 µl. These single-cell dispensers use unique microfluidic cartridges, and these sorters have three different modes, capable of sorting cells at 2 cells per second up to 50,000 cells per second. The instrument is user-friendly, require no special training to operate, and need zero maintenance.

CELLENONE® – CELLENION

The CellenONE® X1 is Cellenion's free standing single cell sorter and isolator using four fluorescent channels and two dispensing channels. Sample processing speed is < 3 mins for a 96-well plate. The imaging-based cell selection allows users to select cells based on morphology or fluorescent signature. Furthermore, it works on any sample volume, from as little as 1µl, meaning it excels at selecting out rare cell populations even in sparse samples.

CYTOFLEX – BECKMAN COULTER

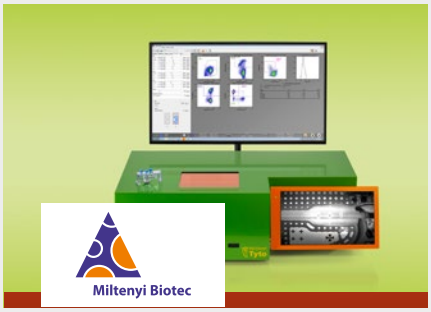
The CytoFlex SRT Benchtop Cell Sorter comes in a variety of different models containing between 2-4 lasers and 5-15 colour filters to match different sorting applications. The most sensitive of these models can sort at <10,000 events per second sorting into plates, slides or conical tubes.

F.SIGHT™ OMICS - CYTENA

The SIGHT Dispenser range from Cytena have a variety of specialisation including the UP.SIGHT™ for colony tracking, the C.SIGHT™ for unlabelled cells and the F.SIGHT™ for dispensing of fluorescent cells. The F.SIGHT™ OMICS is their model specialised for single-cell genomics. This system uses brightfield and fluorescent imaging for precise cell identification and can isolate a 96 well plate in ~ 2 minutes. Single cells are dispensed in picolitre droplets at the centre of conical PCR plate wells. Cell size filters range from 10 µm - 40 µm.

DISPENCELL™ - MMI

The DispenCell™ Single-Cell Dispenser accurately dispenses individual cells into microplates, wells or slides using a disposable tip sorter rather than microfluidics. It operates under severely low pressure - <0.2 psi – similar to a manual pipette. This instrument is low throughput and takes 10-13 minutes to dispense a 96 well plate, but it is extremely gentle, perfect for low throughput delicate studies.



MACSQUANT® TYTO® CELL SORTER
- MILTENYI BIOTEC

With 3 lasers, 8 fluorescent channels and 2 scatter channels, the MACSQuant® Tyto® allows for complex sorting strategies with 10 parameters. Samples are kept sterile in the specially designed Tyto Cartridges, which allow micro-chip based fluorescent cell sorting. The cell flow rate is 55,000 cells per second for the normal cartridge and 110,000 cells for the HS cartridge. The operating pressure is 3 psi for a normal cartridge and <14 psi for the HS. The cartridge loads 100 µl up to 10ml.

GENESIS CELL ISOLATION SYSTEM
– BIO-RAD

The Genesis System with Celselect Slides™ utilize microfluidic channels and 140,000 individual microchambers to efficiently and gently capture 8 µm – 30 µm rare cells based on size. It can accommodate liquid biopsy sample inputs of <10 mL from two samples in parallel. After capture, enriched cells can be recovered for downstream processing (e.g., bulk and single-cell sequencing, digital PCR, etc.), or stained on-slide for immunofluorescent applications such as enumeration and identification of various circulating tumor cells (CTCs) and other rare cell types using microscopy. For successful enumeration, Bio-Rad's Celselect Slides Validated Antibodies for CTC enumeration can be used in tandem with the Celselect Slides Enumeration kit. For more information, visit the Genesis Cell Isolation System Knowledge Hub.

CELL PREP POSSIBILITIES : SAMPLE PREP AND CELL ISOLATION METHODS FOR SINGLE-CELL

SH800S - SONY BIOTECHNOLOGY

The SH800S Cell Sorter uses 4 collinear excitation lasers and 6 fluorescent detectors and 2 scatter channels. It outputs into tubes, well plates and slides. The novel microfluidic sorting chip is available in three sizes including 70 µm, 100 µm, and 130 µm to permit sorting of a wide range of cell sizes. The 70 µm chip can sort 12,000 events per second with >98% purity. Up to 30,000 events per second can be achieved but it will impact purity.

UNO SINGLE CELL DISPENSER™ - TECAN

The Uno system is an automated benchtop instrument designed to improve single-cell workflows. This method does not use FACS but microfluidics to effectively dispense cells. It works with cells ranged 9 µm - 25 µm and dispenses them directly onto well plates. It can dispense a 384-well plate in ~5 minutes with 90% cell viability. The Uno dispenses both cells and reagents across the well plate.

ON-CHIP® SORT - PHCBI

The On-chip® Sort is a disposable microfluidic chip cell sorter. This results in a damage-free, sterile system capable of sorting cells as large as 140 µm. It works at speeds of 1,000 targets a second using 3 lasers with six fluorescent detectors on top of the microfluidic chip.

BIGFOOT SPECTRAL CELL SORTER – THERMO FISHER SCIENTIFIC

The Bigfoot system is fast, capable of sorting a 96-well pate in 11 seconds and 70,000 events per second. It has an impressive 9 lasers and 60 detectors capable of spectral sorting and spectral analysis. A six-position multi-sample loader allows six parallel samples. With configurable output holders and nozzles, this system can be adjusted to many applications, even dispensing directly onto a 10x Genomics chip.

ALERION™ – AKADEUM LIFE SCIENCES

The Alerion™ is an interesting alternative to FACS and MACS, instead using BACS or Buoyancy Activated Cell Sorting. The system uses microbubbles that seek out and bind to target cells (based on a prechosen analyte) that then float to the surface of a suspension while the rest of cells remain untouched at the bottom. It's an exceptionally gentle, targeted technique.

OPENING THE DOOR TO SINGLE-CELL PROTEOMICS.

The quantification of unique proteins within biological samples has traditionally required thousands, tens of thousands or even millions of cells, with no way to identify cellular heterogeneity within the populations. To overcome this issue, researchers at Brigham Young University have been evaluating the potential of single-cell dispensing to allow proteomic analysis of individual cells.

The Kelly Lab at Brigham Young University in Provo, Utah, focuses on ultrasensitive biochemical analysis. Single-cell and high-resolution spatial proteomics are of particular interest. Ryan Kelly, a Professor in the Department of Chemistry and Biochemistry, explained: "Cells are the building blocks of life, and we're looking to develop a more in-dept understanding of physiology – especially pathophysiology – and how diseases such as cancers originate and develop. For example, a solid tumor is not just a uniform collection of cells, but a complex arrangement of different cell types. The arrangement of these cells determines the tumor microenvironment during the initial stages of the cancer, and whether it will be targeted by the immune system for elimination, or protected to help it evade the body's defenses." The Kelly Lab researchers are using single-cell dispensing to allow proteomic analysis of individual cells. The advent of single-cell proteomics (SCP) has enabled researchers to investigate cellular processes in unprecedented detail, providing information that is unattainable via bulk-scale protein measurements or single-cell profiling using other omics approaches. Unfortunately, commercially available platforms for single-cell isolation and sample preparation for SCP have a high cost, require technical expertise to operate, and often suffer from other system limitations, thereby constraining their accessibility. Ryan continued: "Our specialty is the analysis of proteins within individual cells.

Single-cell sequencing approaches have been around for some time and provide a lot of unique insights. However, until recently, there hasn't been a way of directly profiling protein expression within these cells, rather than inferring it based on messenger RNA. Instead, we were performing bulk-level, in-depth measurements, which allowed quantification of thousands of unique proteins from each sample, but could not analyze single cells. To overcome this, we have now developed methods across the entire workflow – sample preparation, separation and mass spectrometry analysis – allowing us to broadly quantify proteins from single cells while increasing the speed of measurement. Instead of taking the whole tumor, putting it in a blender, then measuring an average of the protein expression within the entire sample, we now have the tools to dissect the tumor cell by cell, quantifying the proteins in each of the different cell types present, and sometimes obtain spatial information

"WE NOW USE THE UNO ROUTINELY … IT'S JUST SO EASY AND FAST, AND WE KNOW WE CAN TRUST ITS ACCURACY."

Ryan Kelly, PhD, Professor, BYU Department of Chemistry and Biochemistry

as well." A major focus of the lab is making the methods it develops more accessible to other researchers, and the team started by building robotic nanopipetting platforms to pick up and deliver extremely small volumes of reagents to single cells, as well as fabricating custom nanowells out of microscope slides.

Ryan added: "Business as usual approaches for proteomics simply can't be applied to single cells. The volume of a single cell is about one picoliter, so if a single cell is placed into a standard reaction volume of, say, 100 microliters, the cell contents will be diluted by a factor of 100 million; it's like squishing a grape into a large swimming pool. Miniaturization was, therefore, clearly the way forward."

"We start with intact cells, and the aim is to end up with ready-to-analyze peptides. This requires cell lysis, protein extraction and, potentially, breaking the disulfide bonds that preserve proteins in their native structures. There may also be an alkylation step to prevent the disulfide bonds reforming. The proteins are then exposed to proteases – typically trypsin – to cleave them into smaller peptides that are easier to measure. Miniaturization allows us to keep peptide dilution and surface contact with the reaction vessel to a minimum, as well as to have higher protein concentrations, which makes the kinetics of digestion more favorable. Fortunately, sample clean-up steps are not necessary, as contaminant levels are not sufficient to interfere with the downstream processes, and any gains would be outweighed by protein losses."

Picture 1: The Uno Single Cell Dispenser™ is an automated benchtop instrument designed for ease of use, efficiency, and precision in single-cell workflows. Uno empowers a wide range of single-cell applications across the "-omics" landscape, including MS-based single-cell analysis, iPSC libraries, 3D cell research, cell-line development, and beyond.

The miniaturized workflow, nanoPOTS – Nanodroplet Processing in One pot for Trace Samples – is now used by other SCP labs, with both label-free and multiplexed approaches being actively developed. "Although we use both approaches, our group mostly focuses on label-free proteomics," said Ryan. "We're interested in finding out what proteins are present, rather than targeting specific proteins with this technology. The technique is also quantitative; we can infer the protein abundances based on the mass spectral intensities of the peptides that we measure. Multiplexed approaches are also being explored, where additional reactions are performed to incorporate barcoded tags."

"About a year ago, our lab transitioned to using the Uno Single Cell Dispenser to enhance the accessibility of advanced techniques for biomedical researchers in fields like cancer and developmental biology. This automated benchtop device can isolate single cells into wells very effectively, then prepare them for proteomic analysis. This not only makes SCP easier for us, but also makes the technique more broadly accessible to the general community, due to the low cost of the platform relative to other commercially available solutions."

"It is obviously important that any new technology is fully validated before it is adopted for routine use, and we needed to be certain that a well didn't contain multiple cells when the system software reported the presence of a single cell. We therefore chose to validate the Uno system's single cell dispensing accuracy by mass spectrometry, and were able to demonstrate 97 % accuracy.[1] Reagent dispensing was also validated using fluorescence measurements, and was found to be accurate and reproducible from 200 nanoliters up to 2 microliters.[1] We now use the Uno routinely, as it's perfect for isolating cells from a suspension to give you an unbiased view of the whole cell population, and for low volume reagent dispensing. It's just so easy and fast, and we know we can trust its accuracy. Dispensing is contactless too, so we don't have to worry about cross-contamination. We love the Uno, and think it will have quite an impact in the field of single-cell proteomics," Ryan concluded.

To find out more about the Uno Single Cell Dispenser, visit www.tecan.com/uno

1. Sanchez-Avila, X. et al. Easy and Accessible Workflow for Label-Free Single-Cell Proteomics. J. Am. Soc. Mass Spectrom. 2023, 34, 10, 2374-2380. https://doi.org/10.1021/jasms.3c00240

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

Cell Counting and Quality Control

Following cell sorting, it is always recommended to perform separate cell counting with a reliable cell counter. This is to determine the effectiveness of the cell isolation to avoid wasting time, money and resources on poor quality samples that will produce misleading results. Cell counting also allows cell viability analysis and to determine whether cell clumps have been removed.

Cell counting can be performed manually with a hemocytometer. This is laborious but is preferable in certain situations since experienced cell biologists will be better at excluding debris and sometimes debris can be too overwhelming for automated counters.

However, manual cell counting is prone to error between individuals and between samples. Eliminating user-bias is a key reason why many people turn to automated cell counters. The automated instruments are also faster, a life-saver for when you have a large number of samples and tend to supply richer information on your cells.

Cell counters have a variety of uses in cellular biology. For single-cell sequencing applications, certain instrument features can offer particular benefits. First, it is worth checking the limit of quantification (LOQ) and CV for common sample types. It is also important to establish whether cell size information is generated. Cell size histograms allow one to readily determine the distribution of single cells in a sample and monitor the presence of clumps or debris. The ability to perform high-throughput cell counting can also be vital.

Finally, it is worth remembering that automated counters require proper size gating for your cell type, a sample concentration within the dynamic range of your instrument, and low sample debris for accurate and repeatable cell counts. Ensure your sample meets the specifications of your preferred automated cell counter.

Examples of purchasable automated cell counters, using brightfield (BF) and fluorescent (FL) strategies include:

- The Vi-CELL BLU from Beckman Coulter, which can hold 24 samples and 96-well plates, with reagent packs for trypan blue detection.
- The LUNA-FL™ from Logos Biosystems for BF and FL cell counting and effective debris distinction, with reusable slides for affordable automated cell counting.
- The CytoSMART Exact from Axion Biosystems for BF and FL counting and viability analysis
- The Countess 3 model from Thermo Fisher Scientific for BF and FL counting with reusable slides.
- The C100 from RWD for BF and FL counting taking information such as counts, viability and diameter measures.
- The Cellaca MX, a high-throughput cell counter from Nexcelom. It can count 24 samples in under 3 minutes, with BF and 4 channel FL imaging and simultaneous imaging and analysis.
- The TC20 Automated Cell Counter from Bio-Rad performs counting and viability analysis using auto-focus technology and flexible cell size gating.

Additionally advice on counting cells from leading single-cell sequencing company, 10x Genomics, can be found here.

Chapter 1 references

1. Qu, H.-Q., Kao, C. & Hakonarson, H. **Single-Cell RNA Sequencing Technology Landscape in 2023.** *Stem Cells* **42**, 1-12 (2023).
2. Spits, C. *et al.* **Whole-genome multiple displacement amplification from single cells.** *Nature Protocols* **1**, 1965-1970 (2006).
3. Chapman, A.R. *et al.* **Single cell transcriptome amplification with MALBAC.** *PLoS One* **10**, e0120889 (2015).
4. Blagodatskikh, K.A. *et al.* **Improved DOP-PCR (iDOP-PCR): A robust and simple WGA method for efficient amplification of low copy number genomic DNA.** *PLoS One* **12**, e0184507 (2017).
5. DeLaughter, D.M. **The use of the Fluidigm C1 for RNA expression analyses of single cells.** *Current protocols in molecular biology* **122**, e55 (2018).
6. Brennecke, P. *et al.* **Accounting for technical noise in single-cell RNA-seq experiments.** *Nature Methods* **10**, 1093-1095 (2013).
7. Klein, A.M. *et al.* **Droplet barcoding for single-cell transcriptomics applied to embryonic stem cells.** *Cell* **161**, 1187-1201 (2015).
8. Macosko, E.Z. *et al.* **Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets.** *Cell* **161**, 1202-1214 (2015).
9. Cao, J. *et al.* **Comprehensive single-cell transcriptional profiling of a multicellular organism.** *Science* **357**, 661-667 (2017).
10. Domcke, S. *et al.* **A human cell atlas of fetal chromatin accessibility.** *Science* **370**, eaba7612 (2020).
11. Rosenberg, A.B. *et al.* **Single-cell profiling of the developing mouse brain and spinal cord with split-pool barcoding.** *Science* **360**, 176-182 (2018).
12. Huang, D. *et al.* A**dvances in single-cell RNA sequencing and its applications in cancer research.** *Journal of Hematology & Oncology* **16**, 98 (2023).
13. Pan, Y., Cao, W., Mu, Y. & Zhu, Q. **Microfluidics Facilitates the Development of Single-Cell RNA Sequencing.** *Biosensors* **12**, 450 (2022).
14. Sant, P., Rippe, K. & Mallm, J.-P. **Approaches for single-cell RNA sequencing across tissues and cell types.** *Transcription* **14**, 127-145 (2023).
15. Nguyen, Q.H., Pervolarakis, N., Nee, K. & Kessenbrock, K. **Experimental Considerations for Single-Cell RNA Sequencing Approaches.** *Frontiers in Cell and Developmental Biology* **6**(2018).
16. Hu, P., Zhang, W., Xin, H. & Deng, G. **Single Cell Isolation and Analysis.** *Frontiers in Cell and Developmental Biology* **4**(2016).

PANEL DISCUSSION :

HEAR FROM THE EXPERTS PART 1
SINGLE-CELL TECHNOLOGIES Q&A

THE CONTENT USED HERE IS A SHORTENED, EDITED TRANSCRIPT FROM A SESSION AT THE FESTIVAL OF GENOMICS LONDON 2024. QUESTIONS WERE INVITED FROM THE AUDIENCE OF THE SESSION.

Andrea Corsinotti
Single-cell Multi-omics Facility Manager, Institute for Regeneration and Repair, **University of Edinburgh**

Jan-Philipp Mallm
Head, Single-cell Open Lab
DKFZ German Cancer Research Center

Catia Moutinho
Scientific Adviser
The Single-Cell World

Introduction

Our first expert interlude covers three experts response to questions about single-cell technologies. The first expert is Andrea Corsinotti, who is Manager of the Single-Cell Multi-omics Facility at the University of Edinburgh. The second is Jan-Philipp Mallm, he is Head of the Single Cell Open Lab from DKFZ German Cancer Research Centre. The final expert is Catia Moutinho, who is a Scientific Adviser for the Single-Cell World platform.

The aim is to cover the limitations and challenges that the reader may have with single-cell technologies. When a company describes a technology to you, it will often sound amazing, but when it arrives in the lab and researchers start using it, they find that there are limitations and challenges. As with any technology, there will be limitations and hopefully these experts can help you, or tell you, how they overcame those challenges.

Question: If you start using one single-cell technology, at what point or how easy it is to move to using another one? For example, you may have datasets that you've already created, how do you handle the integration between the two?

Andrea Corsinotti: In my experience, you don't change the horse in the middle of the race. So, if you start a project working on one technology, just with data reproducibility in mind, it is advisable to stick with the same technology. The technologies, even from the same supplier, also evolve over time. So, it's probably a good idea to progress together with the technology and not get stuck with obsolete or laborious methods.

When choosing to move between technologies, this decision should be made before starting an experiment. Ideally, we advise people to come and talk to us experts. If you don't have a facility in your institution, go to people who have been doing the same type of work. And they may guide you through this complex situation of various companies, and various methods. Everything looks amazing on paper, but we are likely to be able identify the better candidates for your project.

Jan-Philipp Mallm: In terms of practical considerations when selecting kits, you need to think about how many cells you can start with. A 1 million cell kit sounds great. However, you need to have a lot more than 1 million cells to really get 1 million cells in your dataset. How do you isolate them and how do you preserve them? These are the bottlenecks that come with your research project or with your sample.

Question: What's your opinion on combining proteomics with transcriptomics for single-cell?

Catia Moutinho: It's already possible, we are doing single-cell proteomics. The quality of it depends. With proteomics, it's about antibodies, so you need good antibodies. There are already commercialised ones that have a specific oligo in the antibody. If you can find a good antibody (whether in single-cell or bulk experiments) this will make your life easier, and it becomes possible to combine both sets of information from the same cells.

Andrea Corsinotti: Talking about proteins, we put so much effort into doing single-cell RNA-seq experiments to study a molecule that is very rarely functional on its own. It's an intermediate of something that is going to be functional. The actually functional RNA, which is non-coding RNA, is not in our data set because it's excluded. So, looking at proteins is really compelling, and looking at problems with proteomics at a single-cell level is even more compelling. Antibodies are the massive bottleneck. It's impossible to do a proper unbiased proteomics approach if you rely on antibodies. There are technologies that are being developed; for example, there was a bioRxiv manuscript published last year in which they managed to do 5,000 proteins per cell. So, it's coming... it's very immature, but it's really a gap that needs to be filled.

Question: I am interested in development and the transition from undifferentiated to differentiated cells. In the UMAPs, if you see cluster A and cluster B, and cells that connect the two, that looks really interesting - perhaps the differentiating cells are going from A to B. Then I've learned about environmental RNA, and how clusters can merge into one another. That is the noise you don't want. How big of an issue is this? Are there ways of separating A from B?

Catia Moutinho: One of the biggest problems that we have is the contaminating RNA that is in the background. Especially if we are working with solid tissues, we have to dissociate the tissues to get the cells in suspension. When we do this, we will have a lot of RNA in the background that will be amplified. Of course, there are ways to reduce this background noise. It depends a lot on the type of tissue. There are also ways of doing the analysis to remove or to decrease this noise and to more cleanly separate the clusters. But if you want to see if an undifferentiated cell goes to a differentiated state, you can do other types of analysis instead of clusters. Clusters are the result of grouping the cells by similar characteristics to find cell types. An alternative is a trajectory analysis.

Jan-Philipp Mallm: One way to do this as well is to find the balance between spliced and unspliced RNA. This would show you much better what kind of directionality you would have.

Andrea Corsinotti: Biology should be louder than noise. So, it's important to validate your result, regardless of the type of analysis - no matter if it's perfect - and regardless of the method that you use to generate your libraries. It's important that whatever results you get, you verify these independently with a different type of assay, in which you can actually functionally differentiate Cell A from Cell B and identify whether the smear of cells between Cell A and Cell B are real or noise.

"BIOLOGY SHOULD BE LOUDER THAN NOISE. SO, IT'S IMPORTANT TO VALIDATE YOUR RESULT, REGARDLESS OF THE TYPE OF ANALYSIS - NO MATTER IF IT'S PERFECT - AND REGARDLESS OF THE METHOD THAT YOU USE TO GENERATE YOUR LIBRARIES."

Question: We now routinely analyse millions of cells with Parse Biosciences (for example). Does that mean we need more storage? And how do we deal bioinformatically with so many cells? Do we need new algorithms, or do we need big clusters to process all the data?

Catia Moutinho: I think storage is always a big problem - there's the cloud or big computers. But I think you are right; we will need more storage to store this data and also to influence the way we are analysing it.

Jan-Philipp Mallm: The tools also need to rise to the occasion. They also need to be able to handle these huge datasets in an efficient timeframe. If you have more cells, you need more sequencing reads and you have more data in general. So that's a given.

Andrea Corsinotti: On top of having more storage, we'll also need to develop the analysis methods to make sense of these more complex data sets, and maybe a different kind of approach to help us understand the datasets that are so complex that our brains just cannot process them.

Question: Different sample types, say PBMCs or HeLa cells lines compared to actual tissues, give you unique challenges when it comes to transcriptomics. Now, we are naturally moving from transcriptomics to multi-omics. What are our challenges there?

Jan-Philipp Mallm: Sample prep is the most important step. You have to go in with high quality sample prep, and sometimes that's painful to optimize, but it really pays off. There are more protocols out there for how to dissociate your tissue and how to get nice nuclei than vendors for any kind of chemistry.

Catia Moutinho: There are a lot of challenges and it's more difficult in general. DNA is more stable. RNA degrades a lot faster, and then we have proteins. There are a lot of things that we need to look at and they're technically more complex. I think this is why there are only a few commercial multi-omics kits. For example, from my experience with RNA and ATAC-seq - to see if the chromatin is opening for gene expression or not - it works, but not always. Also, we did it separated, with RNA from some cells and ATAC-seq with the same sample from other cells. Afterwards we performed data integration, and in this case, we got more robust data. The results from the commercial kits will have a lower probability of failure. But, there's a lot of need for optimization in the wet lab.

Question: Do you think fixation is helping with these kinds of protocols?

Catia Moutinho: There are other problems besides RNA in the background that arise from having to process the samples in a fast way. If we take fresh tissue and that takes four hours, the cells will be stressed and dying. If we fix the cells, it's amazing. We see this technology already for RNA. Finding ways to freeze the cells in their state will give us a lot of time to do the things properly.

Question: What do you wish you had today that you don't have?

Andrea Corsinotti: Cheaper reagents.

Catia Moutinho: Yes, cheaper, please. We have tight budgets.

Jan-Philipp Mallm: When it comes to the read out, many of us do RNA sequencing. There's also DNA sequencing as well as multi-omics. There is proteomics. But what about metabolomics, which is seen as the end result of the genetic dogma - DNA, RNA, protein and so on? Proteins have an output, and this is what you can measure with metabolomics. There is no good way to measure this, at least on this single-cell level, at scale and at the sensitivity needed to read out the cell's state exactly. This is an approach not based on a protein, but based on what the cells are actually doing. For cancer research, making this functional link is important. We hear about genotype-phenotype, and we tend to define the phenotype as just a measurement of RNA, sometimes of protein, but you can go even further. These methods are missing.

Audience Member: A single-cell fixed ATAC kit would be also very nice to have. Especially when it comes to the clinical samples, you will need to work in a fixed space. And now we have single-cell RNA sequencing with FFPE tissue, but we have nothing for ATAC yet.

Audience Member: Just to say that the type of fixation also comes into play. There's not just one way of fixing over the other, there are swings and roundabouts to various types of fixation that will befit certain cell and tissue types. It's always worth assessing everything that's on the market and making sure that that fits your needs in terms of sensitivity, specificity and multi-omics.

Catia Moutinho: Yes, you are right. Something that is also important for us to consider is whether you project is just about RNA or if it will be measuring RNA and protein (or another omic). For any technology, it is worth checking if the company has more types of technology or approaches that we can use. This is something that is important to think through when choosing technology.

Question: I've heard about the importance of a sample prep for every experiment and even more so for single-cell. What are your views on using automated liquid handling versus doing it by hand?

Catia Moutinho: From my experience of sample preparation, automated is better, but again, first you need to optimise the protocol that you will use in the machine. For example, one of the machines that I use is from S2 Genomics. I optimised that for three weeks, first, manually, then I put in a machine, and then we started using the machine full-time. So, that is my advice. With automation, you can control the conditions better. But, for a specific tissue or if you change tissue type, you have to do this process again and start manually.

Andrea Corsinotti: When it comes to sample prep, you need to accept that human samples have inherent variability. Even the same type of surgery, same type of biopsy, from the same type of tissue may, over many samples, produce substantial differences. So, optimisation is really the key. For example, for all these methods, the input makes a difference. What we do in between preparing the samples and making the library is a standard workflow. We change very little. We can do more PCR cycles, less PCR cycles, those kinds of things. But the quality of your data is reflected, essentially, by the quality of your sample before it arrives in the lab.

Catia Moutinho: Yeah, I think my point of view is that the kits that we buy are very optimised because there are R&D teams doing that. The problem is our input. So, the sample preparation, getting the cells that are high enough quality for the commercial kits, this is where we should invest more time and are the biggest challenges there.

Question: A lot of single-cell platforms have an expected amount of cell doublet rates when it comes to the cell partitioning phases. Is there any kind of advice or anything you would recommend trying from a practical standpoint to reduce those rates and get better cell recovery?

Andrea Corsinotti: Strain, strain, strain.

Catia Moutinho: Yeah, strain, strain, strain, and don't overload. If you are using instruments, don't overload them too much. Unless you are using an approach, for example, with hashing antibodies with lipids to distinguish different samples.

Andrea Corsinotti: Although to be honest, doublets are not the end of the world, they are easy to spot in the analysis and very easy to remove. So, if you can choose between a sample that is a little bit lumpy or no experiment, I would say process the sample; maybe use a method that is not based on microfluidics. But even if you have a lot of doublets, you can just sort them out, you will spend a lot of money sequencing doublets that you don't need, but you may still get the data from the experiment.

Jan-Philipp Mallm: And also, some other conditions might help to kind of get rid of the biological doublets, like cells sticking together rather than having them separately in one droplet. So, you can play around with this for sure. This is something that is under your control. However, the doublet rates from the collision of barcodes, etc. I think there's very little you can actually do as a user to reduce that.

Question: What do you think are the pros and cons of doing multi-omics, Say RNA and ATAC, on the same cells? Since this mean you're only looking at the nuclear RNA versus doing the two separately.

Catia Moutinho: I always prefer for different information to come from the same cell. Integrating information is also an option that is being done, but for me, coming from the same cell is preferable. Even if it's only the nucleus because when we do nuclei, we will lose the mRNA information that is in the cytoplasm. However, I think there are already good tools to predict if the nuclear mRNA was also mature in the cytoplasm. This is my opinion; I prefer dual information from one single cell.

Jan-Philipp Mallm: Eventually, if you do this from two different cell preps, you always have to do integration on the cluster basis. And you have to rely on the assumption that ATAC and RNA are very well matched. If you're looking at a disease state, integration might not work perfectly. PBMCs are easy to work with, but whenever you go beyond that, if you try to integrate data, you can struggle heavily - you have to tweak the data a lot. And this will cost you resolution, even though you think, 'I have better resolution if I do this readout separately'. I personally would always go for multi-omics rather than doing them separately.

Andrea Corsinotti: I personally would go for multi-omics as well, as long as you don't compromise on the quality of your experiment, because sometimes, doing both things at once is just either very challenging or almost impossible. As Catia said earlier, if you do them separately, the individual dataset quality will be higher than if you do them together. To do 10x Multiome, ATAC-seq and RNA-seq, you need to preserve both the chromatin structure and the RNA quality, while if you're just doing ATAC-seq, you don't care about losing all the RNA quality. My advice to people, if the biology requires a combination of the two data sets at single-cell level, go for it. If you would like to see both things, then you're probably gonna get a better quality data set if you don't do both from a single cell. I think this probably extends to multi-omics in general.

Question: Do you think single-cell whole genome sequencing, or single-cell whole exome sequencing, is something that we will see in the near future?

Catia Moutinho: There are already approaches; for example, kits from BioSkryb. They sell a kit for whole genome, and they have the ResolveOME - DNA combined with RNA. The problem here depends on the number of cells that you need to get information from for your experiment. If you are doing whole genome, you will need to sequence a lot of cells. If you are thinking of sequencing 10,000 cells, you'll be spending a lot of money. It will always depend on what you want to do and your budget. It's really important. But it's possible. Yes.

Jan-Philipp Mallm: We have used single-cell whole genome sequencing on the individual well plate-based method, and it works really, really nicely. So, Catia just mentioned BioSkryb. We also use a droplet-based method for scDNAseq, you get an overview of many, many cells, but the resolution, of course, is not as good. If you sequence 10,000 or 20,000 cells, you can never get the coverage to really look into individual mutation events, but rather at the level of copy number variations.

For single-cell whole exome sequencing, that also works very well with the BioSkryb technology. You can call mutations de novo from the data set. But then again, the problem comes with the sheer amount of sequencing you would have to do to really do that. But that's an inherent problem. If you know where to look, there are other methods out there that allow you to do targeted approach to look into mutations on amplicons. You can have panels for that. Mission Bio would be one company that provides this kind of technology. So, depending on the level that you would like to look into your data and the prior knowledge that you have, you can select what is the best technology for your specific project. But there are protocols out there that can be used, and they work reliably.

CASE STUDY

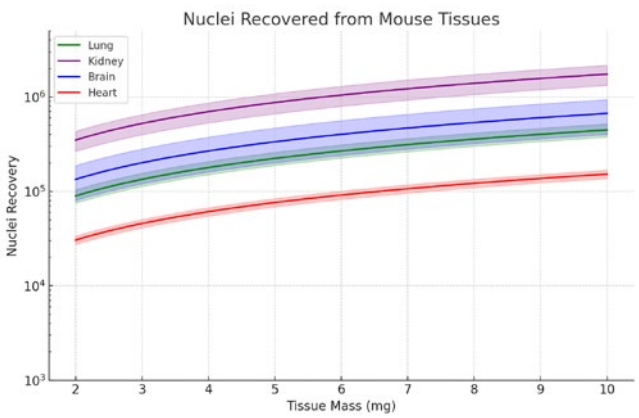
UNLOCKING THE FFPE TISSUE BANK THROUGH AUTOMATED PREPARATION OF SINGLE NUCLEI FROM FFPE SAMPLES FOR SINGLE-NUCLEI RNA-SEQ

FORMALIN-FIXED, PARAFFIN-EMBEDDED (FFPE) TISSUES ARE THE PREFERRED FORMAT FOR PATHOLOGISTS, WITH AS MANY AS ONE BILLION FFPE SLIDES CURRENTLY STORED IN REPOSITORIES. FURTHERMORE, WHILST MANY SPATIAL METHODOLOGIES CAN TAKE ADVANTAGE OF FFPE SLIDES, USING THE TISSUE FOR THE MORE MATURE SINGLE NUCLEI METHODS IS STILL A CHALLENGE.

FFPE tissue samples are difficult to study due to the chemical modifications caused by formalin fixation, which can lead to degradation and sequestration of RNA and DNA. Cracking open this valuable resource for powerful methodologies such as single-cell analysis is crucial but also challenging. While the 10x Genomics Flex scRNA-Seq Kit has enabled FFPE single-cell library preparation, FFPE tissue still needs to be properly deparaffinised, rehydrated and potentially cross-link reversed to get optimal results from downstream single nuclei methods.

FIGURE 1. HIGH NUCLEI RECOVERY EFFICIENCY WITH THE NIC+ CARTRIDGE.

2-10 mg of tissue from mouse lung, brain, kidney and heart tissue had nuclei isolated using NIC+ cartridges, Nuclei Isolation Reagents and the Low Volume Nuclei Isolation protocol. Each tissue type was run on at least two instruments across the input masses. Nuclei recovery is sufficient for targeting up to 10,000 nuclei into a 10x Genomics Chromium Single Cell Assays. Yield was measured by AO/PI assay using the Nexcelom K2.



THE SINGULATOR SYSTEM FOR HIGH QUALITY NUCLEI RECOVERY FOR SNRNA-SEQ

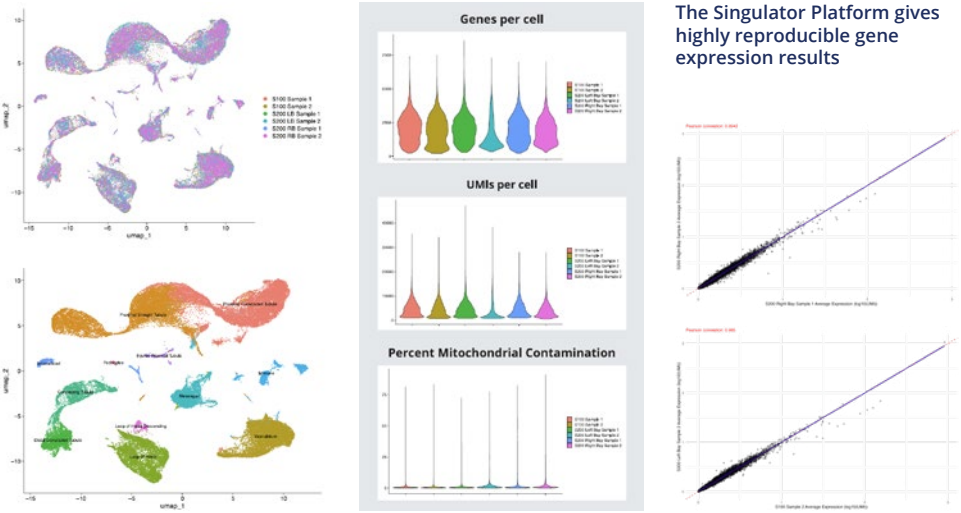
The Singulator™ Platform automates tissue dissociation for single-cell genomics applications, including snRNA-Seq. It works by dissociating solid tissue into single-nuclei by mechanical disruption and chemical lysis, followed by passage through one or two filters to remove large debris within a single-use cartridge. The system is particularly impressive when it comes to low input data.

Nuclei Isolation Cartridges (NIC+) are specially designed by S2 Genomics for optimal recovery of nuclei from extremely low input masses from fresh and frozen material. Nuclei recovery for 10x Genomics Single Cell Assays can be from as low as 2mg of frozen tissue input (see Figure 1).

Importantly, the Singulator Platform is designed to be highly reproducible across instruments and users. To demonstrate this, six nuclei suspensions were prepared from a single mouse kidney using the NIC+ Cartridges on two Singular 100 instruments and two Singular 200 instruments. Figure 2 highlights the highly reproducible gene expression and cell clustering results.

FIGURE 2. THE SINGULATOR PLATFORM ISOLATES HIGHLY REPRODUCIBLE CELL POPULATIONS AND GENE EXPRESSION RESULTS ACROSS TECHNICAL REPLICATES.

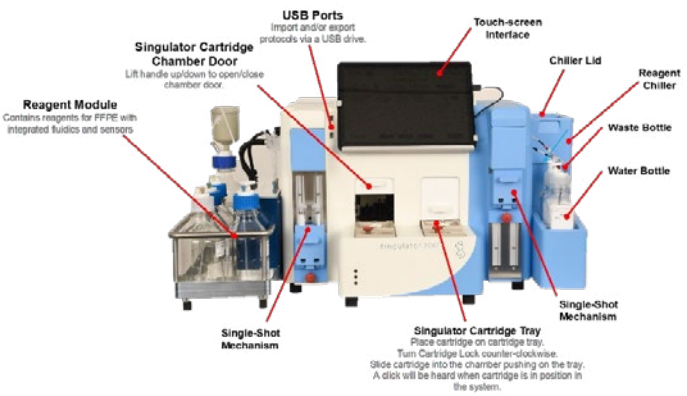
Left. *Integrated UMAP projection of six technical replicates isolated from a single mouse kidney. Nuclei were dissociated with the Singulator 100 and 200 Platform and profiled with 10x Genomics Chromium Next GEM Single Cell 3' v3.2 solution.* ***Centre.*** *The six samples exhibited comparable gene expression, UMI and mitochondrial contamination per cell after down sampling to equivalent reads per cell.* ***Right.*** *Two representative Pearson Correlations demonstrate excellent reproducibility across replicates.*



While the existing Singulator Systems can automatically process nuclei from manually prepared FFPE tissue, the latest **Singulator™ 200+ System** (Figure 3) is a fully automated sample preparation isolation system for single cell formalin-fixed, paraffin-embedded (FFPE) tissues. It contains all the same features as the Singulator 200, with the ability to deparaffinize and dissociate FFPE samples in as little as 40 minutes with no manual intervention. This opens doors to high quality snRNA-Seq from FFPE tissues, which is extremely useful for a number of applications.

FIGURE 3. THE SINGULATOR 200+ SYSTEM.

Including the Singulator instrument, single-use cartridge, chiller for nuclei reagents, SingleShot Mechanism for cell preparation reagents, and Reagent Module for FFPE reagents.



FFPE SPATIAL TRANSCRIPTOMICS ENHANCED BY INCORPORATING FFPE SNRNA-SEQ

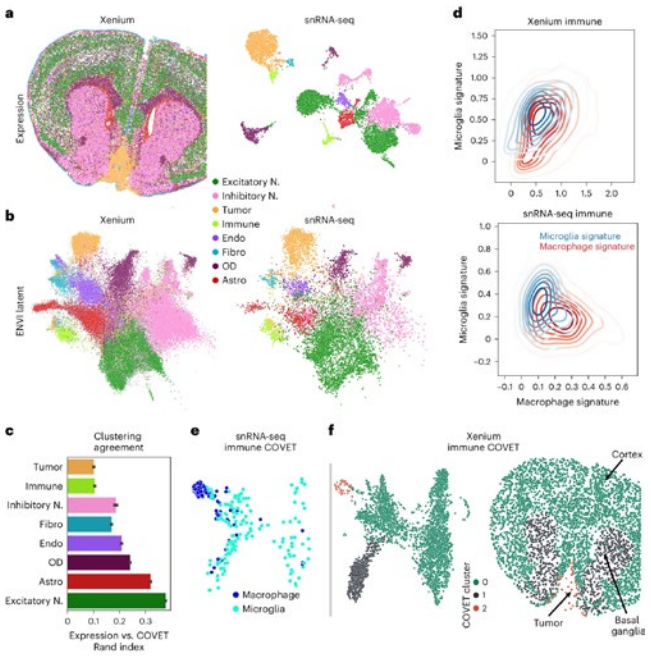
Representing cellular neighbourhoods and niches is a key challenge in spatial profiling, and matching single nuclei methods can be part of the solution. Environmental Variance Inference (ENVI) is a computational method designed to effectively model cellular neighbourhoods by jointly embedding snRNA-Seq data into a latent space, projecting spatial information onto the single nuclei data1. ENVI can confer spatial context to genomics data from single dissociated cells and outperforms alternatives for imputing gene expression on diverse spatial datasets.

In a mouse model of melanoma Leptomeningeal metastasis (LM), adjacent FFPE sections of brain were taken, one for the Xenium platform (10x Genomics) with 243 probes and the other for snRNA-Seq1 (Figure 4A). The 100 µm section of tissue was processed on the Singulator 200+ system in a NIC+ cartridge for deparaffinisation, rehydration and then dissociation in a second NIC+ cartridge.

ENVI successfully harmonised the two datasets into a unified latent space (see Figure 4B), and overall ENVI robustly integrates scRNA-Seq and spatial transcriptomics data. This provided better overall profiling than the limited probe set of in situ platforms and allows the more in-depth clusters from snRNA-Seq to be spatially localised.

FIGURE 4. ENVI INTEGRATES XENIUM AND SNRNA-SEQ DATA PREPARED WITH THE SINGULATOR S200+.

a, *Xenium image and UMAP embedding of snRNA-Seq data from mouse brain bearing a melanoma metastasis, coloured by major cell type.* ***b,*** *UMAP embeddings of ENVI latent space showing cells from the spatial (left) and snRNA-Seq (right) datasets. Similar cell types, including malignant cells, co-embed across modalities.* ***c,*** *Average concordance of technical replicates between the expression of each cell type and its environment in the Xenium data.* ***d,*** *Density plots of microglia and macrophage cell signature expression in immune-labelled cells from Xenium (top) and snRNA-Seq (bottom) datasets. Only snRNA-Seq data measure enough genes to separate cell types.* ***e,*** *UMAP embedding of the ENVI-predicted COVET representation of snRNA-Seq immune cells, coloured by subtype.* ***f,*** *COVET UMAP (left) and spatial coordinates (right) of Xenium immune cells, coloured by COVET clusters representing major immune cell microenvironments. Source: Haviv et al.*[1]



For example, the snRNA-Seq data clearly distinguish microglia and macrophages, based on curated gene sets, whereas the Xenium brain panel lacked the markers to distinguish them (Figure 4D). However, the co-embedded data allowed three distinct immune environments to be spatially mapped, with macrophages found around the tumour and microglia assigned in the basal ganglia and cortex (Figure 4E-F).

This method conferred spatial context to genomics data from single dissociated nuclei and outperformed alternatives for imputing gene expression on diverse spatial datasets. This can only be possible with robust methods for isolating single nuclei from FFPE sections, such as the automated Singulator™ 200+ System that the authors utilised.

ENRICHING NUCLEI EXTRACTED FROM FFPE SNRNA-SEQ

Typical cell enrichment or depletion methods rely on cell surface protein fluorescence-activated cell sorting (FACS). This method is naturally not possible with a nuclei dissociation, such as from FFPE tissues. Instead, the recently developed, PERF-Seq2, uses fluorescent in situ hybridisation (FISH) to label RNAs that identify rare populations, allowing it to work with FFPE nuclei.

FFPE human glioblastoma multiforme (GBM) samples were pre-processed on S2 Singulator 200+ system in a NIC+ cartridge and dissociated to single nuclei in a second NIC+ cartridge, setting the nuclei up perfectly for flow cytometry downstream sorting. DCN, FN1 and VWF were used to enrich for vascular-derived cells (a sparse but disease-relevant population).

The resulting 1,015 enriched vascular cells were sorted into two known major populations and8 subclusters, highlighting the importance of high-quality FFPE nuclei preparation to capture the full cellular profile of rare cell types.

CONCLUSION

Excellent Single-nuclei RNA-Seq from FFPE tissues in storage is an untapped resource that will provide many benefits. Whether combining with FFPE spatial data or enriching and identifying subpopulations of rare cell types, production of high quality single nuclei is essential. The Singulator 200+ is the dedicated platform for working with nuclei from FFPE tissues, with high nuclei recovery and consistency across technical replicates, it provides the basis for excellent snRNA-Seq performance.

References

1. Haviv, D. *et al.* **The covariance environment defines cellular niches for spatial inference.** *Nature Biotechnology* (2024). doi:10.1038/s41587-024-02193-4
2. Abay, Tsion, *et al.* **Transcript-specific enrichment enables profiling rare cell states via scRNA-seq.** *bioRxiv* (2024): 2024-03.

CHAPTER 2

SINGLE-CELL SOLUTIONS: METHODS, KITS & INSTRUMENTS FOR SINGLE-CELL

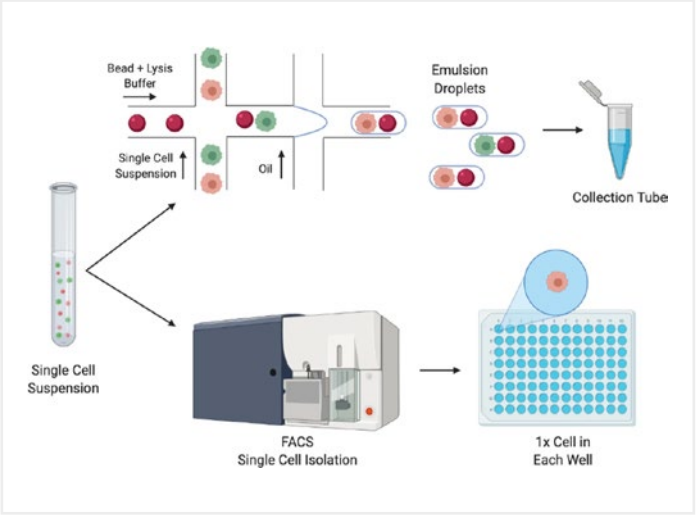
THIS CHAPTER DESCRIBES THE SINGLE-CELL TECHNOLOGY LANDSCAPE AS OF 2024. IT SHOWCASES THE LATEST AVAILABLE COMMERCIAL SINGLE-CELL KITS AND INSTRUMENTS, ALLOWING YOU TO DIRECTLY COMPARE THE SPECS AND PERFORMANCE OF EACH.

The Single-Cell Landscape

Single-cell technologies are around 15 years old, and high-throughput methods have been around for less than a decade. Despite the relative novelty, the single-cell method landscape is dynamic with new expansions, upgrades and methodologies being released yearly. This chapter will explore the nature of change to the commercial landscape. For a broad summary of the single-cell RNA sequencing technology landscape as of the end of 2023, this paper provides a really valuable overview[1].

FIGURE 2.1. WELL-BASED AND DROPLET-BASED SINGLE-CELL METHODS.

Source: Probst, et al. [2]



There are two principal methodologies for sequencing single cells (Figure 2.1). The first methods to emerge were the well-based and plate-based methods, in which cells are placed or guided into separate wells, so that individual cells can be barcoded and sequenced. The newer methods were the droplet-based methods, and these were the ones to open the door to high throughput studies. Here, cells are captured in individual emulsion droplets rather than wells, creating unique reaction vesicles.

Commercial single-cell assays are mostly based on one of these two methods; however, it is worth highlighting that commercial products can't do everything. They are typically designed to do one thing very well; in this case, to be very reliable at sequencing mRNA or another analyte from single cell suspensions. The more out-of-the-box single-cell methodologies tend to originate in non-commercial settings. Hence, before we review the commercial products, let's take a moment to highlight a few of the latest and most exciting non-commercial single-cell methods that are pushing the boat out and diversifying what can be accomplished with single-cell sequencing.

- **FixNCut**[3] and **snPATHO-seq**[4] are new single-cell methods that improve single-cell performance following cell fixation, in the former, and from FFPE tissue in the latter.
- **Live-Seq**[5] is a method for temporal transcriptomic recordings of single cells, escaping the snapshot in time that standard single-cell methods tend to capture.

- Single-cell combinatorial indexing via **sci-RNA-seq**[6] effectively profiles large volumes of cells. into the millions of cells in its latest iteration[7].
- **PERF-seq**[8] for programmable enrichment of very rare cell populations via RNA Flow-FISH by sequencing. This is a refined method for pulling out those rare populations.
- **Perturb-seq**[9] maps the transcriptional effects of genetic perturbations at genome scale. The latest work on this method has leveraged pooled genetic screens and single-cell sequencing to systematically identify the targets of signalling regulators in diverse biological contexts[10].

The rest of this chapter catalogues the commercially-available single-cell technologies within these method categories, including the new category of Hydrogel-based methods. We have included various specs, plus detailed descriptions alongside advice on choosing the right methodology for your application.

DAVID COOK

Scientist, **Ottawa Hospital Research Institute,** Assistant Professor, Department of Cellular and Molecular Medicine, **University of Ottawa**

FLG: What do you think of the single-cell technological landscape right now? The way I see it, 10x Genomics is still the dominant methodology. But what do you think of newer methods, such as the commercial combinatorial barcoding methods?

David: *You're right, 10x Genomics still has a good foothold in the community. I think there's something to be said about the convenience that 10x provides, it's a commercial option, all you need are your cells, it's reliable, you just plug it in, and it tends to work really well. However, I'm thankful that there's competition from the academic space, where groups are continually developing new strategies to generate single-cell data. That's what pushes the field ahead, so I'm glad that people are still committed to doing that. That's where we get things like, SPLiT-seq that led to Parse Biosciences, the combinatorial indexing strategies from Jay Shendure's lab, and various others.*

These other methods excelled in an area I think 10x historically struggled with by being able to accommodate experimental designs with larger sample sizes and lower cell throughput per sample. There are a lot of experimental designs where if you're not in a lab with millions of dollars in annual operating budget, you just can't afford to do the big 10x experiments on dozens and dozens of samples. But these other kits have been designed in a way where the cost is more related to cell yield than sample number. And so, it becomes much easier to scale experimental designs. You can easily do things such as a 96-well drug screen, and for much cheaper than you could by doing 96 independent samples on the 10x Chromium (although multiplexing approaches can help!).

Droplet-Based Methods

The most popular set of single-cell methods are the droplet-based or microfluidics methods. These methods are instrument-dependent in order to operate the microfluidics. The machines works via individualising single cell inside oil droplets and barcoding them. Each droplet works as a unique PCR reaction tube. Certain droplet-based methods, such as the Fluidigm C1 and Dolomite Bio Nadia, are not included as they are no longer purchasable, despite perhaps being accessible from core labs and outsourcers.

The major advantage of the droplet-based methods is the very high throughput and automation. The drawbacks occur from the consequences of high throughput, namely a higher doublet rate and lower genomic yield per cell.

CHROMIUM X/CHROMIUM CONTROLLER – 10X GENOMICS

The most well-known single-cell instrument is the Chromium Controller from 10x Genomics. The Chromium uses advanced microfluidic chips to partition single cells and barcode them on a large scale (up to 128 samples and a million cells). Each chip can be used for multiple single-cell suspensions. Cells are partitioned into tiny water droplets, and reagents and barcoded gel beads are added. These droplets are known as gel beads in emulsion (GEMs). The GEMs are barcoded and so are the individual RNA molecules before the GEMs are pooled to then convert the RNA to cDNA.

The latest generation of the 10x Genomics method includes the Chromium X and the GEM-X technologies. The GEM-X chips improve a variety of factors over the original chips, including a two-fold increase in cells captured per channel, two-fold increase in genes detected, a two-fold reduction in costs per cell and a two-fold reduction in multiplet rates. Furthermore, 10x Genomics technology can be used for multi-omics and on fixed or frozen samples using the Flex kit.

Technology:	Droplet/Microfluidics (Drop-seq)
Instrument:	286 mm x 483 mm x 273 mm, 18.8 kg
Workflow Length:	< 2 day (chip runtime is ~18 minutes)
Cell Throughput Capacity:	160,000 cells per run
Flexibility:	Low and high throughput and multiple assay types
Sample Capacity:	16 samples per chip (8 per row)
Capture Efficiency:	Up to 80% with GEM-X (originally 65%)
Mutliplet Rates:	~4% @ 10,000 cells with GEM-X (originally 8% @ 10,000 cells)
Detected Genes Per Cell:	~3,000 genes with GEM-X (originally ~1,500)

TAPESTRI® – MISSION BIO

Mission Bio's Tapestri® Platform deploys a two-step microfluidic workflow, which makes DNA and protein information accessible from single-cells. The core of the technology is the DNA cartridge. Single cells are partitioned into nanoliter droplets before barcoding beads and PCR reagents are introduced using the Mission Bio Tapestri Instrument and DNA Cartridge. Cell lysis, protease digestion, cell barcoding and targeted amplification using multiplexed PCR occur within the droplets. Droplets are then disrupted, and barcoded DNA is extracted for library amplification. Protein panels can be incorporated for a simple cell staining protocol to turn this into a multi-omics workflow.

Technology:	Droplet/Microfluidics – two step partitioning (DNA)
Instrument:	317.5 mm x 298.5 mm x 313.3 mm
Workflow Length:	2 days
Cell Throughput Capacity:	20,000 – 100,000+ cells
Flexibility:	Customizable panels and protein co-detection
Sample Capacity:	Up to 3 samples in parallel
Capture Efficiency:	>5%
Mutliplet Rates:	<8% rate

DDSEQ SINGLE-CELL ISOLATOR – BIO-RAD

Bio-Rad's ddSEQ Cell Isolator is a compact droplet-based single cell isolation system for single-cell gene expression and gene regulation studies using ddSEQ Library Preparation Kits for 3' RNA-seq and ATAC-seq. ddSEQ droplet technology utilizes disposable microfluidic cartridges to enable simple and fast encapsulation, processing thousands of single cells or single nuclei in under 5 mins. The ddSEQ single-cell NGS library prep kits provide accessible high-quality data, streamlined one-day workflows with convenient safe stopping points, as well as batch sample processing using Bio-Rad's open-source pipeline, Omnition Analysis Software for both 3' RNA-Seq and ATAC-Seq analysis.

Technology:	Droplet/Microfluidics
Instrument:	279 mm x 356 mm x 127 mm
Workflow Length:	7 hours with safe stopping points (Cartridge runtime is <5 minutes)
Cell Throughput Capacity:	20,000 cells per cartridge (500 - 5,000 cells per lane)
Flexibility:	Work with cells or nuclei, open-source pipeline is customizable
Sample Capacity:	4 samples per cartridge
Capture Efficiency:	>50% - 80%
Mutliplet Rates:	<8%
Unique Fragments Per Cell (ATAC):	>20,000 unique fragments per cell; >30 TSS Enrichment Score; >30% FRIP
Detected Genes Per Cell (RNA-Seq) :	>1,500 (PBMCs) or >5,500 (cell lines) median genes detected per cell

DNBELAB C-TAIM 4 – MGI

The DNBelab C method works through loading cells onto microfluidics chips. Cells are partitioned into tiny water droplets together with the reagents and barcoded magnetic beads using DNBelabTM C-TaiM 4 instrument. The TaiM 4 Single-cell droplet generator instrument provides stable and precise cell and cell nucleus separation and labelling. It is equipped with four independent microfluidic channels that can separate single cells and barcode them. This can be done for 4 different samples at a throughput of 20,000 cells per sample at the same time. The TaiM 4 device supports other MGI products for 3' RNA single cell analysis, namely the DNBelab C Series High-throughput Single-cell RNA Library Preparation Set and the DNBSEQ™ sequencing platform series.

Technology:	Droplet/Microparticle
Instrument:	290 mm x 210 mm x 230 mm, 5 kg
Workflow Length:	1 day (on machine 9 min)
Cell Throughput Capacity:	5,000 - 20,000
Flexibility:	Can be run with 5,000 - 30,000 input cells
Sample Capacity:	4 samples per run
Capture Efficiency:	> 50%
Mutliplet Rates:	< 0.4% /1,000cells
Detected Genes Per Cell:	> 2,000 (cells like PBMC)

Plate-Based Methods

The earliest of the single-cell methods to be developed and become available for purchase were the plate-based methods. These kits are deployed following the isolation of individual cells into well plates (either 96-well or 384-well plates). Cell isolation can be done manually under a microscope or automated using a cell dispenser or FACS (see Chapter 1).

Being limited to well plates, these methods typically have a low throughput. In the modern commercial landscape, these kits tend to focus on specialised single-cell applications such as SMART-Seq, which sequences full-length RNA molecules, and whole genome sequencing of single cells. However, companies such as Parse Biosciences and Scale Biosciences have commercialised a combinatorial barcoding technology, which allows experiments to scale rapidly to significantly higher throughputs per plate.

SMART-SEQ® - TAKARA BIO

The SMART-Seq mRNA Single Cell from Takara Bio is a single-cell kit for full length transcriptomic analysis. SMART-Seq works using the Switching Mechanism at 5' end of RNA Template (SMART) technology. This technology ensures that the final cDNA libraries contain the 5' end of the mRNA and maintain a true representation of the original mRNA transcripts. This technology hence allows direct cDNA synthesis from intact cells.

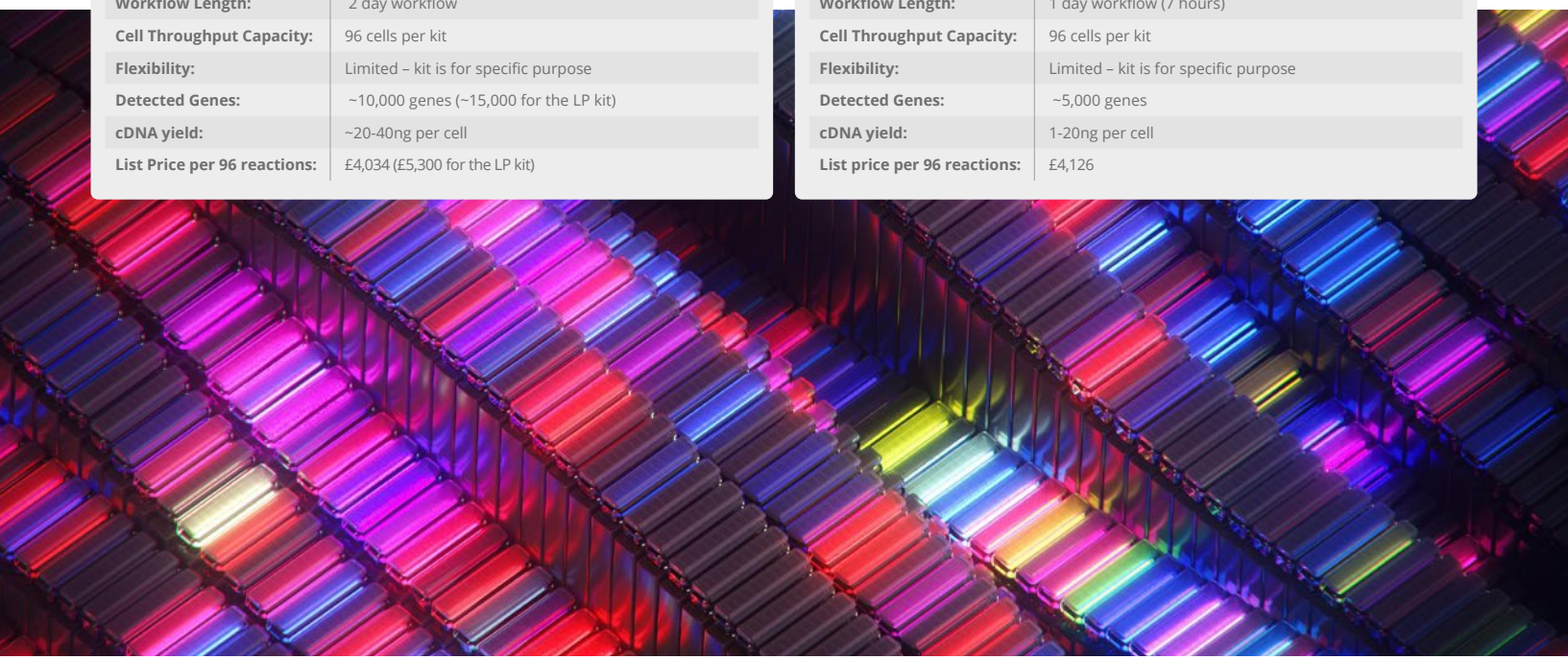
For the kit, cells can be sorted into plates via FACS, or other methods, and cDNA synthesis and library preparation occur within the one tube. Full length methods like this tend to identify more genes and hence work well for cells with low RNA content. This kit also exists in the LP version for cells with very low RNA content. This kit is ideal for getting the in-depth sequencing information from a specific cell target.

Technology:	SMART – full length transcriptome
Instrument:	Instrument-free
Workflow Length:	2 day workflow
Cell Throughput Capacity:	96 cells per kit
Flexibility:	Limited – kit is for specific purpose
Detected Genes:	~10,000 genes (~15,000 for the LP kit)
cDNA yield:	~20-40ng per cell
List Price per 96 reactions:	£4,034 (£5,300 for the LP kit)

NEBNEXT® - NEW ENGLAND BIOLABS

The NEBNext® kit for single-cell/low input RNA library prep creates sequencing libraries from single cells. It is specifically tailored for Illumina® sequencers. It uses a template switching method to generate full length cDNA. These cDNA's are then converted to sequencing libraries through the Ultra™II FS workflow, also provided by New England Biolabs.

Technology:	Low Input RNA Library Prep Kit – full length transcriptome
Instrument:	Instrument-free
Workflow Length:	1 day workflow (7 hours)
Cell Throughput Capacity:	96 cells per kit
Flexibility:	Limited – kit is for specific purpose
Detected Genes:	~5,000 genes
cDNA yield:	1-20ng per cell
List price per 96 reactions:	£4,126



REPLI-G® SINGLE CELL KIT – QIAGEN

The QIAseq solution for both DNA and RNA library preparation delivers either PCR-free NGS libraries or cDNA using the REPLI-g RNA amplification. This method reduce biases that result from typical amplification. Cell lysis is accomplished using alkaline incubation. Library preparation uses QIAseq fragmentation and ligation with Unique Dual Indices (UDI) to move straight over to Illumina NGS instruments.

Technology:	REPLI-g
Instrument:	Instrument-free
Workflow Length:	Libraries in <5.5 hours
Cell Throughput Capacity:	384 cells per kit
Flexibility:	Limited – kit is for specific purpose (24, 96, 384 versions)
Detected Genes:	Unclear
Reads mapping to genome:	Unclear
cDNA yield:	< 40ng per cell
List Price per 96 reactions:	£5,500

RESOLVEDNA®/RESOLVEOME® – BIOSKRYB GENOMICS

ResolveDNA® is a unique offering from BioSkryb Genomics and is specifically designed for complete genome coverage, with bulk sample quality, from a single cell. The kit uses Primary Template-directed Amplification (PTA) to improve coverage and uniformity in whole genome amplification (WGA). ResolveOME® is a multi-omics kit that incorporates the whole genome method detailed above alongside a method for cDNA synthesis to produce a full-length mRNA library in the same tube.

Technology:	Primary Template-directed Amplification (PTA)
Instrument:	Instrument-free
Workflow Length:	1 days (8 hours – DNA, 12.5 hours - OME)
Cell Throughput Capacity:	up to 384 cells per kit
Flexibility:	Kit is specific to purpose
Detected Genes	~3,500
Reads mapping to genome:	>97%
cDNA yield:	1-20ng per cell

EVERCODE™ – PARSE BIOSCIENCES

The Evercode™ technology from Parse Biosciences allows high-throughput single-cell experiments while working from a plate-based assay. Their workflow works through split-pool combinatorial barcoding. Cells or nuclei are fixed and permeabilized, securing transcripts inside and turning the cell into a reaction vesicle. This removes the need for droplet or microwell-based methods.

The cells are distributed into plates and the first set of barcodes are distributed. Cells are pooled and then distributed again for the second barcode, and the same again for the third barcode. The cells are then distributed across several sub libraries and are lysed. A fourth barcode is added by PCR. Now, each transcript can be identified based on its unique barcode. This method leads to poorer capture efficiency due to the cycles of pooling and distributing cells, but it results in high sensitivity in gene detection.

Technology:	Split-pool combinatorial barcoding
Instrument:	Instrument-free (high-throughput)
Workflow Length:	2 days (~14 hours)
Cell Throughput Capacity:	10,000, 100,000 or 1 million cells
Flexibility:	Kits for low - high throughput, fixed cells can be stored for 6 months
Sample Capacity:	12, 48, 96 samples per kit
Capture Efficiency:	Low (< 50%)
Mutliplet Rates:	<3% rate (>100,000 cells)
Detected Genes per cell:	~2,500

SCALE RNA - SCALE BIOSCIENCES

Scale Biosciences' single-cell RNA sequencing kit works through split-pool combinatorial barcoding. Fixed cells or nuclei undergo three rounds of splitting (indexed reverse transcription, indexed ligation and indexed PCR) before being pooled ready for sequencing. This method affixes two barcodes, which can be used to identify transcripts and cells of origin.

Scale Biosciences technology has been applied to methylation, processing 18,000 cells per run, detecting 100's of 1000's of CpG sites, and to chromatin accessibility (ATAC), processing 300,000 nuclei to be loaded into a droplet-based device.

Technology:	Split-pool combinatorial barcoding
Instrument:	Instrument-free (high-throughput)
Workflow Length:	2 days
Cell Throughput Capacity:	125,000 cells. Up to ½ million cells per run with the Extended Throughput Kit (there is a new 2024 two plate workflow for 2 million cells – see below)
Flexibility:	Works with fixed cells, can be stored for 6 months
Sample Capacity:	Up to 96 samples per kit
Capture Efficiency:	Low (< 50%)
Mutliplet Rates:	<5% rate
Detected Genes per cell:	~2,200

Micro-, Nano- or Pico Well-Based Methods

The next set of methods similarly rely on wells, but here the wells are incredibly small -micro, nano or even pico-sized. These wells are based on a chip to which single-cell suspensions are loaded, and gravity does the work of guiding cells into wells from the chips. The big advantage over typical plate-based methods is the throughput.

This type of method is ideal for sensitive cells. By filtering the cells through gravity, this method is much gentler to cells than FACS or microfluidics. Furthermore, methods such as the HIVE CLX allow you to store cells on the chip while collecting them and allow sample collection to extend over a longer period of time. These methods tend to require instruments or specific pieces of equipment to operate.

HIVE SCRNASEQ - HONEYCOMB BIOTECHNOLOGIES

HONEYCOMB

The HIVE CLX, is the picowell technology from Honeycomb Biotechnologies and has 160,000 55µm picowells in their distinctive array. The array is designed for gentle capture of fragile cells. This allows for integration of stable sample storage (up to 9 months) and single-cell profiling without needing specialized instrumentation. Cells are captured quickly and effectively and can be stored as you go - meaning samples can be collected across time without batch effects. While throughput and doublet rate are poorer compared to other methods, >90% viability is common for delicate populations from a fully loaded HIVE, and you can recover delicate cells that would be lost in droplet-methods.

Technology:	HIVE™ Single Cell Solution
Instrument:	Instrument-free (HIVE's are disposed after use)
Workflow Length:	1.5 days
Cell Throughput Capacity:	60,000 cells per HIVE CLX
Flexibility:	Integrated sample storage allows prolonged experimentation
Sample Capacity:	1-24 samples can be processed in parallel
Capture Efficiency:	40-50%
Mutliplet Rates:	9% @ 15,000 cells, 36% @ 60,000 cells
Detected Genes per cell:	~1,300

GEXSCOPE® & MATRIX NEO™ - SINGLERON

Singleron

Singleron's single-cell technology is a microwell-based technology using the SCOPE-chip®, a portable microfluidic chip with microwells. High density microwell chips can capture up to 30,000 cells per chip at a time, while maintaining a low doublet rate. The sCelLiVE® kit is an additional kit from Singleron that can improve cell viability to >90% through dissociating tissues in a preservation buffer to simulate a natural cell environment.

A single-cell suspension is loaded onto the SCOPE-chip®. This chip integrates several processes, including cell partitioning, cell lysis and mRNA capture. Once single cells are captured in the wells, barcoded beads are distributed and are of a diameter that only one bead can cover a well. These beads capture the mRNA, which undergoes reverse transcription to cDNA before sequencing. The Matrix NEO allows for walk-away automation of the Scope-chip®

Technology:	SCOPE-chip®
Instrument:	Matrix NEO (360 mm x 260 mm x 230 mm, 10 kg)
Workflow Length:	1 day
Cell Throughput Capacity:	500,000 cells per run
Flexibility:	Kits for cells, nuclei, microbial RNA, variants and RNA. Each chip can be mix & matched for applications and configurations.
Sample Capacity:	4 samples (1-4 chips per run), up to 16 samples through the CLindex multiplexing kit
Capture Efficiency:	35% - 65%
Mutliplet Rates:	~2% @ 10,000 cells, 7% @ 30,000 cells
Detected Genes per cell:	~2,500

SHASTA™ SINGLE-CELL SYSTEM - TAKARA BIO

The Shasta Single-Cell System uses nanosyringe technology to dispense cell suspensions and NGS library preparation reagents in a fully automated, high-throughput manner at nanoliter scale. The Shasta system supports a variety of NGS applications, such as total RNA-seq or whole-genome amplification (WGA) for detection of single-cell level splicing isoforms, gene fusions, lncRNAs, or CNV and SNV events, respectively. Furthermore, the included Shasta CELLSTUDIO™ software allows customers to design and develop their own single-cell assays

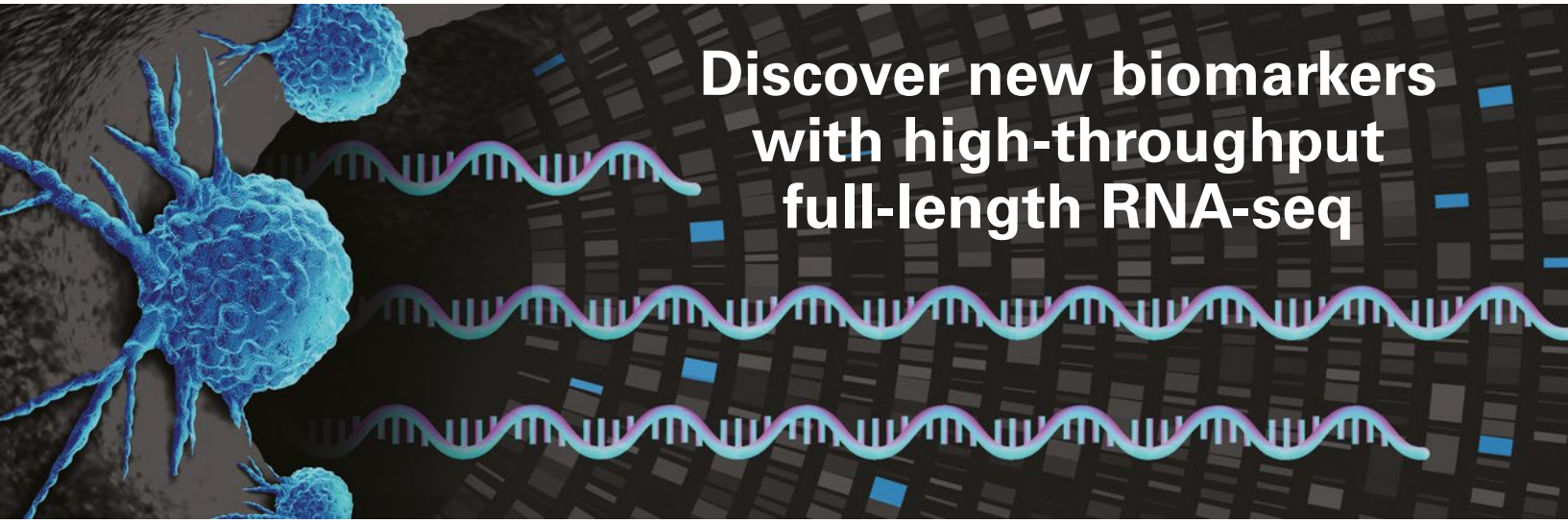
Technology:	Nanowell chips for reaction, Nanosyringe for dispensing
Instrument:	820 mm x 560 mm x 610 mm, 109 kg
Workflow Length:	2 days (7 hours each day)
Cell Throughput Capacity:	1,500 – 100,000 cells per run
Flexibility:	All cells from 5 µm to 120 µm. Very intuitive UI to set up customized assays or chemistries
Sample Capacity:	8 samples per run for mRNA-seq and whole-genome amplification, 96 samples per run for total RNA-seq
Capture Efficiency:	42% with PLUS dispense
Mutliplet Rates:	<1% rate
Detected Genes per cell:	3,000+ for PBMCs; ~10,000 for K562 and A549 cells

BD RHAPSODY™ HT – BD BIOSCIENCES

The BD Rhapsody system from BD Biosciences is comprised of several units. Cells are loaded onto the BD Rhapsody cartridge, which uses gravity-based microwell technology to capture single cells from suspensions. Beads are loaded onto the cartridge (one bead per well) before lysis and mRNA capture on the bead. The cartridge is loaded onto the HT Xpress System for several steps in the protocol including automated bead capture and cell lysis. The BD Rhapsody™ Scanner scans the cartridge at multiple points to ensure accurate capture information and QC.

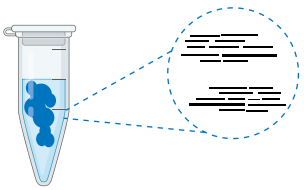
Technology:	Microwell – AbSeq (mRNA and surface protein)
Instrument:	410 mm x 410 mm x 210 mm, 5.5 kg
Workflow Length:	~10 hours
Cell Throughput Capacity:	100,000 - 440,000 cells per cartridge
Flexibility:	8 different experiments per cartridge and partial use of cartridges is possible
Sample Capacity:	12 samples pooled with the multiplexing kit
Capture Efficiency:	Up to 80%
Mutliplet Rates:	2-3% @ 10,000 cells, 8-10% @ 40,000 cells
Detected Genes per cell:	1,100–1,300

Shasta™ Total RNA-Seq Kit

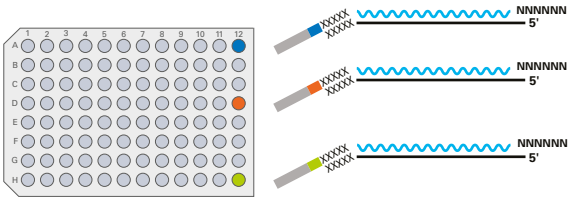


Single-cell, full-length, RNA-seq on the Shasta Single-Cell System

1. Cell fixation and RNA fragmentation

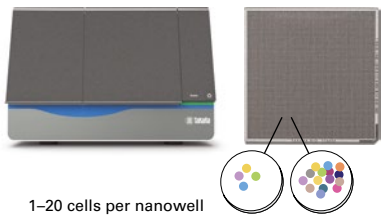


2. In situ reverse transcription
(96 different sample-specific barcodes)



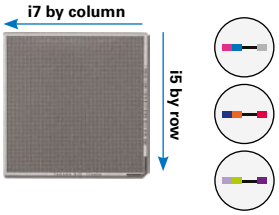
4,000–10,000 cells per well

3. Harvesting and dispensing
into 5,184-nanowell chips



1–20 cells per nanowell

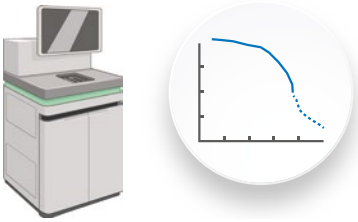
4. Cell lysis
(72 x 72 different barcodes)



5. Pooling and library purification
(off-chip PCR to add adapter)

Sequencing-ready libraries

6. Sequencing and data analysis



5,000–100,000 single cells per run

What can you expect with the Shasta Total RNA-Seq Kit?

- **High cell-throughput**—sequence up to 100,000 single cells
- **Powerful chemistry**—sequence full gene bodies, including non-polyadenylated RNA, with random priming-based chemistry
- **More data**—discover novel biomarkers such as lncRNAs, isoforms, and gene fusions
- **Convenience**—skip laborious pipetting with automated barcoding
- **Comprehensive workflow**—effortlessly analyze data with user-friendly Cogent™ tools.

Visit takarabio.com/shasta-totalrna to learn more

that's GOOD science!®

Takara Bio USA, Inc.
United States/Canada: +1.800.662.2566 • Asia Pacific: +1.650.919.7300 • Europe: +33.(0)1.3904.6880 • Japan: +81.(0)77.565.6999

Hydrogel-Based Methods

The new set of methods emerging in single-cell studies are the hydrogel-based methods. These happen to be the most convenient methods to implement as the entire reaction occurs within a single tube. Just like the plate and well-based methods, mRNA is captured with beads, but the difference here is the solution that the cell suspension is stored in. Cells are captured in a hydrogel mix, which mimics the natural conditions of a cell. The hydrogel partitions cells into individual reaction chambers within the tube without the need for plate, chips or instruments.

This methodology is very new but is already showing great promise for allowing middle-throughput experiments with next-to-no instruments or specialised kits. These kits represent the closest technology to single-cell experiment democratisation since they involve flexible, fast and easy workflows at a low cost.

PIPSEQ™ – FLUENT BIOSCIENCES

PIPseq from Fluent BioSciences is a hydrogel-based high throughput method. Beginning with a cell suspension, the suspension is vortexed with core template particles to form Particle-Templated Instant Partitions (PIPs). Each cell is a contained vesicle, they are lysed on a thermal device and the resulting mRNA is captured by barcoded oligonucleotides incorporated with the template particles. Standard cDNA libraries and NGS are then performed.

This hydrogel-based method enable a higher capture efficiency compared to the other high-throughput instrument-free methods but at a reduced throughput per kit.

Technology:	Barcoded-Hydrogel (PIPs)
Instrument:	Instrument-free (high throughput)
Workflow Length:	2 days
Cell Throughput Capacity:	16,000, 80,000 or 200,000 cells per kit
Flexibility:	Kits for low, middle and high throughput, any cell size
Sample Capacity:	8, 4, 2 samples per kit
Capture Efficiency:	Up to 85%
Mutliplet Rates:	<5% rate
Detected Genes per cell:	~3,500

ASTERIA™ – SCIPIO BIOSCIENCE

The Asteria™ kit is a hydrogel-based kit, which uses a completely reversible hydrogel technology to accomplish cell and mRNA capture within one tube. Cells are gently mixed with barcoded beds, which form 1:1 pairs, before dilution in the liquid hydrogel. Hydrogels are polymerized on ice, which isolates cell-bead pairs before cells are lysed to capture cytoplasmic mRNA transcripts onto the local barcoded bead. Thanks to the hydrogel, cells are thus isolated without freezing, fixation or vortexing requirements, reducing cellular stress and preserving the original transcriptome. The hydrogel is dissolved, and the beads are collected for reverse transcription and NGS.

Technology:	Barcoded-Hydrogel (RevGel-seq™)
Instrument:	Instrument-free (high throughput)
Workflow Length:	2 days (libraries in 14 hours, cDNA in 10 hours)
Cell Throughput Capacity:	Up to 15,000 cells per sample
Flexibility:	Kit is specific to purpose, any cell size can be sequenced
Sample Capacity:	4 samples per kit (~10,000 cells per sample)
Capture Efficiency:	45%
Mutliplet Rates:	<3% rate
Detected Genes per cell:	4,500 - 5,000+

Which Technology Should You Choose?

How should one choose a single cell method? One option is to use an initial discriminator to help you pick, based on the size and scale of your study. This determines whether your experiment would benefit from a low-throughput but versatile plate-based assay, or a high-throughput droplet-based approach.

Are you focused on rare cell populations? Or on exploratory sampling? Because methods with high dropout and high multiplet rates will be poor for approaches focusing on hard-to-find cells.

While newer methods like Parse Biosciences and Scale Biosciences are struggling to compete with 10x Genomics on cell retention, they excel in sample throughput . Given the increasing drive for atlas-level sequencing. Large-scale projects should not overlook these Split-Seq style approaches.

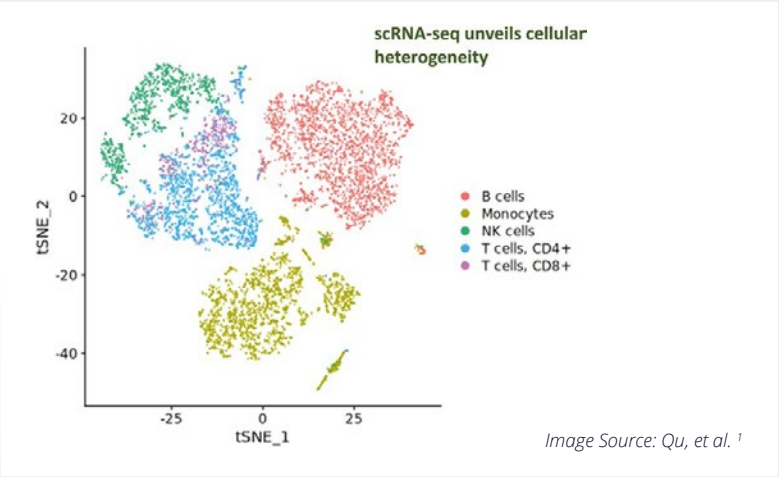


Image Source: Qu, et al. [1]

The revolutionary impact of the hydrogel-based methods has not been fully established. Yet, the promise of these methods is a democratized single-cell landscape like never before. How well these assays will do out in the wild is something to pay close attention to.

One way to help you choose between technologies is to look at the specs as outlined earlier in this chapter. However, looking at the specs provided by companies (in which the machines are run by experts with optimal samples) is one thing. Looking at how the methods perform at the 'regular' bench is another. While small in number, studies that directly compare single-cell methodologies in regular lab environments are invaluable to finding robust methods.

One very recent study[11] compared the performance of Parse Biosciences, which uses sample multiplexing, and 10x Genomics, without the multiplexing, on PBMCs. The general findings were that both platforms produced high-quality data with the recovery of the desired number of cells and genes. However, there were differences. Parse exhibited a two-fold lower cell recovery rate and 13% fewer valid reads. The mechanical forces of the Parse method, involving repeated transfer of cells between wells and tubes, was likely to blame. Reassuringly, Parse's method resulted in a higher sensitivity in gene detection, meaning it could detect rare cell types well, despite the high cell drop out. Parse ultimately performed similarly enough when compared to 10x Genomics to justify its use where large-scale multiplexing is desired.

Another comparison from 2021 compared 4 popular single-cell methods of the time[12] – 10x Genomics, ICELL8, ddSeq and Fluidigm C1/HT. Overall, the platforms showed a high degree of specificity of protein-coding genes, While each platform had strengths and weakness, such as the advantage of the ICELL8's agnosticism to cell size, it was the 10x Genomics workflow that proved the most robust across the measures, providing the best balance between number of reads and number of cells.

We spoke to the lead author of that 2021 study - **Dr. John Ashton**, Director of the Genomics Research Center at **University of Rochester**. His unique experience of directly comparing single-cell platforms in a research publication, alongside directing a genomics core facility, led to a stimulating conversation around navigating the single-cell landscape.

Latest Commercial Announcements

Below are a selection of the single-cell commercial announcements that have been made so far this year:

10x Genomics announced several new developments across their portfolio, but for single-cell, they launched the GEM-X technology for the Chromium X. This technology is set to increase the sensitivity, allow two-fold more cells per channel and two-fold reduction in cost per cell – Press Release.

Scale Biosciences announced the plan for a new two-plate workflow enabling the preparation of up to 2 million cells - Press Release.

Parse Biosciences announced **Evercode Version 3** with improvements in sensitivity and workflow speed – Press Release.

Fluent Biosciences announced early access of their PIPseq-CRISPR screen method alongside a partnership with Ultima Genomics for a low-cost million-cell PIPseq kit – Press Release.

Takara Biosciences announced the planned launch of a new single-cell kit – **Shastal™ Total RNA-Seq**. This new kit would allow full-length profiling of up to 100,000 cells per run. It would be the first high-throughput full-length method to hit the market – Press Release.

Singleron launched **AccuraSCOPE®**, a multi-omics single-cell method for the entire genome and full-length transcriptome of single cells – Press Release.

Mission Bio introduced a sample multiplexing feature to the Tapestri Platform, reducing the per-sample cost by up to 60% - Press Release.

INTERVIEW:

JOHN M. ASHTON, PHD, MBA

ASSOCIATE PROFESSOR, DEPARTMENT OF BIOMEDICAL GENETICS, DIRECTOR, GENOMICS RESEARCH CENTER
WILMOT CANCER INSTITUTE, UNIVERSITY OF ROCHESTER

***FLG:* Can you introduce yourself to our readers?**

John: My name is John Ashton. I'm the Director of the University of Rochester Genomics Center. I'm an Associate Professor in biomedical genetics and my research focus is cancer stem cell biology. I transitioned to this role as director of the genomics centre from a postdoc and have been in role for about 11 years now. In terms of what I do, I use genomics, high throughput sequencing, transcriptomics, and single-cell biology for cancer research. However, another part of my goal here is to continually push the envelope on emerging technologies, particularly, single-cell, spatial transcriptomics and multi-omics.

***FLG:* Given your history with single-cell, could you give us your perspective on how single-cell and spatial technology have been developing over the last few years?**

John: Single-cell, along with a lot of technology in this space, is based around 10x Genomics and their platform. But the push right now is the integration of single-cell and spatial omics. We're seeing an emergence of technology from 10x Genomics – the Visium technologies and the Xenium technology – but also from newer players in the field, like Singular Genomics, who are trying to integrate spatial imaging, proteomics and epigenetics together. I think we are going to be using single-cell for upfront characterization for an experiment, followed by trying to relate the cell type data from that into the spatial context. This allows you to provide a multi-omics approach where you're looking at protein, phosphoprotein, epigenetics, etc.

***FLG:* The Genomics Research Centre offers single-cell services, namely the 10x Genomics suite of options, and Smart-Seq. Do you find that this selection is enough to meet your service user needs?**

John: Yes, for the majority of users. We do have other single-cell methods not listed on our website that are more internally focused. A lot of these are still cell-sorting-based, such as SCRB-seq and CEL-seq and other types of things where it's not just Smart-Seq. We are always evaluating emerging methods and technology, so in the future we hope to expand our single-cell portfolio even more.

***FLG:* What are the options for spatial platforms, you've got 10x Genomics and NanoString, right?**

John: Yes, we have Visium, Xenium and GeoMx right now, but we're exploring other methods such as AtlasXOmics, and other types of methods where we're partnering with the Human Biobank here to try to make use of those types of specimens more valuably.

***FLG:* Are you picking your next platform based on your users or just what you think is going to be the most useful going forward?**

John: We use a mix of data to drive decision-making, but a major focus is on strategic choices that will position us to be the most impactful for our current and future investigator needs. We evaluate platforms and methods constantly. From the single-cell space, we've evaluated Parse Biosciences, Fluent Biosciences, Honeycomb, etc. For us, a major focus of research that we support is cancer-related, and a lot of it is primary patient samples, we find there's a stark contrast in performance when you're dealing with cell lines versus primary specimens. So, that's why we've we are using 10x Genomics mostly, and Smart-seq when we can.

It's the same thing on the spatial side, we're trying to identify technologies that are going to add value and can be utilised on the specimens that we have. We are always primary specimen focused, while xenotransplants and things like that are important, we want to leverage the resected tumours that are already in FFPE blocks - that's not always a universally applied technology.

FLG: Missing from your single-cell portfolio right now are the combinatorial barcoding or the hydrogel methods. You mentioned you have evaluated some of them. What issues did you identify that made you decide not to include them?

John: A lot of the issues we had were drop-out events. These kits worked great on cell lines, but when using primary tissues, the fixation methods, and the other methods that they use, are not very efficient and they damage a lot of the material. Honeycomb was concerning because we recovered very few cells recovered from that in our hands, even using a cell line.

A lot of these methods are still emerging and not many people are using them. There's less support and they're not being vetted in the field as thoroughly as they need to be. That's been our experience with Parse Biosciences and Mission Bio. We've done tests on the Tapestri platform, but it didn't work well for our specific use cases. The low capture efficiency means you don't end up with a plethora of cells and you take a hit on being able to effectively recapitulate the full diversity of cells that are in a sample.

The combinatorial approach is tricky because it has to work perfectly, otherwise you get barcode collisions etc. Our experience trying to build a combinatorial approach with single-cell ATAC has been less successful. This is because on primary tissues you have hundreds of cell types, and you can't be sure that you are binning the right cell into the right cell type. It's the same thing with 10x Genomics when we're trying to do multi-omics; it's not very clear that you've got the RNA data and the ATAC data from the same cell because the barcode collisions are quite frequent. So, that's still a work in progress. But generally, we evaluate the tools, we predict what our user base wants and is going to need, and then we try to get ahead of that.

FLG: In your 2021 paper, you compared four major single cell technology platforms. Could you talk me through the rationale of why you compared those platforms and the sample you chose to compare them across?

John: At that point, the platforms from the leading competitors for single-cell were the SureCell from Illumina, Takara's ICELL8, 10x Genomics Chromium, and then the Fluidigm system. We compared them all to Smart-seq. For the sample, we decided to keep it somewhere in the middle, so we chose to use a cancer cell line. If you use a cell line that's very homogeneous you might get a different answer, but we also understood that there are complexities to trying to do primary tissue on these systems. We set out to compare each, to show the pros and cons, and I think we did a decent job of that.

As of today, at least two of those platforms really aren't in the mainstream, the ICELL8 really isn't used that much and the SureCell/ddSeq is pretty much gone. So, that also gives you a sense for how rapidly evolving the field is.

FLG: What did you see in terms of reproducibility between the platforms?

John: That's a good question. 10x Genomics seem to be the most consistently reproducible, although there are caveats to that. Each platform does have its own value proposition. The ICELL8, for example, had no upper limit on cell size, so if you're looking at large cells most of the other methods are not going to work. Overall, 10x Genomics performed fairly well, and Fluidigm performed well in certain aspects. The main benefit of Fluidigm was that it was full length sequencing, compared to the other methods that we used. They all had decent resolution, but overall 10x Genomics seemed to perform well and was probably the most cost effective for everything. Although the others weren't far behind, they were less correlative.

Chapter 2 references

1. Qu, H.-Q., Kao, C. & Hakonarson, H. **Single-Cell RNA Sequencing Technology Landscape in 2023.** *Stem Cells* **42**, 1-12 (2023).
2. Probst, V. *et al.* **Benchmarking full-length transcript single cell mRNA sequencing protocols.** *BMC Genomics* **23**, 860 (2022).
3. Jiménez-Gracia, L. *et al.* **FixNCut: single-cell genomics through reversible tissue fixation and dissociation.** *Genome Biology* **25**, 81 (2024).
4. Vallejo, A.F. *et al.* **snPATHO-seq: unlocking the FFPE archives for single nucleus RNA profiling.** *bioRxiv,* 2022.08.23.505054 (2022).
5. Chen, W. *et al.* **Live-seq enables temporal transcriptomic recording of single cells.** *Nature* **608**, 733-740 (2022).
6. Domcke, S. *et al.* **A human cell atlas of fetal chromatin accessibility.** *Science* **370**, eaba7612 (2020).
7. Martin, B.K. *et al.* **Optimized single-nucleus transcriptional profiling by combinatorial indexing.** *Nature Protocols* **18**, 188-207 (2023).
8. Abay, T. *et al.* **Transcript-specific enrichment enables profiling rare cell states via scRNA-seq.** *bioRxiv,* 2024.03.27.587039 (2024).
9. Replogle, J.M. *et al.* **Mapping information-rich genotype-phenotype landscapes with genome-scale Perturb-seq.** *Cell* **185**, 2559-2575.e28 (2022).
10. Jiang, L. *et al.* **Systematic reconstruction of molecular pathway signatures using scalable single-cell perturbation screens.** *bioRxiv,* 2024.01.29.576933 (2024).
11. Xie, Y. *et al.* **Comparative Analysis of Single-Cell RNA Sequencing Methods with and without Sample Multiplexing.** *International Journal of Molecular Sciences* **25**, 3828 (2024).
12. Ashton, J.M. *et al.* **Comparative Analysis of Single-Cell RNA Sequencing Platforms and Methods.** *J Biomol Tech* **32**(2021).

CHAPTER 3

SPATIAL SELECTION: COMPARING SPATIAL TRANSCRIPTOMIC WORKFLOWS

WE HAVE SEEN AN ASSORTMENT OF SPATIAL TRANSCRIPTOMIC INSTRUMENTS AND WORKFLOWS BE MADE AVAILABLE IN THE LAST FEW YEARS, BUT WHICH ONES LOOK THE MOST PROMISING FOR YOUR WORK? THIS CHAPTER ADDRESS THAT QUESTION HEAD-ON BY COMPARING THE PERFORMANCE AND APPLICATIONS OF COMMERCIAL SPATIAL TRANSCRIPTOMIC METHODS.

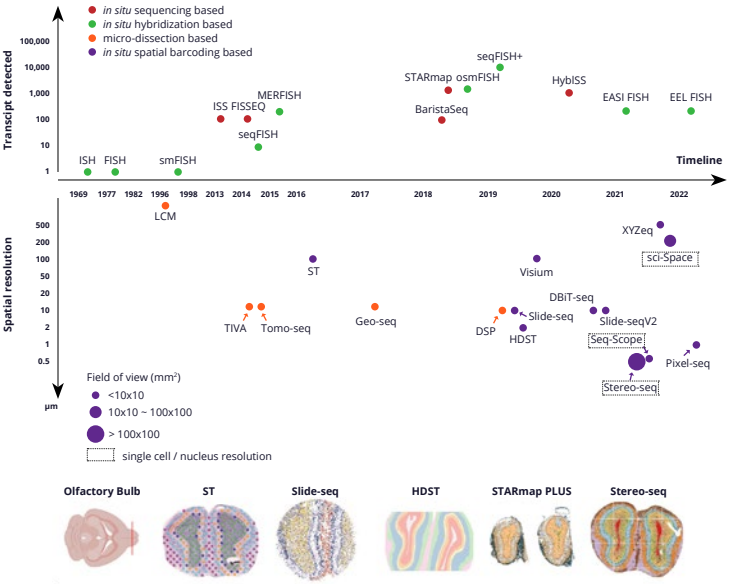
Spatial transcriptomics is a field undergoing radical development. Unlike single-cell, there are distinct methodologies that work differently under the one umbrella of spatial transcriptomics. This makes it harder to keep track of the field. Academically, there are several reviews from the last year[1-6] attempting to map out this development (that we recommend the reader takes the time to read). For our part, we will briefly introduce the major types of method before directly comparing the commercially available instruments and kits in the market right now.

Mapping Spatial Transcriptomic Methods

Spatial transcriptomic methods can be sorted into four camps, which are shown in Figure 3.1. Two of the camps rely on microscopy and imaging to produce spatial data, known as the imaging-based methods. One of these is the classic in situ hybridisation (ISH) methods, in which target RNAs are tagged with fluorescent markers to directly image them in tissue sections. The other method is in situ sequencing (ISS), in which RNA is targeted, amplified and barcoded using padlock probes and rolling circle amplification. The sample is then imaged. These two methods tend to achieve an impressive spatial resolution because they directly rely on microscopy, but the limiting factor tends to be the number of RNA targets that can be effectively visualised in one go.

FIGURE 3.1. SPATIAL TRANSCRIPTOMIC METHODS PLOTTED BY YEAR AND EITHER NUMBER OF TRANSCRIPTS OR SPATIAL RESOLUTION.

The type of method is represented by the colour of the points. Below are representative images of the olfactory bulb using different methodologies highlighted in the graph. Source (Adapted From) : Cheng, et al.[6]



The other two methods are both referred to as NGS-based methods because they rely on NGS to profile the transcriptome, while still recording spatial information. This includes the in situ barcoding methods, which work off arrays that attach millions of oligos to RNA targets within a small area on the array. The sequencing library is made in that spot on the array to get a pixel-based gene expression measure without using an imaging system. The other method is the micro-dissection based method, which involves manually ablating a tiny area of tissue, which is then sequenced, to get a spatially-known specific transcriptome profile. These methods are able to profile mRNA for 10's of 1000's of genes by using NGS, but the major drawback of these methods is spatial resolution. It has only recently been possible to get subcellular resolution with these types of methods.

DAVID COOK

Scientist, **Ottawa Hospital Research Institute,** Assistant Professor, Department of Cellular and Molecular Medicine, **University of Ottawa**

FLG: What's your general perspective of the 'spatial world'? You're in the early stages of setting up your new lab, which of these methods are you looking at using?

David: *I can discuss spatial, but I don't want to discount the importance of single-cell profiling techniques. There is discussion around this in the spatial field. With the new spatial platforms, do we even need single cell profiling methods anymore? Or can these atlas studies be generated and done with just spatial platforms alone? Ultimately, I don't think the spatial technology is at the point to completely replace single-cell for that. A lot of the tools that have been developed to augment single-cell profiling assays are incredibly valuable, and I don't see them becoming obsolete anytime soon. Even with starting my own group, we're still developing strategies for single-cell profiling with the aim to integrate that with spatial biology.*

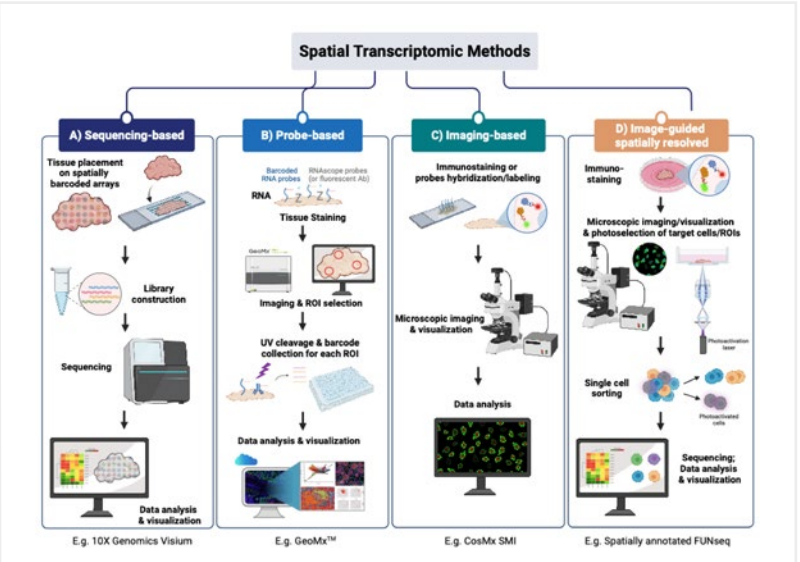
We're really interested in using in situ imaging based spatial platforms for a lot of our work. The resolution these platforms provide is best for digging down into high resolution tissue structure. In their current state, they're limited by the number of genes that we can measure simultaneously. Yet, with well-designed probe sets, I think you can cover a lot, if not all, of the biology you're interested in. Again, it's about coupling single-cell unbiased whole transcriptome analysis with these narrower spatial methods. We've had a lot of success with the 10x Genomics Xenium platform and that's what we plan to use moving forward.

The Instruments and Kits

We will now review the platforms for spatial transcriptomics - first the sequencing-based methods, followed by the imaging-based methods. An overview of the basic workflows for the different methods is captured in Figure 3.2.

FIGURE 3.2. FOUR MAIN TYPES OF SPATIAL TRANSCRIPTOMIC METHODS.

Sequencing-based methods (A) use barcoded DNA arrays to capture polyadenylated RNA transcripts from tissues followed by next-generation sequencing. Probe-based methods (B) capture user-defined targeted transcripts in manually selected regions of interest (ROIs), using corresponding, barcoded oligonucleotide-conjugated probes and can be demultiplexed accordingly afterwards. Imaging-based methods (C), similar to probe-based methods, rely on in situ hybridization but with complimentary fluorescent probes, and the targeted transcripts can be detected in a cyclic manner. Image-guided spatially resolved scRNAseq methods (D) can select spatially different single cells in ROIs (i.e. photoactivation of single cells in ROIs) followed by fluorescence-activated cell sorting and scRNAseq, thereby preserving their native spatial information as well as retaining in-depth profiling of scRNAseq. Credit: Chen, et al. [4]



Sequencing-Based Methods

VISIUM CYTASSIST – 10X GENOMICS

The Visium is a compact instrument that automates the transfer of transcriptomic analytes to Visium Spatial Slides, which have capture areas with thousands of barcoded spots and millions of capture probes within each spot. Within the Visium Instrument, the tissue sections of interest have transcriptomic probes applied, they are aligned on top of the capture areas on the Visium slide and the transcriptomic probes transfer to the Visium slide to bind to the capture probes. Spatial barcodes are added to the bound genetic data, which are then passed along for standard sequencing.

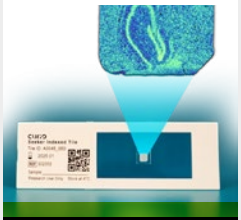
Instrument Dimensions:	203 mm x 313 mm x 333 mm, W = 8.3 Kg
Internal Storage:	Needs to be connected to external storage
Tissue Compatibility:	Fresh/frozen, fixed/frozen and FFPE
Tissue Thickness:	10 µm (5-35 µm range), 5 µm - FFPE
Scan Area per Slide:	Each slide has two 6.5mm² capture areas (max 11mm² capture area)
Slides per Run:	1 Visium slide (2 regular slides)
Turnaround time (prior to sequencing):	2 days (machine run = 30-90 minutes)
Number of Targets per Run:	~18,000 RNA targets
Resolution:	55µm (standard slides – 1-10 cell resolution), (the new Visium HD slides will enable 2 µm resolution for FFPE samples)
Multi-omics compatibility:	Protein (IF) & H&E imaging pre-barcoding and library construction for FFPE samples only

GEOMX® DSP – NANOSTRING

The GeoMx Digital Spatial Profiler (DSP) works through a set of ~20,000 gene-specific probes that are each linked to a unique barcode through a UV cleavable linker. The pool of probes is hybridized with the tissue and the slide can then be stained for imaging. The DSP barcodes are cleaved from an ROI, which can then undergo library construction and sequencing (via classic NGS or the NanoString nCounter). These reads are mapped back to the ROI through the barcodes.

Instrument Dimensions:	760mm x 730mm x 610 mm, W = 100 Kg
Internal Storage:	8 TB
Tissue Compatibility:	FFPE
Tissue Thickness:	5 µm
Scan Area per Slide:	500 mm²
Slides per Run:	4
Turnaround time (prior to sequencing):	2 days (Instrument run ~ 2 hours)
Number of Targets per Run:	~20,000 RNA targets
Resolution:	10µm
Multi-omics compatibility:	570+ proteins simultaneously using protein and RNA probes.

CURIO SEEKER – CURIO BIOSCIENCE



The Curio Seeker is the commercialised method based on the Slide-Seq technology[7,8]. This technology uses a tightly packed monolayer of beads within a Curio Seeker tile. Each individual bead captures and hybridises mRNA molecules from the tissue area placed on top of it. Reverse transcription occurs and the beads are then dissociated from the tile and the cDNA amplified, ready for the NGS library preparation and sequencing. Curio Seeker's bioinformatics pipeline can take the FASTQ files and reconstruct the transcriptomic map based on the original location of the beads on the tile. No extra instrumentation is required to run this assay and no tissue optimization is needed (tissue/species agnostic)

Instrument Dimensions:	Kit with slides provided
Tissue Compatibility:	Fresh/frozen (species and tissue agnostic)
Tissue Thickness:	10µm
Scan Area per Slide:	3mm² or 10mm²
Slides per Run:	Can easily run 2-10 tiles per day
Turnaround time (prior to sequencing):	8 – 8.5 hours
Number of Targets per Run:	Transcriptome-wide (unbiased - >25,000 RNA targets)
Resolution:	Single-cell (~10µm)
Multi-omics compatibility:	No

STEREO-SEQ® – STOMICS

STOmics® have commercialised the Stereo-seq[9] technology through a purchasable kit. STOmics® Stereo-seq® is the only spatial multi-omics solution available that can access the entire transcriptome at true single-cell resolution. With subcellular data achieved through Ultra-HD resolution, this solution provides field of view options over 160 square cm. Built on DNA nanoball technology, Stereo-seq® is designed for in situ RNA capture and RNA localization. Tissue sections are placed on a capture chip, which comprises tens of billions of probes with spatial barcodes (CID, coordinate ID) in a patterned array. Molecular information from the tissue section hybridizes with the spatially barcoded probes during permeabilization. Then reverse transcription produces spatially barcoded cDNA.

Instrument Dimensions:	Kit with capture chips provided
Tissue Compatibility:	Fresh/frozen, fixed/frozen, FFPE
Tissue Thickness:	10µm (Frozen), 5 µm (FFPE)
Capturing Chip Sizes:	0.5cm x 0.5cm, 1cm x 1cm, 1cm x 2cm,2cm x 2cm, 2cm x 3cm
Slides per Run:	1 (4 tissue sections per slide)
Turnaround time (prior to sequencing):	1.5-2 days
Number of Targets per Run:	Transcriptome-wide (>20,000 RNA targets)
Resolution:	Subcellular (~500nm)
Multi-omics compatibility:	Yes

In addition to these in situ barcoding methods, unbiased, transcriptome-wide methods can be performed without a specific kit using Laser Capture Microdissection (LCM) methods. Here, UV lasers are used to cut away a small amount of tissue to isolate specific spatially-localised cells of interest. Commercial LCM rigs can be purchased from suppliers such as the Leica Microsystems LMD6 and LMD7 and the Accuva Cellect from Laxco.

Imaging-Based Methods

XENIUM – 10X GENOMICS

The Xenium is a single-cell spatial imaging platform that utilises padlock probes that are incubated on the slides and bind to their target RNA transcripts at two points, forming a loop. Padlock probes undergo rolling circle amplification at the spot where the molecule was bound, and fluorescent oligos binds to that product to identify the transcript during imaging. Successive rounds of fluorophores and imaging build the final image.

Instrument Dimensions:	1333 mm x 685 mm x 787 mm, W = 192.7 Kg
Internal Storage:	8 TB (Xenium Analysis Computer)
Tissue Compatibility:	Fresh/frozen or FFPE
Tissue Thickness:	5 µm for FFPE and 10 µm for fresh frozen
Scan Area per Slide:	236 mm^2 (up to 3 samples)
Slides per Run:	2
Time per Run (excl. analysis):	3-5 days (machine run < 50 hours)
Number of Targets per Run:	480 gene custom panel
Resolution:	Subcellular (~200nm)
Multi-omics compatibility:	Post-workflow Xenium slides can be stained – H&E or IF (proteins)

COSMX™ SMI – NANOSTRING

The CosMx Spatial Molecular Imager (SMI) is an in situ analysis platform in which RNA-specific probes hybridise to different positions on the target genes. High specificity is achieved by having multiple probes with the same readout domain for each gene target. The instrument performs automated cyclic rounds of hybridisation, imaging and cleaving to build the final image.

Instrument Dimensions:	890mm x 740 mm x 610 mm, W = 127 Kg
Internal Storage:	Designed to be exported immediately to the AtoMx Spatial Informatics Platform (Cloud-based storage).
Tissue Compatibility:	Fresh/frozen and FFPE
Tissue Thickness:	5 µm for FFPE and 10 µm for fresh frozen
Scan Area per Slide:	300mm^2 per slide
Slides per Run:	4
Time per Run (excl. analysis):	2- 4 days
Number of Targets per Run:	Up to 1000 RNA targets (up to 6000 targets with the new ultra-high-plex panel)
Resolution:	<100nm (subcellular) with < 50 nm transcript localisation precision
Multi-omics compatibility:	Oligo-labelled antibody detection can be merged with SMI chemistry (100-plex proteins)

MERSCOPE® / MERSCOPE ULTRA – VIZGEN

The MERSCOPE® and MERSCOPE Ultra are platforms using the MERFISH technology10. MERFISH technology allows multiplexed single-molecule imaging through combinatorial labelling, error robust barcoding and sequential imaging. Each target gene is assigned a unique binary barcode, which undergoes sequential imaging to create a combinatorial barcoding scheme.

This barcoding process allows for an error robustness since a barcode with an error can still be matched to the closest correct one. MERSCOPE is differentiated from other spatial technologies for it's best in class sensitivity, and now with the MERSCOPE Ulta it's extra-large imaging area, enabling larger tissue samples to be run, or even multiple tissues in one run.

Instrument Dimensions:	430 mm × 200 mm × 560 mm, W = 50 Kg
Internal Storage:	64 TB
Tissue Compatibility:	Fresh/frozen, Fixed/frozen, FFPE
Tissue Thickness:	Up to 12µm
Scan Area per Slide:	1cm^3/3cm^2 (fall 2024)
Slides per Run:	1
Time per Run (excl. analysis):	1000 genes, 3cm^3 in 2 days (fall 2024)
Number of Targets per Run:	1000 RNA targets
Resolution:	Subcellular (< 100nm)
Multi-omics compatibility:	Proteins (6 simultaneously) and cell boundary staining

GENEPS – SPATIAL GENOMICS

The GenePS platform is based on the seqFISH technology. Transcripts are identified by a fluorescent barcode that is built up over multiple images, which ends up uniquely identifies the gene among 1000's of targets. The GenePS platform couples this technology with automation and high-quality imaging.

Instrument Dimensions:	1050mm x 630mm x 830 mm, W = 145 Kg
Internal Storage:	24 TB
Tissue Compatibility:	Fixed tissue or cells
Tissue Thickness:	10µm
Scan Area per Slide:	1cm²
Slides per Run:	1
Time per Run (excl. analysis):	Depends on number of rounds
Number of Targets per Run:	10's to 1000's of RNA targets
Resolution:	Subcellular (~100nm)
Multi-omics compatibility:	Proteins

MOLECULAR CARTOGRAPHY™ – RESOLVE BIOSCIENCE

UNDERGOING TECH UPDATE

The Molecular Cartography Platform is adapted from seqFISH technology that uses a three-probe hybridisation strategy. A primary probe binds to the RNA transcript, a secondary probe binds to the first with a barcode, and a tertiary probe binds with the fluorescent marker. After imaging, the secondary and tertiary probes are stripped, new ones bind, and eight consecutive rounds of this process occur. The platform is currently undergoing an immense tech update and an image of the new look platform is not available yet.

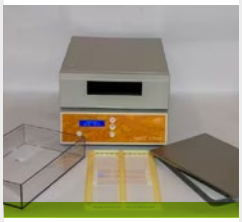
Instrument Dimensions:	Benchtop
Internal Storage:	8 TB + expansion options
Tissue Compatibility:	Fresh/frozen, fixed/frozen, FFPE, suspended cells
Tissue Thickness:	FFPE: 5µm, FF: 10µm
Scan Area per Slide:	up to 750 µm
Slides per Run:	2 slides
Time per Run (excl. analysis):	1-4 days depending on run mode
Number of Targets per Run:	300+ RNA targets, 10s of proteins, H&E staining, all on the same tissue section
Resolution:	Subcellular
Multi-omics compatibility:	10s of proteins, 300+ RNA targets, membrane staining, & H&E staining. All assays can be performed on the same tissue section.

REBUS ESPER™ – REBUS BIOSYSTEMS

The Rebus Esper™ is distinguished by its next-generation fluorescent microscopy – Synthetic Aperture Optics (SAO). Utilising illumination from outside of the lens produces a synthetic numerical aperture (NA) much larger than the physical NA. Coupled with smFISH technology, this imaging system can produce incredible high-quality images. The imaging system, on-system chemistry and analysis software are all contained in on instrument.

Instrument Dimensions:	Not stated
Internal Storage:	Not stated
Tissue Compatibility:	Fresh/frozen, FFPE
Tissue Thickness:	12µm
Scan Area per Slide:	3cm² (large enough to fit multiple samples on one slide)
Slides per Run:	1
Time per Run (excl. analysis):	> 2 days (30 targets for 1cm²)
Number of Targets per Run:	1230 RNA targets (30 for original kit)
Resolution:	Subcellular (< 100 nm)
Multi-omics compatibility:	Proteins

RNASCOPE™ – ACD BIO-TECHNE



RNAscope™ was one of the earliest commercialised spatial transcriptomic technologies. It allows RNA target detection within intact cells using specific probes and associated fluorophores. RNAscope stands out with a unique Z-probe design that has an RNA binding site and an amplification site. Signal is visualised as punctate dots that each represent a single RNA molecule.

Instrument Dimensions:	280 mm x 405 mm x 205 mm, W = 9.5 Kg
Tissue Compatibility:	Fresh/frozen, fixed/frozen, FFPE
Tissue Thickness:	FFPE – 4-6 µm, Fixed/Frozen - 7-15 µm, Fresh/Frozen – 10-20 µm
Scan Area per Slide:	19mm² tissue area per reaction (multiple tissue sections can be introduced per slide)
Slides per Run:	12
Time per Run (excl. analysis):	~ 2 days
Number of Targets per Run:	1, 2, 3, 4, 12 or 48 targets
Resolution:	Subcellular
Multi-omics compatibility:	Protein co-detection simultaneously with antibodies added prior to ISH probes or via Imaging Mass Cytometry for 40 protein markers (see Chapter 5).

Additionally, RNA ISH probes can be purchased from a number of suppliers (such as Agilent Technologies, Thermo Fisher Scientific, the HCR FISH set from Molecular Instruments and Oxford Gene Technology) to wrangle your own workflow.

OLIVER BIEHLMAIER, PHD

Head of the Imaging Core Facility at the Biozentrum
University of Basel

FLG: You chose to invest in the MERSCOPE spatial platform for your core. Was there a particular reason why you chose that one and are you looking to diversify into more spatial omic methods?

Oliver: *I did write a report to our board on which system we could acquire in late 2022, early 2023. Our decision was defined partly by availability, because at the time the system with the shortest lead time was the MERSCOPE. We also had three groups in our facility that wanted to use such a system and were interested in investing. Those groups were interested in spatial transcriptomics on the subcellular level, and from the systems that I looked at the time - which included the NanoString systems, the MERSCOPE, and the 10x Genomics systems - this was the only one which had a 63x objective with subcellular resolution.*

Compared to MERSCOPE, most other platforms essentially involve doing single-cell sequencing with the content of the cell that you collect with a specific technique. They involve labelling a certain cell to identify the transcripts that are in there, but the other systems use 20x objectives so were not at subcellular resolution. We do have another custom built small-scale spatial transcriptomics system in our facility, which is based on a standard Nikon microscope with Nikon software, microfluidics and a strong bleaching laser. That is set up to do sequential FISH, in which you do fluorescence in situ hybridization, and then you quench/bleach the bindings and repeat that process. As this is a rather slow iterative process one can only go up to a maximum of 100 genes.

FLG: Other than computational challenges, what's your take on the issues with spatial omics right now?

Oliver: *One very big problem of spatial omics at the moment is the pricing - it's something that should be heavily considered for those going into spatial omics. If you're using a new technology, you need to be prepared to encounter issues, which is fine. This has happened with imaging before – when super resolution started 10 or 12 years ago, it was pretty cumbersome at the beginning. Also, with Lightsheet, the first terabyte data sets could hardly load on any machine. This is not a problem; I'm quite optimistic that this can be solved. We can work together with the companies to better develop this and produce better results.*

The thing that I think is a huge problem in spatial omics now, especially for countries that do not have a lot of research funds, is the cost of the consumables. I think most people do not realize this right at the beginning. It doesn't matter whether you have the MERSCOPE, the CosMx, or the 10x Genomics Xenium, you always have to buy existing panels or you have to order one. The range of these panels is limited with most panels designed for human or mouse and focused on medical targets such as tumour targets. Let's say you're a neurobiologist and you need to create specific panels. Depending on the size of the panel, it tends to be a huge amount of money that the companies want from you. Still, if you compare this to doing it yourself, you're probably still better off than if you have people working on this that might not succeed. Then you have all the other consumables. In the end, for any of the systems, I would say depending on the size of the panel, you will pay between €3,000 up to €12,000 - €15,000 per run, which is not sustainable at all.

If the panels do not work, then it gets even more absurd. You can imagine the postdoc doing an experiment, and they've already failed once and now they have exactly two attempts to fill the microfluidic chamber. If they fail again, it's another €12,000... I like to compare this situation to inkjet printers. You buy a very cheap printer, and once that first ink cartridge is empty, replacing it actually costs more than the original printer. I think the scientific community is basically trying to "find the syringe to fill the cartridge", because this situation is ridiculous.

MICHELLE OCANA

Managing Director, Neurobiology Imaging Facility
Harvard Medical School

FLG: As someone who is very well read and immersed in bioimaging, what's coming next?

Michelle: *All these spatial technologies, these high-plex processes, they are super-hot right now. Everybody's talking about them, and they really like the data that comes out. It's probably the thing that I'm the most excited about. It's such a new technology and it's a new way of thinking about how we do science. The bottleneck to this is data management and analysis. We can run high plex all day long, and we can take your samples, we can elute, we can re-stain, and repeat that 50 times. At the end, how in the world do we visualise this data? I think that's coming quickly, it's going to redefine the ways we design their experiment and will determine the direction of the science.*

However, right now, it's really expensive, it's complicated, and it's not for the faint of heart. I expect within say, seven or so years, this will be as easy as clicking on an app on your phone. I think with the excitement about AI, something else will be doing the computations and the analysis and will be able to tell you what's happening. I think that RNAScope, new types of in situ, and new spatial experiments that are coming down the pipeline are really exciting.

FLG: What's your experience with the spatial systems, are you considering purchasing a spatial transcriptomic system?

Michelle: *Evaluating these systems have become a bit of a hobby for me. We have talks where we bring in vendors to talk about their systems. It's a bit like a recipe but nothing is fully cooked yet. I do like the look of some of the systems - I like the CosMx system - but I'm not sure any of these systems will be the system of the future. Maybe with the exception of NanoString. Their stuff seems to be pretty well put together and versatile enough to be used for a very long time. It's exciting right now, and there are lots of new things popping up, but I'm reluctant to say, "oh we need this box imager", because none of them can do a lot of imaging. I think we're heading into this realm of high content imaging with slides. It won't necessarily be high volume, it's just one slide with thousands of probes on it and we're going to look at it over and over again, but none of them can really do that yet. So, I'm sitting on my hands. There are a couple of systems that I really like out there, but I will not consider them right now.*

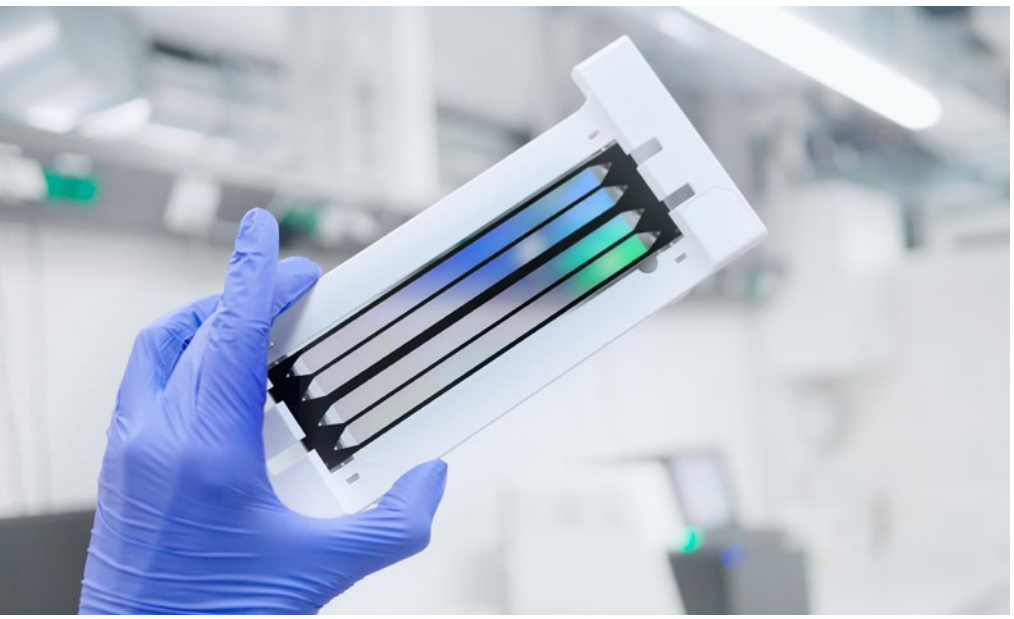
Comparisons Between Platforms

Comparing the specifications between machines is useful, but perhaps more valuable to selecting a platform is the feedback from scientists in the field who are directly comparing the performance of these instruments in their labs. Several formal comparisons between instruments have recently been published as academic preprints. To finish this chapter, we will briefly review the major insights from these papers.

Two of these papers perform direct comparisons of the new in situ imaging methods. One preprint compared the 10x Genomics Xenium, NanoString CosMx and Vizgen MERSCOPE on the same FFPE tissue type[11]. They found that, on matched genes, the Xenium had a higher transcript count without compromising specificity, the MERSCOPE and Xenium had high specificity in general, while the CosMx showed a high false discovery rate.

It is important to bear in mind that the MERSCOPE samples in this study were not run via standard protocol and likely fell below the recommended RNA values for analysis. Significantly, they were run at a 5µm imaging depth rather than a 10µm depth, as recommended, and with a reduced panel size, which are both known to diminish transcript recovery. Despite the inability to fairly compare the platforms to MERSCOPE, Xenium was found to have the highest sensitivity overall and the broad conclusion was that the reduced panel and conservative approach of the Xenium produced a more robust platform across tissues when compared to platforms such as the CosMx.

A second paper performed the comparison between the Xenium and CosMx and found a similar but stronger performance of the Xenium over the CosMx[12]. We spoke to the first name author, **Dr. David Cook**, Assistant Professor at the **University of Ottawa**, about his paper and his opinions around spatial transcriptomics more generally.



Latest Commercial Announcements

Below is a selection of the spatial transcriptomic commercial announcements that have been made so far this year:

Curio Biosciences have announced a new product, the Curio Trekker, based off of the recently published slide-tags chemistry. This product is the first true single-cell mapping kit, able to merge with single-cell workflows to provide spatial mapping of each nuclei. The early access program will begin in the second half of 2024 – Press Release.

Vizgen have very recently unveiled the MERSCOPE Ultra Platform, which is set to improve the imaging area, imaging acquisition speed and data quality acquired from MERFISH imaging. This platform will use MERFISH 2.0 chemistry – Press Release.

10x Genomics has announced advancements for both of their spatial offerings. The **Visium HD Assay** was launched in March 2024, which significantly improves a number of aspects over the original assay – Press Release.

10x Genomics has also announced new updates for the **Xenium Platform**. This includes an improved multi-modal cell segmentation kit and a new 5,000-plex panel – Press Release.

NanoString has launched the **6,000 plex RNA assay for the CosMx** alongside expanded capabilities such as superior cell segmentation, genomic breadth, and capacity to profile genes across nearly every biological pathway and over 400 ligand-receptor pairs. This accompanies the announcement for a 2025 release of the Whole Transcriptome (19,000 gene) panel for the CosMx SMI - Press Release.

STOmics has announced pre-order of the **Stereo-Seq OMNI** solution. This allows the Stereo-Seq level of resolution for FFPE tissue samples with a free probe design for capturing total RNA and microbial RNA – Press Release.

INTERVIEW:

DAVID COOK

SCIENTIST, **OTTAWA HOSPITAL RESEARCH INSTITUTE,** ASSISTANT PROFESSOR, DEPARTMENT OF CELLULAR AND MOLECULAR MEDICINE, **UNIVERSITY OF OTTAWA**

***FLG:* Can you begin by introducing yourself and your background to our readers?**

David: My name is David Cook; my background was originally in physiology as a cell and molecular biologist. Throughout my graduate studies, I became fascinated by applying computational biology to the difficult questions in genomics. How can we use these datasets to answer questions that we haven't been able to answer before? That's where I started gravitating towards single-cell biology and eventually to spatial biology.

During my PhD, I was interested in cellular plasticity in cancer, understanding the phenotypic dynamics of cells as they respond to changes in their environment. This involved using single-cell profiling strategies to track the phenotypic evolution of cancer cells. For my postdoctoral studies with Dr. Jeff Wrana, I wanted to understand how a cell's phenotype is related to its environment and how tissue structure dictates cellular phenotypes. Obviously, that's what brought me into the 'spatial world', to try to map cell populations in space and understand how they're interacting with each other within high-level tissue structure. That's where I developed experience of working with these new spatial platforms.

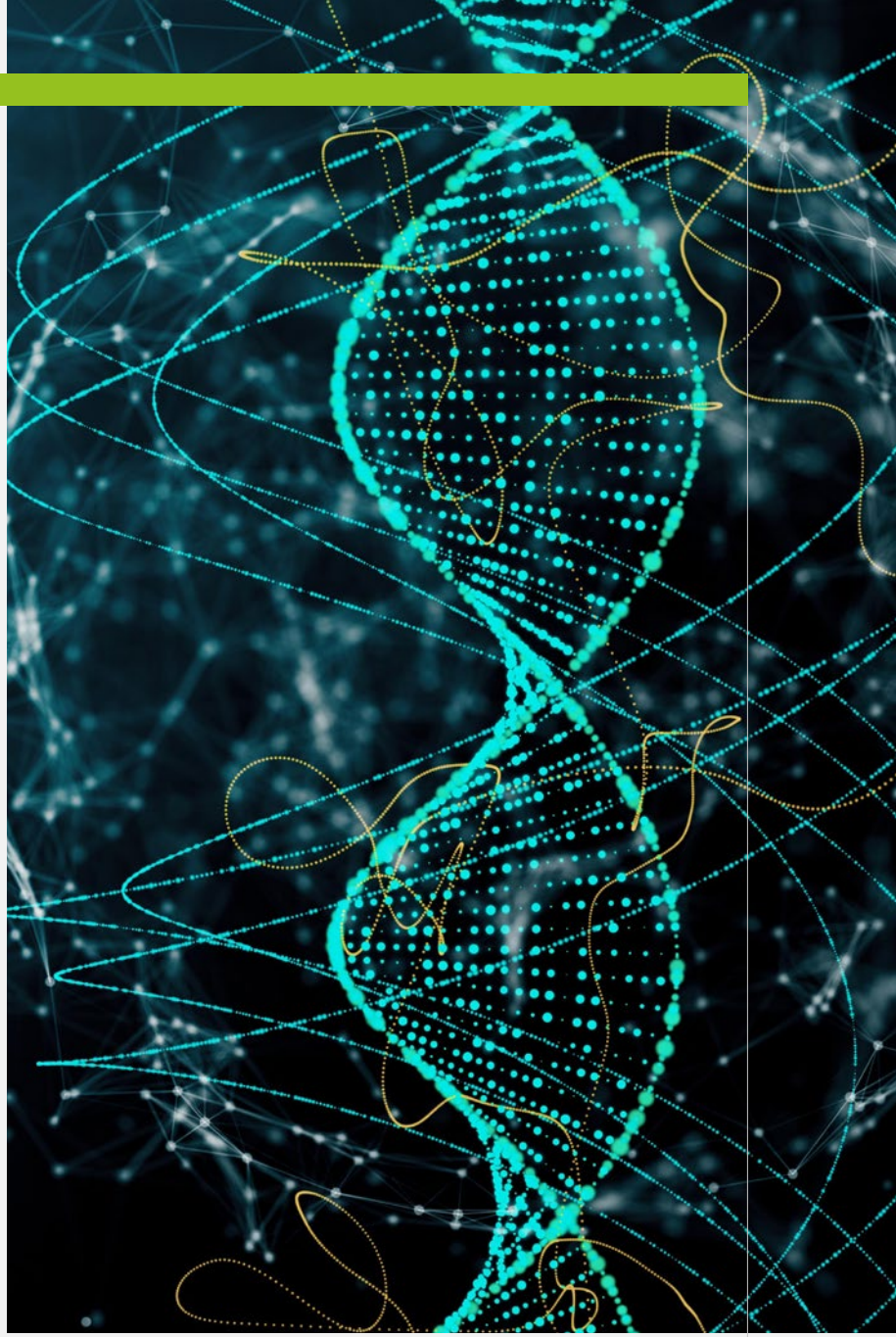
This January, I started my own independent position as a scientist at the Ottawa Hospital Research Institute, and the University of Ottawa. We're studying drug resistance and cancer progression for ovarian cancer, applying a lot of these same principles to that challenge.

***FLG:* You published a preprint last year that directly compared 10x Genomics Xenium to NanoString CosMx. Could you describe the rationale behind that paper?**

David: First, I should say I have no formal affiliation with any of these companies. I've never received financial support. I'm just an academic scientist curious about how these platforms perform. I have no skin in the game in terms of their legal battles. And similar to with single-cell platforms, I think it is beneficial for the community for there to be competition in the market.

However, I'm very invested in the actual performance of these platforms, because as an academic scientist (particularly as a new one that can't afford to buy everything), I want to make sure I'm making good choices. Mid last year, NanoString commercialized their platform, CosMx. 10x Genomics also commercialized the Xenium, early shipments of them were going out to labs. We were hearing rumblings in the field about how they were performing. Both companies were putting out public datasets that they were presenting at conferences and the data looked great! But then you start hearing rumours that some data isn't as good as it looks. As a scientist, I wanted to see whether we could make a formal comparison.

That's when I had some good conversations with Luciano Martelotto, the senior author on that preprint, and who's also invested in this. We wanted to do a head-to-head comparison, just like the other groups that have also recently released similar studies. NanoString and 10x were the big commercial players in the space along with Vizgen's MERSCOPE, but we were mostly interested in the NanoString/10x comparison. We designed a study where we could, in a controlled setting with serial sections of the same tissue, replicate a run on both platforms. Luciano has developed protocols for doing single nuclear RNA sequencing from FFPE tissue, so we have a nice single nuclear reference to compare the data to. Within this nice little contained design, we were able to make a comparison between the two of them.



***FLG:* Can you briefly summarise what you saw in that comparison?**

David: Let's first look at the considerations when choosing a platform. NanoString has made a higher plex platform allowing you to do more genes per run. Their initial offering was 1,000 up to 6,000 genes, and they are soon releasing a whole transcriptome panel. Their big push and their argument is that 'more genes is more biology'.

10x Genomics took a more conservative approach. They said, 'Okay, we're only gonna do a couple hundred genes upfront, because we believe that there are technical issues that arise when you push beyond that'. Their approach focused on smart probe-set design: pick the genes that are relevant to your system and study those rather than casting a wide net, and that way you'll get better data from it. That was an inherent difference right away.

Hence, we wanted to ask, given these differences in the platforms, how do they perform? The high-level takeaway from the comparison was that we consistently saw that the Xenium had better sensitivity across a wider dynamic range. In the Xenium, you can get reliable detection that correlates with single nuclear RNA sequencing across multiple orders of magnitude. In the CosMx, the general trend was that there was a higher level of noise (false positive signal) in lowly expressed genes, and highly expressed marker genes were detected at lower-than-expected levels, reducing the dynamic range. The correlation plots look scary, but if you translate the background signal to transcripts per cell, it's not that egregious. But it becomes difficult to interpret a cell with 1-2 transcripts for a given gene, particularly when considering that "real" markers are also detected at lower amounts. Sample quality can vary quite a bit, but we consistently observed this pattern in a range of samples.

There are practical implications of this too. If you try to annotate the data-based expression of cell type-specific markers, they are detected at much lower levels, and much less reliably in the CosMx than in the Xenium. Cell type identification became problematic for cell types that are isolated in heterogeneous tissue, such as the tumour stroma. In the tumour we analysed, we could not confidently find T cells in the CosMx data, but it was very clearly defined in the serial section Xenium data. There's a figure of that in the preprint.

If the promise of this technology is digital pathology, confident biomarker detection is important. When we see things, we want to be confident that they're real.

I do want to say, because this sounds very pro Xenium right now, and it largely is, but there are limitations with the Xenium. The common one that people talk about in the field is their initial segmentation approach based on nuclear expansion (i.e., first segment nuclei using DAPI signal and then expand the boundaries in all directions 15um). Obviously, this does not produce realistic morphologies and can make it challenging to assess tissue structure, but one could make an argument that the primary goal is to ensure the cell boundaries contain measured transcripts, so inflated boundaries aren't the end of the world.

However, if you define the cell boundaries wrong, there is a risk of misassigning a transcript to neighbouring cell. There was certainly evidence of this in our analysis. Though it is worth mentioning that they have developed on this approach and are now adopting a segmentation approach that leverages protein staining to define the cell boundaries.

> "THE HIGH-LEVEL TAKEAWAY FROM THE COMPARISON WAS THAT WE CONSISTENTLY SAW THAT THE XENIUM HAD BETTER SENSITIVITY ACROSS A WIDER DYNAMIC RANGE."

***FLG:* Is there an argument that if your scientific question was a catch all, throw genes at the wall, type of exploratory approach then the high-plex of the CosMx could still be justified to use over the Xenium?**

David: This is where I'm a little bit more pessimistic because you would imagine - yes. But given the limitations I mentioned before, I suspect it would be hard to be confident about much of the signal in that data. So, the extra genes aren't really getting you anywhere, unless you can work with the small number of them that are expressed highly enough to be confidently detected. And I think in a "discovery" experiment like this, you want to have more confidence in the data because the intention is likely to follow up on the findings, and you don't want to waste time and resources chasing a false positive.

I think many of the goals we have in spatial analysis are highly dependent on the sensitivity and specificity of transcript detection. This includes things like cell-cell communication i.e., which cell has a receptor for a ligand that's being pumped out by an adjacent cell. You want specific detection of those markers. If it's noisy, you might just get both cells expressing receptors and ligands. You need sensitive and specific detection of those genes.

***FLG:* How does your study compare to the other study comparing spatial platforms?**

David: The other study came out of the Broad Institute, and they did a comparison of Xenium, CosMx and Vizgen's MERSCOPE platform. Essentially, the data and the high-level conclusions are the same. The Xenium seems to have better sensitivity, better dynamic range. Their CosMx sample wasn't as poor as the one we had. Admittedly, for the CosMx sample that we had in our study, across the range of all CosMx samples I've seen, it was on the lower end of the distribution of quality. That's been a fair criticism. But the concern upfront is, we had no predictor of how that could have happened. It was processed in the same way, the DV200 percentage was high on that tissue. There's no reason to expect that the tissue was bad.

Another two comparison studies focused on the sequencing-based methodologies. The first carried out a head-to-head comparison of 10x Genomics Visium and NanoString GeoMx DSP[13]. Broadly, they found that the GeoMx was more sensitive but had a high degree of non-specific detection (false positives). This comparison exposed the advantages of the two platforms, namely that the GeoMx excels in deep molecular profiling of closely located cell populations, while the Visium is better for unbiased profiling across larger areas since it provides good coverage at high resolution.

The second of these studies compared 10x Genomics Visium, Slide-seq, Stereo-seq, DBiT-seq and Pixel-seq[14]. Several interesting observations were made, such as that Stereo-seq requiring a much higher sequencing cost to generate high-quality data and Slide-seq produces the best capture efficiency with normalised sequencing depth. There was also a gene capture bias on the Visium platform, with marker genes that were consistently captured by other technologies not showing up in the Visium data. Please refer to the paper for the full comparison.

Another set of preprints chose to compare pre-generated data. This has the advantage of comparing data from platforms produced using recommended protocols, avoiding the issue discussed previously. One such study[15], compared data from 9 methods and platforms and compared sensitivity of spatial methods to matching single-cell data. Both Xenium and MERSCOPE produced high sensitivity scores with MERSCOPE performing the best. Finally, another study assessed the datasets of six different in situ based methods (including Xenium and MERSCOPE) using publicly available data[16]. This study highlighted some unique issues with the various methods considered earlier in the chapter, but ultimately found that 'Vizgen's MERSCOPE datasets exhibit the best performance', with the optimal trade-off between sensitivity and specificity while featuring a large panel size. The Xenium was second but had segmentation issues that we will hear more about below. We spoke to the first author of this study, **Austin Hartman**, who is currently a PhD student at **Stanford University**, but previously worked as a computational biologist in the Satija lab, where he performed the comparison.

INTERVIEW:

AUSTIN HARTMAN

PHD CANDIDATE
STANFORD UNIVERSITY

***FLG:* Can you begin by introducing yourself and your background?**

Austin: My background is in computer science. I came out of my undergrad and worked in the genomic tools space at 10x Genomics. I realized I wanted to be more involved in the research side, so I then moved to New York to join the Satija lab. I ended up focusing more on spatial transcriptomics methods, while also generally assisting with the analysis and development of computational methods for different assays. I spent a couple of years there and have now moved back to the Bay Area to start my PhD.

***FLG:* I've reached out to you about your preprint comparing spatial in situ profiling methods[16]. Within that paper, you were not comparing data you had generated, but data from the commercial on in-house demonstration studies. Could you describe the rationale for your experiment from this paper and what you were hoping to see?**

Austin: As you mentioned, we didn't actually run these assays, or use these machines ourselves. We saw an opportunity, since there are all these datasets out there that have profiled the mouse brain, which people working in spatial transcriptomics love, yet no one had compared these data. There are probably 10's of different methods out there right now and nobody has the expertise to actually run all of those tools on tissue slices from the exact same animal and do that comparison. So, we used what was available to profile a number of different technologies and started benchmarking.

***FLG:* You compared six technologies in the paper, could you briefly overview them and tell me how comparable the datasets that you pulled from all six were?**

Austin: We compared three commercial methods - 10x Genomics Xenium, Resolve Biosciences' Molecular Cartography and Vizgen MERSCOPE, which is the commercial version of MERFISH. We also included MERFISH, STARmap PLUS, and EEL FISH.

Each of the datasets that we compared were from mouse brains. They differ, of course, in the fact that they are each from different mice. They also differ in that they are different slices of the mouse brain. Some of them are coronal sections, there are a couple of sagittal sections, and some are from different regions of the brain. That is certainly a caveat but there are ways we can control for that by annotating cell types and making our comparisons at a cell type level, rather than at a bulk level.

> HOWEVER, IF YOU JUST LOOK AT THOSE GRAPHS, IT JUST LOOKS LIKE ALL OF THESE METHODS ARE EXQUISITELY SENSITIVE. WE QUICKLY CAME TO THE REALIZATION THAT THERE WAS THIS SEGMENTATION PROBLEM."

***FLG:* What were some of the broad observations that you saw between the datasets for these methods? I think you were looking at both sensitivity and specificity of these methods.**

Austin: If we take a step back to when we first began this study, we had a relatively simple plan for how we thought we could do this comparison. The plan was to take each of these datasets in which you have 100s or maybe up to about 1,000 profiled genes. The gene sets between technologies were very different, but we thought that if we found a good single-cell mouse brain reference, we could compare all of the datasets to that. We could plot them on a graph, where each point represented the molecules per cell in single cell space and in the spatial method space. From there we could draw a trendline and figure out how sensitive, relative to single cell, each method was.

We did that. And our paper, as well as these other preprints that have come out, helped us to quickly realize we can't just take the data from each of these methods at face value, because there's so much difference upstream. This includes segmentation, spot calling, all of these important but complicated things that end users might not be able to fine-tune themselves. These things were all very different in a way that I don't think we see in the single-cell methods, where everyone's doing poly(A) capture and you always have either a whole cell or a nucleus; there's no drawing of segmentations, which affects how much material you assign to each cell.

That was the initial plan. However, if you just look at those graphs, it just looks like all of these methods are exquisitely sensitive. We quickly came to the realization that there was this segmentation problem where oftentimes you can draw segmentations, which are too large, and molecules are misassigned from neighbouring cells. Or you can draw segmentations, which are too small, meaning you miss out on molecules that actually come from that cell. And we had to do a lot of fine tuning to be able to make comparisons, after having that initial plan in mind.

It turned out that there are a couple of default methods that produce super large segmentations. And that resulted in a lower level of specificity, even if it looked like sensitivity was very high. For example, when using the default segmentations from the Xenium platform, we noticed that oftentimes, we would see a gene that is only expressed in astrocytes, expressed in cells that we were annotating as neurons. It was only after we modified the segmentations that the problem became a little bit less challenging to deal with. It was still there though, and remains an open challenge.

For the sensitivity, MERSCOPE did really well. The one thing I'll caveat with is, of all of these recent studies, ours looks at fresh frozen tissues. The other studies looked at FFPE tissues and actually saw somewhat different results. We saw what we saw, but there remains this question of how preserving tissue for so long, and how these different methods actually capture individual genes, may alter the ability to perform optimally under different tissue preservation conditions.

***FLG:* Could you elaborate on the segmentation problem. What tool were you using to improve it?**

Austin: Our initial hope was to perform a benchmark study using the processed outputs provided by each company or lab, because a majority of users are not going to necessarily want to do the segmentation or spot calling themselves. These are things that, I believe, should eventually be abstracted away so that the users can actually do the science and ask the questions. Similar to how we never think about base calling from Illumina machines, a somewhat analogous task, hopefully companies can start to abstract away some of the upstream computation and reliably spit out high quality data, although it is unclear if the methods are there yet to do so.

However, the initial segmentation results are clearly very different and so we used this tool called Baysor to standardize things. It considers the position of molecules themselves to draw boundaries around those molecules to segment cells, rather than what's conventionally done, i.e., drawing boundaries based on a DAPI stain or a cytoplasm stain. It's a bit of a different strategy, but I think it is slightly closer to the ultimate goal of what people want to generate segmentations for, which is to construct a good cell by gene matrix. If you're drawing that on the basis of stains, which I think are absolutely vital, you could miss out on molecules that are missed by the stain. I'll mention that Baysor can additionally use both stains and molecules for segmentation, which I think is a great strategy.

***FLG:* Segmentation seems like quite a basic problem, yet it still persists. Is this the most challenging aspect of spatial biology computation?**

Austin: I am not an expert on some of the things upstream, such as spot calling and the decoding of combinations of fluorescent signals into gene identities, but I think segmentation is still a problem. It's a complicated thing because all of the available technologies do it slightly differently and thus have slightly different results.

In an experiment, one of the things that's really challenging on the segmentation front is the fact that we are working with three dimensional slices of tissue.

They're very thin, but they are still 3D. So, by definition, you cannot just draw one boundary in this abstracted 2D space that perfectly describes the cell. It needs to be done in several different planes, which even then may be imperfect. A thing we've thought about is perhaps staining for specific cell types to better segment out that cell type. It's still a challenging problem because you have to be so accurate and there isn't a ton of room for error. Even assigning 5% of molecules incorrectly probably creates a larger specificity issue than what we have in a typical single cell experiment.

FLG: You said in your paper that you saw it as a way of broadly identifying high quality spatial datasets, technologies and methods. Could you just explain what you mean?

Austin: When benchmarking papers come out, the tendency is to look at the results and find a method that is better than the other methods. That's not our intention. After our paper came out, multiple companies announced that they were coming out with new panels, looking at more genes and profiling different tissues. It's unclear if the results from this paper are going to extrapolate to those new data sets, those new panels, and those new tissue conditions.

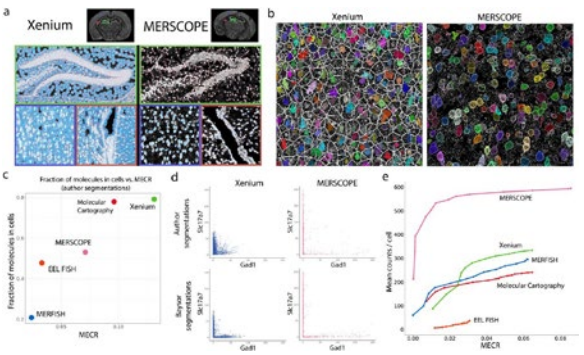
Here in this paper, we've presented ways to measure sensitivity and specificity, which we think work better than the traditional metrics that work well in the single-cell field. You can look at unique molecules per cell or unique features per cell, which I think are decent proxies of things like sensitivity, but those metrics don't work especially well in the in situ spatial transcriptomics field. We were trying to come up with metrics that do a good job of measuring sensitivity regardless of which technology comes out on top right now, because, frankly, these technologies are going to change a ton over the next few years, and we'll have to continue to monitor the ones that are performing well and the ones that aren't.

FLG: What are your thoughts on the rapidly developing single-cell and spatial landscape?

Austin: I am very curious to see if these spatial machines can develop to the point where they are as reliable as an Illumina sequencer, or the Chromium from 10x. I think we want to be in a situation where you can hand off your library or your sample to someone else and feel confident that they can generate good data to give back to you without the machine running into too many hiccups or the computation not performing well. I don't know if the commercial offerings are at that level yet. In the academic space, typically, it's just the creators of the method and maybe a couple of other labs that have the expertise to run them at that level. And commercial companies are certainly trying to push these machines to make them mature, reliable, and widely-adopted. But that's the big question for me. Will they become as reliable as a sequencer is?

FIGURE 3.3. SEGMENTATION SIZE AND QUALITY AFFECTS MOLECULAR SENSITIVITY AND SPECIFICITY.

Credit: Hartman & Satija[15]



Chapter 3 references

1. Moses, L. & Pachter, L. **Museum of spatial transcriptomics.** *Nature Methods* **19**, 534-546 (2022).
2. Zhang, X. *et al.* **Dissecting mammalian reproduction with spatial transcriptomics.** *Human Reproduction Update* **29**, 794-810 (2023).
3. Du, J. *et al.* **Advances in spatial transcriptomics and related data analysis strategies.** *J Transl Med* **21**, 330 (2023).
4. Chen, T.-Y., You, L., Hardillo, J.A.U. & Chien, M.-P. **Spatial Transcriptomic Technologies.** *Cells* **12**, 2042 (2023).
5. Wang, Y. *et al.* **Spatial transcriptomics: Technologies, applications and experimental considerations**. *Genomics* **115**, 110671 (2023).
6. Cheng, M. *et al.* **Spatially resolved transcriptomics: A comprehensive review of their technological advances, applications, and challenges.** *Journal of Genetics and Genomics* (2023).
7. Stickels, R.R. *et al.* **Highly sensitive spatial transcriptomics at near-cellular resolution with Slide-seqV2.** *Nat Biotechnol* **39**, 313-319 (2021).
8. Rodriques, S.G. *et al.* **Slide-seq: A scalable technology for measuring genome-wide expression at high spatial resolution.** *Science* **363**, 1463-1467 (2019).
9. Chen, A. *et al.* **Spatiotemporal transcriptomic atlas of mouse organogenesis using DNA nanoball-patterned arrays.** *Cell* **185**, 1777-1792. e21 (2022).
10. Chen, K.H., Boettiger, A.N., Moffitt, J.R., Wang, S. & Zhuang, X. **RNA imaging. Spatially resolved, highly multiplexed RNA profiling in single cells.** *Science* **348**, aaa6090 (2015).
11. Wang, H. *et al.* **Systematic benchmarking of imaging spatial transcriptomics platforms in FFPE tissues.** *bioRxiv,* 2023.12.07.570603 (2023).
12. Cook, D.P. *et al.* **A Comparative Analysis of Imaging-Based Spatial Transcriptomics Platforms.** *bioRxiv,* 2023.12.13.571385 (2023).
13. Wang, T. *et al.* **An experimental comparison of the Digital Spatial Profiling and Visium spatial transcriptomics technologies for cancer research.** *bioRxiv,* 2023.04.06.535805 (2023).
14. You, Y. *et al.* **Systematic comparison of sequencing-based spatial transcriptomic methods.** *bioRxiv,* 2023.12.03.569744 (2023).
15. Salas, S. M. *et al.* **Optimizing Xenium In Situ data utility by quality assessment and best practice analysis workflows.** *bioRxiv* 2023.02.13.528102 (2023)
16. Hartman, A. & Satija, R. **Comparative analysis of multiplexed in situ gene expression profiling technologies.** *bioRxiv,* 2024.01.11.575135 (2024).

PANEL DISCUSSION :

HEAR FROM THE EXPERTS PART 2
SPATIAL TECHNOLOGIES Q&A

THE CONTENT USED HERE IS A SHORTENED, EDITED TRANSCRIPT FROM A LIVE SESSION AT THE LONDON FESTIVAL OF GENOMICS & BIODATA, IN JANUARY 2024. QUESTIONS WERE INVITED FROM THE AUDIENCE.

Andrea Corsinotti
Single-cell Multi-omics Facility Manager, Institute for Regeneration and Repair
University of Edinburgh

Catia Moutinho
Scientific Adviser
The Single-Cell World

Jonathan Coxhead
Genomics Core Facility Manager
Newcastle University

Sam Jackson
Tools and Technology Platform Manager
UK Dementia Research Institute (UKDRI)

Catia Moutinho: I'm Catia Moutinho. I'm the Founder of the Single-Cell World, a platform that aims to simplify single-cell and spatial technology.

Andrea Corsinotti: My name is Andrea Corsinotti. And I am the Single-Cell Facility Manager at the University of Edinburgh.

Sam Jackson: My name is Sam Jackson. I'm the Tools and Technology Platform Manager at the UK Dementia Research Institute. We've got a couple of omics platforms that we have set up and are making available for our researchers.

Jon Coxhead: I am Jon Coxhead, I run the Genomics Core Facility at Newcastle University. We specialise in single-cell and spatial applications.

Question: I've heard through the grapevine that approaches like CosMx and the Xenium platform recognise low molecule counts per individual cell. Leaving the segmentation issues to the side for now, do you think that Visium, which works fairly robustly, and with the HD approach that they've just announced, will be a potential killer of the imaging platforms? Because they've got the image, they've got, hopefully, a good amount of information per region as opposed to the low capture probes?

Jon Coxhead: I think it probably comes down to the two different ways of collecting spatial data and how each can help answer your biological questions. One being a sequencing-based approach where you're looking tissue-wide at a lot of different genes, but probably not in a great deal of detail. But then once you've gone through that discovery process, you might want to home in on uncertain regions and then move to an imaging-based approach where you can go into detail.

Andrea Corsinotti: I think it also depends on your question. If you have preliminary data from single-cell RNA-seq or you are interested in some particular pathways, then you can design a panel that focuses on the particular transcripts in which you're interested. You get a degree of imaging resolution from a probe-based method that you're not going to get from a transcriptomic-based method. You can set boundaries; you get to see these pretty looking pictures.

So, I think it always goes back to the question. If you're going for more of a 'fishing expedition' in which you want to be unbiased in your measurements, being constrained by 500 probes, 1,000 probes, 5,000 probes, whatever is in the panel, may be a limitation. And in that case, you may want to opt for a transcriptomic-based approach.

Catia Moutinho: Yes, I agree. One approach is more of a discovery tool, and the other is more to validate or to go to the target that we want to study. If you have the money, do both!

Question: We have these new technologies to look at the spatial distribution of cells. For the past three or four years, we have seen some tools developed that use data from single-cell RNA-seq to try to tell us which cells are interacting with which cells. Have you seen, among the data that you're getting in your facilities, a concordance between these single-cell predictions and the actual spatial data, or is it completely wrong? Should we put trust in these computational tools?

Andrea Corsinotti: I think all these spatial tools with some degree of single-cell resolution are still in the very, very early stages. We are still, at least in our facility, interested in knowing if they work at all, rather than asking more sophisticated questions.

Sam Jackson: I'd agree with that. Certainly, in our hands, it has been very difficult to tell which of the techniques is better. We are trying to identify use cases for the different spatial techniques and see where they might fit into a niche better than another technique. But right now, it's very difficult to tell that from the information that's out there. And obviously, it's almost impossible to get all of the instruments in yourself and make a comparison. I think this is one reason why, in the academic community, we need to link together better over spatial techniques and share our information, if not our data. Data is often precious. But how we're using techniques, and which ones are more or less useful for one or another use case; I think this is something that we as a community need to lead on - especially given that the companies aren't going to do this, because they have a commercial interest.

Catia Moutinho: I think it's important to mention that, yes, there are a lot of researchers comparing technology. But for me, I don't think it's fair to compare them in this way. To start with, the technology works differently, the number of genes that we can detect is different - even with probes, we are comparing three hundred genes on one with one thousand genes on another. So, we need to be careful. And as you mentioned, it's important for all of us to start sharing more results from these new technologies, but it's still just the beginning. Just like single-cell sequencing technology was some years ago, so we still need to work on it.

"I THINK ALL THESE SPATIAL TOOLS WITH SOME DEGREE OF SINGLE-CELL RESOLUTION ARE STILL IN THE VERY, VERY EARLY STAGES."

Question: I was wondering if we can talk about the big challenges that come with these huge datasets from different spatial platforms, such as data sharing between users and standardising file types - from imaging to metadata and other types of input files - and how you can actually make them publicly available and share them between researchers?

Jon Coxhead: You've raised a really important point there. All these different technologies have different file formats and they've used different ways to go about analysing the data. And there's no real central repository to put this information in. The is going to have to be a really strong community driven endeavour. I don't know if any of you are aware of the GESTALT initiative that's going forward?

Catia Moutinho: Yes, it's on Twitter/X. You can fill out a form in, and you will get access to Zoom meetings, and a WhatsApp group for researchers. We discuss new problems or challenges with these new technologies. Companies offer software for us to analyse, but that tends to not be enough for what we want to see - we always want to see more, we want to play with the data, and that's the limitation. It's been very difficult to get the data out of spatial instruments that we can automatically analyse with software and various tools.

Andrea Corsinotti: One thing that I can add is that the data analysis tools provided by manufacturers are very automatic. We get to see beautiful presentations, beautiful pictures, we get to see how they work and the things that you can do. However, if you really want to do an unbiased analysis, if you really want to download the coordinates of the cells, look at what they do in an unbiased way, look at their interactions, all these kinds of things, you need to do that off the platform that is supported by the manufacturer, regardless of the manufacturer. We have experience with other technologies as well and the take home message is the same.

It reminds me of the Loupe Browser V1, the one from when people started doing single-cell RNA-seq, and the manufacturer said you can use the Loupe browser to analyse your data. Now you can use the browser to 'look' at your data. These tools are good for some preliminary checks, looking to see if the markers are expressed in the cells that you are interested in. But that's as far as they go. The whole data analysis part of these spatial technologies and the integration with single-cell experiments is getting started now, because until a few months ago, we didn't have the technologies we have today.

Audience Member: Yeah, I completely agree with that. We did our own research, and, in that paper, we used 10x Genomics tools as much as we could, but then you get to a point where you really need to have your computational team take over. The point I wanted to make, when you were talking about public data, we released all of the data associated with our paper and we submitted the Xenium data to GEO, and they didn't know what to do with it. We had to work with them to go through the process of 'how would you even name this file type?', because it was brand new data and file formats.

Sam Jackson: I think this is another place where we can have community impact as academics, because, again, companies are not going to. Whereas we, as academics and as data scientists, could come up with methods to do that quite effectively.

Question: Coming from a core service perspective, we're looking at setting up spatial transcriptomics. But obviously, now, we're starting to cross different cores - you're going from genomics core to histology core to bioinformatics core to imaging core. How have you found that in your experience? Is there anything you've done, any advice, any pitfalls? And how do we start crossing these cores? Because everything is now starting to coalesce into one big core essentially.

Jon Coxhead: Do you already work with other elements within your organisation? Do you work with biobanks and things like that already? Are you in a sequencing core?

Question Cont: Yes, at Queens in Belfast. So, we will be working with biobanks and things like this.

Jon Coxhead: If you've done single-cell, are you aligned with flow cytometry? Because that's the experience that we've had. We've not done anything where we've tried to work outside of our comfort zone, if you like. If there's been a need for histology, tissue dissection and slide preparations - we've teamed up with the histology team in the biobank because that's what they do every day. In the same way that we work closely with the bioinformaticians as well. So, in Newcastle we've bolted expertise together. When people come in to work with us, they see one thing, which is spatial transcriptomics, but behind that there are different facilities working towards one goal. I think it works quite well because it gives us strength in depth across disciplines and everyone's got their own responsibilities.

Andrea Corsinotti: In our specific case, by chance, we set up a service that is solely dedicated to these single-cell and spatial technologies. We send people to regional facilities, for histology or imaging. I think it pays to really focus on the technologies because it gives you the time and the opportunity to really develop them and develop your expertise. It helps deepen your understanding of the pros and cons of the various methods without too many distractions. But what Jon just said, this is the most important aspect. We cannot rely on one person being able to do everything anymore. We really need to bring technology experts together, and then move in synchrony. And if you need to prepare a sample then, okay, it's histology or flow cytometry imaging. One example is flow cytometry as a validation of a single-cell experiment, with new spectrum analysers and spectral sorters. So, you can see how all the various technologies take their place in the ecosystem. You cannot have one person that does too many things, because otherwise, you lose the edge on doing one thing or a few things really, really well.

Catia Moutinho: I think the secret to a successful spatial and single-cell experiment is a multidisciplinary team. This is something that we need to highlight. It shouldn't be in the middle of an experiment when you tell the bioinformatician about it. It should be from the beginning, before we start, even when we start thinking about the project or the experiment. We need to put everyone in the same room and share ideas, because, as a wet lab researcher, when I started with single-cell, I didn't know much about data analysis. When someone told me, 'No, you need more cells, because if you want to analyse 30% of a cell population, you will need to have more cells for statistical power', I didn't know that as a web lab researcher. So, it is very important to start thinking of experiments with a multidisciplinary team. I think this is crucial.

Question: I actually just wanted to go back to the question about integrating all the different data, and

the GESTALT initiative that you mentioned. Where are the software engineers in all of this? The guys who are going to look at all the different data types that you've got and marry it all together. A year ago, I was at a spatial omics conference and there were people talking about Python frameworks for combining multi-omics data, which I'm sure isn't miles away from combining into spatial omics data. But if you start from the bottom up and plan on how you will bring all this data together, then you can start telling people, 'When you make your data, it needs to interface with this framework, or you need to make it possible for me to do this'. It's those software engineers who are going to build that system for you, to combine that data in a way that not only makes analysis possible from the front end, but makes it efficient at the back end, and empowers all sorts of other analyses as well.

Jon Coxhead: I think in some ways, it's a difficult thing, conceptually, to answer, because we're not at the end of the curve on this by any means; we're right at the beginning. I think even for me, as a wet lab scientist, keeping up with the technology developments that come out every six months is difficult. So, trying to strategize a uniform data format is going to be a moving target for a while.

Question Cont: It's more about setting up a framework that allows that flexibility and makes it possible to take multiple data types, but actually streamline the analysis.

Sam Jackson: I think that there are multiple fragmented efforts to do this, in multiple different places. And I agree with you that the methods to do this are probably out there, but they need to be brought together. It's similar to my other point relating to bringing together the community to work on this, because it's a problem we have as academics that we could solve ourselves if we had the money and time.

Catia Moutinho: Yeah, and what you mentioned is the bottleneck right now. Because, yes, the technologies are at the beginning, but we can figure those out, right? It's the analysis that I think is the bottleneck.

Andrea Corsinotti: A serious answer to your question is exactly what Sam just said - there are a lot of fragmented initiatives, and even individuals, that are trying to push in this direction. My hope is that, as often happens in science, eventually the demand will drive the initiative, and then there will be enough push to create networks, a greater group of motivated people to invest their careers in these directions.

Question: I was just wondering, because you mentioned that GESTALT initiative - have you got people who are computationally minded?

Catia Moutinho: There are wet lab researchers like me, and then there are computational people that are more used to analysing single-cell data, but now we are seeing single-cell people shifting to spatial. But, to be honest, I am in the GESTALT WhatsApp group, and people are lost. For example, yesterday, there was a question like, 'How can I share this data format for a publication? How can I upload this?' And what happened is that we discussed this among ourselves - what is the best way? I think the point is trying to reach a consensus and establish guidelines and rules. In single-cell research, this has started happening earlier, so there has been more time to develop and analyse the data. But to be honest, as a wet lab researcher seeing data analysis, I think it's still a jungle. And for me, it's very difficult to understand how the analysis is being done. Because if I talk with different people, they analyse the data in different ways. So, as Sam said, we need to make an effort as academics.

PANEL DISCUSSION :

Question: I'm from a clinical background. I don't know spatial technologies particularly well - my background is histopathology. And there's been a big drive, as I'm sure everyone knows, towards diagnosis by molecular means. I'm trying to think 10 years ahead from now... given the heterogeneity of a lot of tumours, do you think things like spatial technologies - assuming that we manage to jump over all the hurdles that you're talking about now - may become a diagnostic tool? Do you think there's a future, 10-15 years from now, where diagnosis is driven by cell niches, and how tumours are responding in certain ways?

Andrea Corsinotti: The best example that we have is single-cell technologies, which are more or less 10 years old. I don't think we see any single-cell technology becoming a diagnostic tool in proper terms. Whenever you see these presentations, I don't know if you've ever noticed, at the bottom of the slide, it says 'not for diagnostic use.' And this is because these technologies are increasingly noisy. They have extremely high power, and because they are so powerful, they pick up a lot of noise. And by definition, diagnostic tools need to be nearly noise-free, so that you are just brutally assessing the situation. It's difficult, but I hope that it will come. There were a few pharma talks today about using these technologies more for drug discovery and pre-clinical approaches than as actual diagnostic tools, like a histopathologist would like to use. So, I hope that we might get there, but I think we are very far away from it.

Catia Moutinho: Firstly, you are lucky because you have knowledge of histology - this is what you need for spatial technology. Secondly, spatial is already being used in clinics for diagnostics, but it's more protein-based spatial proteomics. Right now, we can use a lot of antibodies. I think that will be a tool for diagnostics. I think it will be immunohistochemistry or immunofluorescence with a lot of markers. Spatial transcriptomics? I think that will be a little bit tricky.

Audience Member: I second what you said about single-cell sequencing not making the jump into the clinic as a diagnostic tool. But for spatial, I'm pretty sure we will have standardised workflows that are used in the clinic for diagnostics, simply because it has been used for quite some time. And, yes, there might be noise in the data, but I think there's even more information than noise that could guide you. The one problem is that we may have to endure a lag phase of three to four years, in order to have enough data generated that we can come up with a cheap way to detect what we want to detect and detect the molecular detail that will guide us. And it has to be robust, fast and also cheap if you want to bring this to the clinic - if you want the NHS or whoever to pay for it. For genomic sequencing, the prices have dropped significantly. I'd like to see the same for spatial. But I'm pretty sure that this is the way that companies will also push it. Illumina has pushed it, eventually, and this will come, I'm pretty sure, for spatial. Not for single-cell, but for spatial, for sure.

Andrea Corsinotti: I think something that will help is the introduction of methods, both for single-cell and spatial, that are compatible with FFPE samples or fixed samples - like Flex from 10x Genomics. I think it's the least noisy workflow, the data was robust and reproducible. And hopefully, now that we can work with fixed cells or fixed tissue, we can get rid of some of the noise and then maybe move forward in this direction.

Jon Coxhead: I think we're making some good progress in what we're discussing here today, because one of the obstacles to successful clinical adoption is education. If we're going to deploy this diagnostically, people need to know what the metrics mean in order to write things like clinical reports. Another huge barrier at the moment is throughput. Whether you're using sequencing-based or imaging-based methods, we're nowhere near where we need to be.

CHAPTER 4

PROTEOMIC PREDICAMENT: FINDING THE RIGHT SPATIAL PROTEOMIC APPROACH

ALONGSIDE THE ARRAY OF TRANSCRIPTOMIC PLATFORMS, AN EQUALLY ABUNDANT SELECTION OF PROTEOMIC PLATFORMS IS ALSO AVAILABLE. THIS CHAPTER WILL REVIEW THESE PLATFORMS, HOW THEY WORK, AND SHOWCASE WHICH ONES MIGHT BE BEST FOR YOU.

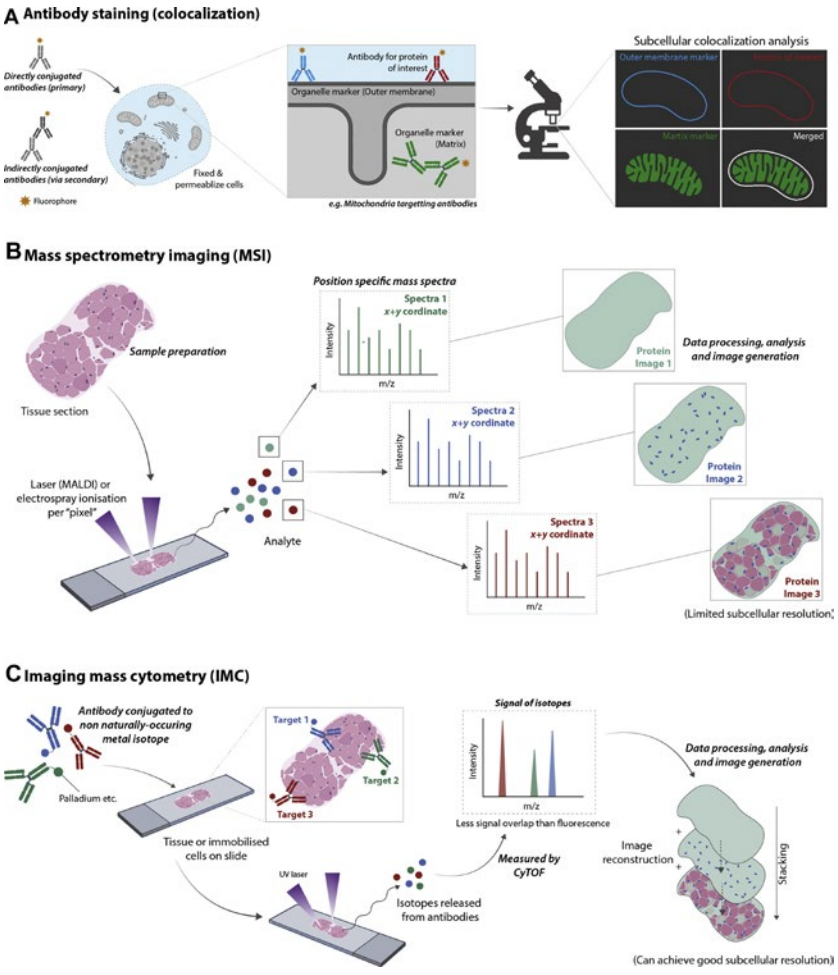
Navigating Spatial Proteomics

The current methodology in spatial proteomics can be divided into two categories and two further subcategories each. The first distinction is the division between methods that use fluorescent antibodies for protein detection and those that rely on mass spectrometry. The further division between the fluorescent methods comes down to whether the system uses conjugated antibodies (tagged with a molecular barcode for the fluorescent imaging) or regular antibodies (Figure 4.1A). The division in the mass-spec methods comes from whether they used metal-tagged antibodies to detect targeted proteins in cells (Figure 4.1C), or whether they use one of the varieties of spatial mass spectrometry imaging (SMI) methodologies for unbiased spatial proteomics (Figure 4.1B). SMI can also be used for metabolomics approaches.

This chapter will review the basic methodology and instrumentation for each of these approaches, starting with the fluorescent-antibody based approaches.

FIGURE 4.1. OVERVIEW OF THE MAJOR SPATIAL PROTEOMIC METHODS.

(A) Fluorescent-antibody based approaches for targeted profiling. (B) Mass Spectrometry Imaging approaches for unbiased spatial profiling. (C) Imaging Mass Cytometry approaches for targeted, fluorescence-free profiling. Source (Adapted From): Christopher, et al. [1]



Fluorescence-Based Spatial Proteomics

Historically, spatial proteomics has been limited to fluorescent-based approaches, namely a few fluorescently-labelled antibodies via immunohistochemistry (IHC). While this approach is limited, by taking advantage of high-resolution microscopy, it has been the backbone of spatial proteomics for decades. The principal limitation with fluorescent proteomic approaches is that one is limited in the number of proteins that can be probed in one sample by the number of available fluorochromes that can be used without producing interference or bleed through. This number is typically between 4 and 6.

Recent developments in cyclic immunofluorescence and conjugating antibodies with oligos or molecular barcodes has allowed for much improved multiplexing. Many of the available and popular spatial proteomic platforms work through this sequential imaging method. Tissue is stained either with a small selection of antibodies or a larger set of conjugated antibodies. The fluorescent marker is then added for a small set of targets, binding to the molecular barcode of specific targets or to the small selection of antibodies added. The image is then taken and the fluorescent marker or the antibodies are quenched to allow another round of this process. This process is non-destructive, and while there tends to be only a certain number of runs that can be accomplished before problems emerge, this number is quite high, allowing for high-plex studies.

Below, we catalogue some of the most popular methods for fluorescence-based spatial proteomics, including a range of throughputs, sample capacities, marker resolutions and general applications.

PHENOCYCLER® FUSION – AKOYA BIOSCIENCES

The PhenoCycler®-Fusion 2.0 (formerly CODEX) streamlines tissue staining by employing molecular barcoding chemistry, enabling the combination of over 100 antibodies into a single reaction. This is facilitated by the PhenoCycler fluidics instrument, which automates iterative cycles. During each cycle, three PhenoCycler reporters, each tagged with a distinct dye, are applied to the tissue, binding to their corresponding barcode. After imaging, these reporters and dyes are removed, making way for the next set of three reporters in the cycle. The PhenoImager Fusion instrument captures images from these cycles, storing data in the Akoya QPTIFF file format. This format incorporates a unique processing algorithm that preserves image quality while managing file size effectively. The PhenoCycler system handles biomarker tagging on whole-slides (left), while the PhenoImager Fusion system manages image acquisition (right).

Virtual training and resources for the PhenoCycler Fusion system can found at the Akoya Academy.

Note: *Akoya Biosciences also offer PhenoCode panels that are 6-plex. These panels address more targeted questions at a greater speed, with a 3 times shorter workflow. This can be done on the bench and then loaded into a whole slide scanner, such as the PhenoImager HT, for a rapid end-to-end spatial proteomic workflow.*

Chemistry:	CODEX (Co-Detection by Indexing)
Dimensions: PhenoCycler®	710 mm X 560 mm X 370 mm, W = 30.3 kg
Dimensions: Fusion	635 mm X 508 mm X 600 mm, W = 54.4 kg
Internal Storage:	Not Specified
Tissue compatibility:	Whole-slide; Tissue sections and Microarray; FFPE and Fresh-Frozen
Tissue Thickness:	4-10 µm
Antibodies:	These are antibodies tagged with PhenoCycler oligo Barcode matched to a PhenoCycler reporter with a fluorescent dye. Ready-to-use discovery panels are available, or there are resources to build your own panels using the list of antibodies.
Scan Area per Slide:	630 mm^2
Slides per Run:	2 (multiple samples possible per slide)
Time per Run:	Under 50 mins per cycle (22 mins for fluidics, 25 mins for scanning)
Number of Targets per Run:	100+ targets (with iterative staining)
Resolution:	Single-cell - Subcellular (600 nm – 250 nm)
Multi-omics compatibility:	Downstream staining - H&E, IHC, In situ RNA

CELLSCAPE™ - CANOPY BIOSCIENCES

The CellScape™ Precise Spatial Multiplexing platform uses cyclic multiplex immunofluorescence to iteratively stain and image tissues. Tissue samples are placed in compatible microfluidic chips or slide adapters. Reagents are then delivered to the sample through microfluidics. Proteins are identified through automated, fluorophore-based antibody staining (open-source commercially-available) and imaging, with high dynamic range microscopy that allows capture of high- and low-expressing protein targets without exposure optimization. This cyclic process can be repeated for more markers, and images can be layered on top of one another to achieve virtually limitless target acquisition. Data are available in the standard OME-TIFF format for compatibility with any image analysis software. Once a sample is stored on the chip, it is reusable for up to 2 years.

More information on CellScape can be found at the Canopy Biosciences Resource Centre.

Note: *With the FalconFast™ upgrade, a much larger FOV image can be acquired at 365 nm resolution.*

Chemistry:	ChipCytometry™
Dimensions:	570 mm x 380 mm x 320 mm W = 60 kg
Internal Storage:	Not Specified
Tissue compatibility:	FFPE, Fresh/Frozen, Cell Suspensions
Tissue Thickness:	5-10 µm
Antibodies:	The platform is built to use any commercially available, fluorescently labelled antibodies (find validated markers here) but predefined panels are available.
Scan Area per Slide:	up to 710 mm^2
Slides per Run:	4 (multiple samples possible per slide)
Time per Run:	24 hours for 4-8 slides at 20-plex and 20mm^2 scan area
Number of Targets per Run:	5 targets per cycle (unlimited cycles are possible)
Resolution:	Subcellular (182 nm)
Multi-omics compatibility:	Downstream Profiling (sample is undamaged from ChipCytometry)

COMET™ – LUNAPHORE

The Lunaphore COMET is a platform that uses sequential immunofluorescence (seqIF) technology to stain and image 4 slides concurrently. Tissue sections are loaded onto a microfluidic chip and the pressure-driven system delivers reagents to the section (patented FFex™ technology). This technology does not use conjugation or barcodes, but standard antibodies. Images are acquired with the integrated microscope; the signal is removed with elution buffer and another cycle of markers occurs. Images are stacked and exported automatically.

More information can be found at the Lunaphore resource centre.

Note: *The COMET can be purchased as part of the multiplex immunofluorescence (mIF) portfolio, including the SPYRE™ antibody panel kit and the HORIZON™ image analysis software.*

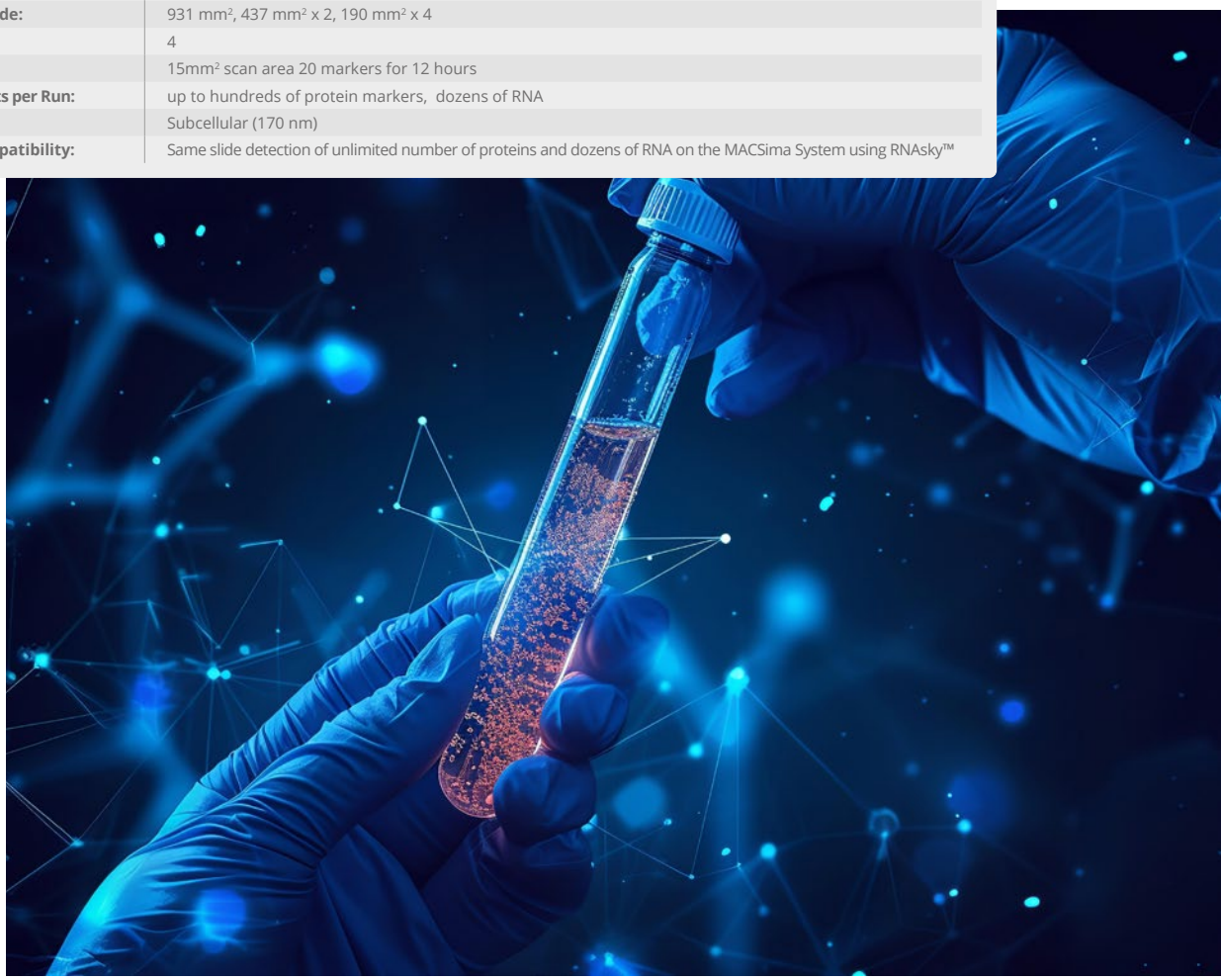
Chemistry:	seqIF™
Dimensions:	Benchtop
Internal Storage:	Not Specified
Tissue compatibility:	FFPE (optimised for), Fresh/Frozen
Tissue Thickness:	3-10 µm
Antibodies:	Standard, commercially available label-free antibodies but premade SPYRE™ panels are available.
Scan Area per Slide:	157 mm^2
Slides per Run:	4
Time per Run:	20 slides of 20 plex per week or 12 slides of 40 plex per week
Number of Targets per Run:	40 targets
Resolution:	Subcellular (280 nm)
Multi-omics compatibility:	Direct RNA co-detection with ACD Bio-techne RNAscope HiPlex Pro on the COMET (12-plex RNA with 24-plex Protein)

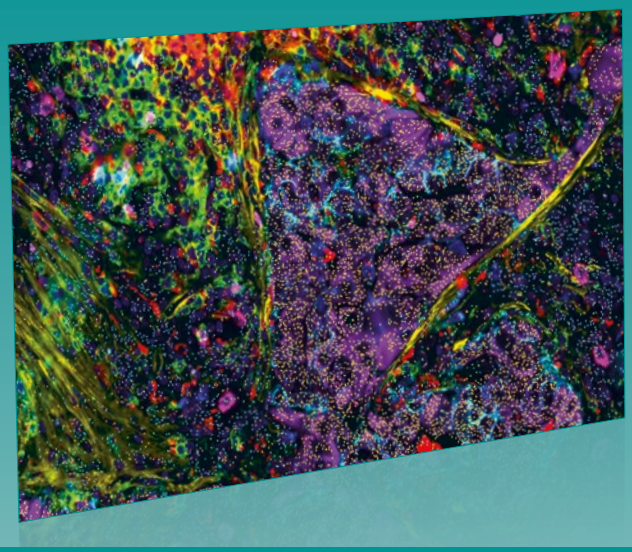
MACSIMA™ IMAGING SYSTEM - MILTENYI BIOTEC

The MACSima Platform is an end-to-end solution comprised of the MACSima System, MACSwell™ Imaging Frames, pre-tested antibodies and panels, and MACS® iQ View Spatial Biology Software. Standard microscopy slides containing frozen or FFPE tissue sections are mounted on MACSwell Imaging Frames, which are available in various sizes to fit small to large tissue samples. A single tissue or up to four slides are stained with the conjugates and then images are acquired. Fluorochromes or antibodies are enzymatically removed with REAlease® reagent or signals are erased by photo bleaching then samples are stained with the next cycle of targets. The user can choose from a broad selection of pretested antibodies or use their own. Miltenyi offers small to large pre-defined and customized antibody and RNA panels. The MACSima performs the staining, imaging, and removal in a fully automated cyclical process, which can produce an unlimited number of individual marker images without compromising the sample for downstream analysis. sample. During Acquisition, the system also processes the data on-the-fly for stitching, alignment and background subtraction. Data stacks are analysed using MACS iQ View – Spatial Biology, which includes: a multichannel viewer, segmentation, advanced plotting tools such as dimensionality reduction and clustering methods, spatial analysis such as distance and density mapping, all encompassed in a workflow designer for deeper insights.

Miltenyi Biotec is a full solution provider with decades of experience in cancer and immunology research and global support. The Miltenyi University offers webinars, on-demand videos and files on how to get the most out of the MACSima.

Chemistry:	MICS (MACSima Imaging Cyclic Staining)
Dimensions:	780 mm x 1220 mm x 650 mm, W = 170 kg (The MACSima is delivered with a table designed to support the system and additional 4K monitor)
Internal Storage:	25 Tb
Tissue compatibility:	FFPE, frozen (PFA-, or acetone-fixed), Adherent and Suspended cells
Tissue Thickness:	3-10 µm dependent on fixation
Antibodies:	Fluorochrome-conjugated antibodies. A selection of pre-tested antibodies are available as well as predesigned panels. The MACSima is an open system and can be used with antibodies from other sources.
Scan Area per Slide:	931 mm^2, 437 mm^2 x 2, 190 mm^2 x 4
Slides per Run:	4
Time per Run:	15mm^2 scan area 20 markers for 12 hours
Number of Targets per Run:	up to hundreds of protein markers, dozens of RNA
Resolution:	Subcellular (170 nm)
Multi-omics compatibility:	Same slide detection of unlimited number of proteins and dozens of RNA on the MACSima System using RNAsky™





Same-section spatial multiomics

Automated RNA and protein detection – deeper insights with RNAsky™ Detection Probes on the MACSima™ Platform

RNAsky Detection Probes: robust and high-performance RNA detection

Target essentially any gene of interest with our proprietary RNAsky Technology and ensure reproducible results with our robust and automated MACSima Platform. The RNAsky Workflow features a protease-free protocol, enabling consecutive protein detection on the same section while preserving epitope integrity and tissue morphology.

The RNAsky Technology complements the MACSima Platform's spatial biology workflow, ideal for various applications such as:

- Cross-validation of upstream screening methods
- Detecting secreted factors like chemokines or cytokines
- Protein-RNA co-expression analysis
- Replacing unavailable or poorly performing antibodies

MACSima Platform: accelerate your multiomics research

Get everything you need for seamless and rapid assay development from one source:

- Pretested antibodies: 600+ off-the-shelf antibodies and plug-and-play antibody panels.
- RNAsky Reagents: easy sample preparation workflow with predefined and custom probe panels.
- MACSima System: walk-away solution for automated spatial biology experiments.
- MACS iQ View Software: visualization and analysis of multiomics data in record time.

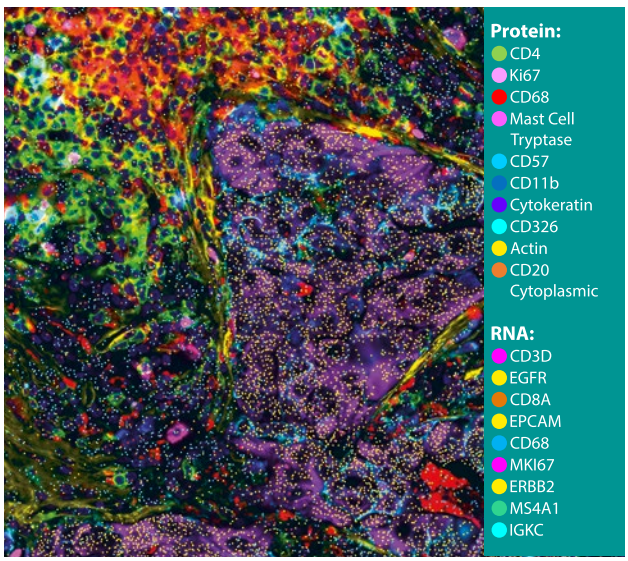


Figure 1: Simultaneous detection and analysis of 40 protein markers and 27 RNAs in non-small cell lung cancer (NSCLC) FFPE tissue.

LEARN MORE

To learn more about the MACSima Platform and see what you can discover, scan this code or visit the URL

▸ **bit.ly/3VP4mh2**

▸ miltenyibiotec.com

Miltenyi Biotec, Inc | 6125 Cornerstone Court East | San Diego, CA 92121, USA | Phone 800 FOR MACS | macsus@miltenyi.com | **www.miltenyibiotec.com**

Miltenyi Biotec provides products and services worldwide. Visit **www.miltenyibiotec.com/local** to find your nearest Miltenyi Biotec contact.

Miltenyi Biotec

ORION™ – RARECYTE

RARECYTE

The RareCyte Orion™ platform has been adopted and recognized worldwide as the only spatial biology platform to address the needs of the translational market for discovery applications through to clinical trial support. Orion enables rapid, sub-cellular resolution scanning and analysis of patient biopsies for up to 20 biomarkers, with turnaround times to support large cohort studies and to inform patient responsiveness in immune therapy trials. Exquisite sensitivity and a broad dynamic range support comprehensive cellular phenotyping, utilized by drug developers to determine mechanism of action for candidate therapeutics. The Orion platform includes a broad array of validated biomarkers, allowing immediate access by researchers and clinicians to advance their studies across immuno-oncology, immunology, neurology, infectious disease, and other critical disease areas.

The ORION Brochure and Spec Sheet can provide further information.

Chemistry:	Multiplexed IHC
Dimensions:	534 mm x 560 mm x 600 mm
Internal Storage:	9 TB scan capacity
Tissue compatibility:	FFPE, Fresh/Frozen
Tissue Thickness:	5 µm
Antibodies:	Flexible, can use commercially available antibodies, conjugation kit or predesigned panels
Scan Area per Slide:	Whole slide - 1875 mm^2 (25 mm x 75 mmx 1 mm)
Slides per Run:	2
Time per Run:	75 minutes per 1cm^2 at 20X
Number of Targets per Run:	Up to 20 biomarkers
Resolution:	Subcellular (200 nm)
Multi-omics compatibility:	Same section H&E staining

CELL DIVE™ – LEICA MICROSYSTEMS

Leica MICROSYSTEMS

Cell DIVE is a multiplexed imaging solution using a patented workflow of dye-conjugated antibodies are added to tissue sections, imaged and then removed using a chemical dye-inactivation to allow another four antibodies to be imaged in an iterative process. Images undergo a patented post-acquisition processing pipeline to remove autofluorescence and stitch them together.

The Cell DIVE brochure can be found here.

Chemistry:	Multiplexed tissue immunofluorescence (MxIF)
Dimensions:	Benchtop – 660 mm x 1128 mm x 635 mm, W = 105.7 kg
Internal Storage:	Images stored externally
Tissue compatibility:	FFPE
Tissue Thickness:	5 µm
Antibodies:	350+ validated antibodies but flexibility to source and validate commercially available antibodies independently from preferred supplier
Scan Area per Slide:	900 mm^2 (45 mm x 20 mm)
Slides per Run:	2 (Up to 15 with the BioAssemblyBot 200)
Time per Run:	Not provided
Number of Targets per Run:	60+ markers (4 per cycle)
Resolution:	Subcellular (~400nm)
Multi-omics compatibility:	Downstream (sample is undamaged from dye inactivation)

NAVENIFLEX™ (TRIFLEX CELL MR) - NAVINCI DIAGNOSTICS

The Naveni TriFlex kit is a novel small-scale in situ spatial proteomic solution. It allows the user to study the interplay of proteins and their interactions in situ. The kit detects 3 signals, the total amounts of two distinct protein targets, and a third signal, their interaction. The interaction signal necessitates that the proteins are within <40 nm of each other. Two user-determined primary antibodies target the protein targets, followed by the addition of Navenibodies that enable the proximity reaction to occur. These three signals are amplified and generate a readout that is detectable using fluorescence microscopy (see Chapter 5).

Note: Navinci Diagnostics provide other kits, such as the NaveniFlex kits. These allow the identification of protein-protein interactions in various types of samples, including cultured cells (NaveniFlex Cell), fresh-frozen samples and FFPE samples (Naveni Flex Tissue) as well as with chromogenic detection (NaveniBright).

A selection of resources can be found in the Navinci Resource Centre.

Chemistry:	In Situ Proximity Ligation Technology
Dimensions:	Antibody Kit – 2 boxes – reagents for 100 reactions
Tissue compatibility:	Fixed cells
Antibodies:	User Primary Antibodies are coupled with Navenibodies, the secondary antibody system. The NaveniLink kit is available for users to craft their own Navenibodies.
Scan Area per Slide:	Variable - 40 µl of reagents per 1cm^2
Slides per Run:	As many as your set up allows
Time per Run:	6 hours if performed all in one day
Number of Targets per Run:	2 (+ 3rd interaction measurement)
Resolution:	Depends on microscope used
Multi-omics compatibility:	Downstream – if using primary antibodies, they need to be a different species.

SCSP KIT – PIXELGEN

Another novel spatial proteomic solution is the Spatial Surface Proteome (SCSP) kit. The workflow begins with antibody-oligo conjugates that bind to proteins on the surface of fixed cells. The uniqueness of this kit is in the pixelation. DNA-pixels are added in two rounds. DNA-pixels A binds to many conjugates in proximity, which generates small, connected protein neighbourhoods. DNA-pixels B are then added to link the local neighbourhoods to construct one global protein map of the cell surface. A standard PCR-based (Illumina compatible) library is generated. After NGS, this kit allows 3D analysis of abundance and co-localisation of surface proteins through graph construction.

A selection of resources can be found in the Pixelgen Resource Centre, including the user manual and product sheet.

Chemistry:	Molecular Pixelation (MPX)
Dimensions:	Pixelgen Kit – 3 boxes – reagents for 8 reactions
Tissue compatibility:	Fixed Immune cells
Antibodies:	Immune Panel of 80 targets provided by Pixelgen
Cells per Run:	Up to 1000 cells
Time per Run:	2 days (~15 hours prior to sequencing)
Number of Targets per Run:	80 (76 targets + 4 controls)
Resolution:	Subcellular
Multi-omics compatibility:	Not stated

To complement this overview of the platforms and kits, we recently spoke to **Carolina Oses Sepulveda**, Researcher and Lab Manager of the Spatial Proteomics Unit at **SciLifeLab**. We asked about her experiences with fluorescence-based spatial proteomics and the instruments in her facility, namely the Akoya Biosciences PhenoCycler and the Lunaphore COMET.

INTERVIEW:

CAROLINA OSES SEPULVEDA

RESEARCHER AND LAB MANAGER, SPATIAL PROTEOMICS UNIT IN SCILIFELAB (KTH ROYAL INSTITUTE OF TECHNOLOGY)

***FLG:* Thank you so much for being here. Can you first begin by introducing yourself and your role at SciLifeLab?**

Carolina Oses: I am Carolina Oses Sepulveda; I am a Researcher and a Lab Manager in the Spatial Proteomics Unit. This unit has a selection of spatial biology platforms, and it is all contained within SciLifeLab. I would like to highlight SciLifeLab here, it is a Swedish national centre that contains a large scale of unique technologies focusing on biomedicine, ecology and evolution research. SciLifeLab that was formed in 2010, bringing together state of the art dedicated staff scientist and technologies, between four different universities – KTH Royal Institute of Technology, Karolinka Institute, Stockholm University and Uppsala University.

***FLG:* Let's start quite broad. What are the advantages of using spatial proteomics?**

Carolina Oses: Everyone who works with human samples knows that the samples are very precious. However, just a few years ago, it was very difficult to identify more than five antibodies simultaneously on a single sample slide. Today, with spatial proteomics technologies, we can detect far more than 5 antibodies. In our facility, we have successfully analysed 45 antibodies in just one sample. Not only do you get to visualize proteins in your tissue, but with spatial proteomics, you gain another level of information: spatial distribution. With this protein expression, it is possible to identify populations within a tissue and understand their distribution in a reference tissue. Are they clustered together? Are they sufficiently distant from a hotspot? The questions here can be endless.

> "THIS IS A BIT CONTROVERSIAL, BECAUSE THE BIGGEST BARRIER IS THE ANTIBODIES THEMSELVES. IN MOST CASES, ANTIBODIES WILL NOT PERFORM THE SAME EVEN WHEN TREATED THE SAME WAY WITH THE SAME KIND OF PREPROCESSING."

***FLG:* What kind of projects do clients approach you with?**

Carolina Oses: We have a very diverse group of users. Some of them have a very specific biological question, so we engage in a targeted approach that we call diagnostic. We also have more general research users; they come to us with broad research questions. We call them the discovery users. For the discovery users, we naturally we use the higher-plex methods, while the more specific clients engage with a middle or low plex method. I would say that around 90% of our users are focused on human cancer, but if they not looking for cancer, they are focused on chronic diseases, wound healing or infections. Two, perhaps three percent at most, are looking at other non-human organisms.

FLG: ***Your unit has two spatial proteomic instruments, the PhenoCycler from Akoya Biosciences and the COMET from Lunaphore. Can you describe how those two technologies work, what their strengths are and why you might choose one over the other for a project?***

Carolina Oses: As I said before, we have these two groups of clients. For the discovery ones, we usually guide them to use the PhenoCycler. The PhenoCycler is designed to be a highly throughput phenotypic screener of cells. In our facility, we typically can use up to 45 markers [You can use up to 100 markers], that is the maximum amount that we have performed in our facility. This instrument uses barcoded antibodies. All of these antibodies are stained in the sample at the same time. Then the instrument goes through cycles, and in every cycle, three different reporters with three different fluorophores are applied, the image is taken, after image just the fluorophores are eluted, the cycle ended. A new cycle start and three more reporters with fluorophores are added, image, elute and over and over until all the markers have been captured, this cycles are automatically done. The instrument uses software to take all these pictures and merge them together in just one file. You can then evaluate all the different markers in just one sample. It's very fast if you want to evaluate many different type of cells, which positions it towards discovery research.

Then we have the COMET, and this instrument has a higher throughput for samples. PhenoCycler allows you to run two samples at the same time, but the COMET allows you to run four samples at the same time. The other major difference is the COMET allows you to use off-the-shelf antibodies with no conjugation, just primary and secondary antibodies as a regular indirect immunofluorescence. The COMET is fully automated, I don't need to do anything other than just put the sample and the antibodies inside. All of these spatial proteomic instruments work in cycles of staining, imaging and elution. For the COMET, at each cycle the primary and the secondary antibodies are added consecutively, the whole sample is imaged, and both the primary and secondary antibodies are eluted. And then the cycle starts again. The same as when using the PhenoCycler, all the images from the cycles are integrated together to give you the final file.

The COMET also allows you to be more diverse, since we are using off the shelf antibody. This means you can use antibodies from every kind of organism or to identify any kind of organism. However, you are restricted in the number of markers that you can use. We can run 20 cycles, with 2 different antibodies in each cycle, but they need to be from different hosts, similar to manual immunofluorescence. Here is the bottle neck because most of the antibodies around the world are made in rabbit. Hence, this instrument tends to get used for translational research. If a user has 100s of samples, this instrument allows you to run many more samples faster but in a low-plex manner.

FLG: Have you got any examples of exciting spatial proteomic projects that you've been engaged with?

Carolina Oses: On our website we have all the ones that we have already published, but two come to mind. One of them is related to the identification of four different types of liver macrophages. They identified four different types of liver myeloid cells, which exhibited distinct spatial distributions when comparing lean and obese patients. This research was published in Nature Metabolism[2].

The other project was looking at neuroblastoma in kid[3]. The researchers identified 27 immune cell subtypes in human neuroblastoma. They found that the presence of natural killer cells correlated with a survival benefit, which is not something that is widely known in cancer research. Moreover, this user is really keen to try to perform multi-omics in the same sample. Our users always push us to think out of the box and going further, and that is our next level of evolution, this is something we're working on.

FLG: What tends to be the most challenging part of the spatial proteomics workflow? What's the biggest barrier that is still in the way for researchers to get the spatial proteomic results that they want?

Carolina Oses: This is a bit controversial, because the biggest barrier is the antibodies themselves. In most cases, antibodies will not perform the same even when treated the same way with the same kind of preprocessing. No matter what tissue type you're using (fresh-frozen, fixed-frozen or FFPE), we need to use just one kind of preprocessing and that one kind needs to work for all antibodies. That's very tricky. Prior to starting this facility, our head of unit worked together with the Human Protein Atlas, which gave us an idea of the kind of antibodies we should look for. Even then, it's very difficult to make all of them work perfectly fine in the same conditions.

The other issue is conjugation, which is necessary for the PhenoCycler. Unfortunately, we required antibodies that are BSA free and hopefully glycerol free, which reduces the number of antibodies we can use. After that, we need to conjugate them. Unfortunately, at present, the chemistry that we use to do the conjugation is a little bit harsh to the antibody, potentially reducing its effectiveness in binding to the tissue. In our facility, about 30% of antibodies may not yield the desired results. Then we need to try to find another clone with all the conditions that I've just said.

However, for COMET, the conjugation issue doesn't matter, but the nature of this instrument means that the incubation time is very short. It's just a minute, sometime up to two minutes depending on the antibody. Hence, the antibody needs to be very efficient to attach to the sample. This means you could have found your antibodies, they work together, but maybe they will not work in the COMET.

But on top of all these points from an antibody perspective, the most important challenge for spatial proteomic flow, and I would say for any kind of Omics technology, is the starting material. If our users don't provide a high-quality sample, then no matter what efforts I make, I will not be able to deliver high-quality data.

FLG: There are quite a few spatial proteomic instruments, and from an inexperienced eye, the chemistry looks quite similar. I didn't know if there were any other instruments that your unit has its eyes on that might add a new dimension to what you're doing?

Carolina Oses: The current instruments are roughly divided into two categories: the ones with the barcode-conjugated antibodies, and the ones without barcoding. That is the big difference between them. Right now, we are not actively seeking another instrument. Although we are always looking at the market. All the current instruments in the market have advantages and disadvantages. The ones that we have in-house are quite quick to produce results, which is why we chose them. While we are not actively searching in the market, we are aiming to further diversify our work, particularly into multi-omics. Multi-omics options are also available on our two platforms.

Some of our users don't need these fancy, multi high-plex instrument. Actually, sometimes they only want to use six or eight antibodies. So, we are looking at other technology that doesn't require these cyclic instruments, but instead utilize a fast scanner. We are looking for a high-quality scanner, both fast and capable of multi-spectral analysis. In the spatial proteomics market, I personally haven't seen any instrument that is remarkable above the other ones or something completely new.

"BUT ON TOP OF ALL THESE POINTS FROM AN ANTIBODY PERSPECTIVE, THE MOST IMPORTANT CHALLENGE FOR SPATIAL PROTEOMIC FLOW, AND I WOULD SAY FOR ANY KIND OF OMICS TECHNOLOGY, IS THE STARTING MATERIAL. IF OUR USERS DON'T PROVIDE A HIGH-QUALITY SAMPLE, THEN NO MATTER WHAT EFFORTS I MAKE, I WILL NOT BE ABLE TO DELIVER HIGH-QUALITY DATA."

Imaging Mass Cytometry

Imaging Mass Cytometry (IMC) is the first of the two mass spectrometry based methods. It is a non-imaging-based approach for spatial proteomics that uses a similar set up to mass cytometry. Mass cytometry involves a hybrid of flow cytometry and mass spectrometry using cytometry by time of flight (cyTOF) technology. IMC, and all mass spectrometry methods, get away from the issues of working with the fluorescent antibodies; namely spectral overlap and the signal clashing with background fluorescence.

The system works with heavy metal tagged antibodies instead of fluorescent ones. Once labelled with these antibodies, small areas of tissue sections are ablated, vaporising the sample and metal tags. This cloud can be analysed by mass spectrometry to determine the amounts of heavy metal ions, and hence, the protein target of interest. This approach is still targeted and is limited to a set number of targets, around 40.

There are two platforms for metal-tagged antibody-based approaches; the Hyperion Imaging system from Standard BioTools and the MIBIscope system from Ionpath.

HYPERION™ XTI IMAGING SYSTEM AND CYTOF XT™ - STANDARD BIOTOOLS

For the Hyperion Imaging System, an imaging mass cytometry approach is used. The CyTOF XT instrument performs the automated cytometry, and the imaging system acquires the images using laser ablation. Samples can be mounted on standard glass slides. They are stained with metal-tagged antibodies for the protein markers rather than fluorescent ones, and slides are loaded into the imaging system. The system has a precise UV laser beam, which focuses on 1 µm2 spots of the sample (stained with metal-tagged antibodies) at a time to vaporise the sample and the tags. The resulting cloud is analysed by mass spectrometry by inductively coupled time-of-flight (TOF) technology to determine their identity.

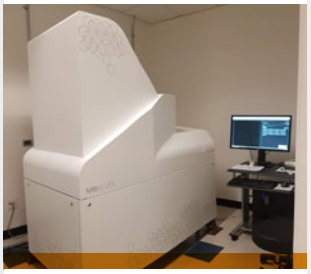
The Hyperion XTi User Guide and Spec Sheet can be accessed for more information.

Note: *Standard BioTools also have the Hyperion and Hyperion+ Imaging Systems attached to the Helios, a CyTOF system. These systems use the same chemistry with many of the same specs but have more internal storage (8 Tb), a smaller scan area (675mm2) and a lower throughput (50% slower scan rate for Hyperion+).*

Chemistry:	Imaging Mass Cytometry (IMC)
Dimensions (Imager):	580 mm x 1320 mm x 580 mm, W = 192 kg
Dimensions (CyTOF XT):	930 mm x 1350 mm x 750 mm, W = 288 kg
Internal Storage:	3.6 Tb
Tissue compatibility:	FFPE, Fresh/Frozen, Cell smears
Tissue Thickness:	7 µm
Antibodies:	Maxpar® antibodies with purified heavy metal labels, see catalogue here
Assessment Area per Slide:	825mm² (15 mm x 55 mm)
Slides per Run:	1
Time per Run:	24 hours for 40 slides with 40+ markers per day across 3 x 0.5 mm² ROIs per slide (30 minutes per scan)
Number of Targets per Run:	Up to 4- 37 protein markers per scan (plus 2 nucleic acid stains and 3 for plasma membrane demarcation)
Resolution:	~ Cellular (1 µm)
Multi-omics compatibility:	Protein and RNA co-detection, by combining with the ACD Bio-techne RNAscope HiPlex workflow (12 targets) using metal-conjugate labelled probes instead of fluorophores.

> THERE ARE TWO PLATFORMS FOR METAL-TAGGED ANTIBODY-BASED APPROACHES; THE HYPERION IMAGING SYSTEM FROM STANDARD BIOTOOLS AND THE MIBISCOPE SYSTEM FROM IONPATH."

MIBISCOPE™ – IONPATH



The MIBIscope platform uses MIBI™ technology - secondary-ion mass spectrometry - to image antibodies tagged with monoisotopic metal reporters. Tissue sections are adhered to MIBIslides - gold-coated conductive microscope slides that prevent interference from chromium and titanium in standard slides. Tissue is stained with the tagged antibodies that bind to their protein target. An ion beam is then scanned across the tissue, releasing the secondary metal isotope ions, which are quantified by a ToF mass spectrometer.

IONPATH

Chemistry:	Secondary-Ion Mass Spectrometry (SIMS)
Dimensions:	Free-standing
Internal Storage:	Images stored externally
Tissue compatibility:	FFPE, Fresh/Frozen
Tissue Thickness:	4 µm
Antibodies:	Conjugated with metal tags (MIBItags). Explore the Ionpath catalogue, or conjugation kits are available for 4 purified antibodies of the users choice.
Assessment Area per Slide:	640 mm^2
Slides per Run:	1
Time per Run:	640 mm^2 scan in 35 mins (staining & imaging = 2 day protocol)
Number of Targets per Run:	40+ markers
Resolution:	Subcellular (400 nm - 1µm)
Multi-omics compatibility:	Potential downstream usage due to not full ablation

To get more detailed information of mass cytometry-based approaches, we recently spoke to **Michalina Mazurczyk**, Manager of a Mass Cytometry Facility at the **University of Oxford** about mass cytometry, IMC and her experiences with the Standard BioTools platforms.

INTERVIEW:

MICHALINA MAZURCZYK

MASS CYTOMETRY FACILITY MANAGER, MEDICAL RESEARCH COUNCIL WEATHERALL INSTITUTE OF MOLECULAR MEDICINE, RADCLIFFE DEPARTMENT OF MEDICINE, **UNIVERSITY OF OXFORD**

FLG: Can you start off by introducing yourself and your experience with mass cytometry?

Michalina: My name is Michalina Mazurczyk and I'm a Facility Manager for a dedicated mass cytometry facility. We only do mass cytometry and we run both modalities, suspension and imaging (IMC). I started in the facility in 2016, with just the suspension method, as that was what was available. When I started here, I had little prior experience with mass cytometry, so I did a lot of testing and development, and I worked with facility Academic Lead to introduce the technology to researchers in the Institute . When IMC became available, we purchased the new instrument and expanded our offering. And again, went through a phase of learning the new tech and in house development, since the technology was very much state of the art. Over the last 8 years I helped many researchers from University of Oxford, other Universities, and private sector to apply mass cytometry to their research. Both suspension and imaging mass cytometry are very powerful techniques that can be used for discovery science and robust enough to be used for clinical studies as well. Now, we are running the facility successfully and regularly advise how to apply mass cytometry to various projects.

As a Facility Manager I am part of technical support at university. I am also a Technician Champion, within Technician Commitment Initiative at University of Oxford, and I am working on increasing visibility and accessibility of Scientific Research Facilities' (SRFs) expertise. I value to have a chance to speak so thank you for the invitation for this interview. As technicians, we can never be as good at research as the researchers – that's not our role but when it comes to technology, quite a lot of experience and expertise lies in the facilities. By crossing fields and bringing peoples' expertise together, we can accomplish much more.

FLG: Could you explain what your facility offers, and the differences between suspension and Imaging Mass Cytometry (IMC)?

Michalina: In my facility we offer a comprehensive support for research projects, from advice on experimental set-up, sample storage and costing for grant application to protocol optimisation, troubleshooting, data acquisition, data QC and data analysis tools.

Suspension mass cytometry allows multi-marker analysis of cells that are in a single-cell suspension. This could be cells from blood, from dissociated tissues, or cell cultures. Most often samples are from human or mouse but also plant samples can be analysed. We can perform analysis for whatever markers we have available antibodies. This can be for proteins on cell surface or proteins within the cells It is also possible to include RNA detection to get insight into cell transcriptome as well as proteome. IMC allows this analysis being done on tissue on a glass slide. The additional information that the IMC provides is the spatial context of cellular interactions within the tissue.

For both methods, we basically end up measuring a cloud of metals using mass spectrometry by time of flight. In single-cell mode, the single cell suspension is nebulised to produce single droplets containing single cells. These droplets are introduced into the instrument, turned into a cloud and we then filter out the biological atoms and leave only the metal tags. So, it works at single-cell level because we are separating each cell in a droplet and analysing each droplet separately.

For imaging, we are ablating one-micron pieces of tissue from the slide at a time to create the cloud. That cloud is plugged into the instrument, and we are analysing one pixel (one micron by one micron) at a time, filtering out everything else apart from the tags. From there, we can localise which metals were in which positions, and then analyse the tissue in that way.

FLG: Could you give examples of typical applications of mass cytometry and imaging mass cytometry that your unit handles?

Michalina: Our institute has two units, the Translational Immune Discovery Unit and the Molecular Haematology Unit, plus some other smaller units. We are also quite close to the hospital, which means our work is a bit more translational. Most of the work is in the field of immunology, immunoncology and haematology, analysing mostly PBMCs or whole blood from projects aimed at identifying cell populations in different conditions, diseases and treatments. IMC is quite an attractive technology for highly autofluorescent tissues, so we had from groups working on projects for Human Cell Atlas and neurobiology, brain tissue is very fluorescent.

FLG: How does IMC compare to other kinds of spatial proteomic approaches for those kinds of questions? What are the advantages of IMC?

Michalina: I am obviously an advocate of mass cytometry because I've worked with it for so long. When we first started with suspension, the main advantage was that we could quantify many more markers at one time. It saved resources, it saved the sample, we did not have to do single stain controls, so it saved your cells for pure analysis. You can also analyse the cells from small samples. Obviously, there are fluorescence-based techniques for tissue analysis with tenths of markers at a time available now. However, it is much easier to optimise an IMC panel because all markers are titrated and stained at the same time. There is no cyclical staining, so we don't need to worry about tissue preservation throughout the process. There is no autofluorescence, so the background is much lower, and the data doesn't need manipulation to retract the autofluorescence. There is one more advantage, the slides can be stained and easily stored in room temperature until the data can be acquired on the instrument. This means the researcher is independent from instrument availability. Standard BioTools provide readymade optimised panels developed that are very robust and can be applied to clinics and clinical trials.

On top of that, all of the markers that you introduce into the panel can be conjugated, which means you have full flexibility on how you design your panel. My facility provides backbone panels to our users on a per sample basis, which works to cut the cost of the first experiment because you do not need to separately buy 25-35 vials of antibodies. Instead, you can purchase antibodies on per reaction basis, and test whether the technology is compatible with your tissue or not.

Obviously, the main advantage is the low autofluorescence, which I already mentioned. You can easily see the targets together that you would otherwise not be able to because of the autofluorescence. Background staining will be present from antibody staining. The user needs to be aware of that because if the background is antibody specific, then we cannot do much about it, but we can always swap the antibody easily.

FLG: When you're engaging with new clients and encountering new projects, where do you find that most of the troubleshooting takes place?

Michalina: We've made a few steps forward in mass cytometry from the early days. Numerous protocols were published, and panels are available, optimised, and both are much more robust. There's guidance and tips and tricks available. That doesn't mean that there will not be problems if you want to do a project. If you're doing something that isn't standard or was not done before, you will probably need some time for optimisation.

"OBVIOUSLY, THE MAIN ADVANTAGE IS THE LOW AUTOFLUORESCENCE, WHICH I ALREADY MENTIONED. YOU CAN EASILY SEE THE TARGETS TOGETHER THAT YOU WOULD OTHERWISE NOT BE ABLE TO BECAUSE OF THE AUTOFLUORESCENCE."

The specific hurdles are still dependent on the project. If we have a tissue the technology was not used on before it may need more optimisation, for example the way the tissue is preserved, or conditions for the staining. For tissue, preservation methods is important, and we aim to standardise it for all samples in a project. For antibodies, it's about choosing the clones that will work and the ones compatible with conjugation protocol.

FLG: What is your experience with the mass cytometry and spatial data analysis?

Michalina: There isn't just one solution for data analysis. For high dimensionality data, it is hard to develop just one solution. Mass cytometry has been around since 2009 and we still regularly see new tools, workflows, pipelines published that improve aspects of the analysis. With so many different tools, you need to find the ones you can use. Some of them will be commercial platforms, which may be expensive. Some of them are free, like R packages, but not everybody will be able to jump in and use them without training. Or the question might be so complex that you need a new approach and a bioinformatics support to develop it.

Spatial analysis gives an opportunity to not only analyse populations of cells and phenotype (name) them. It grants a possibility to analyse the interactions between the cells and therefore gives a real insight into tissue organisation. There are aspects that we are struggling with in IMC and in any spatial analysis – the most pronounced is cell segmentation. Analysing a few markers is not much of a problem because you are not going to need dimensional analysis. When we're using 40+ markers, you have to do a dimensional analysis, which means you have to do segmentation, and the way you do the segmentation will influence your analysis. The better the segmentation is, the better your analysis will be. Computers are not humans; they don't look at things the way we do. They can't decide what is background, what is nonspecific staining, what to ignore or which signal is real. Researcher needs to be aware of this. We need to have knowledge of the tissue and the staining of the markers. We need to really know the biology of the tissue to analyse it and tailor the analysis to get the true information. We need to aid that analysis with annotations, which is purely scientific and human input. It isn't possible to analyse data without interpretation, because it's still just data until you interpret it.

There is another component that is coming into play with this type of data, which is the statistical significance of the cell interactions. We can talk about how many samples we need to analyse, but when we're talking about tissue sections, we also need to think about numbers of cells within the sections and how to analyse the proximity of the cells to have certainty that what we see has a biological meaning.

FLG: What in spatial proteomics and IMC has caught your attention of late?

Michalina: Three things in the field stand out to me right now. The first is the combination of different technologies. Combining genomics, RNA sequencing, proteomics and, recently, metabolomics and lipidomics, is going to result in a huge leap in our understanding of biology. We have a long road ahead before we can do that, but we can see that these efforts are already coming into play.

The second is visualisation of tissue in 3D. This was anticipated with mass cytometry, with microscopy and with radiology techniques. We are currently trained to make conclusions based on a single slice out of tissue that is much larger. We're using a vast experience of pathology to aid the decision about which pieces of tissue are relevant. But these decisions and this experience were also drawn from tissue slices. Now, we're taking a step back and wanting to analyse all of the tissue and see the whole picture.

The last are the robust, replicable, standardised protocols and data analysis pipelines, which allow to move from observations to discovery research, to translational research, to actually informing clinical decisions. IMC was proven to be capable of achieving this level of robustness to inform personalised medicine decisions.

Mass Spectrometry Imaging (MSI)

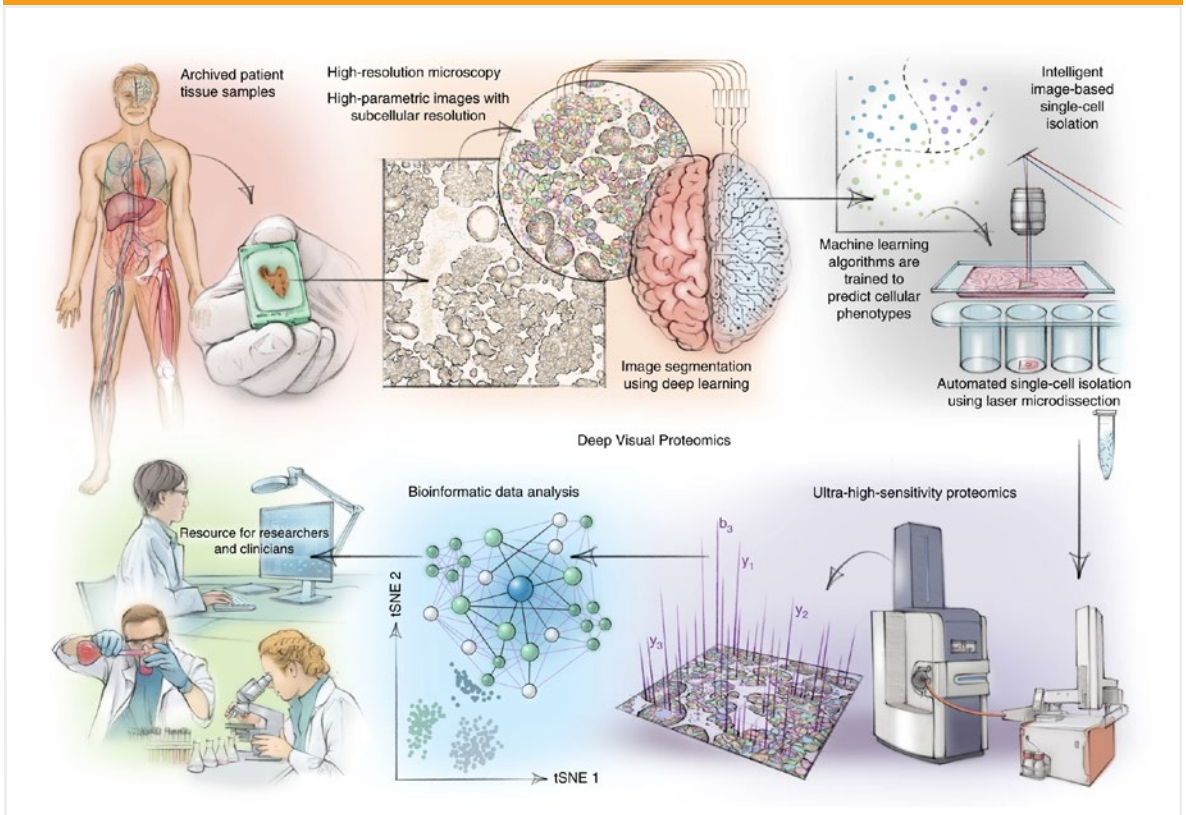
MSI is a hugely valuable approach to proteomics[4]. It is often confused with IMC, but it does not require heavy-isotope derived antibody labelling. Similar to IMC, laser ablation is used to ionise individual pixels of a sample. Every pixel hence has a label-free spectrum, meaning unbiased deeper coverage of molecules compared to the targeted IMC approach. Furthermore, the availability of thoroughly-validated antibodies hampers all other approaches but not MSI[4].

Most successful mass spec approaches use laser microdissection for tissue microsampling. However, methods such as LESA-MS and MALDI-TOF tends to result in a low-efficiency of protein detection. Instead, MS-based bottom-up proteomics, where digested peptides were separated with liquid chromatography (LC), fragmented, and detected by MS, have become the gold standard for protein studies. With state-of-the-art LC-MS instrumentation, nearly the entire human proteome can be detected from cell and tissue specimens[5].

Single-cell or subcellular resolution for MSI has proved challenging and the technique is mostly used for macroscopic imaging. Most successful MSI spatial proteomic approaches utilise hybrid MS set ups to achieve this, such as MALDI with orbitrap analysers. One example is Deep Visual Proteomics[6], which combines artificial-intelligence-driven image analysis of cellular phenotypes with automated single-cell or single-nucleus laser microdissection and ultra-high-sensitivity mass spectrometry. This workflow is visualised in Figure 4.2. This approach still fails to achieve subcellular resolution, hovering around 20 µm. Another example is DUV-LA-nanoPOTS, coupling nanoPOTS[7] methodology with deep ultraviolet laser ablation, which has profiled over 1000 proteins at <10 µm resolution[8].

FIGURE 4.2. OVERVIEW OF THE DEEP VISUAL PROTEOMIC (DVP) WORKFLOW.

DVP combines high-resolution imaging, AI-guided image analysis for single-cell classification and isolation with an ultra-sensitive proteomics workflow. Source: Mund, et al. [6]



While there are limited plug-in-and-play commercial options for such a rapidly developing technology, we have highlighted two technology options below that use mass spectrometry unbiased proteomics.

MALDI GUIDED SPATIALOMX® - BRUKER

MALDI-guided SpatialOMx® is a technique using the timsTOF Mass Spectrometer from Bruker. In SpatialOMx®, a section of tissue is segmented into sub-regions according to similarities in molecular fingerprints measured by MALDI Imaging. Sub-regions of interest are targeted for microextraction by Laser Capture Microdissection (LCM) and 4D-Omics™ analysis. The timsTOF fleX (pictured) is the ideal platform for performing spatially guided 4D-Proteomics™ SpatialOMx®, as it combines the 4D-Omics and MALDI Imaging on a single instrument.

Chemistry	Microdissection into LC-MS
Dimensions:	980 mm x 1400 mm x 2570mm
Internal Storage:	N/A
Tissue compatibility:	Fresh/Frozen, FFPE
Preferred Tissue Thickness:	5 µm - 10 µm
Reagents/Extras:	IntelliSlides and fleXmatrix supplied by Bruker. An LCM system is also needed
Scan Area per Slide:	Variable – Areas of 10 µm^2 per measurement
Slides per Run:	1
Time per Run:	Not Specified
Number of Targets per Run:	1000's – Mass spectrometry based
Resolution:	Cellular (<10 µm)
Multi-omics compatibility:	Yes - Mass spectrometry based

MICROSCOOP™ – SYNCELL

The Microscoop™ platform presents a novel spatial proteomics solution, relying on targeted photolabeling instead of antibodies. By first photo-labelling the sample, the Microscoop™ takes images of small FOVs and uses a real-time deep learning analysis software (Autoscoop™) to segment regions of interest based on predicted protein location. These segments are then illuminated one at a time, which triggers photo-biotinlyation with high spatial precision from the light-sensitive probes. This process occurs for thousands of FOVs. Material from multiple slides are scrapped together (to increase total protein content) and undergoes protein extraction via immunoprecipitation and digestion. These photo-labelled peptides are then analysed using LC-MS and a location-specific proteome can then be constructed.

Chemistry	Microscopy-guided opto-biotinylating
Dimensions Control Unit:	440 mm x 220 mm x 470 mm
Dimensions Optical Unit:	680 mm x 460 mm x 220 mm
Internal Storage:	Instruments connected to external kit and images stored externally
Tissue compatibility:	FFPE, Fresh/Frozen (4-8 sections for a mass-spec run) or Fixed Cells (4x10^5 – 1x10^6 cells)
Preferred Tissue Thickness:	5 µm - 10 µm (FFPE) or 10 µm - 20 µm (Fresh/Frozen)
Reagents/Extras:	Non-antibody based, photo-labelling kit and protein extraction kit are both available through Syncell, with enough reagents for 3 rounds of mass spectrometry.
Scan Area per Slide:	Variable – An illumination spot is 0.1 µm^2, sufficient spots are necessary to tag enough protein for mass spectrometry.
Slides per Run:	1
Time per Run:	Highly variable, < 1 day for 1 round of photolabeling
Number of Targets per Run:	1000's – Mass spectrometry based
Resolution:	Single-cell (1 µm), Subcellular for the upcoming Synlight-Pure™ Kit (~300 nm)
Multi-omics compatibility:	Yes - Mass spectrometry based

Chapter 4 references

1. Christopher, J.A., Geladaki, A., Dawson, C.S., Vennard, O.L. & Lilley, K.S. **Subcellular Transcriptomics and Proteomics: A Comparative Methods Review.** *Molecular & Cellular Proteomics* **21**(2022).
2. Barreby, E. *et al.* **Human resident liver myeloid cells protect against metabolic stress in obesity.** *Nat Metab* **5**, 1188-1203 (2023).
3. Verhoeven, B.M. *et al.* **The immune cell atlas of human neuroblastoma.** *Cell Rep Med* **3**, 100657 (2022).
4. Taylor, M.J., Lukowski, J.K. & Anderton, C.R. **Spatially Resolved Mass Spectrometry at the Single Cell: Recent Innovations in Proteomics and Metabolomics.** *Journal of the American Society for Mass Spectrometry* **32**, 872-894 (2021).
5. Xiang, P. *et al.* **Spatial Proteomics toward Subcellular Resolution by Coupling Deep Ultraviolet Laser Ablation with Nanodroplet Sample Preparation.** *ACS Measurement Science Au* **3**, 459-468 (2023).
6. Mund, A. *et al.* **Deep Visual Proteomics defines single-cell identity and heterogeneity.** *Nature Biotechnology* **40**, 1231-1240 (2022).
7. Williams, S.M. *et al.* **Automated Coupling of Nanodroplet Sample Preparation with Liquid Chromatography–Mass Spectrometry for High-Throughput Single-Cell Proteomics.** *Analytical Chemistry* **92**, 10588-10596 (2020).
8. Piyadasa, H., Angelo, M. & Bendall, S.C. **Spatial proteomics of tumor microenvironments reveal why location matters.** *Nature Immunology* **24**, 565-566 (2023).

CHAPTER 5

FINDING FOCUS: SAMPLE PREP, IMAGING AND MICROSCOPY FOR SPATIAL BIOLOGY

THE END-TO-END SPATIAL INSTRUMENTS FOR SPATIAL ARE EXCITING, BUT SOMETIMES YOU WANT THE FLEXIBILITY TO PREPARE, STAIN AND IMAGE YOUR SLIDES THE WAY YOU WANT. THIS CHAPTER LOOKS AT THE DIFFERENT OPTIONS FOR HISTOLOGY, IMAGING AND MICROSCOPY IN SPATIAL WORKFLOWS.

Spatial Biology - Sample Preparation

In its current iteration, the spatial workflow visualises omics data from tissue sections prepared on slides. This typically takes the form of Formalin-Fixed Paraffin Embedded (FFPE) tissue sections, fresh-frozen or fixed-frozen tissue sections mounted on either glass slides or special slides/chips purchasable with a specific instrument.

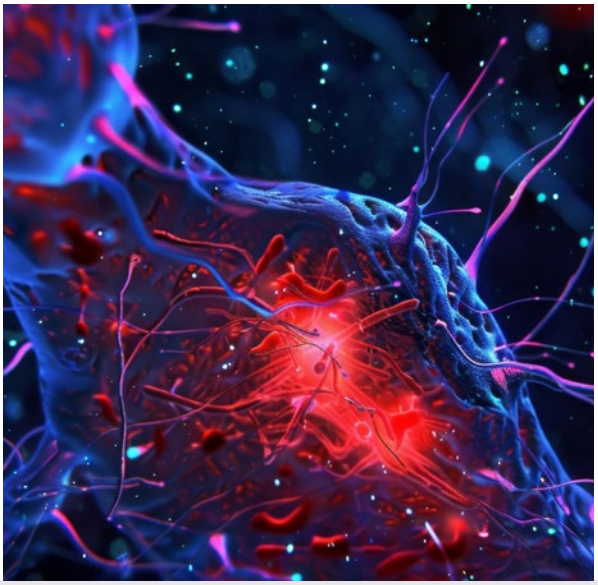
Fresh/Frozen tissue is more convenient to process. It is snap-frozen in liquid nitrogen and sectioned directly on a cryostat at cold temperatures. The nature of frozen tissue means that sections smaller than 10 microns thick are too inconsistent to produce, so sections tend to be 10 – 20 microns in size (which equates to 1-2 cell layers deep).

FFPE tissue sections are prepared through dehydrating and precisely placing a piece of tissue within a paraffin wax block. Preserving tissue in a wax block not only provides stability to the tissue, allowing much thinner sections (as thin as a few micrometres), but it also allows you to section the tissue at room temperature on a microtome. This method requires specific processing using systems such as the HistoCore systems from Leica Biosystems

There are services available to prepare samples for you. Many institutions have a core histology facility that will provide the service at a small cost. Some outsourcing providers (Chapter 7) will also perform this work for you if they are provided correctly preserved and stored tissue.

Spatial Biology - Tissue Staining

Once you have your tissue section on a slide or chip, many of the spatial protocols detailed in the previous chapters then require multiple rounds of slide staining to apply the antigen binding agents and the fluorescent markers. For a manual workflow, this can be accomplished with standard staining jars, or with the specifically designed StainTray™ or EasyDip™ - Simport™. However, there are also automated systems, which not only save time, but reduce errors caused between batches of staining when performed manually.

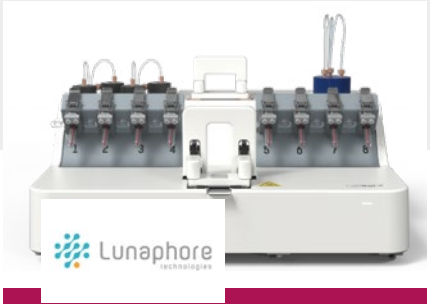


A large number of the instruments considered in Chapter 3 and 4 provide automatic slide staining for a small number of slides within the instrument. However, some workflows do not have an instrument for this. Furthermore, there may be a desire to work at a much higher slide throughput than is possible on a typical spatial instrument. Below, we look at several of the leading slide staining systems currently available in the market.

THE BOND FAMILY – LEICA BIOSYSTEMS
The BOND family of instruments are designed for large-scale high-quality slide staining. The BOND-PRIME (pictured) is the latest system in this family, able to load 72 slides and with space for 70 different reagents. A system such as this dramatically decreases the time for routine staining and specialist staining alike.

ONCORE PRO – CELLPATH
The Oncore Pro allows fully automated IHC/ISH staining with a high-throughput capacity for 36 slides. Independent protocols can be performed on each slide within the processor. Reagents are dispensed onto slides through a port attached to each slide chamber along with gentle agitation.

NANOVIP® - OMICSVEU
The NanoVIP® from Omicsveu is unique in being an automated slide staining system built specifically with spatial multi-omics in mind. The system includes the EZ-AR™ Elegance Line for an all-in-one de-wax rehydration and universal retrieval solution for all proteins and nucleic acids. It has capacity for 10 slides and 24 reagent vials.



LABSAT® - LUNAPHORE
The LabSat® from Lunaphore is an automated single-slide stainer with a consumable microfluidic staining chip that forms a closed chamber over the tissue sections where the staining takes place. It is designed to accompany the COMET®.

TISSUE TEK GENIE® – SAKURA FINETEK
The Tissue Tek Genie System is a level up from the benchtop models covered above, with 30 completely independent staining stations and a fully automated dewaxing to counterstain process. It can process 90 standard IHC slides in an 8 hour period.

Spatial Biology - Microscopy and Image Acquisition

Spatial biology would not be spatial without an imaging system, and high level biological imaging is a topic that would warrant its own detailed report.

Broadly, one can image biological material (tissues, cells and subcellular materials) via two types of microscopy - light microscopy and electron microscopy (see Table 5.1 for an overview of the four main methods). Light microscopy does as the name suggests, passes light through a biological specimen and uses a lens system to view a magnified image. For in situ spatial methodologies, light-microscopy (specifically fluorescent microscopy) is the method of choice.

Electron microscopes use beams of electrons instead of light. The comparatively shorter wavelength of the electrons allows these instruments to produce much higher-resolution images. However, the samples need to be placed in a vacuum. These microscopes tend to be deployed for high-level subcellular structural imaging, and are not suitable for spatial omics in which 10's to 100's of markers are visualised in one sample. However, these methods can be used to accompany transcriptomics for deeper insights[1].

TABLE 5.1. DIFFERENCE BETWEEN MICROSCOPY METHODS. A COMPARISON BETWEEN LIGHT, LASER-SCANNING CONFOCAL, TRANSMISSION ELECTRON AND SCANNING ELECTRON MICROSCOPY.
Source: Study Mind

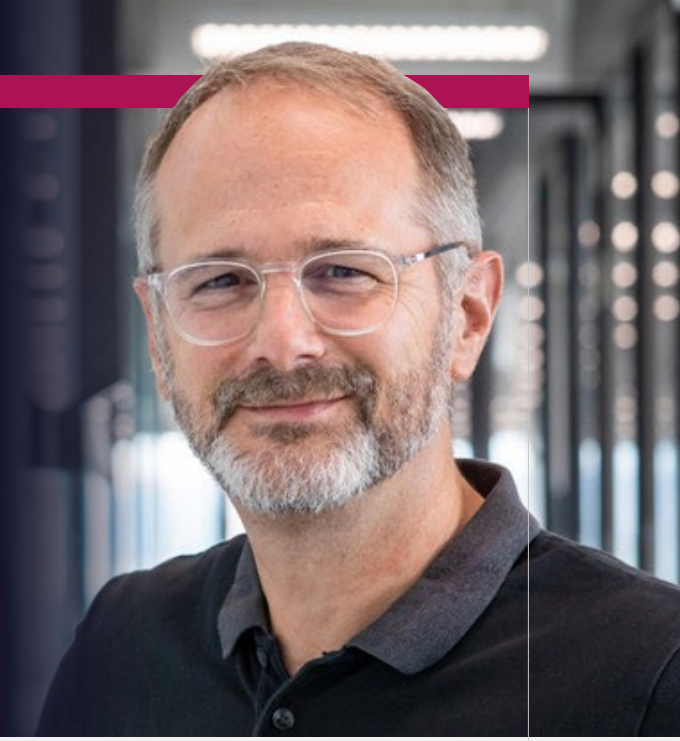
	Light	Laser-scanning Confocal	Transmission Electron	Scanning Electron
Cost	Cheap	Expensive	Expensive	Expensive
Speed	Fast	Time Consuming	Time Consuming	Time Consuming
Specimens	Alive or Dead Thin	Alive or Dead Thin or Thick	Dead Thin	Dead Thin or Thick
Image	2D - Colour	2D and 3D - Colour	2D - Black & White	3D - Black & White
Resolution	200nm	800nm	0.2nm	10nm
Magnification	1,500x	17,820x	500,000x	100,000x

Within this chapter, we have interviewed three heads/managers of imaging cores from leading institutions. We asked them about the latest trends in imaging and microscopy, examples of the projects they are currently involved with, and recommendations when it comes to high-level imaging. Below is the first of these discussions, with **Dr. Oliver Biehlmaier**, Head of the Imaging Core at the **University of Basel**, the other two can be found later in this chapter.

INTERVIEW:

OLIVER BIEHLMAIER, PHD

HEAD OF THE IMAGING CORE FACILITY AT THE BIOZENTRUM UNIVERSITY OF BASEL



***FLG:* Can you briefly introduce yourself?**

Oliver: My name is Oliver Biehlmaier, I'm Head of the Imaging Core Facility at Biozentrum at the University of Basel and have been so for 13 years now. Currently, we have 20 different types of light microscopes, which range from standard widefield to super resolution, Lightsheet, and one spatial transcriptomics system. I'm a neurobiologist originally, but then moved into the field of technology and microscopy.

***FLG:* Could you give some examples of types of projects you get involved with?**

Oliver: There's one large project, the National Competence Center for Research (NCCR) AntiResist, which is trying to fight against resistant bacteria. The scientists involved are designing organoid cultures for bladder or lung and are imaging it over a long period of time to look for what happens during the infection. Eventually, they want to check whether any antibiotics or other medications will help to prevent or clear the infections faster. These projects are usually on spinning disk microscopes and are running for days while they image the tissue and record large datasets.

One of the groups in this NCCR receives patient samples, where they then try to find out where the actual infection sites are. For these biopsies, the infections are usually very sparse, which basically means you have to scan the whole tissue. To make this doable, we developed various smart microscopy workflows. Smart microscopy means that you're not imaging the entire tissue with high resolution, but you instead take images at a smaller resolution, and together with some deep learning algorithms, you try to find out where the bacteria are. Once you identify that, then you will look at this area in much more detail at higher resolution.

We also have projects with scientists from the Zoology department that are looking at the development of different animals, their wings, paws etc. as well as neurobiologists in-house who are imaging brains. These are all large, fixed tissues and they have to be cleared in order to image them. You can then image the entire tissue with Lightsheet microscopy; you can reconstruct it and you can draw your conclusions from that. We get all different kinds of samples and questions, from imaging the inside of the cell to study transport mechanisms, to really large samples imaged under the Lightsheet.

***FLG:* Could you broadly cover how bioimaging has advanced in the last decade?**

Oliver: I've been in the field of microscopy for almost 25 years. If you take it from the beginning, in the last millennium, we were only imaging fixed samples and we usually would only look at one layer, or a very small volume. With the broader availability of confocal microscopes at the beginning of this millennium and the leap in terms of image processing, there was a big jump in development and possibilities. Over the last 10 years, for the optical part of the microscope, there were some key inventions, including super resolution microscopy, which won the Nobel Prize in 2014. We've also seen all kinds of Lightsheet microscopes being built, which then made it possible to image huge samples. Also, the development of GFP in the early 2000s was a game changer, because then on researchers could look at whatever is developing, they can mark it, and they can watch nerves grow or look at mitosis happening live in cells, which was not possible before.

The largest change in the last 10 years has been all the developments in image processing. That comes along with the computational power that we have now, which made it possible to handle huge datasets. It means that today we're handling terabyte-sized files, which five to ten years ago, would have sent you hiding under your desk. Obviously, AI and deep learning is now enabling us to do tasks such as cell segmentation in huge tissues very short time. As an example, in the last century, if you had a brain section and you wanted to count the cells, you would have been there with this fantastic mechanical devices to count them. Now, you just run a deep learning algorithm over it that you have trained, and then you can analyse thousands of brains and get more or less the exact number of cells that you have per section. This is a huge development, and that has effects on all different types of microscopy.

"WHEN I WAS ABOUT TO FINISH MY PHD, AROUND 2000, THE HUMAN GENOME WAS ABOUT TO BE SEQUENCED. NOW WE CAN ACTUALLY SEE IN THE CELL WHERE SINGLE GENES ARE TRANSCRIBED."

The latest addition is all of the spatial omics methods, either with transcripts or proteins. The image analysis part is hugely important since the actual imaging part is relatively simple (in terms of optics/microscopy). The preparation of the samples is also quite a challenge. When I was about to finish my PhD, around 2000, the human genome was about to be sequenced. Now we can actually see in the cell where single genes are transcribed. This, from the biological point of view, is absolutely fantastic, and I think this development is going to continue.

***FLG:** Is there a microscopy technique that you wish more people knew about?*

Oliver: The thing for the last 10 years that we've been dreaming of is this smart microscopy, where you can tackle huge experiments that would not have been possible previously, because you would create such an insane amount of data. If your microscope was smarter, and knew what to image, you could reduce the amount of data and do that analysis.

An example I already gave you was this spatial example. You have a huge piece of tissue, and you can use the smart method to figure out which areas of that tissue you want to look at in greater detail. Instead of a five terabyte scan, you would get half a terabyte, because you only have the – let's say - 12 different regions where the event/effect actually is.

Also, finding a good format to compress data to get rid of the black parts of the images in a smart way would be great. Our largest enemy is the size of the data, in all imaging technologies. We surprise any data scientist in a bad way with that, because they were only used to getting this huge amount of data from astrophysicists but not from biologists. It's really insane, and it's an issue that hasn't been solved.

Fluorescent Microscopy

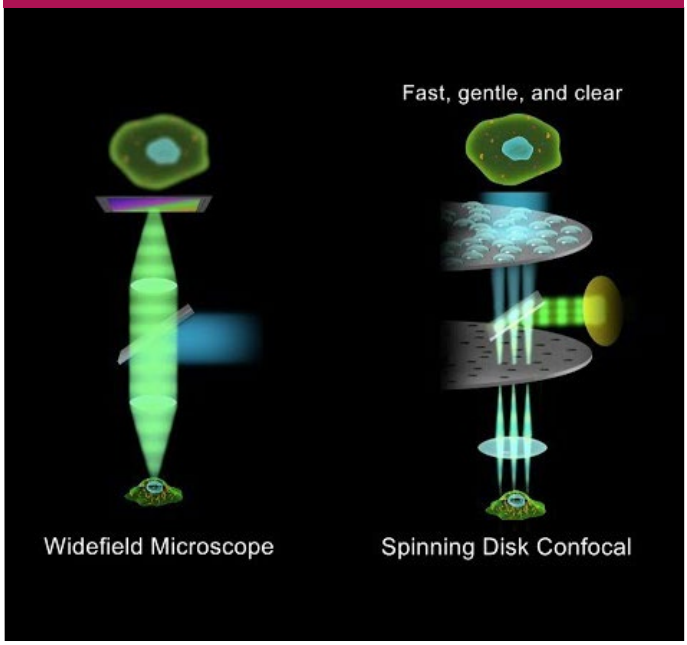
As stated above, fluorescent microscopy is the method of choice for spatial, with molecular molecules tagged with fluorescent markers. When imaging fluorescent probes (as is standard for in situ spatial methods), there are several microscopy options available:

- **Widefield Microscopy.** In widefield. a probed specimen is bathed in light to excite the fluorophores and the fluorescence emitted by the probed specimen is focused on the detector by the same objective that is used for the excitation light. This method is fast and effective but is subject to high background fluorescence and image blur.

- **Point Scanning Confocal Microscopy.** Confocal microscopes use a pinhole to allow only light from the plane of focus to reach the detector. This reduces the acquisition of out-of-focus light, thereby improving image quality and signal-to-noise ratio. It can also be used to construct 3D images. However, this method is time-consuming, complicated and difficult to use for large-scale projects.

- **Spinning Disk Confocal Microscopy.** This overcomes the slow nature of standard confocal microscopy by exploiting the multiplex principle. Rather than a single pinhole, SDCM has hundreds of pinholes arranged in spiral on an opaque disk, which rotates at high speeds. These pinholes scan across the sample in rows as the disk spins. This process massively increases the speed of imaging, allowing for long duration live imaging (see Figure 5.1).

FIGURE 5.1. DIFFERENCE BETWEEN THE OPTICS FOR THE WIDEFIELD MICROSCOPE AND THE SPINNING DISK CONFOCAL.
Source: Screenshot from Evident



- **Lightsheet Microscopy.** Lightsheet fluorescence microscopy functions as a non-destructive microtome and microscope that uses a plane of light to optically section and view tissues with subcellular resolution. This type of microscope works with tissues that have been cleared (a method to make biological samples transparent)[2]. Large volumes of tissue can be explored at micrometer resolution and this method has low photobleaching for fluorescent samples, meaning follow up sectioning and fluorescent imaging or live imaging are easily possible.

Our second interview with an imaging core manager was with **Dr. Debbie Wilkinson** from the Histology and Imaging Core at the **University of Aberdeen**. We covered her experiences with the systems in her facility and her recommendations for microscopy workflows in general.

INTERVIEW:

DEBBIE WILKINSON

CO-MANAGER AND SENIOR MICROSCOPY APPLICATION SPECIALIST, MICROSCOPY AND HISTOLOGY CORE FACILITY
UNIVERSITY OF ABERDEEN

***FLG:* Could you give a brief overview of the typical kind of projects that the Microscopy and Histology Core Facility at Aberdeen gets involved in?**

Debbie: Within our facility we have light, fluorescence, confocal, Lightsheet and electron microscopy, as well as micro-CT and a full histology service, so we deal with a wide range of samples. We have researchers within our institute working in immunology, brain health, microbial pathogenesis, cellular dynamics and neuronal signalling, to name a few. These people are looking at cells on coverslips, tissue sections, organoids, whole organs and tissues, plus organisms such as nematodes and zebrafish. We also assist groups in other parts of the university such as Biological Sciences and Engineering, so we can be dealing with samples from them ranging from nanoparticles to honeybees, no two days are the same!

***FLG:* Can you talk briefly about your experience of how bioimaging has advanced in the last few years, what's really exciting and where it might go next?**

Debbie: There has been a lot of progress in Volume EM over the last few years, that's very exciting. We can do electron tomography in Aberdeen on our TEM, but it would be great to also do things like Array Tomography to look at a larger volume at high resolution.

I recently secured funding for a Zeiss Lightsheet 7 system, something that has been on our 'wish list' for over 5 years. It has just been delivered and installed in our facility, and will enable us to image fluorescently labelled whole large samples such as mouse brains and knee joints, something we haven't been able to do before now. I'm very excited about the work we'll do with that.

***FLG:* Have you got any immediate project plans for the Zeiss Lightsheet 7?**

Debbie: Lots! I currently have people preparing a range of different samples ahead of our first training session in a few weeks. The first batch of samples will include mouse brain, eyes and organoids. Neuroscience research has increased here over the last few years so the Lightsheet 7 will be very beneficial for that.

***FLG:* Most of our readers are typically interested in tissue-based fluorescent imaging. What imaging systems would you typically recommend for a multiplexed fluorescent slide-based study?**

Debbie: Within our facility, I'd most commonly recommend our Zeiss LSM 880 confocal microscope. It also has Airyscan, so we can acquire super-resolution if needed. It produces beautiful images.

***FLG:* Is there a microscope or microscopy technique that you wish more people knew about to use?**

Debbie: I have a soft spot for electron microscopy, having done a lot of it during my PhD. A lot of people don't consider it as an option for their research, which is a shame.

***FLG:* If you were to give some quick advice for optimal sample prep of imaging slides, what you would say?**

Debbie: Think about what you want to get out of the sample. Different microscopes may have different objectives and filters/lasers, so ensure you're using optimal fluorophores for the system you plan to use, and that the microscope is capable of imaging what you want to see.

***FLG:* There appears to be microscopes for every application that one can imagine, but what would you say is the biggest barrier to scientists getting the images they want from these systems?**

Debbie: Poor sample prep, and people insisting on using old equipment in their own labs instead of well-maintained equipment within their core facility.

***FLG:* As someone who has recently secured grant money and purchased a new system, do you have any advice to readers who are looking to invest in or upgrade their imaging equipment?**

Debbie: Think about what your users really need! To secure funding you need to be able to justify the spend to make the most of the equipment.

In general, when choosing a microscope system for spatial experiments, widefield microscopy is sufficient for a number of purposes, providing a good trade-off between time, quality, cost, speed and ease of use. It is great for initial screens and high throughput studies. However, it is best practice to use confocal microscopy for subcellular localisation, due to its superior magnification and image quality in distinguishing puncta.

There are several major suppliers that provide fluorescent microscopes across the widefield, confocal, Lightsheet range. There are also options to custom-build microscopes for specific applications. Below, we highlight some of these suppliers and specific microscope systems that have caught our attention for offering new, valuable or reliable imaging capacities for fluorescence-based spatial imaging.

LEICA MICROSYSTEMS

Leica Microsystems provides a huge variety of microscopes – upright, inverted, stereo, confocal, digital and surgical. These range from basic light microscopes through to the most advanced confocal and digital options.

A highlight from Leica Microsystems is their new Mica system. This Microhub unites widefield and confocal in one system with push-button digital simplicity. 4-colour widefield fluorescent microscopy can switch straight to confocal without having to move the sample.

The second highlight from Leica Microsystems is their STELLARIS confocal microscope platforms. This platform is a fully featured inverted confocal with super-resoluton and lifetime imaging. A resolution of <50nm can be achieved and live cell imaging is also possible on the platform.

ZEISS MICROSCOPY

Zeiss also has a huge variety of microscopes available – widefield, confocal, super-resolution, Lightsheet, stereo as well as digital imaging systems.

The first highlight from Zeiss is their latest confocal imaging system, the Zeiss LSM 980. This inverted confocal imaging system has fluorescent capability between 380 nm – 900 nm and can detect up to 36 simultaneous channels. The Airyscan 2 detector allows super-resolution imaging at much faster speeds than would normally be possible.

Next is the Zeiss Axio Imager 2, an upright widefield microscope with excellent optical performance, making it ideal for fluorescent microscopy. While not benefiting from the enhanced resolution of confocal microscopy, the optical system of this widefield microscope can take high resolution and high contrast images at low magnifications and long working distances - excellent for weaker signals.

Finally the Zeiss Lightsheet 7 is the Lightsheet multiview imaging system from Zeiss, capable of viewing subcellular resolution detail using fluorescent labels in large 3D cleared specimens.

ECHO

The microscopes from Echo are also specialised with an innovative structural design. Their hybrid microscopes works as both an upright and inverted microscope with a revolving platform that positions the microscope as either upright or inverted.

Their microscope design comes in different sizes for different applications, but the Revolution Microscope is the largest with the highest performance for life sciences research. This system allows multi-channel, multi-point images as well as live cell imaging.

The core product here is only widefield but Echo has a Spinning Disk Confocal Microscope that attaches to the Revolution. This enables the higher resolution confocal images on this innovative microscope design with the speed that spinning disk confocal microscopy permits.

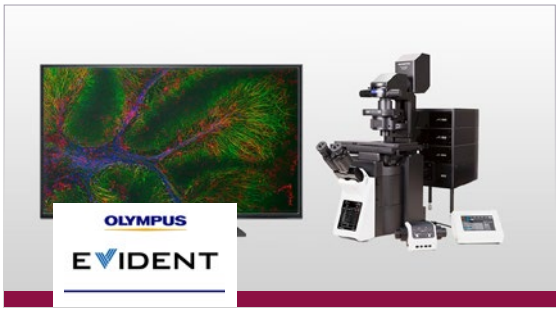
OXFORD INSTRUMENTS

Oxford Instruments are a specialised provider of flexible confocal microscope systems producing field-leading systems.

One highlight from them is the Andor BC43 Benchtop Confocal Microscope, which is a push-button compact imaging system. It can image fixed and live samples, and allows high-speed confocal imaging with a light-tight lid for fluorescent based experiments without the need of a darkroom. It has up to 4 laser lines and 4 imaging modality

The significantly larger Andor Dragonfly series is the more substantial confocal. The 600 model has >10 laser lines, 7 imaging modalities with the focus of delivery outstanding multi-dimensional images with subcellular resolution. Its highly sensitive detectors and background rejection make it well suited for high quality omics analysis.

OLYMPUS/EVIDENT



Olympus is another provider of a wide range of microscopes. Of note are the FV4000, which is the major confocal laser scanning microscope from Olympus. With FLUOVIEW™ and SilVIR™ technology to take images with a combination of low noise, high sensitivity and improved photon resolving capabilities. 6 channels can be imaged simultaneously for high-definition, multiplex fluorescent images.

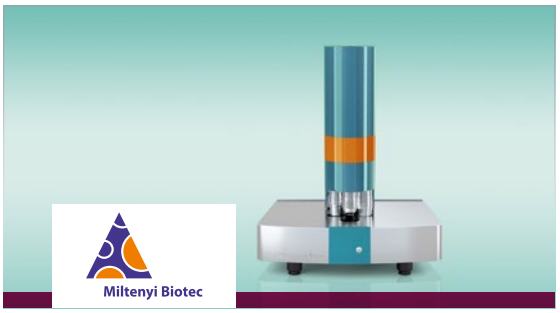
Another valuable system from Olympus is the BX63. This is an automated upright fluorescent microscope with high performance in fluorescence, brightfield and darkfield. It excels in high precision work.

NIKON

AX with NSPARC is the latest confocal microscope from Nikon, with an extraordinarily large field of view and faster imaging times. The 25mm FOV, up to 8192 x 8192 pixels, low noise for high quality imaging with high resolution at any magnification.

The N-STORM provides the highest resolution of the systems offered by Nikon. It reconstructs a super-resolution fluorescence image by combining precise localization information for individual fluorophores in complex fluorescent specimens. An ideal system for multi-colour high-resolution imaging.

MILTENYI BIOTEC



Miltenyi Biotec have recently released their new Lightsheet microscope, the UltraMicroscope Blaze™. It is a user-friendly Lightsheet microscope specifically designed for subcellular 3D imaging of large samples. The new LightSpeed Mode ensures high data quality at high speeds. A standard sample chamber accommodates multiple rodent organs/organoids, while the XXL upgrade expands capacity for human kidney or whole adult mouse models.

THERMO FISHER SCIENTIFIC

The EVOS M7000 Imaging System is the main highlight from Thermo Fisher Scientific. It is a fully automated system with monochrome and colour cameras. It is also inverted, with 5 objectives, up to 100x and four fluorescent channels. It is perfectly suited to cell imaging, taking 96-well plates as input.

LOGOS BIOSYSTEMS

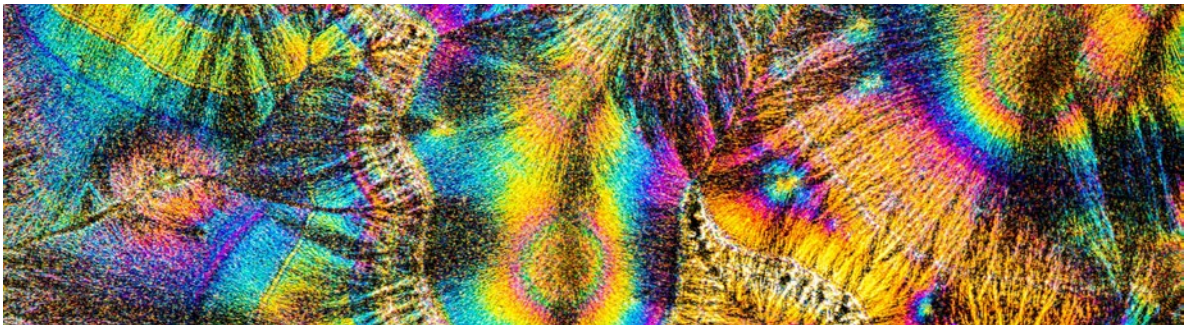
The CELENA® X High Content Imaging System is the main imaging system from Logos Biosystems. It is designed for automated high content image acquisition from both plates and slides with 4 imaging modes, including four channel fluorescence. The system is also flexible, with interchangeable objectives and filter cubes.

OXFORD NANOIMAGING

The Nanoimager from ONI is currently their only microscope. It is a unique benchtop super-resolution microscope, smaller than an A4 page, and can achieve single-molecule localisation microscopy level resolution at up to 100x objective. The lightproof lid means the microscope does not need to be stored in darkness, and the microscope is capable of performing different imaging technologies such as PALM, PAINT, SPT and smFRET.

BRUKER

The main highlight from Bruker's products is the Vutara VXL. This super-resolution microscope incorporates Bruker's single-molecule localisation microscopy technology works multi-modally to allow visualisation of DNA, RNA and proteomic markers at super-structure and subcellular resolutions. The integrated microfluidics allows multiplexed imaging for this multimodal data.



Bruker also provide Lightsheet microscopes such as the MuVi SPIM for both live-sample and cleared-sample imaging. It has best-in-class multiview imaging capabilities.

TISSUEGNOSTICS

The TissueFAXS Q system is a unique system for confocal microscopy, called a Confocal Tissue Cytometer. It is capable for whole slide confocal imaging, both at high-speed and high resolution. The spinning disk unit means a four channel 13 step z-stack of 73.3mm^2 can take 1.5 hours at 20x magnification. The standard configuration has capacity for 8 slides, with objectives up to 100x and 6 fluorescent channels.

Our final imaging core manager that we interviewed was **Michelle Ocana** from the Neurobiology Imaging Facility at **Harvard Medical School**. We discussed her recent confocal, Lightsheet and slide scanner acquisitions, and her views on the future of bioimaging.

INTERVIEW:

MICHELLE OCANA

MANAGING DIRECTOR, NEUROBIOLOGY IMAGING FACILITY

HARVARD MEDICAL SCHOOL

***FLG:* Can you introduce yourself and your imaging core to our readers?**

Michelle: I'm Michelle Ocana, and I direct the Neurobiology Imaging Core at Harvard. Although neurobiology is within the name of our imaging core, we are actually a completely open imaging core; we take in samples of any type and from anywhere. We work with research institutions as well as some biopharmaceuticals and nonprofits.

We originally started out as a small imaging core and then we received a big P30 core grant through the NIH, which we had for many years. We then just kept on growing. We started moving into wet lab work to support the imaging side of things, and now we do wet lab as well as imaging. We also look towards technology-forward imaging techniques. We do have some basic imaging systems like widefield whole slide scanning, but we also support super resolution and volumetric two photon imaging.

***FLG:* What typical workflows do you engage with in your facility?**

Michelle: Most people think that they want to do just one kind of imaging modality - "I just want to collect all of my images on a confocal". Yet, it might turn out that they have serial sections of a whole animal, or many samples, and they're looking for a needle in a haystack. We'll have a discussion about their samples, and it turns out what they want is not solely confocal imaging. That will be part of it, but what they really need is a workflow to find that needle in the haystack. It may be that they need to start out with whole slide scanning - throw 100 slides in and look at all the samples find their target - and then move to the confocal. It may mean that they need to completely change their project and do something like RNAScope. It could be that the confocal shows us that they need super resolution. We could even get what they need faster if you just cleared your tissue and use the Lightsheet. It's really varied. We really do see every kind of tissue from every animal, from insects to humans.

***FLG:* I was interested in the tissue clearing and the Lightsheet microscopy service that you offer. When are people choosing to do that kind of 3D imaging over just serial sectioning?**

Michelle: Speed is one reason. When you have a lot of volume of tissue to work through in the workflow and tissue clearing can often save a lot of your time. If you're looking for, for example, the migration of cells within an organ, or maybe you injected cancer cells and you're looking for a specific mutation and you don't know where it is. You could do hundreds of serial sections and then image them all. But that's going to take time. Maybe you have 20 mice. In this case, you're not looking for subcellular resolution, but cellular resolution or just spatially within an organ, where has the cell gone? A Lightsheet can answer that rather quickly. In a few hours you could have imaged an entire mouse brain, or even a whole mouse if you wanted to. More recently, we've moved into tissue expansion, which is definitely going to change the way we do Lightsheet in the future. So, Lightsheet will have more of the resolution of a confocal and that's pretty exciting. We have a few projects focusing on that.

***FLG:* I was going to ask you about the directions of future change, so I guess that's one of them!**

Michelle: I have to say that Lightsheet is one of those bread-and-butter methods. People have a hard time imagining imaging a whole mouse heart, lymph node, or brain all at once. It's been a little slow to catch on.

If you're a new researcher coming in and you don't know or feel comfortable using Lightsheet, you're going to be a little behind. I think you have to at least become somewhat comfortable and semi-competent at Lightsheet. Because it's the future. It's a fast way to find your cells of interest.

FLG:* *I saw that you've recently purchased the Leica Stellaris Confocal Microscope. It appears that confocal imaging seems to be improving fairly consistently. Why did you purchase that model and how has confocal improved across the years?

Michelle: I'm of the opinion that every confocal is essentially the same. It has a confocal aperture, you shoot lasers at it etc. It's the application itself that means a lot to me. Over time, confocals have improved in function. There are now white light lasers (which is what we carry), which allow us to specifically choose the excitation lines. Our systems are all spectral so we can specifically choose the emission bandpass as well. Under all that, they're all just confocal, they have a confocal aperture that blocks light and gives you an optical plane.

When it comes down to deciding which confocal microscope we want next, and the reason I chose the Stellaris, is because of the software. I really liked the white light laser too - I actually like a lot of the things about that system. In particular, I really liked the configuration of the Acoustic-Optical Beam Splitter. I like the orientation of the Photomultiplier Tubes, which are much closer to the light source. They have these hybrid detectors that are very sensitive. Comparing it to an older system of theirs, it appears as if they upgraded their scanner, so it's a little bit more sensitive at reflecting light. Ultimately, the real reason I bought that is because the software is so easy to use.

I usually spend a good amount of time training people in confocal imaging. I want to make sure that everybody leaves understanding how the confocal works, and how pixel resolution versus optical resolution is configured. You're dipping into lots of resolution ideas - there's bit-depth, pixel resolution, and optical resolution - and you have to know all of those and make good choices when you scan. Otherwise, you're going to take poor images that you won't be able to analyse well; they will just be noisy. Or you will be creating images that maybe look good, but they took you so much time that it wasn't worth it. The confocal can be a time-sink and it can suck up your soul if you set it up wrong. With this system, I can show them how to use it, which doesn't take long, then most of our conversation is just about how to build your confocal image. My training is much shorter and that is a huge win for me.

FLG:* *Does it tend to be that software is the main consideration for the kind of microscope upgrades you go for, or is that only for confocal systems?

Michelle: Ultimately the optics and the software win over everything. We have ~2000 users in our core and we're training multiple people each week, so having a software that works, doesn't crash, is intuitive, and you can look at and say, "Oh, I understand exactly what it is that I want, and this software is helping me get there"; that is really important. If you have to spend a lot of time troubleshooting software, or rebooting systems, it doesn't matter what you're collecting at the end, because you're taking too much time.

FLG:* *I saw you've also recently got hold of two Olympus VS200 slides scanners. Earlier you were talking about using slides scanning as an initial pass when you have a lot of slides in order to find your target slides. I'm assuming there must be drawbacks from a system that works with that level of automation. I know it doesn't achieve the resolution of a confocal, but are there other drawbacks to using slide scanners for primary imaging?

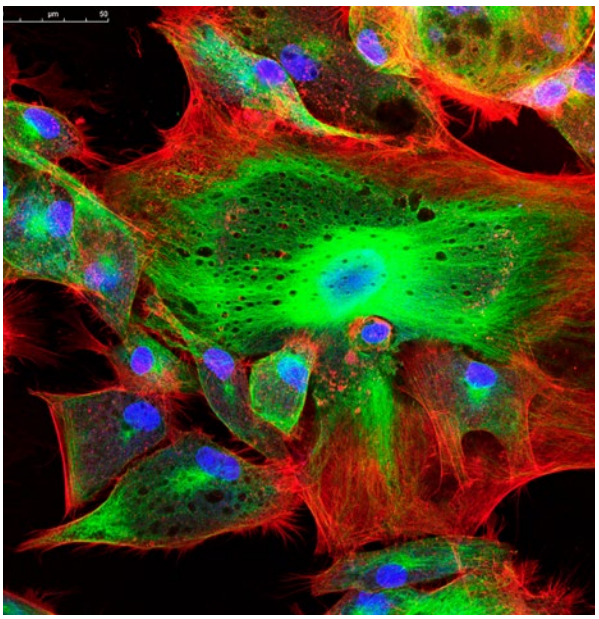
Michelle: Firstly, you're imaging in widefield not in confocal. Because it's widefield, you have fixed filters, unlike our confocals that are all spectral. You're basically stuck with the filter sets that are available to you within the slide scanner for fluorescence. Otherwise, a lot of times people are using the confocal and they don't even really need it. They might be taking five microns sections, and they want to image at 10x, they could just use the slide scanner and be done by the time I have trained them to use the confocal.

The filter combinations can be difficult and detecting weak signal can be limiting, because it's a wide field system. If you have weak signal, you'll have an increased exposure time. You can always add a lot of laser power to your sample on the confocal and still capture your probe of interest. Another limitation of the slide scanner is that it could take you hours and hours to scan. On the other hand, you don't have to be there to do it, you can just hit go and then you walk away.

FLG:* *For the uninformed, there appears to be a microscope for anything that you can imagine when it comes to imaging - variety doesn't appear to be the problem. So, what do you think is the biggest barrier for scientists to get the images they want from these systems?

Michelle: Time. These things just take a lot of time. There's more interest in using them than there are systems to capture the data. There are more users than any of us could ever support, because the time it takes to capture the images is just so large. Then, at the end, it's the question of what do you do with this data? These massive datasets that need to be registered and aligned before you can even look at what you're trying to measure. The registration is hard, everything about spatial is hard.

High Throughput Imaging



The microscopy systems considered earlier in this chapter demonstrate a wide range of impressive features and analysis options. However, they are primarily single-slide systems and are not suitable for high-throughput large-scale experiments. Some of the systems may be digital, allowing some walk-away automation, but most also require a large quantity of hands-on time in focusing and optimising the imaging setting for each slide in turn. For the high-throughput spatial experiments, another set of imaging systems may be more beneficial - the digital slide scanners.

These systems allow you to load 10's or even 100's of slides in one go, with automatic programs for focusing and imaging acquisition. This means that slides can be loaded and left for hours or even days while images are acquired. Naturally, this technique lends itself to more histological imaging, given its simplistic nature, but there are fluorescent slide scanners that capable of producing images for spatial biology experiments. The compromise is on image quality, but some spatial projects do not require the highest quality images and the time saved from using a digital slide scanner may be more essential.

There are also systems that can physically attach to the high-quality imaging systems and automatically remove and load the slides. This simulates a slide scanner and allows for applications such as automated high-throughput confocal imaging.

Some examples of digital slide scanning or slide loading systems are highlighted below.

AXIOSCAN – ZEISS

Perhaps the most widely deployed digital slide scanner is the AxioScan from Zeiss, currently acquirable as the AxioScan 7. This system can load **up to 100 slides** in a single run. Slides are robotically loaded in sets of four and the system can run 24/7 to complete the scans and can rapidly switch between brightfield and 4-channel fluorescence. This is a large-scale digital slide scanning system, not a slide loading system.

BIOPIPELINE SLIDE – NIKON

The BioPipeline SLIDE system from Nikon uses the Marzhauser Slide Express 2 loader to provide the platform for automated slide scanning. This system attaches to an existing microscope, and robotics transfer slides to and from the microscope. **Up to 120 slides** can be loaded at once.

VS200/APX100 – EVIDENT/OLYMPUS

At Evident/Olympus, there are two major research-based slide scanning options. The first is the SLIDEVIEW™ VS200, which can load up to 210 slides using robot arms to transfer slide trays between the microscope and the storage unit. Multiplex fluorescent scanning is an available mode on the scanner.

The other is the APEXVIEW APX100, a user-friendly all-in-one microscope with the same digital advantages of a slide scanner but can only load three slides. This system is a compact accessible scanner with the goal of making imaging easy.

TISSUEFAXS SL – TISSUEGNOSTICS

The TissueFAXS SL attaches to the TissueGnostics whole slide confocal microscopes and works as an automated slide loader. The SL can hold **up to 120 slides**, and the confocal microscope can image at a 100x objective. Since the SL can be attached to the confocal microscope, this uniquely enables high-throughput, whole-slide, multi-spectral confocal imaging.

PHENOIMAGER HT – AKOYA BIOSCIENCES

The PhenoImager® is a slide-scanning digital system **with an 80 slide capacity** and continuous loading technology. It is designed with speed and multispectral imaging in mind, since it has been designed by a spatial biology company. It is a fully automated system, capable of taking a 7 colour fluorescent scan in under 12 minutes for a 15mm2 region. It has up to 40x resolution and is capable of imaging up to 9 colours.

There are a selection of histology-focused whole slide scanning systems, which are focused on high throughput brightfield imaging. This includes systems such as the PreciPoint Series from Nikon, the Aperio system from Leica Biosystems, the Ventana from Roche Diagnostics, the Morpholens slide scanners from Morphe Labs, the MoticEasyScan Infinity and the novel Grundium Ocus® scanners. This last one is compact and designed to be used remotely. These microscope systems allow pathology-based stain imaging such as H&E, but do not allow the fluorescence-based spatial biology solutions.

Imaging Standards

The final consideration that we wanted to raise in this chapter is the need for high-quality imaging data and a set of rigorous imaging ethics. As more and more researchers are discovering the benefits of spatial solutions, for some it may be the first time that they find themselves needing to produce high quality images. It is important to have a set of practices that allow everyone to create images that aren't deceptive, and can be used by the wider community as a resource.

Ultimately, guidance on imaging standards will come from the community[3,4]. Below, in Figure 5.2, you can find a set of community-driven imaging standards represented in a checklist format. It is by following checklists like these that reliable spatial data will be available for the wider community from this point on.

OLIVER BIEHLMAIER, PHD

Head of the Imaging Core Facility at the Biozentrum
University of Basel

FLG: As more researchers are engaging with these spatial methods, a very clear set of standards for imaging data is going to become more important. I've seen there are community-based guidelines for imaging emerging over the last year. But what is your take on this? How can we make sure that this data analysis and images remain high quality and usable?

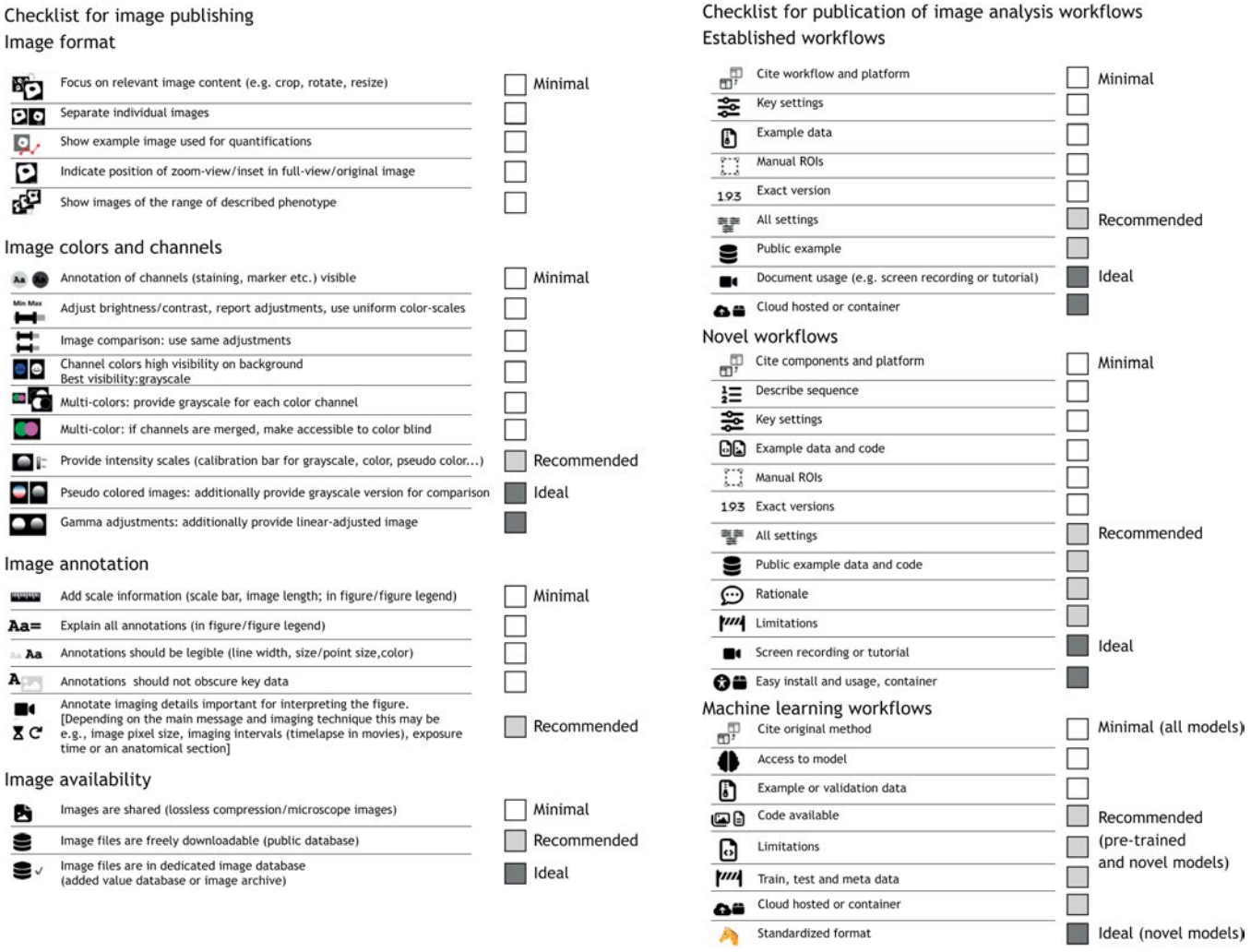
Oliver: *The problem is that you can only try to make sure that the data quality, or the imaging ethics, are in place once the data comes out. At the moment, this is still in development. With your spatial data, it might turn out that you don't actually have a specific transcript, there might be some false transcripts, there might be some normalizations or corrections that you have to do.*

We also found out that the success of the run (and the data) can depend heavily on the person who is doing the sample preparation. At the moment, for the microscopy core facility, we have one staff member who is dedicated to the MERSCOPE, and she spends probably 50% of her time on that system. That has to be clear to anybody who wants to start spatial; you'll need that level of dedication and specialism, at least at the beginning.

For the analysis, it's the same. I have several image analysis personnel in my group, and they are always involved in checking the outcome and looking for alternatives to analyse the data and verify whether what the software or the microscope produces is actually what you see. In some institutions, the sample preparation happens in the histology facility, and then the imaging at the microscopy facility. They work together with the groups mainly involved in spatial omics, who tend to be very good at the analysis. Each of these three steps has one group of specialists and this is something that should be considered here.

FIGURE 5.2. CHECKLIST FOR USING IMAGE DATA AND IMAGE ANALYSIS IN PUBLICATIONS.

Source: Jambor [3]



Chapter 5 references

1. Androvic, P. *et al.* **Spatial Transcriptomics-correlated Electron Microscopy maps transcriptional and ultrastructural responses to brain injury.** *Nature Communications* **14**, 4115 (2023).
2. Richardson, D.S. et al. **Tissue clearing.** *Nature Reviews Methods Primers* **1**, 84 (2021).
3. Jambor, H.K. **A community-driven approach to enhancing the quality and interpretability of microscopy images.** *Journal of Cell Science* **136**(2023).
4. Schmied, C. *et al.* **Community-developed checklists for publishing images and image analyses.** *Nature Methods* **21**, 170-181 (2024).

CHAPTER 6

IN ADDITION: INCORPORATING MULTI-OMICS IN SINGLE-CELL AND SPATIAL WORKFLOWS

ONE THING THAT MAY BE ON YOUR MIND IS HOW TO INCORPORATE ADDITIONAL ANALYTES INTO YOUR SINGLE-CELL AND SPATIAL ASSAY. THIS CHAPTER WILL BRIEFLY LOOK AT THE OPTIONS AND THE WAYS YOU CAN DO THIS WITH COMMERCIAL KITS.

The Method of the Moment

Multi-omics has been a consistent theme throughout this Buyer's Guide for both single-cell and spatial. It is clearly something that is on a lot of people's minds. In fact, when we asked several of our expert contributors what they were most excited about in their particular field, they highlighted multi-omics.

DAVID COOK

Scientist, **Ottawa Hospital Research Institute,** Assistant Professor, Department of Cellular and Molecular Medicine, **University of Ottawa**

FLG: As you're aware, spatial as a field is developing very rapidly. Is there anything that you think is particularly exciting right now that you're keeping your eye on?

David: *I'm still really interested and excited about the single-cell multi-omics space. Being able to do combined measurements of different modalities is huge. For the spatial side, not based on things I've seen but what I feel like is inevitable, I'm really excited about the analytical breakthroughs that are going to happen in the coming year or two. Even just helping with the segmentation issue. It's this horrible challenge in the field right now just to improve segmentation. I'm also excited about better ways to analyse spatial data. The analysis right now is inheriting a lot of the tooling from single-cell. But the data type is inherently different because there's the element of space that can come into all of these analyses.*

CAROLINA OSES SEPULVEDA

Researcher and Lab Manager

Spatial Proteomics Unit in SciLifeLab (KTH Royal Institute of Technology)

FLG: Is there anything in the spatial proteomics field that has excited you recently?

Carolina Oses: *Yes, in the image analysis, we have seen changes every day. One day, a new analysis that is really nice, easy and fast appears, but the next day, something completely different and out of the box emerges. While we are not an analysis facility, we ultimately produce the images, and we need to ensure that our images are good enough to go to the next step.*

Another important aspect is multi-omics, I think this is the holy grail for everyone. Spatial proteomics, transcriptomics, metabolomics all in one tissue sample, doesn't that sound sci-fi for you? More and more scientists want to identify many different molecules in the same tissue to resolve different levels of information.

FLG: For those sparse, important medical samples, multi-omics must be, as you put it, the holy grail.

Carolina Oses: *It's very rare to have a sample from a user that is of really good quality. We need to emphasize the importance of understanding how to prepare high-quality samples. If you provide a sample of poor quality, then you cannot expect high-quality images or data in return. In some cases, we have to inform users that their samples are not good, advising them not to waste their resources. This underscores the importance of utilizing good samples whenever possible.*

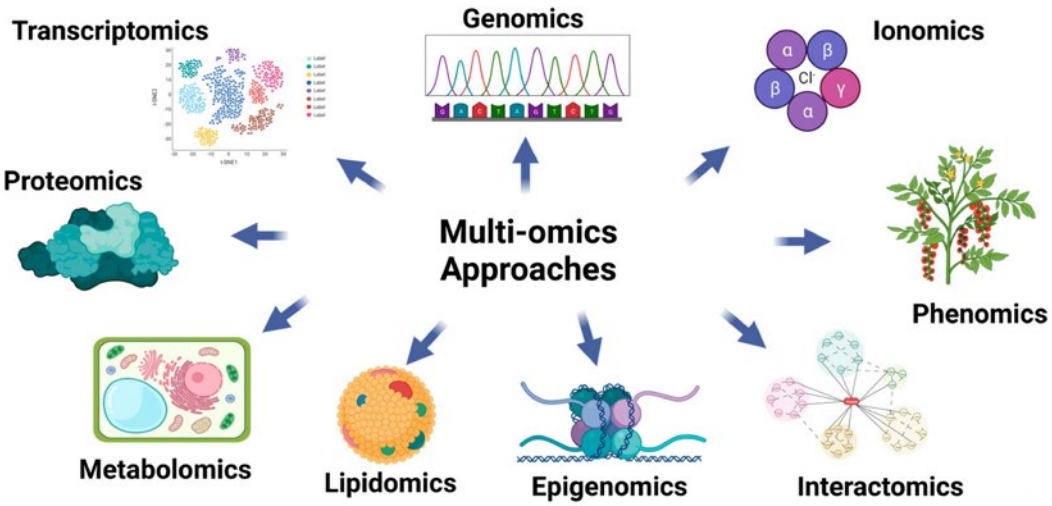
In this chapter, we will shine a brief spotlight on single-cell and spatial multi-omics, taking a moment to highlight some of the emerging commercial options to begin performing multi-omics experiments in your own lab. For an in-depth overview of the latest multi-omics methods, an array of applications and the most exciting integration methods available, please refer to the Front Line Genomics **Multi-Omics Playbook**.

Introduction to Multi-omics

If you want to understand a disease, or define a cell type, is profiling DNA or RNA enough? The answer, quite often, is no[1,2]. This is why multi-omics sequencing is becoming an increasingly valued methodology for scientists[3-5]. Furthermore, these methods are being developed at high resolution, with exciting single-cell and spatial methods released in the last 12-18 months[6-8].

Multi-omics involves collecting multiple 'omics' measurements from a single sample, or even a single cell. This multi-modal data is then integrated together using sophisticated computational methods. Now, relationships between DNA, RNA, proteins and more (see Figure 6.1) can be explored, and we can construct multimodal profiles of diseases. These are more robust than their mono-omics counterparts, and can be used to help untangle heterogeneity of disease progression and treatment response.

FIGURE 6.1. A SELECTION OF THE MULTI-OMICS APPROACHES THAT ARE CURRENTLY AVAILABLE TO RESEARCHERS.
Source: Roychowdhury, et al. [9]



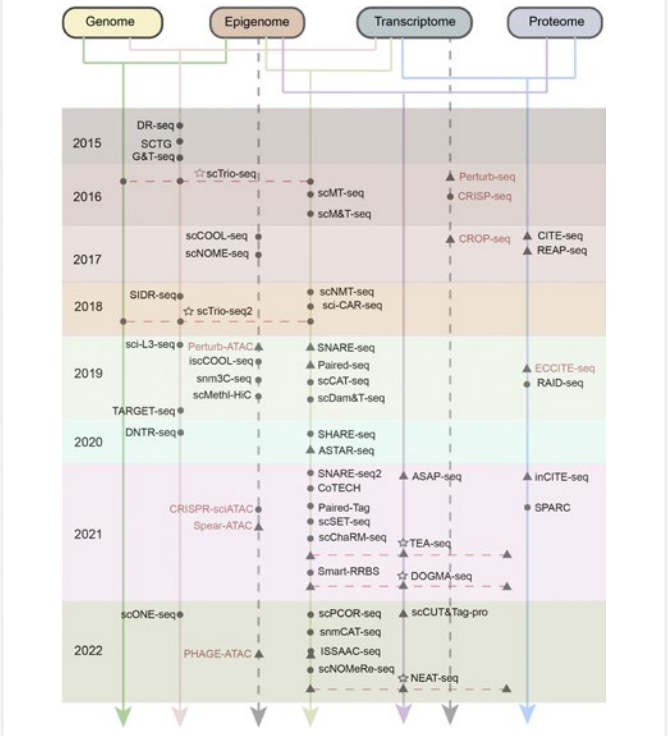
Commercial Offerings

Before we review the commercially available single-cell and spatial multi-omics methods, it is worth pointing out two things.

The first is that most single-cell or spatial mutli-omics methods that are published in papers[3-5,10-13] have not yet been made commercially available. Figure 6.2 shows most of the multi-omics methods that have been released between 2015-2022, which highlights the rate of method development. Most of these methods are only able to be performed by a small selection of groups who have learned how.

The reader may have noticed that almost all of the spatial instruments in Chapter 3 and Chapter 4 have some capacity for multi-omics. This is specifically low-plex proteomic staining on slides prior to being used for spatial transcriptomics (or vice versa). We will not be reviewing those methods here, but instead focusing on multi-omics solutions that do not require separate staining of the same section or have not been covered so far in this Buyer's Guide.

FIGURE 6.2. TIMELINE OF SINGLE-CELL MULTIMODAL METHODS.
Gray dots represent plate-based methods; Gray triangles represent microfluidic assays. The five-pointed star represents the three-omics approach, and the levels involved are connected by dotted lines. Methods denoted with a red mark indicate the incorporation of expression perturbation within the respective techniques. Source: Bi and Weng [10]



IN ADDITION: INCORPORATING MULTI-OMICS IN SINGLE-CELL AND SPATIAL WORKFLOWS

We start with the **10x Genomics** Multiome kit. This kit is perhaps the most widely used single-cell multi-omics method and allows the simultaneous profiling of gene expression and ATAC-based chromatin accessibility, using the 10x Chromium controller. This kit allows for simultaneous detection of two readouts from one cell, and 10s of 1000s of cells per sample. The protocol starts with nuclei isolation and tagmentation, then single-cell barcoding. However, Single Cell Multiome beads now do the barcoding job, and these can capture both mRNA poly(A) tails and the adapters added to accessible chromatin regions by Tn5. From here, the two produced libraries undergo separate amplification, NGS, and data analysis before the results are linked.

The Tapestri® platform from **Mission Bio** has already been covered in Chapter 2. This instrument measures DNA as its primary analyte, allowing the analysis of the genotype of a cell such as CNVs or SNVs. However, TotalSeq protein panels are included at the sample prep stage prior to library prep, resulting in a single workflow for both analytes. This is the only single-cell DNA and protein platform. The new version of the chemistry (version 3) allows up to four times more cells captured per sample, which increases the ability to detect rare cells.

Biolegend produce the TotalSeq™ kit mentioned above. This is a kit with oligo-conjugated antibodies for single-cell surface proteomics, which can be combined with other sequencing based genomics (such as the Tapestri) or transcriptomics (such as 10x Genomics or Drop-Seq) to enable CITE-seq. The protein staining happens prior to single-cell partitioning and the proteomic information can be read out of the sequenced libraries.

The ResolveDNA® kit was covered in Chapter 2, but the **BioSkryb Genomics'** ResolveOME™ kit is a different method for the near-complete survey of the genome and mRNA transcriptome at single-cell resolution. For each single cell, once isolated in a well, two different processes produce an RNA and a DNA library. The cytoplasm is lysed, and reverse transcription prepares the transcriptome. Then, nuclear lysis and genome amplification prepare the genome. Both libraries can then undergo fraction separation and library construction. With their associated BaseJumper™ data analysis software, this setup creates a unified workflow for DNA and RNA interrogation at single-cell level.

Singleron has several multi-omics kits. The ProMoSCOPE™ kit allows the simultaneous quantification of the whole transcriptome, as well as protein glycosylation, at the single-cell level. Relying on the SCOPE-chip® technology, this is the first kit to quantify protein modifications alongside transcriptomics using ProMoSCOPE-Tag technology. This workflow is designed with chemical moieties for specific recognition and covalent binding to N-Acetyllactosamine (LacNAc) on cell surface proteins. This tag includes a cell barcode and poly-A for seamless processing with the entire transcriptome. The new AccuraSCOPE® is another kit for transcriptome and genome library prep. This one works with unique barcoded beads in each well of a plate, which then captures gDNA and mRNA for the respective library preps and sequencing.

IN ADDITION: INCORPORATING MULTI-OMICS IN SINGLE-CELL AND SPATIAL WORKFLOWS

BD Biosciences BD® AbSeq Assay allows whole transcriptome and surface protein detection in single cells (CITE-seq) and is designed to be used with the BD Rhapsody™ Single-Cell Analysis Systems. BD AbSeq enables high-parameter protein expression data simultaneously with single-cell RNA-Seq data to significantly improve your understanding of individual cells and cell populations.

Bruker Cellular Analysis has a kit for transcriptomics and functional proteomics - their Duomic kit. This kit is unique since, on a single chip, you can measure the transcriptomics and in vivo proteomics from the same cell. This includes cytokines, chemokines, growth factors and phosphoproteins rather than the typical cell surface markers that other multi-omics kits profile.

Scale Biosciences have produced two epigenetic profiling kits, one for ATAC-seq and one for methylation. The Single-cell Methylation Kit works with 18,000 nuclei, in which bisulphite conversion and library construction occurs for each one. This leads to the detection of 100s of 1000s of CpG sites. The Single-Cell ATAC Pre-Indexing Kit can process a much larger 300,000 nuclei per run, enabling very high throughput chromatin accessibility studies than other alternatives. These kits do not allow measurements from the same cell, but can be used with pooled cells to make separate multi-omics measurements.

The **AtlasXOmics** Platform is the first spatial multi-omics method considered here. It is a commercialisation of the DBiT-seq technology, and services/kits are available for both spatial CUT&TAG (for spatial histone modifications) and spatial ATAC-seq (for spatial chromatin accessibility). These methods operate at cellular resolution and offer a genome-wide analysis of these epigenomic features (but can also be used for RNA, proteomic and genomic features). The methodology begins with the addition of multi-omic probes that have a known sequence to a tissue section mounted on a standard pathology slide. Sequentially and orthogonally applied microfluidic chips deliver x-barcodes and y-barcodes that ligate to the known sequence, creating an in situ x-y coordinate system in tissue. Tissue is lysed and molecules labelled with our spatial barcodes are extracted for library preparation. The final library is sequenced and analysed using standard single-cell bioinformatics pipelines.

ACD Bio-techne are slightly different from other spatial transcriptomic companies in that they are looking to combine their workflow with existing proteomic instruments/methods to create a multi-omics solution. The recent announcement of their next-generation, protease-free RNAscope spatial multi-omics workflow is the first instance of a commercial spatial workflow specifically for multi-omics, rather than adapted at a later stage. This workflow leverages ACD's RNAscope™ HiPlex RNA detection and Lunaphore's universal multiplex sequential immunofluorescence (seqIF™) technology on the COMET™ platform, which performs protein detection with standard, non-conjugated antibodies. It will be commercially available in Q2 2024.

However, while single-cell and spatial multi-omics is an incredibly exciting field, there are still issues with adopting these methods more widely. Each omic requires its own optimal sample prep conditions, making the act of balancing two protocols at once very challenging. Furthermore, there are issues with aligning barcodes back up to the same cell even when multi-omics is done successfully. Many labs still find the most success from running each assay on separate cell populations or separate tissue sections and trying to draw points of commonality.

JOHN M. ASHTON, PHD, MBA

Associate Professor, Department of Biomedical Genetics, Director, Genomics Research Center, **Wilmot Cancer Institute, University of Rochester**

FLG: For multi-omics, what methodology are you finding the most useful? There's so many methods that people have developed in-house, but not a lot of commercial options. What have you had the most success with?

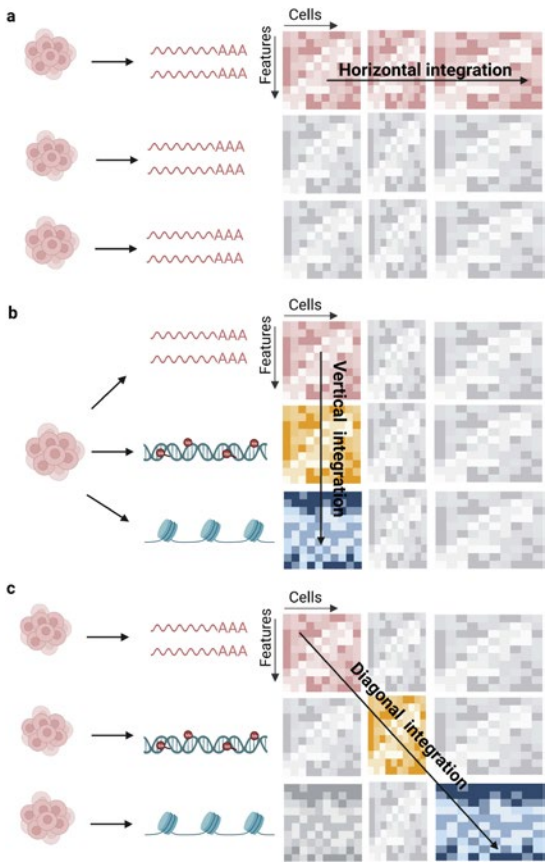
John: *Not much honestly. Most of our experience comes from groups of investigators that are working in large consortia that are using 10x Multiome. We've moved into evaluating single cell transcriptome and proteomics methods, there's a new company called the CellenONE that launched recently that claims that they can do multi-omics on a single cell. There's a lot of hope for that, but this is still very new.*

For us, what we do is we try to capture enough cells in each of those buckets and do multi-omics independently to then correlate them after the fact. I'm hopeful that the new Singular Genomics technology that will do spatial will be a transformative technology. However, the more I see the data, the more I feel like the change is likely to be incremental. I think it will work great for pharmaceutical biotech, but I don't think it's going to be very beneficial for academic medical researchers. The answer is there is no single technology to do that really well yet.

Where multi-omics is performed, among the key problems in getting usable data is the issue of integration. Multi-omics data exist at different data scales, noise ratios and 'completeness' (amounts of missing data). Furthermore, mutli-omics data can be integrated horizontally, vertically or diagonally, each of which requires a different approach (see Figure 6.3). Computational methods for each direction of integration do exist and deploy either algorithm-based or machine learning models to effectively match the omics data within or across samples/cells[14]. The latest of these methods can perform sophisticated mosaic/diagonal integration, linking omics data from the same sample and from different samples alike[15,16].

FIGURE 6.3. THREE STRATEGIES FOR SINGLE-CELL MULTI-OMICS DATA PROCESSING.

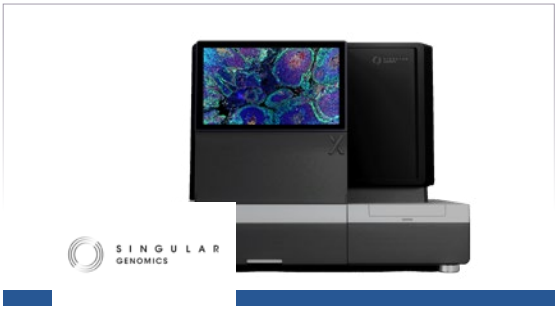
(a) Horizontal integration, which utilizes cells as anchors. (b) Vertical integration, which utilizes features as anchors. (c) Diagonal integration, in which there is no anchor. Source: Bi and Weng [10]



Sequencer Innovation

To finish this chapter, we will review one of the incredibly exciting things that is on the horizon for single-cell and spatial multi-omics, which is advancements in the sequencer space. Two companies stand out here (so far), enabling spatial multi-omic profiling directly on the flow cells of their sequencing instruments.

SINGULAR GENOMICS



Singular Genomics unveiled a new type of sequencer this year, the G4X™ Spatial Sequencer. It is a high-throughput in situ spatial sequencing platform capable of simultaneous direct RNA sequencing (Direct-Seq™), targeted transcriptomics, proteomics and fluorescent H&E from formalin-fixed, paraffin-embedded (FFPE) tissues. The G4X will share the same platform as the existing G4® sequencer and is expected to be the industry's first dual-purpose instrument offering both traditional NGS and tissue-based spatial sequencing capabilities.

With sub-micron resolution and ultra-high throughput capacity, G4X employs rapid 4-color sequencing-by-synthesis (SBS) chemistry to profile RNA transcripts and proteins in FFPE tissue. It generates fluorescent H&E (fH&E) images, producing multi-modal spatial images of 40 cm^2 of tissue across 4 flow cells in less than 24 hours. Transcripts are detected by annealing a padlock probe sequence to a target RNA and additional specificity is conferred by requiring the 3' and 5' end to be adjacent for ligation. Proteins are detected by staining with oligo conjugated antibodies, which are then targeted with padlock probes that use the ab-oligo as a splint. All ligated padlock probes are then amplified by rolling circle amplification. Demo data can be seen by following this link.

ELEMENT BIOSCIENCES

Element Biosciences have also announced a new product, their AVITI24™ instrument. This is another sequencer adapted to combine state-of-the-art sequencing alongside cyto-profiling. This means the simultaneous profiling of DNA, RNA, proteins, phosphoproteins and cell structure within single cells.

Element's new Teton™ chemistry means this multi-omics functionality is captured in one read-out in a run that is under 24 hours in length. The 2024 instrument is planned to sequence 350 RNA targets, 50 protein targets and 3 cell morphology markers at subcellular resolution on fixed cell suspensions directly on the flow cell. This will result in spatial multi-omics over a 10cm^2 area with two independent runs on the two flow cells. Future plans aim to combine an untargeted transcriptome readout with DNA sequencing, up to 20 morphology markers and custom protein panels as well as additional sample type capacity.

Chapter 6 references

1. Babu, M. & Snyder, M. **Multi-Omics Profiling for Health.** *Molecular & Cellular Proteomics* **22**, 100561 (2023).
2. Battle, A. et al. **Impact of regulatory variation from RNA to protein.** Science **347**, 664-667 (2015).
3. Vandereyken, K., Sifrim, A., Thienpont, B. & Voet, T. **Methods and applications for single-cell and spatial multi-omics.** *Nature Reviews Genetics,* 1-22 (2023).
4. Baysoy, A., Bai, Z., Satija, R. & Fan, R. **The technological landscape and applications of single-cell multi-omics.** *Nature Reviews Molecular Cell Biology,* 1-19 (2023).
5. Li, X. **Harnessing the potential of spatial multiomics: a timely opportunity.** *Signal Transduction and Targeted Therapy* **8**, 234 (2023).
6. Zhang, D. *et al.* **Spatial epigenome–transcriptome co-profiling of mammalian tissues.** *Nature* **616**, 113-122 (2023).
7. Liu, Y. *et al.* **High-Spatial-Resolution Multi-Omics Sequencing via Deterministic Barcoding in Tissue.** *Cell* **183**, 1665-1681.e18 (2020).
8. Liu, Y. *et al.* **High-plex protein and whole transcriptome co-mapping at cellular resolution with spatial CITE-seq.** *Nature Biotechnology* (2023).
9. Roychowdhury, R. *et al.* **Multi-Omics Pipeline and Omics-Integration Approach to Decipher Plant's Abiotic Stress Tolerance Responses.** *Genes* **14**, 1281 (2023).
10. Bi, H. & Weng, X. **Single-Cell Epigenomics and Proteomics Methods Integrated in Multiomics.** *Fundamental Research* (2024).
11. Wang, S. *et al.* **The Evolution of Single-Cell RNA Sequencing Technology and Application: Progress and Perspectives.** *Int J Mol Sci* **24**(2023).
12. Flynn, E., Almonte-Loya, A. & Fragiadakis, G.K. **Single-Cell Multiomics.** *Annual Review of Biomedical Data Science* **6**, 313-337 (2023).
13. Zou, Y., Zhao, Z. & Song, Y. **An overview of multiomics: a powerful tool applied in cancer molecular subtyping for cancer therapy.** *Malignancy Spectrum* (2023).
14. Argelaguet, R., Cuomo, A.S.E., Stegle, O. & Marioni, J.C. **Computational principles and challenges in single-cell data integration.** *Nature Biotechnology* **39**, 1202-1215 (2021).
15. Hao, Y. *et al.* **Dictionary learning for integrative, multimodal and scalable single-cell analysis.** *Nature Biotechnology* (2023).
16. Ghazanfar, S., Guibentif, C. & Marioni, J.C. **Stabilized mosaic single-cell data integration using unshared features.** *Nature Biotechnology* (2023).

HEAR FROM THE EXPERTS PART 3
THE RESOLUTION REVOLUTION: GOING BEYOND SINGLE-CELL ANALYSIS AS YOU KNOW IT

THE CONTENT USED HERE IS A SHORTENED, EDITED TRANSCRIPT FROM A SESSION AT THE FESTIVAL OF GENOMICS BOSTON IN OCTOBER 2023.

Mandovi Chatterjee
Director, Single-cell Core
Harvard Medical School

Josh Fienman
Scientist, Genomics (NGS Technology Center)
Pfizer

Linda Orzolek
At time of panel: Director, Single Cell & Transcriptomics Core, **John Hopkins University**
Now: Single cell & Transcriptomics Senior Product Manager, **Psomagen**

Devjanee Swain Lenz
Director, Sequencing and Genome Technologies
Duke Center for Genomic and Computational Biology

Josh Fienman: I'm Josh Fienman. I work in the Systems Immunology group at Pfizer. We use 10x solutions – 3', 5', Flex – and we also have been using Parse's solutions. And we also do a little bit of bulk RNA-seq and a bit of spatial transcriptomics as well.

Devjanee Swain Lenz: I am Devjanee Swain Lenz. I am Faculty at Duke University in the Department of Molecular Genetics and Microbiology. And I'm also the Director of Sequencing and Genomic Technologies at Duke. We work with Duke researchers and other academic researchers for a wide variety of things. For single-cell, we work with various cores throughout the University who have the fancy equipment, and then we do more of the plate-based methods and Tapestri DNA sequencing ourselves.

Mandovi Chatterjee: I'm Mandovi Chatterjee, I'm the Director of the Single-cell core at Harvard Medical School. We offer our services to all the academic labs and industry labs in the Greater Boston area and from other parts of the country and overseas as well. We support single-cell applications across many different technologies - that includes 10x Genomics, Parse Biosciences, BD Rhapsody, Fluent Biosciences. And we're also expanding in the spatial transcriptomics area.

Linda Orzolek: My name is Linda Orzolek. I'm the Director of the Single-cell and Transcriptomics Core at Johns Hopkins University. So, at Hopkins, similar to Harvard, we're offering our services both internal and external. We offer all 10x services from standard 3' assays through to Xenium in situ spatial. We also support Parse Biosciences, Mission Bio, we are introducing Curio Biosciences for their spatial transcriptomics, Bioskryb and other one-cell-per-well-plate methods that are available also.

I know, personally, as a Core Director, both my favourite and most frustrating thing to do is have conversations with our clients, because it's exciting to show people what is now available that they weren't really aware of. ***So, what are some of the most crucial bits of information that you find yourself consistently giving your clients when it comes to preparing for a single-cell or a spatial assay?***

Mandovi Chatterjee: We work with many different kinds of tissues across many different organisms. And in my experience, users come with different levels of expertise. Some people need quite a bit of hand holding, and in which case, from start to beginning, they need a lot of guidance. And some people are very savvy, not in terms of just bioinformatic analysis, but also, they're quite familiar with the sample prep part of the workflow, and associated molecular biology bit. Regardless of what the expertise is, we start with learning about their project and what their experimental design is, what kind of system they're working with and what their needs are. And based on that, we sometimes suggest a suitable approach. Many times, they come decided which platform or technology they want to use, but sometimes we see that there is a better option available out there. Then we try to lay that out in front of them and ultimately, it's their choice what they want to go with.

Devjanee Swain Lenz: I find that the advice I end up giving the most is to define your question. I think a lot of people think 'I want to use this exact platform.' And based on what their actual biological question is, we can give them better advice, sometimes a lot cheaper. For instance, if you're choosing between Nanostring or 10x Genomics, depending on how deep you want to go, and how many samples you have, you're going to choose one or the other. And those are two different price points, right? So, my advice is to define your biological question and talk to your biostatisticians before you actually plan your experiment. Understand that each experiment is unique, every tissue type and every species is different and will come with its own journey.

Josh Fienman: I would echo my colleagues' impressions there that we are very, very big on getting all parties from the scientific team, from the technical team to computational biologists, all in one room together to discuss the question and make sure everyone's on the same page. And then we defer to each other's expertise. We can sometimes guide you to the right platform based on your question – it's our job to put the puzzle together of what best fits your question. But we find it very important and very helpful to have all parties together to hash things out at the very beginning, before you start anything. And also, to determine how much it's going to cost. I started saying, 'I'm not telling you no, I'm telling you how much do you want to spend.' Because these things can be very expensive, and you want to make sure it's worth the investment and that you're going to get out what you need.

Linda Orzolek: Obviously, there's a common theme here - talking to people before you start. We have had – and I'm sure you have too - way too many times where people would show up and say, 'I have my cells, let's go.' And in the end, they're going to be disappointed with the answers. They're going to be disappointed with the waste of money, because the project wasn't properly planned. So, project development and those discussions are some of the most crucial steps.

Now, Josh pointed out a good note there on the cost. We all know single-cell sequencing is going to be expensive. And I've had clients come to me and say, 'It's not fair, we're too small to be able to do this.' As a result, I've had people who have been scared away by the dollar amounts. How do you approach that conversation, and get people to look past the dollar sign and take a chance on single-cell when it's their first assay, and they realise, 'Hey, my pilot study might cost me $20,000'?

Devjanee Swain Lenz: I typically ask if they've actually done bulk RNA-seq upfront. People, I think, just jump into single-cell. If they've done that cheaper bulk RNA-seq project in the past couple of years, that gives them a little bit more confidence that they can actually design their project well. And then with that, and I'm not the biggest fan of doing two replicates, but you can always do two replicates, and then write a grant. And we do offer grant support. So, that usually is where that goes.

Josh Fienman: I think we've also had similar conversations about right-sizing your experiment to, unfortunately, the money and the resources you have at hand. But again, there's sometimes a cheaper way to do something, if that still answers your question. I think that's a common barrier, depending on a department's finances at the time, knowing that it's going to be expensive, but also making sure that you're not sacrificing your experiment just to save money, and then you end up spending a lot of money on nothing. That's the other side of that coin, I guess.

Mandovi Chatterjee: Yes, single-cell experiments are not going to be cheap. That's the bottom line. But in my opinion, the most important decision-making factor should not be the price, but the quality of the experiment. Obviously, some balance needs to be drawn; it's not possible for every single lab to have 10 biological replicates for one condition or profile one million cells per sample. So, that preliminary conversation is very important, where we can understand what your needs are, what your biological question is, whether there are less expensive options available or some trick that can be applied. For example, when working with human samples, you can pool multiple samples in one, without any hash tagging approaches, and you can bioinformatically demultiplex those samples based on SNPs. There are caveats to this approach too. This is a trick that can be applicable to certain cases. However, it's important to know about the project before we can suggest such or other tricks, which is why the initial conversation about the project in detail is crucial.

Linda Orzolek: One of the big things that we also talk about is price per data point. One of the questions we always get is, 'How many replicates do I need?' And I always laugh at that question because five or six years ago, you only needed one replicate because every cell was a replicate. And that argument existed because everything was too expensive for people to justify doing things with the actual number of biological replicates that you need. So, now that sequencing costs and library prep costs are all coming down, we're in a position where we can generate a lot more data for less cost than we did five years ago, and we can justify the need for biological replicates now. But if you look at it and you do stick with that kind of assessment that 'every cell is a replicate', every cell is an entity that you're sequencing and looking at a transcriptional profile for. So, every cell cluster you get is another data point.

"ONE OF THE QUESTIONS WE ALWAYS GET IS, 'HOW MANY REPLICATES DO I NEED?' AND I ALWAYS LAUGH AT THAT QUESTION BECAUSE FIVE OR SIX YEARS AGO, YOU ONLY NEEDED ONE REPLICATE BECAUSE EVERY CELL WAS A REPLICATE."

If you're looking at spatial resolution, like when we talk about in situ sequencing at subcellular levels, and you're looking at 400 transcripts across these large areas, we'll think about some of the older methods because a new method might cost you $5000-$10,000. Whereas an RNAscope might be cheaper than that, but you're getting a much smaller number. So, it's the throughput that you also have to consider, and looking at the big picture when it comes to what the cost is actually going to generate for you. Getting into spatial, I know that, for us, it's been about a year and a half since we started integrating spatial technologies into our services. And we're starting to see that everything ramps up, especially in conjunction with the in situ aspects.

Are you guys seeing the shift towards spatial? Do you see that spatial is reducing the number of single-cell assays you're doing? Are people doing it in conjunction together? Or is it bringing out new researchers who are just getting into these areas?

Mandovi Chatterjee: We are offering both Visium and MERSCOPE. Visium is an NGS-based approach and MERSCOPE is an image-based approach. We actually adopted it pretty slowly. The interest is there, but I feel that people are slightly deterred by the cost of spatial transcriptomics. It's still early days in spatial transcriptomics; a lot of things are not very well understood, there are challenges and the bioinformatic solutions require a lot of improvement too. So, I think people are slightly hesitant about it.

Devjanee Swain Lenz: : Duke has Visium [in the Molecular Genomics Core Facility] and that took off really quickly. And then [my colleagues said] the people who had grant money used up their grant money on Visium. So, yes, the people who have the grant money are going spend it, and then they can't anymore, and then the people who want to use it, can't.

Josh Fienman: We're also seeing a relatively slow uptake, but for us it doesn't seem like it's cost related. It's almost like we need to market it a little bit better, because we don't get broad 'everyone wants to do Visium' type of projects. But the people who do single-cell will usually want to do spatial, and if you get people in that realm, we kind of view them as complementary, where you'll do single-cell and then spatial, and then someone who did spatial wants to come back and start doing the in situ methods. So, groups that are really, really into it seem to go all in on all three different types of readout. But we do view them as complementary. So usually, we'll try to convince someone to do Visium and single-cell up front, so you get better resolution and spatial context.

Linda Orzolek: ***Do you think that cost or bioinformatics is the most daunting aspect of single-cell or spatial analysis for your clients?***

Devjanee Swain Lenz: I'm going to say it's the bioinformatics. It's not just at Duke, it's not just in academia, it's not just in industry. It's not even just in science. There's a bioinformatics bottleneck with all of the data that we are producing. So, I'm going to go with bioinformatics on that one.

Mandovi Chatterjee: I agree. The experiments are designed by biologists, who are, most of the time, not experts in bioinformatics. Half of the workflow is wet bench work, and the rest is bioinformatic analysis. So, it's important to build a team of both biologists and bioinformaticians before you dive into the experiment. I have seen people who have generated the data, or libraries, and they've been sequenced, but the sequencing data is just lying there somewhere without getting analysed for months.

Josh Fienman: I would actually argue that we're having the opposite problem. We try to make sure we enlist the help of our very talented computational biology colleagues at the beginning. Usually, we keep them informed. So, we don't see the files just sitting in storage waiting to be analysed so much. The cost proposal, I think, is probably the bigger issue for us. But that might just be because we have a lot of very talented computational biology colleagues that are ready to tackle this kind of stuff.

Question: Do you think scalability will be a daunting task for bioinformatics going forward? Is that a concern for the users; they can generate a billion cells, but will they be able to analyse them? Is there infrastructure for that?

Linda Orzolek: If we're debating whether bioinformatics is more unapproachable, then what is the impact of generating these much larger scale studies? For us, unfortunately, we are not a bioinformatics core, we can't support it. We don't have people doing a million cells yet, because we only have a NovaSeq 6000. We're hoping that with a NovaSeq X, we could start supporting more people to do these larger scale experiments, but then, does that hold anything up on the bioinformatics side? Josh, maybe you have a better approach to that, since you seem to have more of that computational support.

Josh Fienman: I do not necessarily have the computational support in this case, but I think the tools are evolving to accommodate those datasets. I feel like a year or so ago, the infrastructure wasn't there. And you could generate the data, but you couldn't analyse it – there was that type of problem. But I think we're seeing tools that are slowly maturing to be able to do these kinds of experiments and our analysts seem to be embracing them as they come along. So, this hasn't been quite as much of an issue. To be fair, we've proposed a couple of million-cell experiments, but I don't think we've actually executed them just yet. So, it's to be determined.

Mandovi Chatterjee: I guess it's the cost factor. That is a bigger challenge than the bioinformatics challenge here, because if you barcode a billion cells, you have to sequence a billion cells as well, which can increase the cost of your experiment quite a bit.

Linda Orzolek: I think the underlying hope would be that as all of these pipelines are being developed that they will be adaptable for the obvious changes that are coming. The expansion of these projects, the time investment, the computational power that's going to be required to address them; this is all going to increase exponentially. But the hope, I think more on the bioinformatics side, is to develop things that are adaptable.

Question: When it comes to the different single-cell platforms that are out there, what are the factors that you should consider if you want to select one?

Linda Orzolek: I think a lot of it comes down to what your sample type is, what your sample availability is, whether you can collect samples at one time and process immediately or if fixation is going to be beneficial. For example, if you're doing a time course study, or you're doing patient studies where you may have one patient in the OR every couple of months. So, it really comes down to the details of your project, as we talked about before, and working out the biological question that you want to ask. And how can you go about setting that up?

So, do you need RNA? Is it just for gene expression? Or is a multi-omics approach going to be necessary for you? Do you need to look at ATAC? Do you need to look at DNA or protein? It's a very complex question. I think, realistically, you should rely on the support of core services and people who have their hands in a little bit of everything. That's what our jobs are, that's what we're here for, to talk about those details. That's my perspective of the main questions that need to be answered.

Question: Sometimes when single-cell data goes through QC, you lose a lot of cells. Maybe because of high doublet, maybe because of ambient RNA. And when using different methods, you get a lot of differences. Moving forward, is there going to be a way to improve this? For example, to get high quality data so that you don't lose a lot of cells, especially the QC level.

Mandovi Chatterjee: How good is your sample quality? Regardless of the technology you use, sample quality plays a huge role in data quality. There's a very common term, 'garbage in, garbage out.' If you have a good quality sample, then usually the data quality is very good. When the sample quality is bad, you cannot expect good quality data. A good quality sample means good single-cell suspension, with high viability, and in case of nuclei, good lysis and intact nuclei.

Linda Orzolek: The other thing to consider when you're looking at data quality is not throwing out data just because it doesn't reach some threshold. We've had projects where people have called us up in a panic, because there's such high mitochondrial levels in all of their cells, and they have filtered it out to the point that they targeted 10,000 cells, and they have 1,000 left. And as a service facility, we're trying to figure out what we did wrong, and we look back and it's muscle tissue. So, there's biological relevance for what you're seeing there. Again, that's where bioinformatics comes in - making sure that we're not setting up standard thresholds, and that the data are actually being analysed correctly, because low quality data can also be very biologically relevant.

I think, in addition to looking at the obvious sample input quality, what have you done to your samples to start with? If you're looking at immune repertoire, have you activated your T cells by some treatment that they have undergone? Are you cell-sorting, or have they been sitting out on ice for a while? Cells undergoing something that is not biological will always influence the data, and it's a matter of processing cells in a gentle and appropriate manner, so that we're not triggering transcriptional changes that are going to ultimately cause your data to be filtered out. For example, triggering apoptosis, so all of a sudden, we're starting to see, maybe not necessarily dead cells, but they're starting to die. So, the transcriptional profiles are changing. All of that ultimately comes down to sample quality, but also evaluating it in within the biological context.

"WE'VE HAD PROJECTS WHERE PEOPLE HAVE CALLED US UP IN A PANIC, BECAUSE THERE'S SUCH HIGH MITOCHONDRIAL LEVELS IN ALL OF THEIR CELLS, AND THEY HAVE FILTERED IT OUT TO THE POINT THAT THEY TARGETED 10,000 CELLS, AND THEY HAVE 1,000 LEFT."

CHAPTER 7

OUTSOURCING OUTLOOK: THE BENEFITS OF OUTSOURCING SINGLE-CELL AND SPATIAL

THIS CHAPTER PRESENTS A DIFFERENT CONSIDERATION FOR YOUR PROJECT; SHOULD YOU OUTSOURCE SOME, MOST, OR POTENTIALLY ALL OF YOUR SINGLE-CELL OR SPATIAL WORKFLOW (INCLUDING ANALYSIS)?

Introduction to Outsourcing

Single-cell or spatial outsourcing is when single-cell and spatial workflows and/or analysis are performed by an external group, department or company. But why would someone outsource their experiment rather than invest in their own set up?

As you will see below, there are several factors at play to determine whether outsourcing is genuinely the better option when it comes to single-cell and spatial. Although many individuals would embrace the opportunity to own the instruments themselves, and to have complete control over the process, it's not always the most practical or economical decision.

Advantages of Outsourcing Single-Cell and Spatial Assays.

REDUCED UPFRONT COST

Whether it is for academic research, drug discovery or diagnosis, single-cell and spatial assays comes at a cost. This is in large part due to the high price of the instruments (and the expensive reagents too!) coupled with the high volume of samples that are typically required for these assays.

While this cost is unavoidable (single-cell and spatial is expensive!), ultimately, investing in a single-cell or spatial set-up is a cost vs. use decision. Will you use the set-up enough to justify the purchase? If not, it generally is far cheaper to outsource the occasional experiment and avoid paying the investment in an instrument.

LEAVING YOUR OPTIONS OPEN

As we have seen in the preceding chapters, there are a wide variety of methods and instruments to use for single-cell sequencing and for spatial omics (transcriptomics and proteomics). Outsourcing services are in the fortunate position to be able to afford a variety of single-cell and spatial assays from different suppliers. This means that the latest novel technologies could be within your reach.

Moreover, single-cell and spatial is still evolving. New instruments are released annually with better yields, better coverage, better throughput and faster turn-around times. In such a dynamic market, you should consider whether it is worth investing in a machine if you cannot afford to keep up with this pace. Outsourcing services continually invest in upgrading their arsenal to include new assays meaning that you can take advantage of the latest technologies with every project.

LONGEVITY OF THE METHOD

One thing we will see in the last chapter is that the influx of spatial technologies has created legal complications between specific vendors. This issue has spilled over into reducing public accessibility to specific platforms and

reagents to certain regions of the world. In this fast moving and unpredictable climate, investing in a new instrument could result in you being unable to purchase reagents in the near future. The tumultuous nature of these affairs will settle. For now, it may be sensible to test new instruments through an outsourcer who can afford to take a risk on a new technology.

TO SAVE TIME

Time is a factor in the decision for two reasons. Firstly, a significant advantage of outsourcing is freeing up your time as a scientist/clinician to engage in other activities. The process of learning to perform these new methods is costly, the workflows can be time-consuming (2 days per run) and, ultimately, the workflow is unlikely to be performed better by you than by an outsourced expert. While you lose the flexibility and ultimate control over the workflow, what you gain back is time to allocate in any way you see fit, knowing that the assay is taken care of.

Secondly, it can actually take less time to outsource an assay compared to doing it within your lab/facility. These assays tends to be performed by junior members of a team, which requires training and often many failed attempts before the process can be reliably performed. And if you choose to use a core facility, there can often be inconsistency in turnaround times. This can depend on machines breaking down, staff illness and absence, and whether this is simply a queue of other samples. For reliable and generally fast turn-around times, outsourcing may be a better option.

TO ACCESS EXPERTISE

As eluded to above, it takes time to become an expert in executing single-cell and spatial assays. Cell isolation, library preparation, slide preparation and spatial workflows are the foundation of a successful experiment, and the value of in-house knowledge cannot be underestimated when it comes to performing these manual tasks. Just one serious mistake in a two-day-long single-cell or spatial workflow could result in costly errors and loss of precious, sometimes irreplaceable, samples.

For reliable results, leaning into the expertise of an outsourcing option can be less ‘painful’ than learning the process and dealing with inevitable failed runs. Furthermore, this accessible service can allow for the scaling of research or clinical sample analysis, allowing many more samples to be analysed than could be achieved in-house with the same machine.

OVERALL

Ultimately, if you plan to perform a lot of single-cell and/or spatial analysis, then investing in a machine for your lab or core may be a logical decision. However, it is likely that many people reading this report will not be performing that much analysis. Even if you were to do a lot of single-cell or spatial, the full economic cost is significantly greater than the advertised price of consumables, so you need to be certain that the space, light, power, staff, extra equipment, failed runs, maintenance contracts, repairs and IT/servers is accounted for. Outsourcing ultimately provides a cheaper, reliable alternative to access the latest single-cell and spatial options.

Outsourcing Data Analysis

Much of the above relates to the outsourcing of wet lab single-cell and spatial experiments. Another equally important part of the process is the dry lab analysis.

The data output by single-cell and spatial workflows is at such a scale that it requires advanced protocols to handle it. There are three major factors when considering whether you are set up to analyse the data:

- Do you have the expertise to choose the right pipeline for the project?
- Do you have familiarity with the right software?
- Do you have access to the necessary hardware?

While core facilities often provide bioinformatics support, this tends to be specialised and might not support the exact goals of your project. For example, single-cell and spatial analysis is rapidly developing, and core technicians may be well versed in standard normalisation and segmentation methods, but to deploy the latest and most advanced options, expert outsourced bioinformaticians with which you can consult may be the way forward.

By engaging an outsourcing service, you can receive a consultation in designing the analysis pipeline and gain access to a large selection of tools and packages for analysis. You will often get options for primary and secondary data analysis as well as downstream data analysis workflows. As can be seen in the second half of Table 9.1, this can take the form of specific types of analysis, specific applications or the use of specially designed platforms to allow end-to-end single-cell and spatial analysis.

Furthermore, by consulting a BioIT service before even starting the wet lab experiment, you can ensure that a number of aspects of your experiment are correct such as number of samples and type of samples for your intended analysis.

A final data outsourcing consideration is storage. The largest single-cell and spatial experiments produce terabytes of data per experiment, leading to the inevitable decision of whether to invest in expensive storage or delete raw data. Outsourcers can provide cloud-based storage and can format raw data into the form you desire.

Outsource or Insource?

Before going straight to outsourcing, it is valid to consider whether it would be better to insource to a different group within your own institution or go directly to an external outsourcing service. Below, we overview the advantages and disadvantages of each.

INSOURCING TO A COLLABORATOR LAB

Pros:

- Can be the cheapest option for the individual researcher
- Personal relationship with the insourcing group
- Easily transfer samples if internal or nearby

Cons:

- Can get stuck with collaborator workflows
- Limited by minimal collaborator single-cell/ spatial options
- No guarantee of expertise
- Have to balance your request with their work – could result in slow turnaround time

INSOURCING TO A CORE FACILITY IN YOUR INSTITUTION

Pros:

- Tends to be cheaper than outsourcing externally
- Easily transfer samples to internal facility
- Likely have a selection of single-cell and spatial options
- More expertise and reliability than an individual research group

Cons:

- Have to share access to the facility with others – slow turnaround times
- Can be limited by facilities' expertise, not all options are available
- Can be difficult to contact and difficult to work things out with
- Can involve multiple core facilities co-ordinating (e.g. histology, single-cell and imaging), which can be inefficient

OUTSOURCING TO AN EXTERNAL RESOURCE

Pros:

- Widest range of single-cell and spatial options and workflows
- Dedicated customer care team to support you
- Guaranteed place in the queue
- Has the fastest turnaround on average – and there is often a 'priority' or rapid turnaround option
- Highest level of expertise – maximise likelihood of success and tends to result in consistent quality
- Can operate as a 'one-stop-shop' with access to a variety of other technologies (See Table 9.1)

Cons:

- Likely to be a more expensive option
- Have to send samples externally using dedicated couriers, which is easy, but not risk-free
- While up-front agreements ensure that customers receive exactly what they are expecting/require, that may make changes to plans more of a challenge
- Hard to find everything you want in one place, meaning multiple outsourcers may be needed

Single-Cell & Spatial Outsourcing Options

Naturally, there are many outsourcing providers that you can seek the services of. To finish this chapter, we have highlighted a selection of outsourcing options in Table 9.1, along with an overview of some of their relevant services.

TABLE 9.1. EXAMPLE SINGLE-CELL AND SPATIAL OUTSOURCERS AND A SELECTION OF THE SERVICES THEY OFFER.
Where are they located? Do they offer single-cell sequencing services? Which spatial assays do they offer? Do they offer histology and microscopy? Do they offer analysis services? Do they offer multi-omics or proteomics options?

Outsource Providers	Location	Single-Cell Sequencing	Spatial Assays	Sample prep and Histopathology	Analysis Support	Proteomics / multiomics
Single Cell Discoveries	Netherlands	10x Genomics, SORT-seq, VASA-seq	Currently selecting one	Sample prep	Basic informatics and exploratory data report	No
Source Genomics	UK	10x Genomics, Parse Biosciences	10x Genomics Visium	Histopathology services, digital pathology, molecular diagnostic services	Basic informatics and downstream analyses	Proteomics (Olink)
GENEWIZ	UK, Germany, USA (MA, IN, CA, NC, WA, NJ), China, Japan	10x Genomics	Nanostring GeoMx DSP	Sample prep, storage and tissue block processing	Basic informatics	Proteomics (Olink), 10x Multiome
Novogene	China, Singapore, USA (CA), UK	10x Genomics	10x Genomics Visium	Sample prep	Basic informatics - 10x Genomics Cell Ranger and in-house pipelines	10x Multiome
Sampled	USA (NJ), UK	10x Genomics	No	Sample storage, management and prep - histology & cell isolation	Basic informatics	Proteomics (Olink)
CD Genomics	USA (NY)	10x Genomics	10x Genomics Visium, Slide-RNA-seq	Sample prep	Basic informatics – specific single-cell analysis service	Proteomics (Imaging & Mass Spec), Slide-DNA-seq, Spatial-ATAC-seq, Spatial-CUT&Tag-seq
Neo Genomics	USA (FL, CA, IL, NC, TX, AZ), UK	No	NanoString nCounter and GeoMx	Anatomical pathology, FISH, IF, flow cytometry	Basic informatics, specialised oncology analysis	Proteomics (NanoString) Akoya PhenoImager

OUTSOURCING OUTLOOK: THE BENEFITS OF OUTSOURCING SINGLE-CELL AND SPATIAL

Outsource Providers	Location	Single-Cell Sequencing	Spatial Assays	Sample prep and Histopathology	Analysis Support	Proteomics / multiomics
MedGenome	USA (CA), India, Singapore	10x Genomics	10x Genomics, SLIDE-seq, MERFISH Vizgen and Nanostring	Sample prep	Basic informatics – custom visualisations, bespoke workflows	Cite-seq, 10x Multiome
Seqmatic	USA (CA)	10x Genomics, Parse Biosciences	10x Genomics Visium Cyt Assist	Histopathology and slide scanning	Single-cell and spatial pipelines and gene expression analysis	10x Multiome
BKMGENE	China, Germany	10x Genomics	10x Genomics Visium, BMKMANU S1000	Sample prep	BKMCloud Platform dedicated cloud platform for end-to-end high-throughput sequencing data	No
Psomagen	USA (MA)	10x Genomics, Parse Biosciences, Fluent Biosciences, BioSkryb	Nanostring CosmX, 10x Genomics Visium & Xenium	Sample prep	End-to-end service - multi-omic singe-cell and spatial analysis	Olink proteomics and epigenomic sequencing methods
IGATech	Italy	10x Genomics, Parse Biosciences	Curio Seeker	Sample prep, single-cell QC and sample management	Purchasable bioinformatic support	Single-cell DNA and epigenomic methods
Singulomics	USA (NY)	10x Genomics, Deep Single Cell RNA-seq	10x Genomics Visium & Cyt Assist	Spatial slide preparation and single-cell cell prep and QC	Basic informatics	10x Multiomics and CITE-seq
3D Genomics	USA (CA)	10x Genomics – single cell and single nuclei	10x Genomics Visium	Histology workflows, confocal imaging, FACS, high content imaging	Custom analysis workflows and informatics	Multi-omics integration services
Propath	UK	No	Nanostring GeoMx, CosMx, RNAScope, BaseScope	GCP compliant histopathology, slide scanning	Image analysis	Spatial proteomics - Lunaphore COMET, Akoya PhenoCycler
sciLifeLab	Sweden	10x Genomics, Parse Biosciences, Smart-seq	smFISH, In-Situ Seq, 10x Genomics Xenium, Visium	Sample prep, advanced FISH, imaging	Dedicated bioinformatics core	Spatial proteomics – Akoya PhenoCycler and Lunaphore COMET. Spatial Mass Spec
Canopy Biosciences	USA (MO, CA), Germany	10x Genomics	NanoString GeoMx, CosMx	Histopathology, IHC, ISH services	Rosalind-supported analysis	Spatial proteomics - ChipCytometry (CellScape)
CellCarta	Canada	No	10x Genomics Visium, NanoString GeoMx	Sample prep	End-to-end data support	Spatial proteomics - CyTOF
AMK Biotech	France	No	No	Sample prep	Basic informatics	Spatial proteomics - Hyperion
Firalis Molecular Precision	France, USA (MA)	10x Genomics	Vizgen Merscope	Sample prep	Basic informatics for single cell and spatial	Proteomics (Olink)
Acela Bio	USA	No	10x Genomics Visium Cyt Assist	Sample prep and end-to-end histopathology workflows	Basic informatics	No
Macrogen	The Netherlands	10x Genomics, Singleron GEXScope	10x Visium	Sample prep	Basic informatics	Epigenomic sequencing
Crown Bioscience	USA (CA)	10x Genomics	NanoString GeoMx, RNAScope	Digital pathology, high content imaging,	Omics specific data analysis	Mass Spec proteomics

OUTSOURCING OUTLOOK: THE BENEFITS OF OUTSOURCING SINGLE-CELL AND SPATIAL

Outsource Providers	Location	Single-Cell Sequencing	Spatial Assays	Sample prep and Histopathology	Analysis Support	Proteomics / multiomics
Molecular Genomics	Singapore	10x Genomics	No	Sample prep	Basic informatics	TotalSeq – RNA & Protein sequencing, SomaScan
BioChain	USA (CA)	No	10x Genomics Visium HD and Xenium, Nanostring GeoMx, Curio Seeker	Sample prep and sample procurement, sectioning and whole slide scanning	Dedicated consulted data analysis service	Spatial RNA & Protein
Sirona Dx	USA (OR)	No	NanoString CosMx	Optimisation, slide staining, Image acquisition	Basic informatics	Standard BioTools Hyperion, Akoya PhenoCycler, Lunaphore COMET
Discovery Life Sciences	USA (CA, WA, MA, PA, AL), Germany, Ukraine, Poland, Romania, Bulgaria	10x Genomics	10x Genomics Visium Cyt Assist	Digital pathology, slide scanners, IHC, ISH	Basic informatics and Data storage	Proteomics – Olink, Mass Spec
Abiosciences	USA (CA)	10x Genomics	10x Genomics Visium	Sample prep	Basic informatics	TotalSeq – RNA & protein sequencing
Theracues	India	Custom	NanoString GeoMx	Sample prep	Basic informatics	No
K2bio	USA (TX)	No	10x Genomics Xenium	Histology Services – processing, sectioning, staining and imaging	Basic informatics	No
Flagship Biosciences	USA (CO)	No	NanoString GeoMx	Histopathology and digital pathology service	Image Analysis and diagnostics	No
Capital Technology	China	10x Genomics	10x Genomics Visium, Nanostring GeoMx	Sample prep	Basic informatics	Single-cell epigenetic and proteomic sequencing
BIOINFORMATICS OUTSOURCERS ONLY						
Ariadne.ai	Switzerland	No	No	No	Cloud-based SPATIAL omics platform with specialised segmentation, registration and analysis options	Multi-omics analysis options
Enable Medicine	USA (CA)	No	No	No	End-to-end spatial analysis platform – Immuno – phenotyping.	Akoya – PhenoCycler. Multi-omics analysis options
Scailyte	Switzerland	No	No	No	Single-cell platform – ScaiVision - harness single-cell multi-omics data with AI for biomarker discovery	Multi-omics analysis options
Sapient	USA (MA)	No	No	No	Specific analysis option: Discovery Proteomics, Metabolomics, Lipidomics, Targeted Proteomics	Multi-omics analysis options
Rosalind Bio	USA (CA)	No	No	No	End-to-end single cell analysis – optimised for 10x Genomics Chromium	ATAC-seq analysis

OUTSOURCING OUTLOOK: THE BENEFITS OF OUTSOURCING SINGLE-CELL AND SPATIAL

Outsource Providers	Location	Single-Cell Sequencing	Spatial Assays	Sample prep and Histopathology	Analysis Support	Proteomics / multiomics
Celsius Therapeutics	USA (MA)	No	No	No	SCOPE platform – targeted for single-cell in human disease analysis	N
BioTuring	USA (CA)	No	No	No	BbrowserX for single-cell data analysis and Lens for spatial data analysis	N
Partek	USA (CA)	No	No	No	Partek Flow software for single-cell and spatial	Multi-omics analysis options
[tile]DB	USA (MA)	No	No	No	TileDB Cloud – partnership with CZI	Multi-omics analysis options
Qlucore	Sweden	No	No	No	Qlucore Omics Explorer for all types of multi-omics data at single-cell level	Proteomics, epigenomics and metabolomics analysis
Panomics	USA (NY)	No	No	No	Single-cell dataset curation and analysis service	Multi-omics workflow analysis platform
AltraBio	France	No	No	No	SCAnIO – single-cell analysis platform. Cytometry and Omics specific platforms	Multi-omics analysis options
NucleiAI	USA (IL), Israel	No	No	No	Nucleai Atom platform for pathology slide analysis	N
Fios Genomics	UK	No	No	No	Standard single-cell workflows and downstream analysis	Proteomics, metabolomic and epigenomic analysis
Nygen Analytics	Sweden	No	No	No	Scarfweb dedicated cloud-based single-cell platform	N

"THE LARGEST SINGLE-CELL AND SPATIAL EXPERIMENTS PRODUCE TERABYTES OF DATA PER EXPERIMENT, LEADING TO THE INEVITABLE DECISION OF WHETHER TO INVEST IN EXPENSIVE STORAGE OR DELETE RAW DATA."

CHAPTER 8

ANALYSIS AID: WAYS TO GET THE MOST OUT OF YOUR SINGLE-CELL AND SPATIAL DATA

IF THE WET LAB WORKFLOWS ARE CHALLENGING, DEALING WITH SINGLE-CELL AND SPATIAL DATA CAN BE OVERWHELMING. THIS CHAPTER WILL REVIEW SOME COMMON DATA ANALYSIS PRACTICES AND SIGNPOST TO USEFUL PLATFORMS AND RESOURCES, BOTH COMMERCIAL AND OPEN-SOURCE, TO ORGANISE AND ANALYSE THAT DATA.

Big Data to Useful Data - The Challenge of Analysis

With multi-step, multi-day wet lab workflows, acquiring data is often thought of as the prime challenge for single-cell and spatial methods. However, a substantial bottleneck occurs within data management, analysis and visualisation.

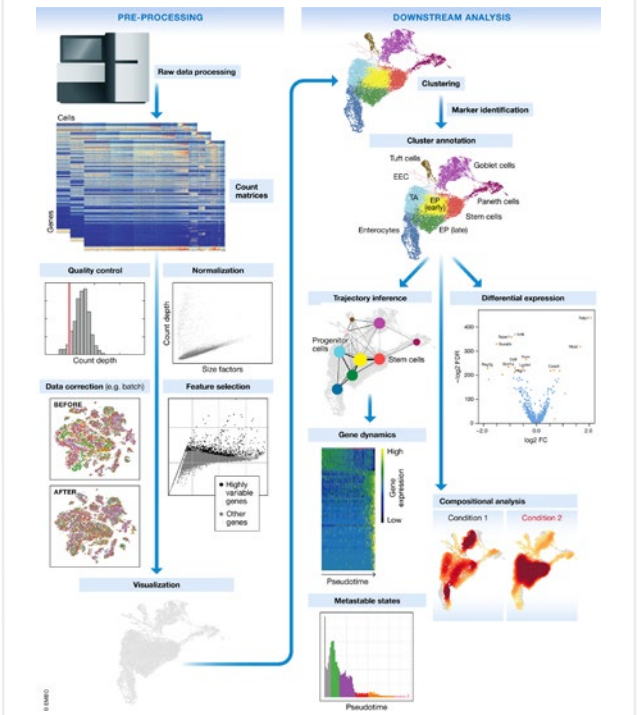
The datasets produced by single-cell and spatial workflows are large, sometimes resulting in terabyte-sized files. These expensive and time-consuming workflows can be rendered useless without the appropriate bioinformatics set-up to work with these files.

For single-cell analysis specifically, a number of challenges overlap with those of NGS - quality control, normalisation and batch correction. With over a decade of development and refinement, this has resulted in the formation of a detailed standard workflow and a set of best practices (Figure 8.1).

For imaging-based spatial methods, aligning images from cycle to cycle is a challenge, in addition to properly decoding the individual spots that represent targets and segmenting cells. NGS-based spatial methods escape these challenges as the data is already in a traditional sequencing format and is matched to a specific location via barcodes. This means it is mostly analysed like single-cell data. However, barcodes do not necessarily match up to specific single cells and so the barcodes need to be assigned biological significance (Figure 8.2).

FIGURE 8.1. TYPICAL SINGLE-CELL RNA-SEQUENCING WORKFLOW.

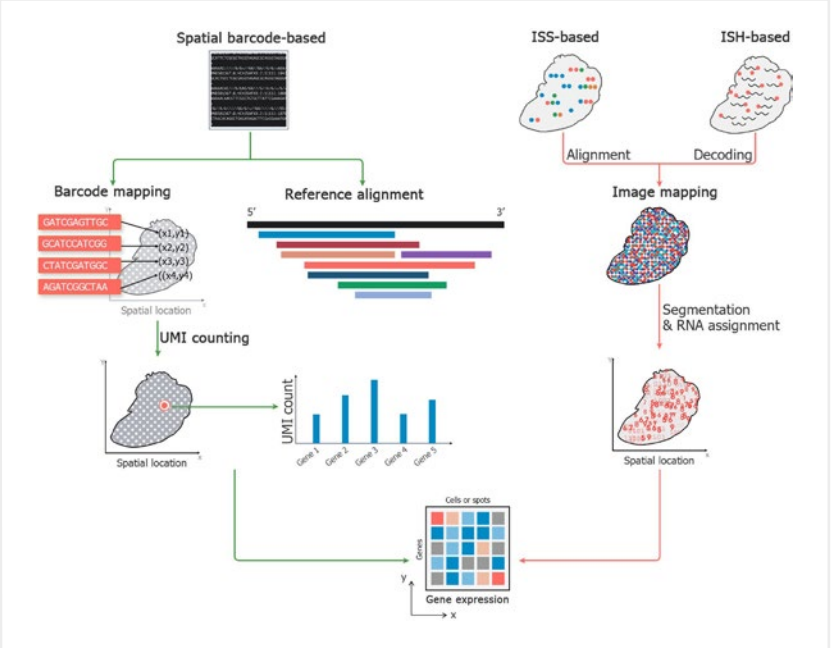
Raw sequencing data is processed and aligned to give count matrices, which represent the start of the workflow. The count data undergoes pre-processing and downstream analysis. Source: Luecken and Theis[1]



There are common challenges that all single-cell and spatial methods have to overcome. Whether sequencing single cells, mapping transcripts to barcodes, or visualising 'spots', all these data are eventually reduced to a simple feature matrix (Figure 8.2) – the number of markers (RNA, protein, etc.) per area of interest (cell, spot, barcode). From this point, there are computational challenges for how to quantify, filter, cluster and visualise this data, and how to turn the results into meaningful biological insights such as variable genes, differential expression measures, cell-cell interactions and cell neighbourhoods.

FIGURE 8.2. DATA STEPS TO CREATE GENE EXPRESSION MATRICES FROM IMAGING-BASED AND NGS-BASED SPATIAL OMICS METHODS

Source: Fang, et al. [2]



An in depth guide on how to analyse single-cell and spatial data is beyond the scope of this Buyer's Guide, but interested readers are directed to our Spatial and Single-Cell Analysis Playbook. Instead, the remaining focus of this chapter will be highlighting some important advances in the areas of single-cell and spatial analysis and highlighting useful resources for formulating a data strategy.

Latest Single-cell and Spatial Analysis Advances

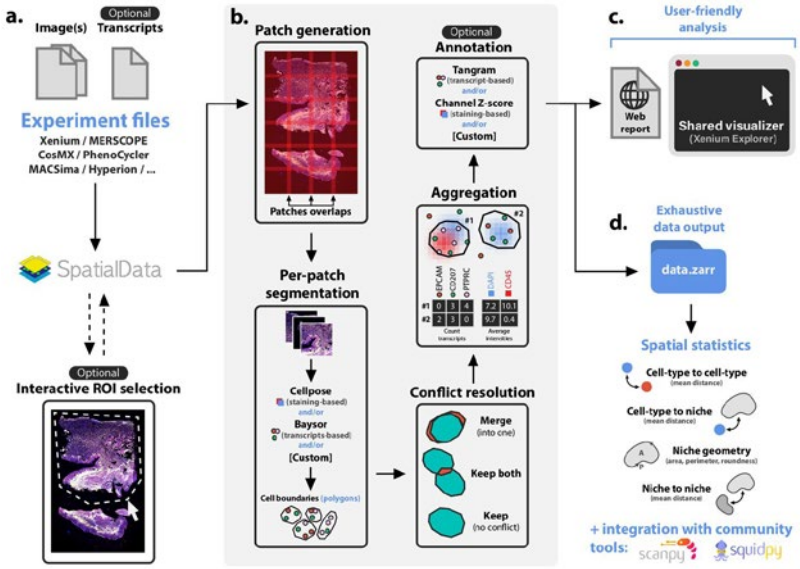
For those looking to build a single-cell or spatial analysis pipeline with reliability and longevity, the packages maintained by the scverse team are good start. These include core data structures such as anndata, mudata and spatialdata, alongside packages such as scanpy, muon, scirpy and squidpy. An additional 49 packages are maintained by the scverse community, all performing essential functions in single-cell and spatial pipelines.

Furthermore, there are concerted efforts to form a set of best practices for both single-cell and spatial data analysis. The best practices for single-cell E-book is an effort co-ordinated by members of the Theis Lab, Saez-Rodriguez Lab and other prominent single-cell computational labs[1,3]. This E-book provides community-led best practices in a number of analysis modules for single-cell workflows. The best practices for spatial transcriptomics analysis e-book is another community-led resource highlighting best practices for analysing Visium data using Bioconductor.

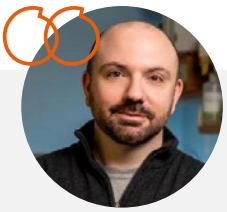
When looking at some of the latest packages, several of them directly build off the scverse core packages. For example, SnapATAC2[4] is a fast, scalable and versatile tool for single-cell data, which captures cellular heterogeneity using nonlinear dimensionality reduction and improves on previous methods by being more time efficient and for scaling with number of cells. Another example is Sopa[5], a novel spatial analysis pipeline that is technology-invariant with a unified visualizer for all image-based spatial omics. It builds off of the spatialdata package of scverse. Given the variety of spatial platforms (Chapters 4 & 5), a technology invariant pipeline is incredibly valuable. An overview of Sopa is highlighted in Figure 8.3.

FIGURE 8.3. OVERVIEW OF SOPA.

(A) The pipeline input is transformed into a SpatialData object, on which an ROI can optionally be selected. (B) Afterwards, the data is split into overlapping patches, and segmentation is run on each patch (for instance, Cellpose, Baysor, or a custom segmentation tool). The next step is aggregation, i.e., counting transcripts and averaging each channel intensity inside each cell. This allows annotation, either based on transcripts (using Tangram) or on channel intensities. (C) Afterwards, Sopa outputs a user-friendly report and files to be opened in the Xenium Explorer. (D) All data files are kept for further analysis in Sopa, such as spatial statistics, or integration with community tools. Abbreviations: ROI – region of interest. Source: Blampey, et al. [5].



Large-scale generative pretrained models could be incredibly valuable for single-cell analysis, and the second iteration of scGPT[6] has now been made available. Pretrained on 33 million cells from 441 studies, scGPT is capable of deducing cell type annotation, predicting genetic perturbation, integrating single-cell datasets and multi-omics, and creating gene regulatory networks. As this model continues to develop, it stands to become integrated into single-cell pipelines as standard practice, in order to leverage the existing knowledge from the expansive repertoire of single-cell experiments.



JOHN M. ASHTON, PHD, MBA

Associate Professor, Department of Biomedical Genetics, Director, Genomics Research Center, **Wilmot Cancer Institute, University of Rochester**

FLG: What's your perspective on the state of single-cell data analysis?

John: *I view data analysis in two steps. Firstly, there's the quality control data processing, which I think can be very standardised and fairly routine. But the second step of interpreting data and doing all the tertiary analyses is where a lot of the work happens. We often don't have enough data to be able to identify correct cell types and that is part of the issue of the databases that are out there, the data to help us make sense of single-cell data, are too sparse and sometime conflicting. Trying to understand the associations between sub Cluster A and B, even if they're different cell types, is very empirical. Some of the tools out there, such as Seurat, are very good for parts of it. However, when you're trying to do trajectory analyses, there are other tools that maybe a better fit.*

I think that part of this struggle is trying to identify batch effects from single-cell data and how do you integrate different datasets that are done on different days. The workflow isn't something that you can do 100's of samples at one time so you accrue a lot of potential batch effects. This brings up the question - how do you integrate the data? I think a lot of the field is still struggling with this. Incorporating antibody tags, and things like CITE-seq, help a lot with cell typing, but how you integrate those is also tricky. I think part of the issue is, as we get more data, we understand more what cell types are doing, and as we get more methods to be able to really interconnect those, I think the tools will become a little bit more standard.

At the end of the day, most don't use single cell to answer a very specific question, they use it for exploratory reasons. That could be exploring pathways or exploring interactions, but that's really where you need spatial context. This is ultimately why I feel the field will go in this direction.

We struggle now with even trying to apply a disease state like COVID infections on lungs to really understand what's going on. Once you deal with a pathology, everything goes out the window, and no cell marker is completely unique. Cell markers are becoming irrelevant. Cell states transition so quickly and you're able to capture that from single cell, but you don't see that in bulk. I think for all those reasons, it's still an empirically determined path and there's no one tool or suite of tools that will apply to everything.

FLG: What is the HyperGen project you've been involved with?

John: *HyperGen is a good example of how we acknowledge that not everyone has the computational skills to perform theses analyses themselves, nor do they have collaborators that that they can necessarily engage. We tried to build an in-house data analysis portal that allowed us to provide some of the analysis pipeline and workflows that we generate to the investigator, in an easy to consume way. If they want to reanalyse data, ask some questions, or explore data, they can do that with HyperGen. It helps accelerate research by getting to the pointed questions. We're planning to roll out additional methods for that. Ultimately, we're trying to train the next generation of scientists to do this themselves, but that's a long haul. We want to enable them now and reduce barriers and that's where HyperGen comes in.*

DAVID COOK

Scientist, **Ottawa Hospital Research Institute,** Assistant Professor, Department of Cellular and Molecular Medicine, **University of Ottawa**

FLG: What do you see as the major areas of development for single-cell analysis? It seems to me that the community is coalescing on a set of pipelines for how to process the data, but there still seems to be new developments happening every month. What's your perspective?

David: *I will back up a little and say, it's interesting that for data generation (sample prep and instrument output), we've not seen massive gains over the last several years. All of the platforms perform comparably in terms of cell throughput, transcripts recovered per cell, noise levels etc. Is this the best we're going to get? Or are we going to see some exponential gain in transcript recovery or sensitivity? Though the recent GEM-X development from 10x Genomics seems to have made some pretty promising improvements.*

For the analytics, there's still a lot of active development. Many of the simpler, common steps in the analysis pipeline have become fairly standardized as the field has matured for over a decade now. I think we have a good sense of the technical characteristics of the data that you need to be concerned with for things like normalization. There have been many different papers tackling different aspects of the challenge – can you do just simple library size scaling and log transformation or do you need to do more complex transformations? Can we develop models that are based on the count data alone? So, for many of the core steps in basic analysis, there seems to be less room for innovation.

Where we are still seeing a lot of developments is in how to handle complex experimental designs or population level data with a large number of samples, often spanning different experimental groups. These designs introduce challenges that weren't there with single-sample experiments, from more comprehensive modelling, batch correction, deriving sample-level representations or embeddings, etc.

This adaption of AI for general utility tools isn't just being applied to downstream analysis. Sapio Science's ELaiN is an AI assistant, based on large language machine technology, which allows scientists to ask simple language questions to facilitate their experiment. This means that while you have a pipette in hand, you can ask ELaiN to write some code for you, to review your inventory or to create a new experiment plan.

Using spatial transcriptomics to draw gene expression predictions from standard histological images (available in abundance from the clinic) is an exciting prospect for spatial analysis. iStar[7] from the lab of Professor Mingyao Li is one of the latest methods for this purpose. This method replicates the hierarchical nature that a pathologist would use to assess a histology image to predict gene expression at high resolution. As a side note, solutions like Sapio's AI scientific assistant ELaiN can be used to talk to these models, in an example of different AI models working together. ELaiN can call these models from within Sapio ELN via a natural language prompt, delivering the gene expression predictions directly into your experiment.

Effective identification of niches and tissue structures from spatial data and allowing analysis between datasets from different technologies are constant areas of progress. A new tool, MENDER[8], is a scalable tissue structure identifier, able to differentiate spatial organisation differences. CellCharter[9] is another new tool to characterise and compare cell niches across datasets from different technologies. Furthermore, tools such as SEraster[10] and STalign[11] from the Fan lab enable increased scalability to spatial analysis through aggregating cellular information into spatial pixel, improving alignment between ST datasets within and across technologies respectively. Professor Fan is a proponent of making these packages accessible and posts videos of herself live-coding using her software.

An incredibly helpful resource for those building spatial analysis pipelines is the Museum of Transcriptomics[12]. This project, organised by the Pachter Lab, collates the history and development of spatial transcriptomic methods. Usefully, they provide a database of all the computational tools developed for spatial analysis. This resource also collates experiments using spatial transcriptomic measures and their associated data. Storing and using this large-scale data is the next consideration of this chapter.

Data Storage and Management

As already discussed, single-cell and spatial methods are now so intricate and expansive that they actually create hardware problems, whilst also requiring advanced software solutions. Where do you store terabytes of data and how do you meaningfully interact with it?

Single-cell and spatial data comes in three flavours: images, sequencing data and the downstream gene-expression matrices. A particular area of advancement is tailoring existing compression algorithms to work for single-cell and spatial data, which can work with the locational information and reduce the burden of this data. This compression format needs to support high input/output seed, have a high compression ratio and high scalability[2].

While we are in that developmental process, it's important to have a reliable means by which to store and manage this large-form data. Local solutions (such as large hard-drives, solid-state drives, or high-performance computing clusters (HPCCs)) may have a high startup cost or high maintenance fees, but you tend to have more control over the data and its security. A major drawback is the cumbersome nature of storing data on-site in a physical format.

An arguably more elegant solution has been developing over recent decades – cloud storage. It is perhaps the only storage model capable of providing a widely accessible storage solution for large-scale omics experiments. Cloud computing is an online backup space that maintains a repository of your data on multiple servers across different locations. This tends to be more cost efficient and keeps your data secure and legally compliant, while remaining easily accessible. Furthermore, it is important to know what type of data storage architecture you are dealing with.

The major systems for data storage are highlighted in Table 8.1. While Data Warehouses are excellent for querying well-structured data, the variety of formats of single-cell and spatial data need structures that can support storing data in its native format. Lakes are good for this, but the data can be hard to query and access. This defeats the point of having data accessible to inform scientific progress. Lakehouse's are the latest solution to this problem, presenting a model that can store data in unstructured formats with a governance layer to still allow users to easily query the data. However, these systems need to be built with scientists in mind to truly see the benefits for single-cell and spatial experiments.

TABLE 8.1. BASIC FEATURES OF DATA WAREHOUSES, DATA LAKES AND DATA LAKEHOUSE'S.

Features	Data Warehouse	Data Lake	Data Lakehouse
Data types	Structured	Structured, semi-structured, and unstructured	Structured, semi-structured, and unstructured
Performance	High	Low to high	High
Scalability	High	High	High
Cost	High	Low	Medium to high
Complexity	Low	High	Medium
Use Cases	Business intelligence, analytics	Data science, machine learning, big data workloads	Business intelligence, analytics, data science, machine learning, big data workloads

To meet this need, scientific data cloud services are available from companies such as Sapio Sciences, Cloud Life Sciences (Google) or Microsoft Genomics. However, clouds like the Sapio Scientific Data Cloud really exemplify a storage solution that is intelligible and usable by scientists. The Data Cloud is a centralised place for all scientific data at all stages of an experiment, with analytics and visualisations, the ability to read any instrument data file and capabilities for specific scientific techniques including flow cytometry analytics and NGS.

Widely available, cloud-based, databases are becoming foundational to single-cell and spatial studies. We will next review how to make use of these resources to improve your single-cell and spatial studies.

OLIVER BIEHLMAIER, PHD

Head of the Imaging Core Facility at the Biozentrum
University of Basel

FLG: Are there unique problems for high-plex fluorescent images at the image acquisition stage, but also at the analysis stage?

Oliver: *From the imaging point, it is among the simplest microscopes that we have. It's a standard fluorescent widefield system connected with microfluidics. This is triggered when the fluids go in and out. That's pretty standard, there's not any high tech involved there. It's more on the analysis side that things are challenging, because depending on the number of transcripts that you have, which can be in the millions, you have to detect them, segment the cells, etc. Just recently, we had an example where, with our current infrastructure, if we used that workflow, we would need 160 days for the analysis of one dataset. This is the big challenge now, and we need to increase the computational power or get better software.*

FLG: What do you think is the biggest barrier to scientists getting the images that they want from these spatial systems?

Oliver: *I think the size of the data and the handling of it. The raw data from these machines is somewhere between two and four terabytes per run. This creates a problem of whether you need to keep the raw data, because there could be a different analysis you want to run in the future. A barrier is also the complexity of the technology. It's very difficult for everybody to choose a technology from what is around at the moment, and to actually be sure that you have the right system for your spatial question. There are some very clear choices, for example, if you're really into proteomics and you want to look into some very specific cells, then you will do something connected with mass spec. But for the newer stuff, this is a very difficult thing to decide, and it's currently a bit of a gamble. To use these systems, it's not very important to have microscopy experience because the microscopy step in spatial is very simple. You tend to be better off if you have histology and sample preparation experience. Or experience working RNase-free if you're working in transcriptomics because that's also a huge challenge.*

MAKING USE OF DATABASES AND ATLASES

Databases are essential resources to collect, integrate and display single-cell and spatial datasets. By engaging with these collections of single-cell and spatial data, one can learn a lot about which technologies may be worth the investment, based on the type of data they can produce and the quality of the data that others have achieved with this technology. Here we will detail the available databases and resources for single-cell and spatial.

For single-cell data, a number of database efforts exist to pull together diverse datasets. These include PanglaoDB[13] (~250 datasets), the Single Cell Portal (650 datasets) from the Broad Institute, the Single Cell Expression Atlas (355 datasets) from EMBL and, perhaps the most impressive, CELLxGENE database (~900 single-cell datasets, ~350 spatial dataset) from the Charles Zuckerburg Initiative. Some databases, such as DISCO[14], presets deeply integrated single-cell data (in DISCO's case, ~ 14,000 human samples) building a living, ever-expanding atlas of a specific tissue or organism with utility, over data storage, in mind.

This level of sophistication has not yet been reached for spatial data. Initial efforts include SpatialDB[15] (24 datasets) and the BioImage Archive, supported by EMBL, hosting biological images of all varieties including large quantities of spatial omics data.

However, we are getting there. As of 2024, the largest, specific and most up to date spatial databases include STOmicsDB[16] (221 datasets) and CROST (182 datasets), which have collected spatial transcriptomic data across multiple platforms and technologies. This year has also seen the publication of tissue-specific databases, for example the Spatiotemporal Cloud Atlas for Neural Cells (SCAN)[17] and the Single-cell and spatially-resolved Cancer Resources (SCAR)[18] databases for both single-cell and spatial data.

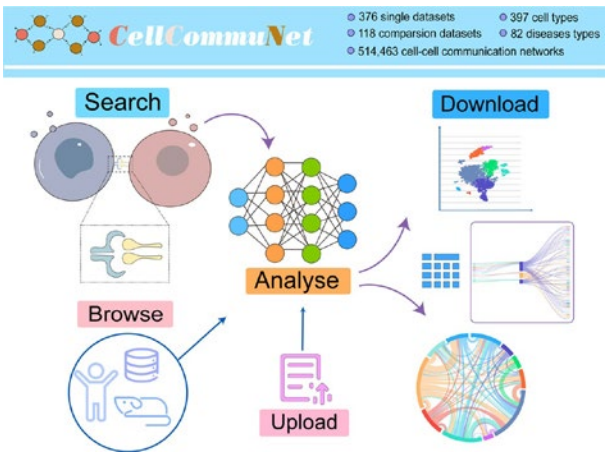
Popular single-cell atlasing projects such as the Human Cell Atlas (HCA) boast an impressive collection of datasets (>400) and are increasingly incorporating spatial and multi-omic measurements into their data. Furthermore, the Human Biomolecular Atlas Program (HuBMAP) boasts over 2000 datasets, with a healthy collection of single-cell, spatial and multi-omic data. HuBMAP is focused on producing these datasets with a heavy spatial multi-omics component.

These databases, projects and atlases are kept up-to-date and are vital resources for those thinking of developing their own single-cell and spatial workflows. However, leveraging this atlas and database data and meaningfully interacting with large-scale single-cell and spatial data is not completely straightforward, and there are tools to assist this.

The latest examples of which include STellaris[19], which uses spatial atlases to assign spatial mapping to scRNA-seq data (potentially saving the need to do the spatial experiment), CellHint[20], which harmonises independently annotated cell types across many single-cell datasets and scPoli[21], an open-world learner to learn sample and cell representations of large-scale data to perform multi-scale analysis.

FIGURE 8.4. OVERVIEW OF CELLCOMMUNET.

An image to display the functionality and specifications of the CellCommuNet platform. Source: Ma, et al. [22]



Large amounts of sc-RNA-seq or spatial data can be leveraged to make specific deductions. A great example is the recently published, CellCommuNet[22] (Figure 8.4), which brings 376 single-cell datasets together to build a comprehensive map of cell-cell communication, drawn from over 4 million cells. Finally, a very recent preprint presents NicheCompass[23], a graph deep learning method that can learn about cell communications and build meaningful cell niches from large-scale spatial data, as well as spatial reference mapping (alongside tools such as ExpiMap[24] and treeArches[25]).

Investing in a Lab Informatics Platform

While this chapter has principally focused on data analysis, managing data is a challenge at every stage of the single-cell and spatial workflow. By their nature, these experiments tend to have a large number of samples, procedures, instruments and produce an excess of data. Laboratory data management can be just as essential as post-workflow data analysis, and an effective solution to organise, store, analyse and share single-cell and spatial data is essential.

Laboratory Information Management Systems (LIMS) and Electronic Lab Notebook (ELN) software are popular commercial options to streamline this process. Through them, experiments and protocols can be standardised and stored, workflow data can be visualised in real-time, and samples, reagents and metadata can be effectively stored and tracked throughout an experiment. Furthermore, workflows and processes can be easily shared, which benefits collaboration and reproducibility.

To get the full rundown of what the Sapio Platform can do to accelerate and simplify single-cell and spatial workflows, we sat down with **Dara McCreary PhD**, Scientific Business Analyst at **Sapio Sciences**. He uses his previous experience in the lab to assist scientists with their transition onto the Sapio Platform and to help them tailor the platform to their specific needs. Our discussion covered the platform's general utility, its specific single-cell and spatial functionality and the exciting AI tools incorporated into the platform.

> SAPIO SCIENCES IS UNIQUE IN OFFERING A TRUE SCIENCE-AWARE™ PLATFORM, WITH THE USER-EXPERIENCE OF SCIENTISTS A MAJOR CONSIDERATION IN ITS DESIGN."

sapio sciences

There are a number of lab informatics providers, but Sapio Sciences is unique in offering a true science-aware™ platform, with the user-experience of scientists a major consideration in its design. For Sapio, data accessibility and utility are paramount. Their platform is not simply a LIMS or ELN, which are essentially distinct islands of data, but instead it is unified platform for input, visualising and analysing all that data while maintaining the core functionality of LIMS and ELNs.

Their Sapio Jarvis (named after Iron Man's AI assistant) works to integrate data at large scales, from a variety of LIMS and ELNs across an institution and directly from hundreds of experimental instruments. Furthermore, there are built in tools for flow cytometry and small molecules, directly aiding with single-cell and spatial workflows.

Ultimately, this platform makes life significantly easier for scientists working with data. Everything is accessible, visualisable and analysable, it will reduce errors that occur in the lab, but, more importantly, allow you to achieve the FAIR principle of data management and keep all your experimental data secure and accessible.

SPONSORED BY

sapio sciences

INTERVIEW:

DARA MCCREARY PHD
SCIENTIFIC BUSINESS ANALYST
SAPIO SCIENCES

***FLG:** Sapio's platform is a lab informatics platform that aims to help single-cell and spatial scientists do their job better, faster, and smarter. How does the Sapio Platform specifically enable this?*

Dara McCreary: First of all, the fact the Sapio Platform allows you to have everything in one place, which makes a huge amount of difference. If you've ever worked in a lab, you'll know that you often have important data spread across multiple spreadsheets and you regularly waste your time looking for things. Having everything in one place enables you to work more efficiently, more collaboratively, and increases transparency across teams.

The platform also has loads of inbuilt tools for different modules; for example, we've got a flow cytometry module and a histopathology module. There's a range of different options in the Sapio Platform. You don't have to do an experiment, get the information from an instrument, pop it on a USB and possibly lose your data. The Sapio Platform can be integrated with nearly any lab instrument so you don't lose your data. It's about working smarter; it's about being more efficient.

For example, I'm working with customers who are doing in vivo studies at the moment. If they have a mouse in one hand and their tablet in the other, they can use our platform to record measurements digitally there and then. Some customers do experiments in the lab, write everything down on paper, and then at the end of the day they go into the office and type everything up. Now, they can just bring their laptop with them. We also integrate directly with many different instruments, so you can get rid of that USB stick!

We also have new tools for the platform, like Jarvis and ELaiN. Jarvis is our science-aware scientific data integration platform. With Jarvis, you can consolidate all of your scientific data in one place, and then search, visualize and analyze it in a way that works for a scientist. It could be archived data from years ago, disparate lab informatics platforms, or data from an instrument in another lab. And because Jarvis is science-aware, you can use Jarvis to look at the data, create graphs, and understand what's going on. For example, you might have two different labs using two different ELNs. You could be doing genetics in one area and proteomics in another, and you want to analyse those results together. Rather than having to go to two different locations to access the data, it's all in one place with Jarvis.

The other tool is ELaiN, our AI lab assistant for scientists. ELaiN allows you to ask natural language questions to your ELN or LIMS and get an instant answer. For example, you can ask ELaiN to create an experiment based on the criteria you give it, and ELaiN will build that experiment template for you. Or you can ask it how much of a particular reagent you've got in stock, or even get it to write a Python script for you. The aim of ELaiN is to reduce the time scientists are spending on manual or repetitive tasks so they can focus on the science itself.

Ultimately, the Sapio Platform is helping scientists to be more efficient, so that they don't have to constantly move between different tools and platforms. In my previous role, I worked in one hospital, and I had a colleague who worked in a different hospital, and we were working on the same patient data. For us to work together, we would have to ring each other up or send each other emails, which became very time-consuming. Whereas if we'd both been on the Sapio Platform, we'd be able to look at the live data coming in and both be aware at the same time of how a particular clinical experiment was going.

On top of that, the clinical doctors that I'm working with can see a patient's progress in real-time. I used to get phone calls at 2:00am saying, 'Where's my sample? Have you got results?'. If they were using Sapio, they'd be able to just log on and see the progress of the experiment and know when they'll have their results.

And that's another reason why the Sapio Platform was built for scientists - this kind of collaboration feature helps take the pressure off them.

I've also had customers from a high throughput NGS company who were taking two days to decide which samples to process, because they had to be done in a specific order. We're solving that problem in an intuitive way with the platform, using the rules that they provided. So, they aren't sitting there on Excel moving samples and saying, 'Oh, this one has priority.' We simply flag the samples, and then all they have to do is click a button and their sample order is made straightaway. We're giving scientists back time to do the actual science, which is important.

FLG: You did briefly mention it, but can you explain what ELaiN is?

Dara McCreary: ELaiN is integrated into the Sapio Platform. ElaiN is an AI-powered scientific assistant, built using large language models (LLM), that can answer questions you've got about your data in the Sapio platform. In other words, it's a virtual assistant that you can ask basically anything about the platform, using natural language prompts. And that means that you can also use voice to interact with ELaiN. For example, imagine you're in a clean room and you have a microphone and mask on - you don't really want to be touching things. Instead, what you can do is talk to ELaiN and say, 'Hey, I've got 96 samples, make me a plate layout with 96 samples of X, Y, and Z'. And there it is. It's really going to make things more efficient and it's a really useful tool in the scientist's arsenal. I think it is going to put us to the forefront of lab informatics.

We're also training ELaiN on new things all the time. You can ask it about writing a Python script. You could also ask it if you have a reagent in-house; it will go look and give you an answer. If you don't have it, it will give you a link to where you can buy it. If you wanted a chart made, you can say, 'Hey, ELaiN, can I get the results from X, Y, and Z', and out pops a chart. But that's only the start - there's lots more planned for ELaiN and it is only going to get better.

FLG: Single-cell and spatial experiments often involve different types of data throughout the workflow, including flow cytometry, histopathology, NGS and multimodal data types. I want to dig a little deeper into the specific modules that your platform has to help with various scientific data.

Dara McCreary: We're constantly adding to our list of modules available in the Sapio Platform. If there's a need, the development team can work to create modules for that need. I'm sure they're working on things that I don't even know about at the moment; these modules are always evolving, based on customer needs and requirements.

One thing I really do like is our flow cytometry module, which we think rivals FlowJo. This means that you don't have to have a USB with a FlowJo license on it to do your experiment. You just need to integrate the Sapio Platform with that instrument and all that data is then stored on the Sapio Platform. This means you don't have to open FlowJo, get the results and extract

them; we have it already inbuilt. From the beginning of the experiment, once you get those sample requests in, you can track them, and you can do placements and assay design - it's as simple as drag and drop.

We have a histopathology module as well. You can track all your samples and see where the sample has gone throughout its history. It's like a pedigree or a tree of life, you can see your sample in the middle and you can see that it's gone to flow cytometry, histopathology, or NGS. So, if a sample hasn't gone somewhere, or if there's an issue, you can look at the history and say, 'Okay, it's gone here, and we're waiting on the results for that.' And then you can bring all this information back and it's all in one place. In a regular lab, you've got Excel sheets for tracking, you've got FlowJo, the histopathology lab, your ELN, and whatever way you're tracking your histopathology. That's all in one place if you're using the Sapio Platform.

There are other useful aspects to the modules too. They are high throughput - we have customers with vast amounts of samples, and we can track all of them within the platform. Also, when it comes to producing results, if you're using R, we have that integrated into the platform itself. Anything you're doing in R we can do within the platform, there's really no need to go and process your results elsewhere. The process can be modularized too. The data can be sent off to the next person whose expertise is flow cytometry, and they'll go do that work. The results can then be sent back, and the bioinformatics guy can then work on it. It's very collaborative.

These modules are constantly evolving. We work with the scientists themselves, if there's something that they see isn't right, or isn't working correctly, we can tweak it.

***FLG:* For single-cell and spatial analysis, there are a lot of specific packages and software that are mostly in R, but some are in Python or other languages. Can Sapio integrate that analysis alongside this more general experiment management?**

Dara McCreary: Yeah, there are ways of bringing it in. I'm not a technical person, so that's usually where I leave it up to a Solution Architect! We very much integrate with instruments; we integrate with other platforms. We do Python binding; we work with Java as well. So, there are ways around it.

This is the great thing with this team, when you're working with Solution Architects, they use their experience and give you the best options. They'll handhold you if you need it, and if you don't need it, they'll just give you the documentation. But they're always on hand to help. It is a real team and customer-focused effort.

***FLG:* Single-cell and spatial is quite a crazy area of development right now, and new instruments are appearing yearly. What is Sapio's capacity to integrate these?**

Dara McCreary: The Sapio Platform already integrates with over 200 machines, and we're adding new ones all the time; if there's a way in, we will find it. For example, we have a customer who's working with the new 10x Genomics Chromium workflow. We hadn't worked with that before. When we spoke with one of our Solution Architects about it, they said, 'That's fine. We can make that work'. As part of this, it was important to the customer that they could switch between doing the experiment manually or automate it, depending on the number of samples they need to run that day. So, we have both types of templates set up for the client. - If they don't have high throughput they can do it manually, or, by switching a button, they can take an automated approach and put it on the Chromium.

And you can even us ELaiN to help with integrations. At our recent SapioCon event, we demonstrated the 'give me the Python script for...' prompt and ELaiN was able to generate a script to speed up integrations with instruments. The integration is only going to get quicker and better. If you ask if there's any instrument we can integrate with, the answer will be yes. It just takes a bit of time. But I think it's going to get quicker as we continue to build on our advanced AI capabilities within the company.

***FLG:* I wanted to highlight the platform's ability to visualize the data in real time and analyse it as well. Could you touch on Sapio's ability to do that?**

Dara McCreary: For example, if I'm using a sequencer, we can track what's happening on that sequencer. You can have real-time monitoring within the platform. So, if you're at home and you want to know how your sequencing run is going, that information is there. As a lab manager, you may want to see how efficient a certain worker or a certain workflow is, and the platform allows you to track that visually. You can see it's taken 'X' length of time for whole genome sequencing, and you can delve into it and say, 'Okay, this section is taking a little bit longer, we have to think about how to fix that.'

"THE SAPIO PLATFORM ALLOWS YOU TO HAVE EVERYTHING IN ONE PLACE, WHICH MAKES A HUGE AMOUNT OF DIFFERENCE. IF YOU'VE EVER WORKED IN A LAB, YOU'LL KNOW THAT YOU OFTEN HAVE IMPORTANT DATA SPREAD ACROSS MULTIPLE SPREADSHEETS AND YOU REGULARLY WASTE YOUR TIME LOOKING FOR THINGS."

There are also graphs. So, if you have data, you can use a pre-built graph within the Sapio Platform, or with a click of a button, you can change that graph. For example, if you've got a lot of QC data, and you want to build a standard curve or a histogram, they can be built automatically or generated when needed. You don't have to go in and drag and drop samples, although you can do that as well. But you can just highlight the data you want and click a button, and the graph is produced.

Everything in Sapio is running on our own platform. The time I track for working with customers is actually tracked back into Sapio, so my line manager can look and say, 'Dara has worked with XYZ,' but it's on a graph that is prebuilt. Within Sapio, the project managers can track the project, the timescale and how it's going. We like the usability of it so much that we use it ourselves! Any graph building or anything like that, it's all done within Sapio.

FLG: As a scientist yourself and in your job making sure that the scientists who come to you get what they want from the platform, what do you think is the main reason to switch to using Sapio?

Dara McCreary: My main reason - and this is me personally - is that it's collaborative. Everything you're working on is stored in the Sapio Platform, so it is much easier for teams to work together on a project. If I need something I can find it easily, rather than having it spread across all these Excel files, across different computers, across USBs. Everything is in one place, and it's searchable. And it gives transparency across the team - a lab manager can see what everybody else is doing and all scientists can work collaboratively.

With this in mind, my favourite feature in the app is the '@' button. You can '@' someone, and they can see exactly what you're working on so you can work together. You could be at a conference in Singapore, and I could be working in the lab here, and we could work on a document or experiment together at the same time. An example would be if I was working on a rare genetic disease, and I found a variant. I would need to find out if anybody else had seen it. If I'm working at an organisation with a number of different labs and they're all using Sapio, I can search, 'Is there any patient with X?', or 'Has anybody else found X?'. I can take a look at their data to see if the profiles match and whether I discovered something of interest. We want to help identify those rare events, or interesting findings to make a real impact in patients' lives. Sapio's platform enables you search and find those insights inside your own data in a fast and effective manner.

Chapter 8 references

1. Luecken, M.D. & Theis, F.J. **Current best practices in single-cell RNA-seq analysis: a tutorial.** *Molecular Systems Biology* **15**, e8746 (2019).
2. Fang, S. *et al.* **Computational Approaches and Challenges in Spatial Transcriptomics.** *Genomics, Proteomics & Bioinformatics* **21**, 24-47 (2022).
3. Heumos, L. *et al.* **Best practices for single-cell analysis across modalities.** *Nature Reviews Genetics* **24**, 550-572 (2023).
4. Zhang, K., Zemke, N.R., Armand, E.J. & Ren, B. **A fast, scalable and versatile tool for analysis of single-cell omics data.** *Nature Methods* **21**, 217-227 (2024).
5. Blampey, Q. et al. **Sopa: a technology-invariant pipeline for analyses of image-based spatial-omics.** *bioRxiv*, 2023.12.22.571863 (2023).
6. Cui, H. *et al.* **scGPT: toward building a foundation model for single-cell multi-omics using generative AI.** *Nature Methods* (2024).
7. Zhang, D. *et al.* **Inferring super-resolution tissue architecture by integrating spatial transcriptomics with histology.** *Nature Biotechnology* (2024).
8. Yuan, Z. **MENDER: fast and scalable tissue structure identification in spatial omics data.** *Nature Communications* **15**, 207 (2024).
9. Varrone, M., Tavernari, D., Santamaria-Martínez, A. & Ciriello, G. **CellCharter: a scalable framework to chart and compare cell niches across multiple samples and spatial -omics technologies.** *bioRxiv*, 2023.01.10.523386 (2023).
10. Aihara, G. *et al.* **SEraster: a rasterization preprocessing framework for scalable spatial omics data analysis.** *bioRxiv*, 2024.02.01.578436 (2024).
11. Clifton, K. *et al.* **STalign: Alignment of spatial transcriptomics data using diffeomorphic metric mapping.** *Nature Communications* **14**, 8123 (2023).
12. Moses, L. & Pachter, L. **Museum of spatial transcriptomics.** *Nature Methods* **19**, 534-546 (2022).
13. Franzén, O., Gan, L.-M. & Björkegren, J.L. **PanglaoDB: a web server for exploration of mouse and human single-cell RNA sequencing data.** *Database* **2019**, baz046 (2019).
14. Li, M. *et al.* **DISCO: a database of deeply integrated human single-cell omics data.** *Nucleic acids research* **50**, D596-D602 (2022).
15. Fan, Z., Chen, R. & Chen, X. **SpatialDB: a database for spatially resolved transcriptomes.** *Nucleic Acids Res* **48**, D233-d237 (2020).
16. Xu, Z. *et al.* **STOmicsDB: a comprehensive database for spatial transcriptomics data sharing, analysis and visualization.** *Nucleic Acids Res* **52**, D1053-d1061 (2024).
17. Deng, Y. *et al.* **SCAN: Spatiotemporal Cloud Atlas for Neural cells.** *Nucleic Acids Research* **52**, D998-D1009 (2023).
18. Deng, Y. *et al.* **SCAR: Single-cell and Spatially-resolved Cancer Resources.** *Nucleic Acids Research* **52**, D1407-D1417 (2023).
19. Li, X. *et al.* **STellaris: a web server for accurate spatial mapping of single cells based on spatial transcriptomics data.** *Nucleic Acids Research* **51**, W560-W568 (2023).
20. Xu, C. *et al.* **Automatic cell-type harmonization and integration across Human Cell Atlas datasets.** *Cell* **186**, 5876-5891.e20 (2023).
21. De Donno, C. *et al.* **Population-level integration of single-cell datasets enables multi-scale analysis across samples.** *Nature Methods* **20**, 1683-1692 (2023).
22. Ma, Q., Li, Q., Zheng, X. & Pan, J. **CellCommuNet: an atlas of cell–cell communication networks from single-cell RNA sequencing of human and mouse tissues in normal and disease states.** *Nucleic Acids Research* **52**, D597-D606 (2023).
23. Birk, S. et al. **Large-scale characterization of cell niches in spatial atlases using bio-inspired graph learning.** *bioRxiv*, 2024.02.21.581428 (2024).
24. Lotfollahi, M. *et al.* **Biologically informed deep learning to query gene programs in single-cell atlases.** *Nature Cell Biology* **25**, 337-350 (2023).
25. Michielsen, L. *et al.* **Single-cell reference mapping to construct and extend cell-type hierarchies.** *bioRxiv*, 2022.07.07.499109 (2023).

CHAPTER 9

BROADER CONSIDERATIONS: ESG, LAWSUITS AND THE WILD WEST

OUR FINAL CHAPTER WILL OUTLINE SOME OTHER FACTORS THAT YOU MAY WANT TO BEAR IN MIND WHEN THINKING ABOUT WHICH SINGLE-CELL AND SPATIAL WORKFLOWS TO INVEST IN. THIS INCLUDES THE ENVIRONMENTAL, SOCIAL AND GOVERNANCE PLEDGES OF COMPANIES IN THE SINGLE-CELL AND SPATIAL MARKET, THE ISSUES THAT ARISE FROM THE CURRENT SPATIAL LAWSUITS AND SOME PERSPECTIVES ON THE DIRECTION THAT THE FIELD IS MATURING.

ESG in Single-cell and Spatial

As a society, we are becoming more conscious of where our products come from and the environmental and social impact of what we consume. Why would this be any different for single-cell and spatial? When looking for equipment and reagents, the environmental, social and corporate governance (ESG) policy of the manufacturer (see Figure 9.1) can help you make informed consumer choices. To reflect this, in a section below, we have detailed the areas in which single-cell and spatial companies have made ESG positions and pledges.

It is worth pointing out that, some newer companies who are still finding their feet in this space are unlikely to have made such pledges, and so are not featured on this list. This is not a reflection of that company's stance on ESG issues. In some ways, it is a privileged position to feel secure enough to focus substantial company resources into more ethical practices. With time and growth, many of these newer companies will likely make pledges of their own.

FIGURE 9.1. AN ILLUSTRATION OF THE TYPICAL TYPES OF INFORMATION THAT MAY BE USED TO ASSESS AN ESG CRITERIA

For example, if a company's social contribution was being measured, human rights and child labour might be considered. Likewise, compliance and shareholder democracy may be considered if an organisation's governance was being analysed. Source: anevis

10X GENOMICS

10x Genomics has a page dedicated to governance, highlighting business conduct ethics and corporate governance guidelines. This involves a well fleshed out set of corporate responsibility policies to handle insider trading, bribery and corruption. There are currently no formal policies regarding environmental and social responsibility. The board of directors is 29% female and 43% underrepresented minority, and there are dedicated Employee Resource Groups to foster safe spaces for diversity.

NANOSTRING

NanoString have a dedicated ESG strategy and report, which tracks their progress. Environmentally, all their sites encourage and engage in composting/recycling, they encourage the use of public transport and electric vehicles, they use energy-saving lamps and have engaged in water conservation activities. Socially, they are committed to Equality,

Diversity and Inclusion (EDI) practices, 44% of the board is female and 34% of full-time employees are members of underrepresented minority groups. They are also part of WERE and participate in the Hiring Our Heroes scheme. Their report explains that they take pride in their employee development, compensation and benefits.

BIO-TECHNE

Bio-techne also have a clear ESG strategy, with each major component represented as a chapter. Environmentally, Bio-techne have recently built a thorough assessment of all emission data across their 38 worldwide sites covering electricity, gas, vehicle miles etc. This level of transparency is rare and Bio-techne have a set of pledges to help reinforce these values through industry-leading sustainable packaging development, renewable energy resource and effective water management.

For social and corporate governance, Bio-techne's board is 22% female and 11% underrepresented minority. Culturally, they are committed to advancing their employees and nurturing career growth evenly, with women acquiring 53% of promotions. Employees are polled for their feedback, highlighting the strengths and weaknesses of the company, and 77% of feedback about the company is currently positive. Bio-techne also prides itself in investing in the community and charity, with a large list of events and sponsorship.

MILTENYI BIOTEC

Miltenyi Biotec has a dedicated webpage to document their corporate social responsibility. Environmentally, this includes moving towards greater sustainability by replacing EPS insulation with straw insulation packaging, which can be recycled and is more sustainable. Miltenyi Biotec values their workers seriously with flat hierarchies and a campus-style modern workplace.

STANDARD BIOTOOLS

Standard BioTools (previously Fluidigm) have been producing ESG reports for a few years now, with the latest one released in April 2023. Their guiding principles are human rights, environmental responsibility, labour rights and anti-corruption.

Environmentally, Standard BioTools is committed to reducing its impact, and engages in sustainability initiatives such as: schemes for recycling and waste management on site, using low-energy lighting systems, using sustainable materials and using water dispensers and reusable water bottles. Socially, the company is committed to EDI and considers itself an equal opportunities workplace. The board is 14% female and 14% underrepresented minority. It also engages in affirmative action practices. There is also a clear code of conduct for business partners, including ethical and sustainable practices.

BRUKER

Bruker have an ESG report for 2023, a code of conduct document that outlines their ESG stances alongside a supplier code of conduct. Socially, the company embraces diversity and inclusion, striving to create a harassment-free work environment. The board is 30% female and the company has supported the setup of employee affinity groups to promote EDI within the business. Bruker also demonstrates a clear commitment to its communities. This includes charitable contributions and a commitment to the environment and protecting human rights.

Environmentally, Bruker has been engaged in a number of sustainability practices, such as identifying illegal logging, enhancing landfill mining and safeguarding global food supply chains. Bruker has kept direct measure of their energy usage and carbon goals, with 39% of their energy currently coming from renewable sources. They have dedicated efforts for recycling, water usage and waste management.

BD BIOSCIENCES

BD Biosciences have a dedicated ESG page with an overview of their impact on global public health, sustainability, social investing and EDI. Environmentally, BD have a clear set of 2030 targets for a 25% reduction in energy, 40% reduction in water use and a 50% waste reduction. This goes alongside goals of producing sustainable packaging, products and responsible supply chains.

Socially, BD Biosciences is committed to investing in social initiatives such as developing healthcare systems, responding to disasters and encouraging its employees to give back to the community. This is coupled to their commitment to improve global public health through tackling antimicrobial resistance and infectious diseases such as TB, HIV and AIDs. EDI is also a focus for BD Biosciences, with dedicated resource groups to meet the needs to different groups. The board is 36% female and 9% underrepresented minority.

BIO-RAD

Bio-Rad has a clear vision for sustainability, with a set of 2030 goals such as aiming for a 46% reduction in carbon emissions, and using 100% renewable energy. Furthermore, there are social goals, such as aiming for 60% of their US workforce to be comprised of underrepresented groups and for at least 45% of leadership roles to be taken up by women. Bio-Rad take social responsibility seriously, with a community involvement program to promote science education and a clear focus on employee health, safety and diversity and inclusion. The board is 33% female.

LEICA BIOSYSTEMS - DANAHER

As a division of Danaher, Leica Biosystems' ESG position is tied to Danaher's. Leica Biosystems is having substantial impact on medical advancement, with over 1.6 million patients diagnoses enabled per week. Environmentally, Danaher has a goal of reducing greenhouse gas emissions by 50.4% by 2032. This involves a specific decarbonisation model and water stewardship. Leica Biosystems has specifically reduced water consumption by 22% since 2019. Socially, there are initiatives for EDI, pay equity, wellness, mental health and employee learning and development. The board is 29% female and 21% underrepresented minority.

AKOYA BIOSCIENCES

Akoya Biosciences also has a dedicated page for corporate governance, with a code of business ethics and corporate governance guidelines. This involves a well fleshed out set of corporate responsibility to handle insider trading, bribery and corruption, but no formal policies regarding environmental and social responsibility. The division of the board is 28% LGBTQ+ and 14% female.

Lawsuits in Spatial

Another consideration to bear in mind when purchasing a spatial instrument is the longevity of the method and the company that provides it. The technology space is fast paced and dynamic, with new companies emerging every year and longstanding companies diversifying into different methods. With that level of transformation, inevitable legal issues around intellectual property (IP) arise.

In particular, headlines for spatial have been dominated by legal disputes between the major players (see here and here). The question of which company may succumb to legal pressures is getting as much attention as the capabilities of the latest methods. While most people agree that IP should be respected, this commercial environment is stifling for emerging companies with new spatial solutions, because of the fear of potential lawsuits. A recent appeal from scientists has asked for fair competition, antitrust laws and for scientific discovery to not be forgotten.

The stability and IP of the company that you plan to purchase a spatial instrument from clearly matters. It would be disastrous to invest in a spatial platform that is suddenly removed from the market for legal reasons. Despite the interest in NanoString products, they had been banned from selling the CosMx or the reagents in Europe since September 2023 due to patent infringement, preventing access to many labs who had started engaging with NanoString in 2022 and early 2023.

While NanoString filing for Chapter 11 Bankruptcy in February 2024 came as a shock, it did result in the overturning of the ban in Europe and NanoString products are again accessible in Europe. However, the fragile state of the company placed numerous projects relying on NanoString products in severe jeopardy. With the company being acquired by Bruker in May, these projects are currently safe. However, the concern remains for other companies in similar positions. This needs to be held in mind for potential buyers.

Getting Out of the Wild West

Development in spatial has been coming from multiple angles:

- From the biotechnology companies developing streamlined methods.
- From the innovative labs developing new methodologies and shortcuts.
- From small groups and large scale consortia alike.

It is not uncommon to hear this scenario described as a scientific Wild West, with rules, standards and regulations hard to come by.

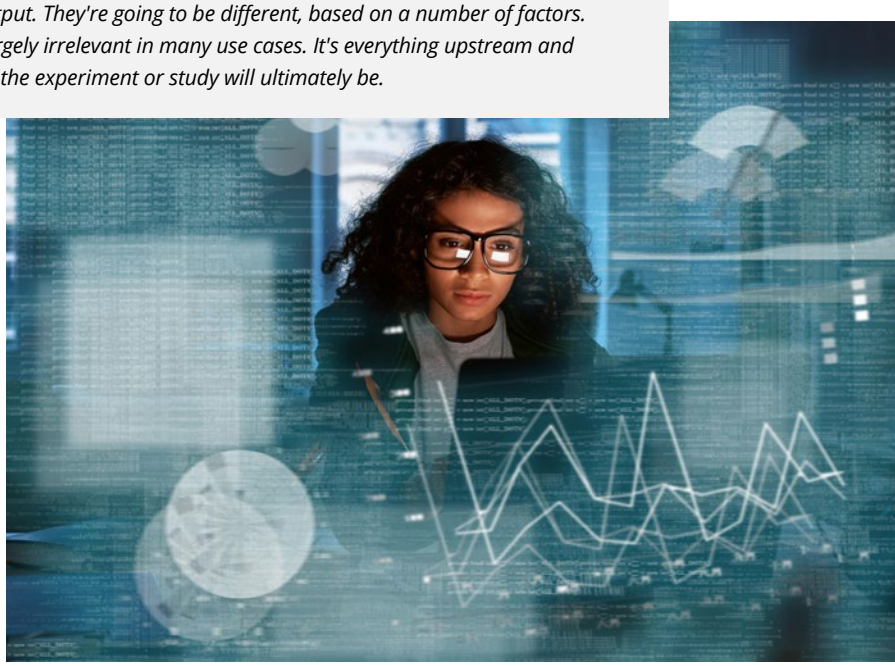
With single-cell sequencing, and bulk sequencing before it, the maturation of the field took (and is still taking) time. It can take decades of working with new instruments and data types to get anything that looks like a gold standard. For single-cell, we are already seeing multiple efforts to standardize the data analysis such as the Single-cell Best Practices[1] and Open Problems in Single-Cell Analysis initiatives. However, the process is by no means over.

JOHN M. ASHTON, PHD, MBA

Associate Professor, Department of Biomedical Genetics, Director, Genomics Research Center, **Wilmot Cancer Institute, University of Rochester**

FLG: The single cell community is looking for standardised procedures and gold standards, and your paper is part of the way toward getting there. In your opinion, what do you think is the path towards finding standardised, reproducible single-cell practices?

John: *I think that all comes even before you use the platform. It's all the preparatory work and I think that's something that people don't appreciate enough. There needs to be a standardised way to process the cells, process tissue, and to quality control the material. We didn't put it in our comparison paper, but we had a very low correlation of differential expression across any of the platforms. You can't necessarily validate a 10x Genomics experiment and a Fluidigm experiment and expect to get the same output. They're going to be different, based on a number of factors. Which method you use and which tool you use is largely irrelevant in many use cases. It's everything upstream and downstream from that which drives how successful the experiment or study will ultimately be.*



For spatial, the question is, can we fast forward the slow and painful process of maturation? There are several individuals and groups working towards this goal now, and it is on everyone's mind. The GESTALT initiative was created partially as a response to this community concern. Pioneered by Luciano Martelotto, Jasmine Plummer and Ioannis Vlachos, GESTALT is the town square of spatial, in which researchers discuss the latest methods and communicate in a way that hasn't happened before for a technology-based group. There is hope that this community will rapidly accelerate the process of maturation.

DAVID COOK

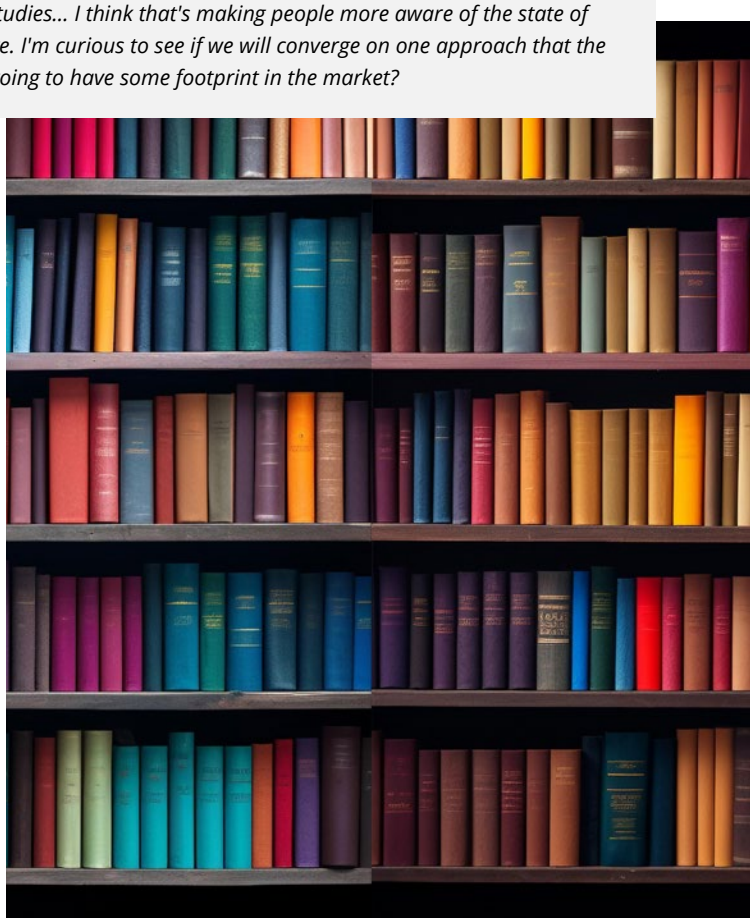
Scientist, **Ottawa Hospital Research Institute,** Assistant Professor, Department of Cellular and Molecular Medicine, **University of Ottawa**

FLG: Studies comparing the performance of these spatial platforms appear to be coming thick and fast; it is a very prominent discussion in the field. Furthermore, you've got the GESTALT initiative that has been set up to discuss these developments. What do you think is the path forward to identify effective spatial strategies for researchers who can't invest in everything? Do we need to do more of these studies on a larger scale, or is there a different path?

David: *It's something I'm curious about. It's funny, I'm thinking back to when single-cell boomed and trying to remember if there was a similar energy. I was still new to it, and I don't remember if there was controversy about whether 10x Genomics' data was good, because it was such a new concept and there weren't really competitors in that space. Moving forward with spatial, I think it's something that's going to be born out of more data from each of these methods and from the first movers on their respective platforms. Maybe it'll just come naturally when more papers come out from different labs that use two different approaches on a common tissue and see something on one platform that you couldn't find on another.*

What I do like is with things like GESTALT is there's conversation happening. It's trying to pull the community together so that we don't have to wait years for the knowledge from the first rounds of publications in the space to spread. To get the community discussing early benchmark studies... I think that's making people more aware of the state of things. I think that's going to be how we slowly evolve. I'm curious to see if we will converge on one approach that the field decides is optimal? Or are all these platforms going to have some footprint in the market?

We recently spoke to one of the original GESTALT pioneers, **Ioannis Vlachos**, who is an Associate Professor and is also the Director of the Spatial Transcriptomics Technologies Unit (STU). We spoke to him about GESTALT and the need for maturation in spatial, and we also asked him about the STU. This unit is uniquely set up in that it is entirely focused on spatial methods. In the following interview, Ioannis introduces his background with spatial and the STU, discusses his experience with the various instruments in the STU, and closes with a discussion of GESTALT and the maturation of the spatial field.



INTERVIEW:

IOANNIS VLACHOS

ASSOCIATE PROFESSOR, **HARVARD MEDICAL SCHOOL,** DIRECTOR OF THE SPATIAL TRANSCRIPTOMIC TECHNOLOGIES UNIT **BETH ISRAEL DEACONESS MEDICAL CENTER**

***FLG:* Can you briefly introduce yourself and your background in the single-cell and spatial fields?**

Ioannis: My name is Ioannis Vlachos, I'm an Associate Professor of Pathology at Harvard Medical School, and at Beth Israel Deaconess Medical Centre. I'm also a PI at the Broad Institute of MIT and Harvard. My lab is a producer and consumer of large quantities of bulk, single-cell, and since 2020 of spatial data. I direct two facilities: the Bioinformatics Unit for Precision RNA Medicine Core and the Spatial Technologies Unit.

I started working with next generation sequencing (NGS) about 14 years ago, when the field started growing rapidly, and I became quickly addicted to learning and trying every new library prep method. Every new way of interrogating RNA or DNA with these new technologies could open up a frontier that needed years and years on the bench to solidify. I understood very quickly that the limits of my science are often the limits of the technology that I have access to. I was focused on how to maximize the information you can get from these types of data. That focus moved to single-cell in 2016, which felt like a natural technological progression. We had a lot of questions back then about tissue heterogeneity, about different cell populations, and single-cell came as a natural next step.

As of 2019, we moved to spatial - but spatial was very different. It was a very quick realization - spatial is its own thing. It's not the natural progression of bulk to single cell to spatial, as we see in the presentations. It's a disruptor but also a natural progression of humanity interrogating tissue for hundreds of years, only with different means and different technologies. We started with the naked eye, and then we added microscopy. In the past two-three hundred years, we started with carmine staining, moved to H&E, methylene blue, paraffin embedding, the microtome, and a wide spectrum of techniques and stains. When we reached the 1940s, immunofluorescence and immunohistochemistry became the cutting-edge technique of that time. In the 2000s, multiple antibodies and RNA ISH became mainstream. I think spatial is the natural progression of that. Histology and Pathology v2.0, not a new independent domain.

When immunofluorescence or immunohistochemistry became available, they gradually became a workhorse for innovation. I think that spatial is exactly on the same route now. It's a shiny, new thing that we're trying to understand, but it's practically the future workhorse for discovery and innovation. I think it will reshape how we perform research, but also how we diagnose, stratify and monitor patients, or discover therapeutic targets down the road.

I, and many colleagues, have been exposed to the iterative refinement of bulk sequencing methods. It started as an exotic technology and gradually became mainstream. The focus then shifted on finding ways to bring it into the clinic. Now, it is in every day practice; a prerequisite for many diagnostic and research activities. Similarly, with single cell. It started as an exotic modality, and now we're in the process of standardizing and expanding it in research and clinical trials. Single cell is actually considered rather mainstream in many research fields. However, it hasn't reached the same level of maturity or penetration for clinical practice. For spatial, the question is, do we have to go through the same iterative refinement again? Do we need to wait so many years until it transitions from the wild west situation we are currently in, into an era of maturity and mainstream utility in research and clinical practice?

***FLG:* Your Spatial Transcriptomics Unit (STU) is quite special; can you discuss its inception and its core principles?**

Ioannis: When I was exposed to the benefit and importance of spatial, and I could see the historic significance of this emerging new field, with my colleagues at BIDMC, Drs. Frank Slack and Winston Hide, we went to stakeholders within our institution and at the Commonwealth of Massachusetts with our proposal. If we want spatial to reach the potential that we all believe that it can achieve, and if we want to revolutionize how we perform our research and our clinical work, we need to enable everybody to have access to the best infrastructure and education without each lab or group having to invest millions setting up everything in-house. It shouldn't be that the richest or the labs that develop the technologies the only ones to benefit. Many times, the best questions are from the clinical groups or from experts of domains further away from these technologies. Unfortunately, they rarely have the technical expertise or the funding to internalize such technologies. Thankfully BIDMC, our Institution, the Harvard Medical School Initiative for RNA Medicine, our Department of Pathology, and importantly the Commonwealth of Massachusetts, joined forces to enable us to make it a reality. The Spatial Technologies Unit is currently supporting Academics from across not only Massachusetts and the US, small pilot projects, extensive consortia, clinical trials, start-ups, biotech, and industry.

Dealing with tissue is very different than the requirements for bulk NGS, where you distil everything into RNA or DNA, and then you move on with your NGS assay. Every tissue has its own quirks, its own requirements, its own processing, its own way of being analysed, its own way of being segmented and so on. Even if you had invested in a single instrument, working in diverse tissues is a very big undertaking. So, we pitched – "what if we create, instead of a core facility, a centre of excellence that is outward facing that enables everyone to go as deeply as they want in this field?". The only things you would need to work with that facility is a tissue and a question. Everything else can be taken care of by production grade assays, by experienced specialized personnel that have their careers there, with extensive experience and funds to establish SOPs. I would like to reiterate the importance of career personnel and how fortunate we are to have the exceptional people that work for us. We are always looking for more talent to hire of course.

We have internalised everything, and the STU is that facility; it is a centre of excellence for spatial tissue profiling that has been established with four mandates. One is democratized access to technology and provision of services. The second is provision of education and training, because if we want this new generation of scientists to maximally utilise these tools, we need to explain what they can do. It's not just training people to apply an assay. That's the minimum. We train and educate physicians and PIs on how to use these methods for their research. We educate decision makers, venture capitalists, investors of pharmaceutical companies on what this means for their day-to-day work, and how they can internalize the outcome of this process. The third is supporting entrepreneurship, which means we support startups, spin offs, new technology development, curbs to IP etc. The fourth one is how we can bring these assays into the clinic. Everybody may be thinking we mean that, by tomorrow, we aim to start measuring 1,000 genes per patient and revolutionize clinical practice, but this is not the case. We can actually revolutionize clinical practice with just six genes, compared to what we're doing today. It's a very sobering realization that what you see in papers is not what the patients need, right here, right now.

Within the facility, we're investing in automating the workflows as much as possible. We're trying to use spatial today, in the way that spatial should be used a few years from now. We have the luxury of doing so because we were established, from the get-go, as a spatial tissue profiling centre, not a genomic centre, single-cell centre, or a histology centre that has moved to spatial and now has to be restructured while figuring things out. We also have invested significantly in the data analysis, enabling us to handle a project from its inception, tissue processing, staining, imaging, and spatial and/or single cell data generation, to data analysis, biological insight generation, and manuscript, report, or grant authoring.

***FLG:* Having set up this Center of excellence for spatial, and having got as many platforms as you have under one roof, your exposure to them is unique and valuable. From the selection of technologies you have, are there any that you're particularly impressed with? If someone comes to you with a tissue and a question, are there ones you find yourself using more often than others?**

Ioannis: Everybody is focusing on what an instrument can do, i.e., this instrument can capture this many genes or that resolution. Very often you will see these graphs of plexity (i.e. number of targets) versus resolution and each spot represents an instrument. As a researcher you can look at it and say – 'I want this one because it's got a lot of this and a little bit of that'. But even if the instruments are identical on those measures - they're not. Let's say they were in the same category and identical in terms of capability - they're often very different as products. The product and the technology are not the same thing.

For instance, they may vary on which assays are available, and what the strategy for assay development for that vendor is. There are assays where you have an easy-to-use, pre-select, multi-100 target panel that can be used for any tissue, making it very fast to start. Or there are sequencing-based assays that are also very fast for you to start, or other assays that can be targeted an powerful but you might need to customize them, especially if you're working with non-model organisms, or on specific questions. Every solution has its pros and cons and has to be matched with the right research question.

We usually start our discussions with what I mentioned, a tissue and a question. Often users have a certain number of samples and a specific assay in mind, but we rarely move forward with exactly that. We have a very detailed consultation - what will success for the project look like? What are the populations you're interested in? What are the pathways? This is because there is no single assay that does everything today. We've seen many times that there is no single spatial assay that can answer your question, which means we bring in single-cell at the same time and we can integrate the data. There is currently no hammer that can address every nail that is out there.

We usually choose an approach based on the question. There are some instruments that are used more than others, but I think that's also based on the fact that many of the researchers we support have a cancer immunology background, a domain that is now greatly benefiting from spatial tissue profiling. During the past 4 years we saw initially a lot of CNS work, moved to cancer and cancer immunology. However, we now see absolutely everything to be interrogated with these powerful techniques: from autoimmunity or non-model organism work, to new target identification, engagement or mechanism of action elucidation. We can and use multiple technologies in the same project, at different phases of the project, to address a question. The question should be the one dictating the technology, and not vice versa.

That's why I find that spatial works best as a centralized service, as opposed to, 'I will buy this and use it in my lab'. If you do that then you will be locked into a specific set of questions that you are able to address. We should not focus on the instrument specs. Instead, we should focus on how these specs manifest in an assay. When I'm using a assay, I think: what's the throughput? How long does it take to be performed? What is the imaging area? How is the data analysed? How can I access the data? How is the analysis done? Is it an open system? Is it customizable? There are a lot of questions that can determine whether the approach will be right for your problem and your setup, or not.

Just for an example, let's say I'm submitting a spatial project for funding to an institution, such as NIH or a Foundation. You will need to show pilot results. If you're designing the pilot, you need something that can be performed on a small scale, with full power, so you can show exactly why this is the right approach. But it also needs to not break the bank. You're going to need an assay that can get you started quickly, without requiring an insurmountable investment in time and effort, since you don't have access to that funding.

If your application is selected for funding, you will want to move to the discovery phase of the study. The discovery phase typically needs a lot of information, potentially through the engagement of multiple modalities. Often, your final phase will be to transition to cohort-scale inquiries or to streamlined biomarkers or correlative work, where you want to capture patterns across large numbers of patients. So, you might start your project with an assay that's easy to initiate and run, and you might transition to a very information-rich assay, or a combination of assays, for your discovery phase, and the you could graduate to a solution that can be applied at low cost very robustly at a cohort scale to interrogate very specific patterns that you have prioritized. Within this same application, we might have used three or four different technologies to capture different aspects. You might have been able to do most of the project with a super powerful assay, but it could need more resources, increased effort or time, or both, to reach a similar conclusion. Using these tools in a thoughtful and efficient way enables you to perform research you will easily be able to do in five years, today.

***FLG:* Can we talk about GESTALT, the international community for spatial, and how it fits into the path forward to finding standardized practices for spatial? How do we reach maturity for these methods?**

Ioannis: First of all, GESTALT, the Global Alliance for Spatial Technologies, was not formed based on a moment of inspiration.

It was the result of an evident pressing need that reached a tipping point. GESTALT is the next step for our vision, the vision that manifested initially with the Center (Spatial Technologies Unit) for anyone to be able to access standardized and best practice spatial technologies. Luciano (Martelotto), Jasmine (Plummer) and I have had joint projects focusing on shared resources. The Spatial Technologies Unit is also a tissue mapping centre for Human Biomolecular Atlas Program, an NIH initiative that apart from aiming to redefine human anatomy with cutting-edge single cell and spatial tissue interrogation, it has a very clear focus on standards and best practices. However, without the Community actively adopting these standards and utilizing them in their everyday work, publications, data deposition, or grant submissions, the progress we all want to see will not happen. You need the whole community to come and work together.

"WITH GESTALT WE WANTED TO CREATE NOT JUST A SAFE SPACE, BUT A FORUM OR THE EQUIVALENT OF THE TOWN SQUARE FOR SPATIAL. WITHIN GESTALT, YOU CAN MEET COLLEAGUES, USERS OR INVENTORS OF ALL AVAILABLE TECHNIQUES, PITCH OR LISTEN TO PROPOSALS FOR NEW PROJECTS, DISCUSS CURRENT OR FUTURE PROBLEMS, AND DISCUSS TOGETHER WITH EVERYONE ON HOW WE CAN ADDRESS THESE CHALLENGES."

With GESTALT we wanted to create not just a safe space, but a forum or the equivalent of the town square for spatial. Within GESTALT, you can meet colleagues, users or inventors of all available techniques, pitch or listen to proposals for new projects, discuss current or future problems, and discuss together with everyone on how we can address these challenges. There are currently large-scale efforts trying to bring standards from bulk or single-cell to spatial, but to my opinion it's not an optimal solution. This would be like making spatial wear a shoe that doesn't fit. We can seed the process with those ideas, but we need to tailor specific solutions that are specific to this new field, and for that you need everybody's input.

GESTALT's overarching goal is to help us transition from this 'Wild West' phase of organic growth into a field that enables true discovery, that reaches patients. Where instrumentation and assays are robust, they have the throughput and the speed, and the community has access to resources and education. However, it's also about the right here and now. If you want to start a new assay and there are 50 different protocols for the same tissue for that assay, which one should you use? Of course, you could test 30 or 40 protocols, but you could save time, effort and resources by just asking. So, having access to this community for a field that is moving so fast enables you to jump on that fast moving train, even if you're only in that first wagon.

For GESTALT, we have a WhatsApp group, and we interact daily. There are ~16 channels on the WhatsApp - from data analysis to new technology development. Apart from that meeting space, there are dedicated Working Groups to Data Analysis, Imaging, Sequencing, or Protein-based Assays, to Non-Model Organisms, or Clinical Spatial Tissue Profiling. Regarding best practices, the Standards Working Group, liaises with all other groups within Gestalt to establish standards from the whole lifecycle of each technology. From how a new assay can be effectively assessed, how to evaluate whether a run was successful, how results can be presented in a manuscript in a transparent and accurate way, what a journal could or should be asking from authors using spatial tissue profiling during submission, review, or publication, how pilot data should be reported in a grant submission, or even how spatial data and metadata should be deposited for the community to access. All these questions are still open, and of course, GESTALT shouldn't be doing it by itself because then it will be another consortium trying to set the field up in the way that it wants to. We're reaching out to every other consortium, and to every other team doing standards, so we can set these standards as a field. These are exciting times and we see history in the making. It's a great opportunity to come together and enable these great advancements to reach every corner of research and clinical practice, benefiting scientists, but most importantly patients, and the Community as a whole.

Chapter 9 references

1. Heumos, L. *et al.* **Best practices for single-cell analysis across modalities.** *Nature Reviews Genetics* **24**, 550-572 (2023).

Front Line Genomics

frontlinegenomics.com

info@frontlinegenomics.com
@FLGenomics
@frontlinegenomics

Festival of Genomics: festivalofgenomics.com