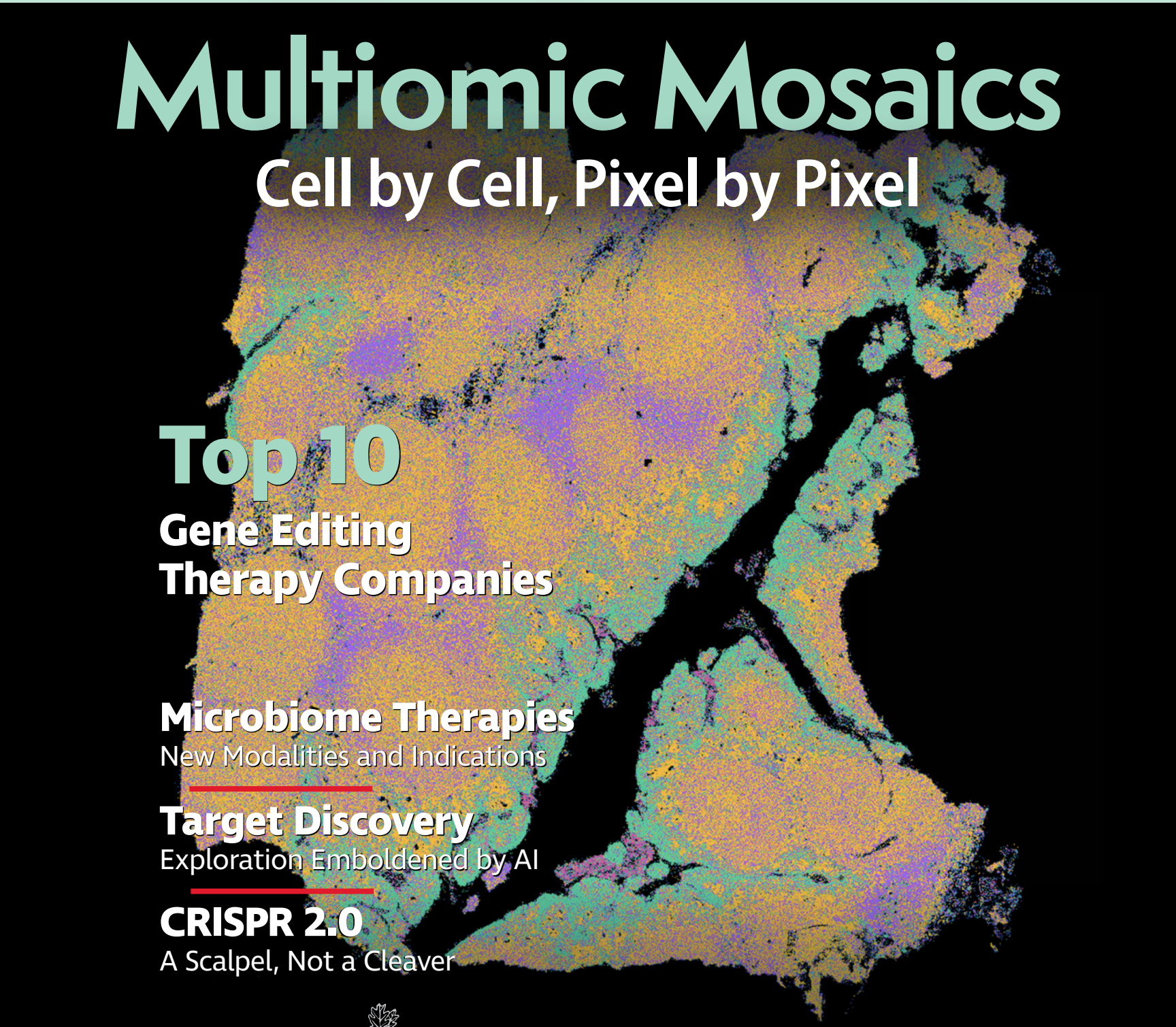
GEN Genetic Engineering & Biotechnology News

JUNE 2024

Multiomic Mosaics

Cell by Cell, Pixel by Pixel



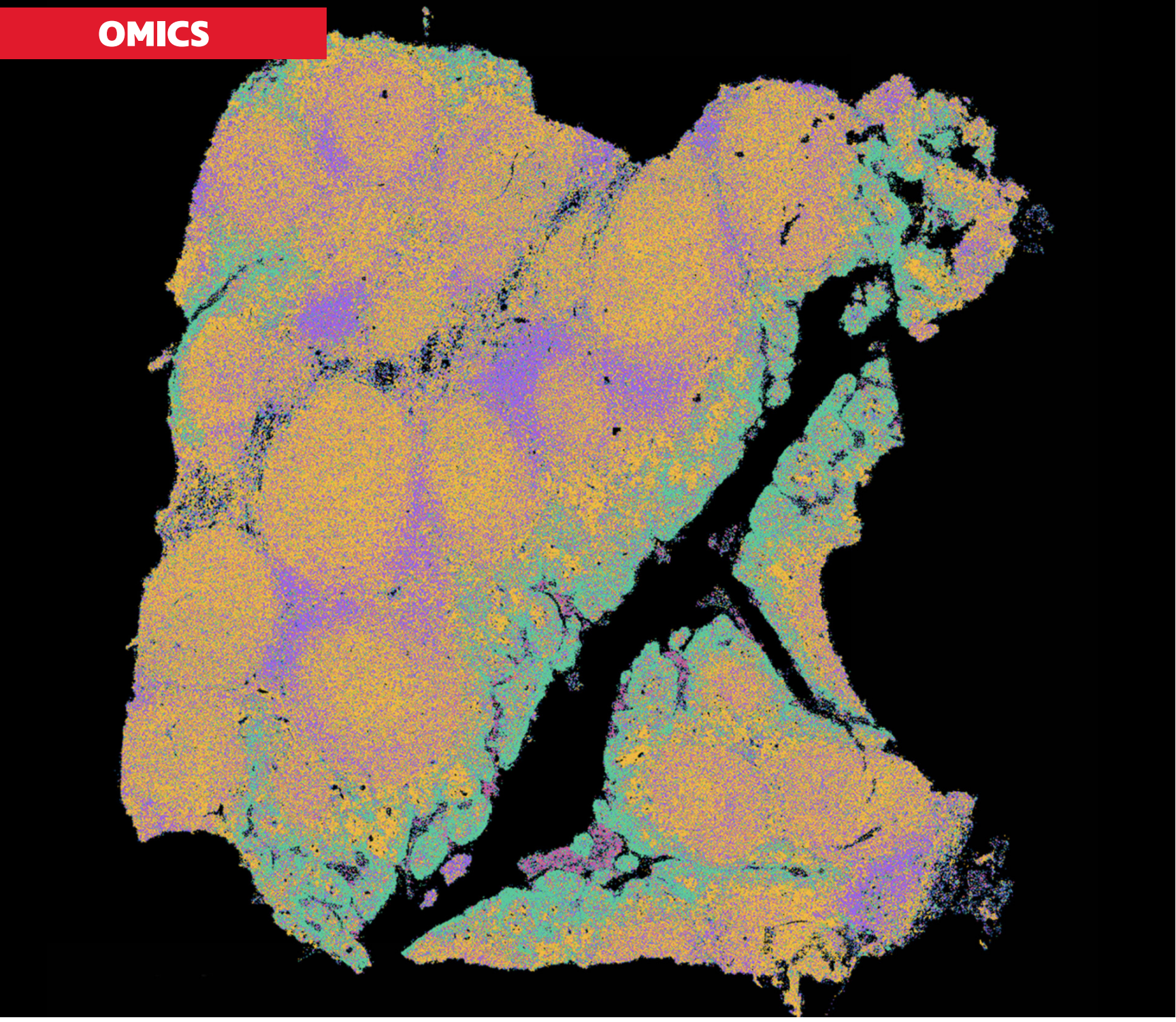
Top 10 Gene Editing Therapy Companies

Microbiome Therapies
New Modalities and Indications

Target Discovery
Exploration Emboldened by AI

CRISPR 2.0
A Scalpel, Not a Cleaver

Mary Ann Liebert, Inc. publishers



Multiomic Mosaics

Cell by Cell, Pixel by Pixel

By Mike May, PhD

Single-cell analysis and spatial omics technologies are being combined in various ways to distinguish biomolecular processes and determine where they take place

Investigating the heterogeneity of biology is difficult. Basic methods, such as the averaging of data, can hide variability. Even fairly advanced methods, such as bulk sequencing, may do the same. As the "bulk" in bulk sequencing implies, it combines cells of varied types into a single sample for analysis, accomplishing another kind of averaging. Fortunately, investigations of biological heterogeneity are starting to take advantage of single-cell sequencing and spatial omics technologies. Indeed, biological heterogeneity has fewer places to hide now that these technologies can penetrate different "omes" such as the genome, the transcriptome, and the proteome. There are even technologies that encompass multiple omes.

Care must be taken, however, to ensure that biological heterogeneity won't be missed simply because the technologies for probing it are overlooked or underused. Several such technologies are considered in this article. Most are suitable for in-house deployment. Some may be more conveniently accessed via service providers. In either case, they can help investigators overthrow the tyranny of averages that reigns over so many fields in biology.

Separating cells

Before working on single cells, scientists must collect them. Some techniques separate cells with high pressure, and others use tags to enable the identification and extraction of specific cells. Scientists at **LevitasBio** in Menlo Park, CA, have developed a platform that uses a lighter touch. The platform is called LeviCell. (This name, like the name of the company, takes "levi" from "levitation.")

"[LeviCell] uses magnetic fields to enrich cells without directly labeling the cells or using any other direct manipulation of the cells," says Kevin Travers, PhD, senior vice president of R&D at LevitasBio. "This is a very gentle process."

A cell sample and a paramagnetic solution—basically a weakly magnetic concoction—are introduced to a LeviCell cartridge, directed down a separation channel, and subjected to a "magnetic density field." In the separation channel, cells of different types levitate at different heights.

"It's really an indirect effect," Travers emphasizes. "The cells are different from the solution, and they get pushed up and away from the edges of the magnetic levitation chamber." A real-time view lets a scientist observe the cells as they levitate within the LeviCell platform and then determine how the cells get separated into wells.

"The flow channels in our system are very large, which allows us to manipulate very large objects," Travers explains. "So, the same technology can work with everything from bacterial cells up to things as large as organoids." The system is designed to collect objects that are up to 350 µm in diameter. Moreover, LeviCell works with a wide range of cell numbers. Scientists at LevitasBio have successfully

Left. In February, Singular Genomics Systems revealed its G4X Spatial Sequencer at the Advances of Genome Biology and Technology conference in Orlando, FL. The G4X is a high-throughput in situ platform that can simultaneously provide direct RNA sequencing, targeted transcriptomics and proteomics, and fluorescent H&E from formalin-fixed, paraffin-embedded tissues. The platform produced this image, which shows the distribution of RNA transcripts in a sample of tonsil tissue.

In July 2023, contract research organization Single Cell Discoveries moved from incubator space at the Hubrecht Institute to its own facility in Utrecht, the Netherlands. The company's new sequencing laboratory includes a NovaSeq X Plus from Illumina, and a Chromium X from 10x Genomics (*shown here*). With its collection of platforms, Single Cell Discoveries can sequence RNA from single cells at a range of scales.

worked with as little as a few thousand and up to millions of cells.

From these samples, scientists can then "analyze the gene expression state of cells in as close to a native state as possible," Travers asserts. "It ultimately allows researchers to increase their confidence that what they're looking at is the cleanest data possible by removing the noise that comes along with dead cells and debris that are inherent to a lot of these cellular samples."

Sequencing single cells

After collecting cells, a scientist might want to sequence them one by one. This can focus on single-cell DNA sequencing or RNA sequencing (scDNA-seq and scRNA-seq, respectively). That's not easy to do, especially at high throughput. Instead of performing those processes in-house, a company might work with a contract research organization (CRO).

For example, **Single Cell Discoveries**—a CRO located in Utrecht, the Netherlands—operates a laboratory developed specifically for single-cell sequencing. "We are platform agnostic and will offer the highest quality and most requested single-cell technologies," says Dylan Mooijman, PhD, head of R&D at Single Cell Discoveries. "Our technologies are divided between 384-well-plate-based scRNA-seq methods that we develop ourselves—SORT-seq and VASA-seq—and commercially available high-throughput scRNA-seq methods, such as those from **10x Genomics**, **Parse Biosciences**, and **Scale Biosciences**."

Using fluorescence-activated cell sorting (FACS), Single Cell Discoveries puts one cell in each well of a plate. "Afterward, the RNA from these single cells can either be directly reverse transcribed or manipulated to yield either 3′ (SORT-seq) or full length (VASA-seq) single-cell RNA sequencing," Mooijman explains. "High-throughput scRNA-seq methods are either droplet based or rely on fixation and combinatorial barcoding." He adds, "VASA-seq allows for the detection of the full transcriptome," including small nuclear RNAs (snRNAs) and small nucleolar RNAs (snoRNAs).

Single Cell Discoveries can sequence various quantities of RNA in as many as a million cells in just one run. "High-throughput methods generally have a slightly lower sensitivity but make up for it in sheer numbers," Mooijman notes. "These high-throughput methods ensure coverage sufficient to sequence multiple guide RNAs targeting about 16,000 genes in a single-cell experiment."

In many situations, scientists want to explore the immune system in basic and applied research. To help with this, Single Cell Discoveries is developing a method that utilizes 384-well plates to profile a sample's immune cells. With this method, scientists will be able to sequence full-length T-cell and B-cell receptors from a small number of cells.

"Many immune cell–selection methods end up with only 100 to 300 cells, which is not sufficient for any current immune profiling method," Mooijman points out. "By offering a method tailored to this need, we hope to further research on antibody discovery and antigen-specific T-cell receptor research."

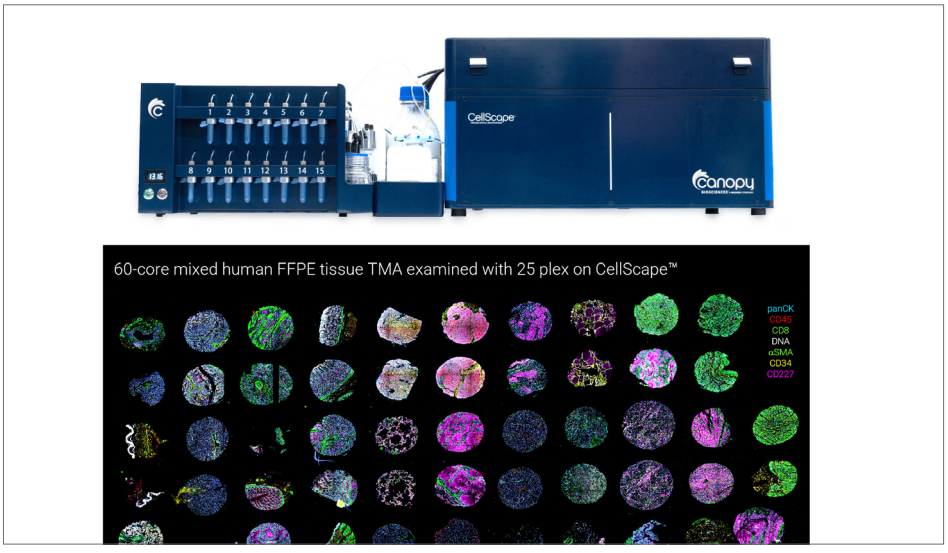
Adding imaging to omics

It's one thing to analyze a single cell's omics, and another to simultaneously determine that cell's location. To do both, scientists must add some form of imaging.

As one example, **Singular Genomics**, based in San Diego, CA, recently announced its G4X Spatial Sequencing Platform. To offer this technology, the company combined its sequencing-by-synthesis (SBS) chemistry and high-speed imaging platform with novel methods to enable spatially resolved sequencing inside of cells and tissue. "We've figured out how to apply SBS sequencing to in situ read out of RNA transcripts and protein in a FFPE tissue section at high throughput with resolution of half a micron." says Drew Spaventa, CEO of Singular Genomics. Because the G4X can collect more than two billion pixels per second while covering an imaging area of 40 cm^2, it can process dozens of samples in a single day.

"By sequencing RNA in situ, researchers

Continued on page 31



In April, Canopy Biosciences launched its CellScape Whole-Slide Imaging Chamber, which works with the company's CellScape Precise Spatial Multiplexing platform. In the example shown here, scientists used CellScape to provide high-resolution imaging of a collection of targets in human tissue microarrays.

Simple, Scalable Absolute Concentrations in Untargeted Metabolomics

Rapid determination of the concentrations of endogenous metabolites in biological systems is of unquestionable scientific value across applications such as drug discovery and development, microbiome science, synthetic biology, health, nutrition, and agriculture. Yet, it remains a holy-grail problem in the analytical sciences.

Absolute quantitation offers objective measurements that connect directly to translational biology, kinetics, and phenotype in a way that relative quantitation simply does not. Furthermore, grounding measurements in absolute concentration units inherently provides seamless comparability across data sets acquired at different times, on different instruments and across different studies or experiments.

To address these challenges, we harnessed recent breakthroughs in machine learning to reimagine how LC-MS–based untargeted metabolomics is performed. Our approach enlists large-data semantic models and transformer-based architectures, to learn the quantitative relationship between raw MS data and the concentrations of the molecules present in the sample.

Pyxis is an AI-enabled technology for absolute quantitation of untargeted analytes that: (1) standardizes analysis to remove custom method development; (2) allows non-experts to rapidly determine absolute analyte concentration by eliminating highly specialized tasks; and (3) return concentrations for an increasingly broad set of analytes.

Quantitative Metabolomics—The Problem

LC-MS is a spectacularly powerful tool for the detection and identification of biologically critical metabolites. While peaks in an untargeted LC-MS chromatogram can be associated directly with molecular identity, integrated peak area is only indirectly associated with concentration.

The functional output differences between relative and absolute quantitation are demonstrated in *Figure 1*—which shows relative quantitation in panel A and absolute concentration in panel B. The x-axis in both panels is the true concentration of 50 analytes. In *Figure 1A*, changes in concentration are captured as increases in peak area, yet both the experimentally observed abundance (peak area), and the scale of the change (slope), are deeply analyte dependent.

In contrast, the y-axis in *Figure 1B* is in micromoles per liter (µM)–a universal and reproducible scale. Even more powerful than the experimental utility of global comparability, absolute metabolite concentration is a deeply biologically relevant quantity. Absolute concentration can be used to gain better understanding of enzyme kinetics to model pathways and in the development and validation of mathematical models of metabolism.

While the power in directly reporting absolute concentrations of metabolites is evident, in conventional workflows, converting peak area to concentration is achieved by explicit analyte-by-analyte calibration. Absolute concentration is thus limited to narrow panels of selected molecules for which a method is developed, and standards are available.

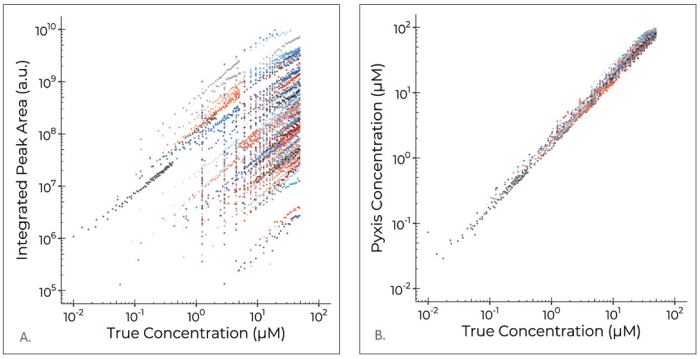


Figure 1. Pyxis absolute concentrations (*right*, B) versus LC-MS peak areas (*left*, A) for a 4-log concentration range spanned by 50 untargeted analytes selected from primary and secondary cellular metabolism. For each analyte, peak area scales with concentration, but the quantitative relationship between area and concentration varies >1000-fold for different analytes, due to structure-dependent MS ionization efficiency, making inter-analyte comparisons generally impossible from relative area analyses. Pyxis employs a machine-learned model of MS ionization as a functional of molecular structure to provide absolute analyte concentration directly from the raw LC-MS signal, without the need for analyte-matched calibration standards or peak area integration.

Introducing Pyxis

The first application of the Pyxis technology is focused on the identification and quantitation of polar metabolites. It includes five critical technology pillars—(1) universal calibrators (StandardCandles™) to represent chemical space of polar metabolites, (2) a turn-key LC-MS method optimized for speed and broad, sensitive detection, (3) an immense, well labeled internal training data set used as input for (4) the unique Pyxis ML model, and (5) the secure cloud-based software platform.

The model is constructed to predict absolute concentration based on generic chemical structure, not specific analyte behavior and is thus functionally an untargeted assay designed to identify and quantify a broad range of analytes, rather than a targeted assay where the results are applicable to a small subset of analytes.

Matterworks has recognized the immense power of absolute quantitation to metabolomics and has reimagined what is possible from raw LC-MS data. In a truly interdisciplinary effort, we are taking powerful AI techniques developed for language and image processing and applying them to LC-MS data, enabling the direct transformation of uninterpreted, raw LC-MS data to a biologically actionable list of the identities and concentrations of detected metabolites.

Get to Your Data Faster: Overcoming Challenges in Spatial Biology

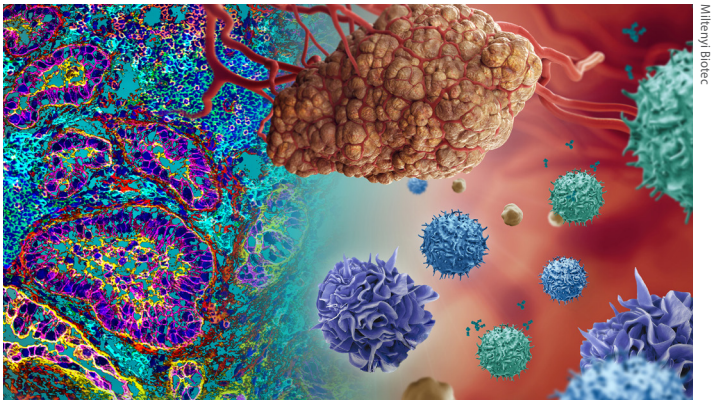
The power of spatial biology utilizing a variety of "omics" is undeniable. Scientists across the globe have demonstrated its utility in tissue mapping, deep phenotyping, and biomarker discovery to gain a deeper understanding of human health and disease. There are certainly advantages to applying multiple spatial omics to pose relevant research questions. While spatial transcriptomics gives us a view of what appears to be happening, spatial proteomics provides additional context on what makes it happen by profiling protein expression to elucidate where cells are located in the tissue, their biomarker co-expression patterns, and how they organize and interact to influence the tissue microenvironment.

Solid tumor cancers are particularly challenging as they can contain many different cell types, both malignant and nonmalignant, in a dynamic local environment. To better understand the role of these cells in cancer biology, a thorough characterization of the spatial context of the heterogeneous tumor microenvironment (TME) is needed. This requires a vast number of markers to identify the location and relationships between immune infiltrates, tumor-specific markers, and structural components of the tumor.

The MACSima™ Platform was developed to address the challenges of navigating complex tissue environments such as the TME. This system is unique in its ability to automatically stain and image a virtually unlimited number of targets using MACSima Imaging Cyclic Staining (MICS) technology. This nondestructive approach enables the visualization of hundreds of markers while leaving the tissue intact for additional staining or downstream applications.

Antibodies and panels

While traditional immunohistochemistry or immunofluorescence relies on just a few antibodies, high-plexed immunofluorescence requires up to dozens of antibodies that can stain a particular tissue. Before an experiment even begins, one must optimize the individual antibodies and panel. Miltenyi Biotec provides a broad array of hundreds of immunologically relevant antibodies that are qualified for formalin-fixed, paraffin-embedded (FFPE) and frozen tissues on the MACSima System. These reagents are directly conjugated with common fluorochromes and are commercially



Miltenyi Biotec

available as individual antibodies or preconfigured plates. Panel optimization may take mere days or weeks, empowering scientists to begin experiments sooner. The MACSima Platform is an open system and can readily be used with antibodies from other sources.

The REAscreen™ Immuno-oncology Kit includes a convenient and standardized staining panel dried down in a 96-well plate. With 61 essential markers, this predefined panel for human FFPE samples enables the identification of 12 potential immune cell subsets within the TME as well as the activation or checkpoint status of those populations. These subsets consist of immune cells and tumor stroma—including blood and lymphatic vessels—and malignant epithelial cell populations, along with their proliferative or apoptotic state. To complement this proteomic panel, Miltenyi Biotec recently developed the 24-plex RNAsky™ IO Explore Panel to enable true spatial multiomics with the detection of both protein and RNA on the same tissue slice. This investigative panel consists of well-characterized immuno-oncology markers that are broadly applicable across various cancers.

Multiplex data analysis

Hyperplexed imaging experiments generate large amounts of data that contain a treasure trove of complex information. MACS® iQ View Software is an advanced informatics tool specifically designed for spatial analysis. This user-accessible data analysis package enables interpretation of large, multidimensional data sets and allows deep dives into the spatial relationships among multiple markers. It provides effective cell segmentation, flexible gating, and advanced plotting tools that help to reveal insights held within large and complex data sets.

In conclusion, many of the challenges in spatial biology can be addressed using tools that increase the availability of markers that can be analyzed, reduce panel optimization time, and enable simple, yet sophisticated, data analysis.

Visit the **MACSima Spatial Biology Homepage** to learn more.

Miltenyi Biotec

can analyze gene fusions, single nucleotide polymorphisms, and insertions/deletions in spatial context with other multiomic data," Spaventa remarks. "This application will be particularly powerful for cancer and immunology research, providing a more comprehensive picture of the tumor microenvironment and cell interactions." He also asserts that the unprecedented throughput of the system is uniquely suited for large-scale retrospective spatial characterization of tissue banks for new biomarker identification and therapeutic response stratification.

Locating protein biomarkers

In addition to technologies for determining DNA and RNA locations, there are technologies for determining protein locations. For example, a protein-localization technology called CellScape has been developed by **Canopy Biosciences**, a **Bruker** company in Saint Louis, MO. "CellScape is an imaging platform for precise spatial multiplexing," says Thomas Campbell, PhD, Canopy's associate director of product management.

"Historically, if people wanted to look at many protein biomarkers on a tissue sample, you had two options," Campbell continues. "You could take multiple sections of your tissue and look at two or three markers per section, or you could grind that tissue up and then do a higher-plex assay, such as flow cytometry, but then you would lose spatial context—how different cells are interacting with each other, and how they're spatially distributed within the sample."

With CellScape, a scientist can analyze dozens and dozens of protein biomarkers in one tissue section, all without disrupting any spatial information. "And we can do that with single-cell resolution," Campbell asserts. "So, you can phenotype every individual cell really deeply and start to understand all the different cell types and subtypes that are present within a sample and how they're interacting with each other—all in a really quantitative fashion."

To accomplish that, CellScape uses the company's proprietary high dynamic range (HDR) image acquisition, which provides an 8-log detection range and is combined with a digital resolution of 182 nm/pixel. "We take the same image multiple times using different exposure times," Campbell explains. "That allows us to quantify really bright signals using short exposure times and to quantify really dim signals using long exposure times. We can create one composite image that has a broader dynamic range and enables you to truly quantify both high-expressing and low-expressing cells within the same sample, and to really differentiate cells in situ across the entire expression range." Plus, because the workflow is nondestructive to the sample, the same tissue section can be reexamined again the next week, month, or even year to study additional protein biomarkers.

From this combination of capabilites, says Campbell, "We have data that is more representative of the biology that's happening, which will help fuel discoveries across a broad range of applications." GEN

A Universal Framework for Spatial Biology

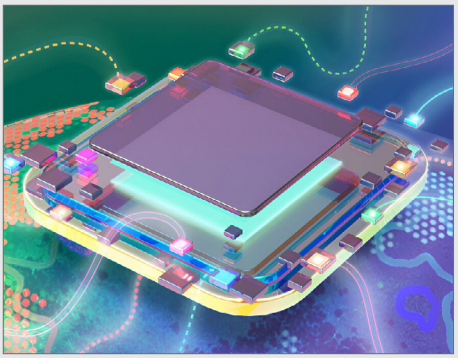
Spatial omics technologies can give detailed information about the molecular makeup of individual cells and their spatial arrangement. However, different spatial omics technologies focus on different characteristics of a cell, such as RNA or protein levels, and they generate heterogenous datasets.

To integrate different forms of spatial omics data—and permit holistic insights into health and disease—a collaborative project led by Oliver Stegle, PhD, of the European Molecular Biology Laboratory developed SpatialData, a data standard and software framework. Stegle and colleagues recently shared details about SpatialData in a *Nature Methods* article (DOI: 10.1038/s41592-024-02212-x).

"SpatialData [is] a flexible, community standards-based framework for storage, processing, and annotation of data from virtually any spatial omics technology available to date," the article's authors wrote. "The ability to flexibly create common coordinate systems by aligning datasets is critical [unlocking] new analysis approaches that facilitate robust comparison and reuse of samples across studies."

The Stegle group used SpatialData to reanalyze a multimodal breast cancer dataset generated by 10x Genomics technologies as a proof of concept. The study, which processed two in situ sequencing datasets generated by Xenium and one spatial transcriptomics dataset generated by Visium CytAssist, demonstrated the complementary nature of these technologies.

In the future, a patient's tumor might be analyzed with different technologies commonly used in the clinic, with the data then unified by SpatialData to gain a holistic understanding of the tumor. Furthermore, the interactive interface would allow the doctor to annotate the data, thus enabling detailed analysis of specific tumor regions and characteristics, potentially leading to personalized treatment approaches. ■



Isabel Romero Calvo/EMBL (CC BY-NC-SA)