Cell Systems

CellPress

Voices

What is the main bottleneck in deriving biological understanding from spatial transcriptomic profiling?

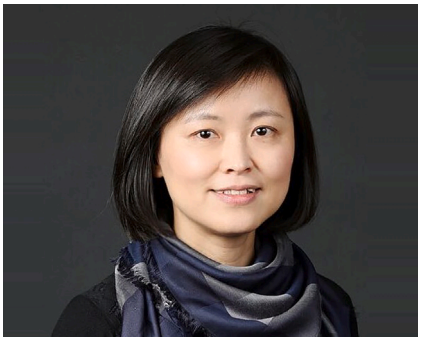
Jeffrey R. Moffitt
Harvard Medical School

Quality in, quality out

Image-based approaches to single-cell transcriptomics offer the exciting ability to directly image gene expression at the transcriptome scale within intact tissue slices, defining, discovering, and mapping the cells within. The discovery promise of these methods emerges, in part, from their merger of two powerful but historically distinct methods: high-resolution optical microscopy—which provides intracellular organization, morphology, and tissue organization of cells—and genomic-methods—which quantify the transcriptome-scale molecular complexity of cells.

Yet in merging these strengths, spatial transcriptomic methods also merge their challenges. High quality measurements require both good histological protocols—intact tissue slices, strong coverslip adherence, and minimal cellular morphology perturbations—and good RNA preservation—limited fragmentation of this easily degraded molecule. Yet best practices—both in histology and RNA preservation—vary from tissue to tissue, and most histological preparation methods were not developed to preserve RNA. Critically, variation in sample quality can dramatically shape the biological resolution of these measurements. Low quality measurements might define only coarse cell types (e.g., lymphoid versus myeloid) while high quality measurements promise not just more subtle cell type divisions (e.g., cDC1 vs. cDC2) but different activation states within.

The challenge is that established, best preparation practices for a wide range of tissue types simply do not yet exist and optimization efforts face imperfect sample and measurement quality metrics. Nonetheless, these challenges are most likely temporary. As more researchers adopt these methods, good practices will emerge. An openness to share not just the protocols that worked but also the ones that did not will speed this process. Until then, researchers looking to leverage these methods to answer questions in their tissue of interest should be prepared to roll up their sleeves and help establish these tissue-specific best preparation practices. The reward is the biological discoveries only high-quality data can provide.



Mingyao Li
University of Pennsylvania Perelman School of Medicine

Overcoming tissue size limitation

Since being named "Method of the Year" in 2020 by *Nature Methods*, spatial transcriptomics (ST) has revolutionized biomedical research by providing spatially resolved molecular insights beyond the reach of traditional methods. Current ST platforms fall into two categories: sequencing based and imaging based. Sequencing-based platforms such as Visium and Visium HD offer whole-transcriptome coverage but are limited by small tissue capture areas. In contrast, imaging-based platforms such as Xenium, MERSCOPE, and CosMx analyze larger areas but sacrifice gene multiplexing.

To advance biomedical research, there is a growing need for platforms capable of analyzing large tissue samples—particularly in human studies—with single-cell resolution, whole-transcriptome coverage, and cost efficiency. A promising solution lies in high-resolution histology images, which are readily available in most ST datasets. Whole-slide images, spanning up to 2.5 cm by 7.5 cm, are significantly cheaper than ST platforms and strongly correlate with spatial molecular patterns.

Computational approaches are essential for addressing the tissue capture area limitations of ST. By integrating histology images with ST data, these methods can extend spatial measurements beyond the physical capture areas of current ST platforms while enhancing spatial resolution and gene coverage. These innovations not only maximize the utility of ST data but also enable cost-effective large-scale studies, significantly

expanding our ability to extract meaningful biological insights and shaping the future of ST.

Qing Nie
University of California, Irvine

Uncovering spatial information flows

In multicellular tissues and organs, the intricate and dynamic flow of spatial information from one region to another drives the diversity and complexity of biological functions. Tissue function is governed not only by the diverse cell states and their spatial arrangements within local regions but also by the intercellular interactions. Through ligand-receptor interactions, communication signals diffuse from cells to cells, regulating downstream genes and creating new signals that propagate to neighbor cells. How do these inflows and outflows of cellular signals draw the patterns of spatial communication flow that affect biological functions? Reconstruction of such ''cause and effect'' flows using single-cell and spatial omics data (e.g., via FlowSig analysis) has started to address some of the questions. Uncovering how spatial communication flows evolve over time or vary under different biological conditions remains a major challenge.

Beyond short-range ligand-receptor interactions, long-range signals, such as metabolites and hormones, transported through the body fluid circulation create a larger-scale spatial communication flow. Identifying metabolic communication flows from spatial multi-omics data requires incorporation of tissue structure and transport mechanisms based on prior knowledge. On an even larger spatial scale, molecular and cellular information may be transported from organ to organ, leading to communication flows that go beyond individual tissues. Spatial omics data present unprecedented opportunities and challenges for scrutinizing information flows across multiple spatio-temporal scales.

Kazumasa Kanemaru and Sarah A. Teichmann
University of Cambridge

Life in 3D: Insights and challenges

Spatial transcriptomics encompasses a suite of methods that preserves spatial context while interrogating the entire transcriptome. These techniques enable us to pinpoint specific cell types and gain an integrated view of tissue anatomy. The bottlenecks are the flip side of the coin to the excitement. Chief among them is the need for deep biological expertise—from experimental design to data analysis—paired with robust computational skills. When those elements align, spatial transcriptomics can yield remarkable insights; without them, deriving biological meaning becomes a challenge.

One technical challenge is that spatial transcriptomics platforms rely on thin sections with limited fields of view (e.g., 1 cm^2). Even microanatomical structures often exceed this volume, so the assembly of multiple sections is needed to gain a full picture of the biology of the tissue niche, the tissue, or the organ. We addressed this by assembling planar sections as a first approximation with the VisiumStitcher (PMID: 38057666; https://github.com/Teichlab/visium_stitcher) and using quasi-3D modeling via a symmetry approximation to build a whole-thymus cell atlas with our OrganAxis approach (PMID: 39567784). Once we overcome the technical hurdles of assembling a full 3D cell atlas for tissues and organs, the next challenge is interpreting molecular and cellular mechanisms. The most obvious opportunity is to decipher paracrine signaling between adjacent cells, as illustrated recently through our CellPhoneDB.org application to the sinoatrial node of the human heart (PMID: 37438528).

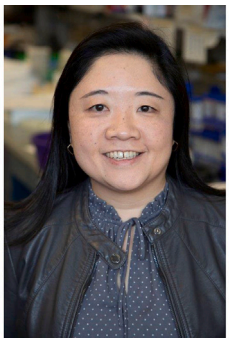
Daniel Dar
Weizmann Institute of Science

Spatial transcriptomics meets microbial complexity

Spatial transcriptomics offers transformative insights into microbial behaviors, from biofilm formation to microbiome structure-function relationships, *in situ*. Yet, adapting these technologies, originally designed for mammalian systems, to the world's tiniest and mightiest organisms presents significant challenges. Bacteria lack mRNA polyadenylation, complicating transcript capture. Their small size—1/1000th the volume of a mammalian cell—means few transcripts per cell, further shielded by sturdy cell walls. Microbial communities are also compact, spanning only tens of microns and often comprising numerous species, complicating spatial analysis.

While current multiplexed imaging techniques achieve single-molecule resolution and can profile a few hundreds of genes at the relevant physical scale, it falls short of the comprehensive insights provided by RNA sequencing. Applying these techniques in natural microbial communities, where genomic sequences are often unknown, adds another fundamental challenge. These limitations create a critical bottleneck: how to infer global transcriptional states from limited spatial data.

One current solution is leveraging marker genes to report on essential biological states, such as growth rate, stress responses, or antimicrobial/anti-eukaryotic activity. Such markers can provide a coarse yet invaluable snapshot of microbial and host-microbe interactions. Alternatively, two-step strategies that combine single-cell dissociation and sequencing with targeted spatial analysis offer another potential path to reconstruct specific contexts. Overcoming these challenges with widely accessible technologies will help unlock a new understanding of microbial systems.



Luciane T. Kagohara
Johns Hopkins University School of Medicine

Imprecise cell boundaries

Spatial transcriptomics profiling has emerged as a cutting-edge technology that allows for gene expression analysis while maintaining tissue architecture. This capability has revolutionized the understanding of complex and heterogeneous diseases where intercellular interactions are crucial, such as in cancer. Despite continuous progress in increasing spatial resolution, a significant limitation persists: the inability to precisely detect individual cells' boundaries. This bottleneck hinders the granularity and specificity of biological interpretation derived from spatial transcriptomics.

The lack of cell boundaries detection results in the blending of gene expression signals across neighboring cells captured by the same or adjacent spatial spots, blending the contributions of individual cellular types and states. The ambiguous data interpretation and biased inference of cell states and intercellular interactions limit the discovery of relevant mechanisms that are cell type specific. Consequently, detailed individual cellular heterogeneity, which is crucial for understanding complex biological processes and disease mechanisms, remains unclear. While some approaches can deliver whole transcriptome profiling without individual cell detection, others allow relatively accurate cell segmentation but with limited transcript measurement.

Therefore, experimental developments are needed to combine both: whole transcriptome profiling with true single-cell detection. Despite quick advancements in computational methods and imaging approaches for spatial spot deconvolution, they are not perfect as they still rely on the spatial data generated or single-cell sequencing datasets for reference. Overcoming this bottleneck in spatial transcriptomics is fundamental to unlocking the full potential of this technology and enabling accurate understanding of the spatial dynamics of gene expression in health and disease.

Shalev Itzkovitz
Weizmann Institute of Science

Exploring the third dimension

Spatial transcriptomics is a powerful tool for reconstructing high-resolution expression profiles across defined regions, typically spanning several millimeters. This technology allows for detailed snapshots of spatial gene expression in uniform tissues such as the healthy liver and intestine. However, its application is limited when dealing with heterogeneous tissues, such as tumors or inflamed areas, which include necrotic cores, fibrotic zones, and inflammatory infiltrates spanning large distances. In such tissues, distances to landmarks such as blood vessels and immune infiltrates might strongly dictate gene expression but can be misrepresented without considering the three-dimensional context. Furthermore, many tissues consist of repeating 3D structures (such as thyroid glands and pancreatic islets of Langerhans) that are not captured in thin 2D sections. This omission can overlook critical factors, such as the overall size of these units, that may influence gene expression patterns.

By extending spatial transcriptomics to encompass three dimensions, we can significantly enhance the granularity of data obtained even from small samples, such as biopsies typically measuring 1–2 cm in diameter. This advancement would not only refine our understanding of morphology, such as unit sizes, but also provide accurate measurements of distances to critical tissue landmarks. Since applying current spatial transcriptomics on serial sections would be prohibitive in terms of price and potential batch effects, advancements in techniques such as tissue clearing and probe permeabilization, alongside innovative new methods, are essential for developing reliable and high-resolution three-dimensional gene expression maps.

Roser Vento-Tormo
Wellcome Sanger Institute

Defining biological questions

Recent experimental and computational advances enable the profiling of cells within their tissue context, driving new discoveries in how cells interact and make functional structures.

But the field faces several significant challenges. Most high-throughput methods are designed for 2D analysis, whereas biological mechanisms operate in 3D. Overlapping cells in 3D space may thus appear as single entities in 2D, leading to misidentification or the inclusion of doublets when segmenting cells. Another challenge is that spatial transcriptomics samples cover only small snapshots of the organ, so multiple sections must be integrated to gain a comprehensive view. While we and others (PMID: 39567784; https://doi.org/10.1101/2024.10.30.621114) have used common coordinate frameworks to map cells along tissue axes, these are limited in their generalizability across different samples.

A final and fundamental issue is the overwhelming volume of data generated in these studies, which makes it hard to decide what phenomena to focus on. To address this, it is essential to define research questions before your experimental and computational plans begin. This helps streamline hypothesis generation and exploration. Interdisciplinary collaboration can help ensure that biological questions drive the selection of technologies and the design of subsequent analyses. By tailoring tools to specific biological questions, researchers can maximize the potential of spatial transcriptomics to deepen our understanding of complex cellular systems.

Itai Yanai
NYU School of Medicine

Exploratory data analysis

Spatial transcriptomics datasets are rich and vast but teasing out novel discoveries from them requires an attitude of exploration. This is because—beyond "day science" hypothesis testing—spatial transcriptomics allow us to apply "night science" in the form of exploratory data analysis, as originally championed by John Tukey in the 1970s. While many useful standard pipelines and packages are available and are being developed for this purpose, we must also explore the data directly: instead of going straight to high-level summaries (averages of averages), we must visualize intermediate steps and meditate on how things look.

This process has no straightforward protocol for discovery in spatial transcriptomics datasets; there are only analysis tools and operations (select, characterize, score, relate, and cluster), and applying each one can lead to a lot of thought on what the result means and what might be the next step. As we explore, we look for new patterns and imagine what may explain them, and we also consider what is absent; that is, what expectations are being violated by the data?

By exploring a dataset, we recognize that it does not come with labels marking what is new and exciting about it. Figuring that out is not simple and cannot be automated. Rather, discovery is more of an act of self-expression and creativity. Different people will make different discoveries with the same dataset. Overall, taking control of the creative process in analyzing a large dataset actually means losing control of any initial direction and following the data wherever they lead.

Elana J. Fertig
University of Maryland School of Medicine

The 4D nature of tissues

Tissues are multi-scale; function arises from the interplay between cellular and molecular phenotypes. Spatial transcriptomics technologies have unlocked our ability to measure these interactions with a single measurement. Hypothesis testing in human tissue is now possible. Realizing the power of these measurement technologies requires software for their analysis. In many cases, the fundamental biological questions we seek to answer with these technologies lack appropriate computational methods for their analysis. As a result, the invention of new algorithms and software has emerged as a new pillar of biomedical science, requiring the support of the research community.

The next frontier will arise as both the experimental and computational techniques expand to account for the full four-dimensional nature of tissues. I look forward to a future research community that can watch in real time how cellular and molecular landscapes change in time, enabling prediction of the future state of tissues using techniques similar to those employed for weather prediction. The emergence of spatial transcriptomics movies will ultimately enable us to forecast the impact of diverse interventions to personalize continued human health in a new generation of predictive medicine.

Fabian J. Theis
Helmholtz Munich and Technical University of Munich

Unlocking structure in spatial transcriptomics

How can we compare, understand, and also communicate recurring structures in high-dimensional spatial transcriptomics (ST) data? Let's look at the successes of its disassociated counterpart, single-cell RNA sequencing (scRNA-seq). There, understanding and real use (in particular in atlas-scale data) came initially mostly from clustering and annotating cell profiles into groups, denoted as cell types or states. These discrete structures don't exist in biological reality—instead there are gradients, dynamics, and cell plasticity on many scales. But these are extremely useful human constructs to give names to bounded areas in expression space and make sense of the complexity of a 20k-dimensional, sparsely populated vector space. Such metaphors are more difficult to come up with in ST data.

ST comes with two axes of variation: gene expression, akin to scRNA-seq, and spatial organization of cells. These two axes are more difficult to visualize, so to find 'understanding' we need to come up with ways to describe and communicate patterns in the combined spatial-expression space. Just as we decompose images into meaningful, physical objects, we can intuitively interpret groups of cells as structured entities. In order to deal with the fact that we cannot interpret 20k colors at the same time, we

may use the annotated cell state/type cluster labels in ST to concisely define and construct similar local structures called tissue niches or microenvironments.

Current approaches (ours included) mostly rely on fixed-radius niches and ignore scales and continuous structures. The challenge is to develop approaches that infer scale-dependent, potentially continuous data-driven niches of varying size and structure in a way that makes these identifiable, comparable, and sharable across samples and labs. This would provide a structured abstraction for ST images beyond simple aggregation and visual inspection, ultimately making ST as interpretable and actionable as scRNA-seq.

DECLARATION OF INTERESTS

M.L. is a co-founder of OmicPath AI LLC. E.J.F. has been on the scientific advisory board of Resistance Bio and is a paid consultant for Mestag Therapeutics. She is on the advisory board of *Cell Systems*. J.R.M. is a founder of, stakeholder in, and advisor to Vizgen, Inc. J.R.M. is an inventor on patents associated with MERFISH applied for on his behalf by Harvard University and Boston Children's Hospital. S.A.T. is a scientific advisory board member of ForeSite Labs, OMass Therapeutics, Qiagen, Xaira Therapeutics, a co-founder and equity holder of TransitionBio and Ensocell Therapeutics, a non-executive director of 10x Genomics, and a part-time employee of GlaxoSmithKline. F.J.T. consults for Immunai Inc., CytoReason Ltd, Cellarity, BioTuring Inc., and Genbio.AI Inc. and has an ownership interest in Dermagnostix GmbH and Cellarity.