

IN FOCUS

Accelerating Drug Development Using Spatial Multi-omics



Richard J.A. Goodwin¹, Stefan J. Platz², Jorge S. Reis-Filho³, and Simon T. Barry⁴

Summary: Spatial biology approaches enabled by innovations in imaging biomarker platforms and artificial intelligence-enabled data integration and analysis provide an assessment of patient and disease heterogeneity at ever-increasing resolution. The utility of spatial biology data in accelerating drug programs, however, requires balancing exploratory discovery investigations against scalable and clinically applicable spatial biomarker analysis.

INTRODUCTION

Spatial biology is not a new concept. Researchers have been utilizing spatial biology data for over a century, building from classical histology and expanding through to protein or RNA-based markers for target expression, cell phenotype, and biomarkers of efficacy or safety. What has driven the recent revolution in spatial biology is the ability to multiplex the number of measured biomarkers in a single experiment and the computational capabilities that allow us to interrogate the patterns and relationships between biomarkers. This has delivered unprecedented views on complex biology and therapeutic interventions within tissues.

In recent years, advances in technology have allowed the spatial biology field to evolve beyond single omic-specific multiplexing experiments to studies with multimodal imaging biomarker data. This has led to a reinvigoration of the field as researchers begin to layer and cross-correlate increasingly complex biomarker datasets. These data can be used for broad hypothesis-targeted biomarker analysis, delivering improved insight and greater confidence in the conclusions drawn, or for hypothesis generation, revealing novel insights into disease heterogeneity. Here we consider how spatial-omics platforms and data can be utilized most effectively in the medium to long term to inform development of new therapeutic concepts. Key to this is the appropriate deployment and efficient use of spatial biology and associated challenges in scaling and distilling data to actionable insights. We describe the concept of an exploratory and discovery-minded spatial biology “sandbox” that uses “fit for purpose” panels of assays that adapt to constantly evolving analysis platforms.

¹Imaging and Data Analytics, Clinical Pharmacology and Safety Sciences, R&D, AstraZeneca, Cambridge, United Kingdom. ²Clinical Pharmacology and Safety Sciences, R&D, AstraZeneca, Cambridge, United Kingdom. ³Cancer Biomarker Development, Early Oncology, AstraZeneca, Gaithersburg, Maryland. ⁴Bioscience, Early Oncology, AstraZeneca, Cambridge, United Kingdom.

Corresponding Authors: Richard Goodwin, AstraZeneca R&D, The Discovery Centre, Cambridge Biomedical Campus, 1 Francis Crick Avenue, Cambridge CB2 0AA, United Kingdom. E-mail: richard.goodwin@astrazeneca.com; and Simon Barry, simon.t.barry@astrazeneca.com

Cancer Discov 2024;14:620–4

doi: 10.1158/2159-8290.CD-24-0101

©2024 American Association for Cancer Research

Assays are validated for each system but not constrained by cross-platform and interstudy variation. This approach delivers rapid, powerful exploratory capabilities to derive mechanistic insights in early discovery and clinical development. Separately, it can also be used to inform on decision criteria for scalable and deployable clinical diagnostics and supporting treatment selection. Outputs from a spatial biology “sandbox” can also inform and accelerate both classical biomarker selection strategies but also aid the development of artificial intelligence (AI)-based biomarkers.

OMICS DRIVING ONCOLOGY PHARMA RESEARCH AND DEVELOPMENT PROJECTS

Oncology drug development was transformed with the genomic revolution (1); using RNA-based signatures enabled greater segmentation of disease into differentiated functional subgroups, providing additional insights that also allow more faithful alignment of models to specific disease segments (2). Bulk tissue-based multi-omic approaches such as genomics, transcriptomics, and proteomics are now routinely used in developing new treatment concepts (3). Indeed, in all drug development projects, there is now routine application of multi-omics approaches for preclinical and clinical data generation to understand the molecule, define mechanisms, and select patients. However, with the recognition that tumors are heterogeneous (4) and the drive to develop better combination therapy regimens for specific tumor subsets, spatial multi-omics approaches have become transformative, providing the required contextual insights to validate mechanisms of action and probe whether diverse responses are being observed within tumors.

Most biotech and pharma organizations assess projects versus a data framework that seeks to define whether right molecule, right exposure, right patients, and right safety profile criteria are being achieved to enable delivery of a successful project (5). A foundation of this is being able to understand drug levels and distribution in normal tissue or tumor tissue and the impact on proof of mechanism biomarkers informing on target modulation. Over the past 20 years, IHC has been the primary method to explore proof of mechanism biomarkers in tissue. This approach gives limited insights because biomarkers are based upon specific pathway or target biology and may not deliver decisions informed by

a holistic analysis of the tumor. However, advances in multiplex antibody-based imaging platforms allow us to move beyond using a maximum of six to seven specific biomarkers to assess target and complementary pathway biomarkers in context with cellular phenotype and outcome biomarkers. Moreover, multi-omic biomarker guided segmentation of tissues with multiplex antibody-based biomarkers (IHC/multiplex immunofluorescence/multiplexed ion beam imaging by time of flight/imaging mass cytometry), RNA biomarkers (RNAscope), or metabolic biomarkers (mass spectrometry imaging) can guide regional spatial RNA or DNA analysis to assess whether these biomarkers define genomic or transcriptional heterogeneity within the tumor. Routinely deploying these technologies in parallel as a component of a spatial-omics platform has the power to deliver unique insights but also to give increased confidence in the conclusions or decisions.

DEPLOYING SPATIAL BIOLOGY OMICS PLATFORMS AND STRATEGIES

The choice of technologies available to probe or image tissues, *in situ*, for a spectrum of omic biomarkers is extensive and rapidly developing. Utility of any omics platform is determined by technical parameters including capacity, speed, robustness, resolution of analysis, and measurements as direct or surrogate using a probe (which itself might be direct or amplified) and have been regularly reviewed (6). As a technology becomes more routinely deployed, it offers the opportunity to apply them in much more complex ways, for example, building up 3D images of a tumor from sequential sectioning (7) or integrating multiple platforms together. We also need to continue into other dimensions beyond 3D, namely a measurement of temporal spatial and flux. It is important to be able to link these increasingly complex and high-resolution images of biology to outcome data or to other biomarkers that can be more routinely readily measured over longer time periods. This can be achieved using established or emerging technology or approaches; however, considering the compatibility of different platforms should not be ignored. How we develop a spatial biology view of disease and link it to a circulating biomarker and to noninvasive clinical imaging will be critical in maximizing the value of omics analysis.

Spatial data and spatial-omics approaches are currently most effective when using either a single platform to analyze large cohort studies or multimodal approaches to deeply interrogate and characterize a selected subset of samples. To have value, the analysis strategy needs to enable the most efficient use of precious samples or reduce study timelines by maximizing the contextualization of different correlative or causative biological events from the same study. There are forward-looking risks that need to be considered. Improving analytic platforms or upgrades to current platforms (including routine factors such as detectors or reagent changes) may mean that when transitioning from proof-of-concept studies to reanalysis or further analysis of separate or larger cohorts may effectively be acquired on different specification machinery requiring constant revalidation. In addition, selection and availability of assays or antibodies may change. For example, the gold standard for a given endpoint may evolve, or reviewers or investigators may have preference to specific anti-

body clones. These transitions may require extensive ongoing validation, data processing, and standardization. Platform-specific data preprocessing, thresholding, automated analysis, and complexity of data integration risk spatial biology studies being used to generate volumes of data and not increase meaningful insight long term. Although the results can be exciting and generate disruptive hypotheses, there is a significant risk that the propagation of bias or errors leads to a lack of interpretability of the terabytes of data collected. It has already been highlighted that researchers face a decision in balancing and prioritizing capacity, resolution, and sensitivity of technology platforms. These considerations must also be weighed against the cost of deployment. It is widely recognized that the deeper or more extensive the analysis, the higher the expense with respect to time, sequencing costs, antibody or reagent usage, and downstream data storage and processing. It is worth noting that there are emerging technologies that are presenting solutions where in operational costs are considerably lower compared with capital expenditure. For instance, imaging techniques using mass spectrometry offer a relatively broad detection of metabolic biomarkers, ranging from small-molecule therapeutics to lipids, peptides, and proteins. The specificity and sensitivity achieved with mass spectrometry imaging (MSI) might not meet the requirements for a specific biomarker assay. However, the compatibility of the MSI to a range of other analysis technologies means it can provide orthogonal data complementing classical morphologic histology and further contextualizing antibody-based biomarker endpoints (8).

If the cost of deploying and running each spatial-omics platform is estimated at approximately \$1,000,000 per system, it is understandable that institutions and academic groups would rely on specific omics platforms and expertise. This strategy enables the generation of more reliable longitudinal data from studies. However, it could restrict insights and limit interpretation because of bias inherent in any technology platform used in isolation. For larger institutions and pharmaceutical companies, there is opportunity to combine spatial-omic platforms to deliver a multimodal spatial biology approach that works across cross-informing technology platforms. However, the complexity of the data generated and the rapid pace of technological advancements make it challenging to define routes to standardized platforms or processes to validated assays and to integrate data to give consistent results between organizations. This challenge, which should be surmountable, is currently throttling the full potential of multi-omic spatial biology beyond exploratory biomarker studies. Therefore, even greater impact on therapeutic development would be fit-for-purpose deployment of multi-modal spatial-omics. Data generated must align to robust decision criteria for clinical decision making. This requires omic and imaging platforms to support deployable assays in routine clinical practice and development of companion diagnostics.

SPATIAL-OMICS “DISCOVERY SANDBOX” FOR MECHANISTIC AND PATIENT SELECTION BIOMARKER INSIGHT

Increasingly, it is becoming clear that spatial biology experiments can deliver impactful studies. For example, multiplex

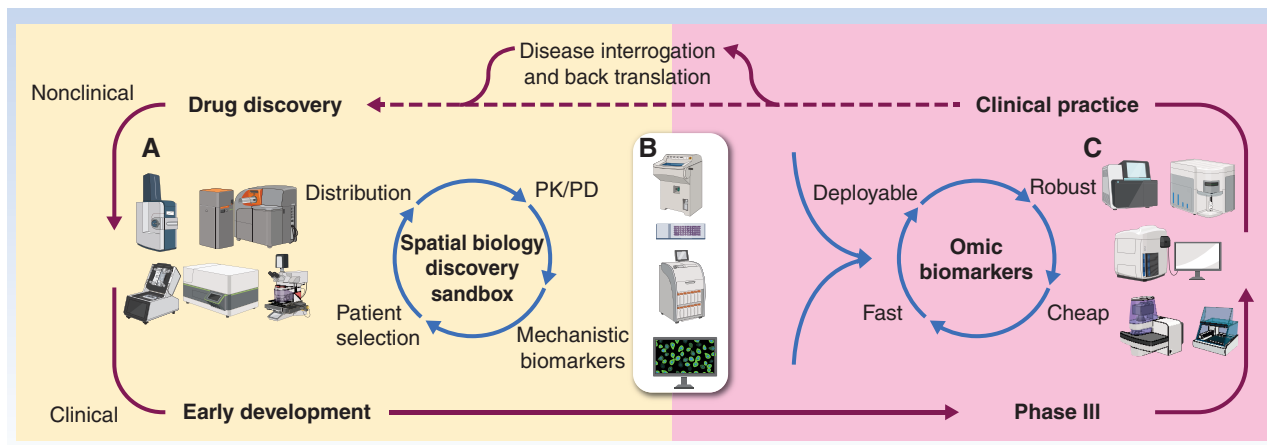


Figure 1. Applying a multi-omics toolbox to therapeutic development. Biomarker insights supporting drug discovery, development, and routine clinical practice have two phases. In discovery and development, the Omics toolbox examines distribution of a molecule in tissues, pharmacokinetic/pharmacodynamic (PKPD) relationships, and mechanistic biomarker changes and develops patient selection hypotheses using preclinical models and clinical trial material (typically from small standalone samples sets). As a therapeutic concept progresses to phase III trials, biomarker assays need to be robust, rapid to use, deployable in large cohorts of patients, and in different labs. For routine clinical utility, biomarkers need to be deployed in general as part of a suite of robust simple biomarker assays that help clinician make choices between therapies. The Omics toolbox can then be reapplied to large cohort of patients treated in phase III or with a routine therapeutic regimen to refine mechanistic understanding, explore resistance, and drive the next round of drug projects or treatment hypotheses. Spatial omics: approaches to generate when multiple spatial biomarkers data. Multi-omics: integrating different single omics platforms. **A**, Systems generating complex omic data (including mass spectrometry imaging, imaging mass cytometry, spatial transcriptomics, laser micro dissection). **B**, Platforms common to tissue processing and analysis (including cryostat and microtome, tissue microarray, tissue processes, image analysis). **C**, Technologies delivering robust, automated data used in clinical practice and phase III (including sequencer, flow cytometer, multiplex imager, autostainer, slides digitization). Image created with BioRender.com.

spatial proteomic characterization for HER2-positive breast tumors has demonstrated robust biomarkers that can enable the stratification of sensitive tumors early during neoadjuvant HER2-targeted therapy (9). In addition, transcriptomics is now performed spatially at a single-cell resolution (10), and platforms such as imaging mass cytometry can acquire high-dimensional images that are also spatially resolved at a single-cell level and able to characterize intratumor phenotypic heterogeneity in a disease-relevant manner (11). The expectation of spatial biology data, especially from multimodal spatial-omic studies, is that they provide such rich orthogonal datasets that it is possible to answer any question or identify “unseen” relationships in the data by applying machine learning, deep learning, and generative AI. To achieve this, there is an important potential role for foundation models (in particular foundation models that include spatial-omics data in addition to histology, clinical data, and large language models) in the standardization of the outputs from different platforms, as well as in the potential reduction of the sample sizes needed for a biomarker signal to be detected (12). This is a future possibility, but in the short term, this expectation can potentially undermine confidence in spatial-omics approaches. Currently, the field does not have datasets with sufficient sample sizes ($n =$), depth (resolution/sensitivity), and endpoints (different biomarker platforms) to move beyond classical scientist-led and interpreted outputs. Therefore, in the near to medium term, the spatial-omics platforms can show value and experience/confidence developed as a broader discovery or investigational “sandbox” (Fig. 1). This allows multiple questions to be asked at different levels of spatial resolution, with orthogonal platform enabling some form of cross validation of relationships derived or to give additional perspective to rapidly refine hypotheses. For drug discovery, this

“sandbox” approach can readily support in nonclinical models through to early clinical phases of drug projects. Moreover, once a therapy is established and well-annotated cohorts of samples are available, it adds insights derived from clinical outcome data to segment responders or nonresponders or understand emerging resistance. In the longer term, as technology becomes more robust and experience with data analysis increases, it would be exciting to envision deployment of a suite of multiplexed spatial-omics assays in conjunction with genomic profiling to guide more personalized treatment options.

As discussed, single-cell omics and similar approaches describe tumor heterogeneity within the tumor cells and tissue microenvironment but are limited by lack of context. An important application of the multi-omic spatial platforms is the ability to provide a different lens on the heterogeneity question and enable targeted analysis to understand the context in different neighborhoods. When coupled with AI digital pathology, it allows us to redefine heterogeneity at a level that goes beyond that classified by pathology or traditional biomarker imaging (13). A key aspect of spatial biology use in drug discovery is the capacity to apply lower-cost, high-throughput scanning technology. This approach can be used to screen across disease models or patient tissue microarrays, pinpointing those samples and regions that necessitate more in-depth tissue profiling, significantly reducing costs and timelines for data production.

Although we can acquire high-dimensionality complex data, it does not mean that the initial analysis needs to incorporate all that data. This is particularly critical when considering multimodal imaging datasets, where in disparities in acquired resolution across the sample or variability in sample quality or reproducibility may arise. Instead, the experimental design should first address the primary research question before

adding more complex data layers. There is also an intersection between the AI-driven exploration of large-scale routine digital pathology and exploratory multimodal spatial biology. How can researchers take digital pathology from large-scale real-world studies, augmented with genomic and outcome data, and use that to target deep spatial and omic analysis from a small cohort to allow us to recapitulate the complex insights back to the large-scale real-world data? This requires approaches that avoid the requirement for spatial biology analysis on everything. As the quality of the datasets, annotation, and metadata improve, AI approaches and foundation models will allow us to achieve this more efficient use of the technology platforms.

WHEN IS ENOUGH EVER ENOUGH? REUSE, RESOLUTION, AND A MORE-PLEX FALLACY

Balancing spatial biology data reuse versus the option to reacquire it using the latest emerging advancement (increase spatial resolution, sensitivity, or higher plex) is also a key question for the community. This is in part driven by the perennial challenge of the confidence in legacy data. Reproducibility and reliability between different platforms are a question multisite science initiatives invest significant effort to mitigate through protocol harmonization and quality control. Analytic systems are launched with ever-increasing spatial resolution or number of biomarkers simultaneously detected used as a clear metric to trumpet the new platforms over legacy or competitor systems. Therefore, researchers need to propose and implement unifying approaches to maintain expensively acquired spatial assays, usually collected in small cohorts, and reuse that data most effectively.

It is desirable to define a standard against which future technology innovations can be accurately compared. Some standalone omic biomarker platforms, for example, genomic technologies, can have a describable ground truth that all similar platforms are driving toward (albeit with more accurate, faster, or more sensitive measurements). Spatial biology encompasses an ever-expanding field with new technology bringing in suites of omic biomarkers or signatures at ever-changing resolutions. Each new platform claims to enable more precise characterization of the tissue microenvironment and cellular interactions; however, there is no definable ground truth on the appropriate ultimate resolution these approaches could map, especially when integrating data across these technologies. What we need to consider when deploying spatial biology platforms (either as standalone or multimodal/multi-omics) is what resolution or number of markers are required to capture sufficient complex biology to address disease of patient heterogeneity. These datasets need to have sufficient resolution or number of markers to allow a continuum of the same degree of measurements from preclinical efficacy and safety studies to clinical trial execution and ultimately inform on companion diagnostics. With such comparisons, spatial biology can provide an approach to rapidly identify and interrogate patterns of heterogeneity across the omic spectrum that both define the disease tissue landscape and identify features modulated by therapeutic intervention.

In terms of spatial resolution, the length scales at which biomarkers can be measured, the goal has traditionally been to continually increase resolution with the hope of tracking or

measuring biology in ever-greater detail. However, technologies often rely on surrogate measurements at these extreme resolutions, which can lead researchers to mistakenly interpret the measured distribution as informative of events at a subcellular resolution. This lack of specificity, for example, in understanding protein-bound versus released drug or the phosphorylation state of proteins, can result in misleading conclusions. Furthermore, with some new imaging approaches, we can mathematically segment a tissue based on complex biomarker profiles into smaller regions that go beyond what can be biologically or therapeutically meaningful. Indeed, algorithms and statistical analysis performed on spatial biology biomarker data can mathematically describe greater and greater levels of heterogeneity that may confuse the data consumer or scientist trying to gain insight from the findings. Therefore, understanding at what point resolving greater complexity is biologically useful is an important question. For example, in the immunotherapy field, both T-cell and myeloid cell biology analysis using classical single-cell omics approaches have produced many studies describing highly complex subphenotypes of individual myeloid cell types. Although it is important to appreciate the basic biology, the data have had limited impact on improving the positioning of current therapies or target identification for future therapies (14, 15).

Establishing confidence in the biomarkers used to guide decision-making is critical. Therefore, using a range of cross-validating omics technologies including classical omics, wherein tissues are dissociated or homogenized, with spatial-omics that are providing the spatial context of the biomolecules can be beneficial. This allows data from these technologies to help validate the biomarkers and provide more robust evidence for their role in disease processes as well as offering more readily deployed, robust, and readily usable assays. When using multiplex or hyperplex histology technologies, the validation for each biomarker and subsequent incorporation into a panel of assays can be laborious and expensive if attempting to maintain the rigor expected for isolated biomarker assays. An advantage of large panel imaging is that prospective probes can be added and in part validated against other markers in the same panel. Agility, speed, and a pragmatic approach offer the most effective operational model. Moreover, for hyperplex histology, there is a need to determine where the tipping point between reagent cost versus interpretable data is. Analytic strategies should focus on selecting a subset of samples for maximum analysis, with the preliminary data then being used to select informative number of biomarkers for use across the full experiment.

CONCLUSION

In conclusion, it is crucial to remember that although spatial biology, with its visually striking images and large data volumes, is easily marketable, the true actionable insight might be obscured behind the data complexity. We are transitioning from using machine learning/deep learning approaches to replicate pathologist insights to enlisting data scientists to uncover biological complexity beyond what was previously comprehensible. This paradigm shift carries the risk of interpreting correlations as biologically relevant, when in reality the true actionable insight might still be concealed within the data.

Authors' Disclosures

R.J. Goodwin is a full-time employee of AstraZeneca and holds company stock. J.S. Reis-Filho reports personal fees from AstraZeneca during the conduct of the study; personal fees and other support from Grupo Oncoclinicas, Repare Therapeutics, and Paige.AI; personal fees from Goldman Sachs, Saga Diagnostics, and MultiplexDX; and personal fees from Bain Capital outside the submitted work. S.T. Barry reports other support from AstraZeneca outside the submitted work (AstraZeneca shareholder and employee). No disclosures were reported by the other authors.

Published first April 4, 2024.

REFERENCES

1. Suehnholz SP, Nissan MH, Zhang H, Kundra R, Nandakumar S, Lu C, et al. Quantifying the expanding landscape of clinical actionability for patients with cancer. *Cancer Discov* 2024;14:49–65.
2. Pleasance E, Bohm A, Williamson LM, Nelson JMT, Shen Y, Bonakdar M, et al. Whole-genome and transcriptome analysis enhances precision cancer treatment options. *Ann Oncol* 2022;33:939–49.
3. Liu Q, Zhang J, Guo C, Wang M, Wang C, Yan Y, et al. Proteogenomic characterization of small cell lung cancer identifies biological insights and subtype-specific therapeutic strategies. *Cell* 2024;187:184–203.
4. Dagogo-Jack I, Shaw AT. Tumour heterogeneity and resistance to cancer therapies. *Nat Rev Clin Oncol* 2018;15:81–94.
5. Morgan P, Brown DG, Lennard S, Anderton MJ, Barrett JC, Eriksson U, et al. Impact of a five-dimensional framework on R&D productivity at AstraZeneca. *Nat Rev Drug Discov* 2018;17:167–81.
6. Elhanani O, Ben-Uri R, Keren L. Spatial profiling technologies illuminate the tumor microenvironment. *Cancer Cell* 2023;41:404–20.
7. Kuett L, Catena R, Özcan A, Plüss A; Cancer Grand Challenges IMAXT Consortium; Schraml P, et al. Three-dimensional imaging mass cytometry for highly multiplexed molecular and cellular mapping of tissues and the tumor microenvironment. *Nat Cancer* 2022;3:122–33.
8. Sushentsev N, Hamm G, Richings J, McLean MA, Menih IH, Ayyappan V, et al. Imaging tumor lactate is feasible for identifying intermediate-risk prostate cancer patients with postsurgical biochemical recurrence. *Proc Natl Acad Sci U S A* 2023;120:e2312261120.
9. McNamara KL, Caswell-Jin JL, Joshi R, Ma Z, Kotler E, Bean GR, et al. Spatial proteomic characterization of HER2-positive breast tumors through neoadjuvant therapy predicts response. *Nat Cancer* 2021;2:400–13.
10. Longo SK, Guo MG, Ji AL, Khavari PA. Integrating single-cell and spatial transcriptomics to elucidate intercellular tissue dynamics. *Nat Rev Genet* 2021;22:627–44.
11. Jackson HW, Fischer JR, Zanotelli VRT, Ali HR, Mechera R, Soysal SD, et al. The single-cell pathology landscape of breast cancer. *Nature* 2020;578:615–20.
12. Lipkova J, Chen RJ, Chen B, Lu MY, Barbieri M, Shao D, et al. Artificial intelligence for multimodal data integration in oncology. *Cancer Cell* 2022;40:1095–110.
13. Shmatko A, Ghaffari Laleh N, Gerstung M, Kather JN. Artificial intelligence in histopathology: Enhancing cancer research and clinical oncology. *Nat Cancer* 2022;3:1026–38.
14. Veglia F, Sanseviero E, Gabrilovich DI. Myeloid-derived suppressor cells in the era of increasing myeloid cell diversity. *Nat Rev Immunol* 2021;21:485–98.
15. van der Leun AM, Thommen DS, Schumacher TN. CD8+ T cell states in human cancer: insights from single-cell analysis. *Nat Rev Cancer* 2020;20:218–32.