BMB Reports - Manuscript Submission

Manuscript Draft

Manuscript Number: BMB-22-014

Title: Recent advances in spatially resolved transcriptomics: challenges and opportunities

Article Type: Mini Review

Keywords: Spatially resolved transcriptomics; single-cell RNA sequencing; Integrative computational algorithm; multimodal data analysis; scRNA-seq

Corresponding Author: Jungmin Choi

Authors: Jongwon Lee[1], Minsu Yoo[1], Jungmin Choi[1,*]

Institution: [1]Department of Biomedical Sciences, Korea University College of Medicine, Seoul 02841, Republic of Korea,
[2]Brain Korea 21 Plus Project for Biomedical Science, Korea University College of Medicine, Seoul, Korea,
[3]Department of Genetics, Yale University School of Medicine, New Haven, CT 06510, USA,

Manuscript Type: Mini Review

Title: Recent advances in spatially resolved transcriptomics: challenges and opportunities

Author's name: Jongwon Lee[1,2*], Minsu Yoo[1*], Jungmin Choi[1,3]

Affiliation:

1. Department of Biomedical Sciences, Korea University College of Medicine, Seoul 02841, Republic of Korea
2. Brain Korea 21 Plus Project for Biomedical Science, Korea University College of Medicine, Seoul, Korea
3. Department of Genetics, Yale University School of Medicine, New Haven, CT 06510, USA

* These authors have contributed equally to this work

Running Title: Trends in spatially resolved transcriptomics

Keywords: Spatially resolved transcriptomics, single-cell RNA sequencing, Integrative computational algorithm, multimodal data analysis

Corresponding Author's Information:

+82) 2-2286-1469

jungminchoi@korea.ac.kr

ABSTRACT

Single-cell RNA sequencing (scRNA-seq) has greatly advanced our understanding of cellular heterogeneity by profiling individual cell transcriptomes. However, cell dissociation from the tissue structure causes a loss of spatial information, which hinders the identification of intercellular communication networks and global transcriptional patterns present in the tissue architecture. To overcome this limitation, novel transcriptomic platforms that preserve spatial information have been actively developed. Significant achievements in imaging technologies have enabled *in situ* targeted transcriptomic profiling in single cells at single-molecule resolution. In addition, technologies based on mRNA capture followed by sequencing have made possible profiling of the genome-wide transcriptome at the 55~100 μm resolution. Unfortunately, neither imaging-based technology nor capture-based method elucidates a complete picture of the spatial transcriptome in a tissue. Therefore, addressing specific biological questions requires balancing experimental throughput and spatial resolution, mandating the efforts to develop computational algorithms that are pivotal to circumvent technology-specific limitations. In this review, we focus on the current state-of-the-art spatially resolved transcriptomic technologies, describe their applications in a variety of biological domains, and explore recent discoveries demonstrating their enormous potential in biomedical research. We further highlight novel integrative computational methodologies with other data modalities that provide a framework to derive biological insight into heterogeneous and complex tissue organization.

INTRODUCTION

Single-cell RNA sequencing (scRNA-seq) technique has enabled us to investigate deeper into the functions and characteristics of each cell, looking into the details of its genetic compositions of mRNA that are transiently active at the time of the data acquisition (1). Compared to bulk-RNA sequencing, this is a tremendous step forward that has allowed scientists to gain information about the functional differences between individual cell types (2). However, scRNA-seq loses spatial information of individual cells, interactions among adjacent cells, and local signaling networks. For example, the tumor microenvironment has remained elusive (3, 4). In order to overcome this limitation, transcriptomic techniques that capture the spatial information for tissues of interest have been actively developed (5, 6).

Spatially resolved transcriptomics has achieved significant progress in the biomedical research field with advances in imaging and next-generation sequencing (NGS) technology. Spatially resolved transcriptomics has two main categories: 1) image-based *in situ* hybridization (ISH) and *in situ* sequencing (ISS) methods; 2) technologies based on mRNA capture, including laser capture microdissection (LCM) that are spatially barcoded on slides or beads. The different spatially resolved transcriptomic methods have their key features, such as the number of transcripts detected, cellular resolution, and size of the region captured. While image-based methods require prior knowledge of the genes of interest, NGS-based methods capture the whole transcriptome from tissue sections unbiasedly. Also, the spatial resolution of image-based methods for tissue imaging is limited only by the optical diffraction limit and, therefore, suitable for analyzing a subcellular organization. In contrast, the spatial resolution of NGS-based methods is limited by the physical size of a spot that captures cells. Although both methods have improved considerably in the past few years, efforts to refine the sensitivity and increase the cellular resolution of these methods are still ongoing. The overview of two main spatially resolved transcriptomic techniques is described in Figure 1.

In this review, we describe how different spatially resolved transcriptomic techniques have evolved based on the needs resulting from distinct tissue environments (Table 1) and how those methods are combined with the high-throughput scRNA-seq data to complement each method's technical and functional constraints (Table 2).

RESULTS

Development of spatially resolved transcriptomic techniques

Image-based spatially resolved transcriptomics There are two major approaches that visualize targeted genes of interest by fluorescent labels. First, *in situ* hybridization (ISH) is a technique that detects mRNA molecules in thick tissue samples to find gene expression patterns in single cells within the native tissue environment. In detail, it detects target sequences using fluorescent images generated by hybridization of a complementary probe. This approach has advantages in visualizing RNA molecules directly in the original locations of the biological environment. However, the high autofluorescence background of tissue samples present limitations in detecting a large set of different transcripts simultaneously.

In 2016, Shah *et al.* developed single-molecule fluorescence *in situ* hybridization (smFISH) methods with multi-color and multi-RNA imaging in deep tissues using single-molecule hybridization chain reaction (smHCR) (7). The central innovation of smHCR was the fluorescent amplification of probes complementary to the target mRNA to increase the sensitivity for gene detection. This technique was essential to ensure that puncta generated by the fluorescence were bright enough for high sensitivity. Using this approach, deeper tissue was successfully visualized with higher sensitivity as a single mRNA was detected *in situ,* even in thick tissues in the intact vertebrate embryo of zebrafish. However, they were only small enough for diffraction-limited resolution. Similarly, unamplified cyclic-ouroboros smFISH (osmFISH) visualizes each RNA molecule by binding 20 nucleotide-long fluorescently labeled DNA probes (8). Subsequently, the expression of 33 marker genes was successfully analyzed in mouse somatosensory cortex and hippocampus by conventional epifluorescence microscope.

To understand the spatial context of cells, direct imaging of individual RNA molecules within intact cells and tissues is vital. The multiplexed fluorescence *in situ* hybridization (FISH) technique has been widely used for this purpose. MERFISH (multiplexed error-robust

fluorescence *in situ* hybridization) identifies RNA through a combinational labeling approach that consists of RNA molecules with error-robust barcodes followed by sequential rounds of smFISH to read out these barcodes (9). With this approach, an imaging process with an ensemble of fluorescent spots that define binary barcodes (on or off fluorescent signal) allowed the characterization of 130 genes in 100,000 cultured U-2 OS cells. The same group improved MERFISH-based imaging where scRNA-seq was combined with the MERFISH-based images for *in situ* cell-type identification and mapping (10). This upgraded analysis platform successfully identified functionally important 155 marker genes from ~70 neuronal populations and spatial organizations in the hypothalamic preoptic region of mice. Another strategy to increase the number of distinguished mRNA molecules was achieved by applying four colors with eight barcoding rounds into pseudo-colors. Using seqFISH+, images of mRNA for 10,000 genes in a single cell were acquired with high accuracy and sub-diffraction-limit resolution using a conventional confocal microscope (11). This method identified a subcellular localization pattern of 10,000 mRNAs in the brain, and ligand-receptor pairs from potential cell-cell interactions were suggested by the observed mRNA expression patterns in the cells. This proves that seqFISH+ can overcome not only optical crowding and provide a tenfold or greater improvement over existing methods, but also enables super-resolved imaging with conventional confocal microscopes.

Higher intensity of FISH signals is important to improve the gene detection performance. Signal amplification by exchange reaction (SABER) technique was designed to amplify the intensity of quantitative FISH signals (12). In brief, DNA and RNA FISH probes were first chemically synthesized with a primer sequence of their 3' end extended into primer-exchange reaction (PER). In SABER, a multitude of PER-concatemerized probes sets can be hybridized to their targets simultaneously and read out in sequential rounds of imaging. With this approach, 18,000 probes in total targeting a 3.9-Mb region of human metaphase spreads and interphase

cells were mapped to three colors, which all colocalized as expected. However, a key challenge of multiplexed FISH is that it is difficult to detect individual target probes in complex tissue environments due to the noisy background of tissue. To address this issue, Goh *et al.* adopted an alternative approach to reduce off-target background fluorescence by integrating split-probe strategy with multiplexed FISH (13). By improving the bridge probe design to hybridize with sufficient complementary base pairing, split-FISH is capable of accurate transcriptomic profiling on four mouse tissue types, even in opaque tissue. However, taking sequential images with multiple hybridization rounds for several target genes can sometimes be misaligned because of possible technical or experimental errors. Therefore, correcting these misalignments is an essential step; thus, tools for precise alignments of multiple images have been developed (14).

Another imaging-based approach for spatially resolved transcriptomics is *in situ* sequencing (ISS). The main advantage of ISS over ISH is that ISS can detect more genes (~10,000) than ISH in a non-targeted manner. In ISS-based methods, transcripts were sequenced directly in fixed tissues. Specifically, mRNA molecules are reverse transcribed and amplified by rolling circle amplification (RCA). The micrometer- or nanometer-sized RCA products are then subjected to sequencing-by-ligation (SBL), and the barcode within the probe is simultaneously decoded.

Wang *et al.* devised the spatially resolved transcript amplicon readout mapping (STARmap) methods to detect and quantify 160 gene sets simultaneously in the mouse primary visual cortex. STARmap incorporates hydrogel chemistry, improved padlock-probe technology, and error-robust SBL methods (15). Also, the extended STARmap can image a larger number of genes successfully with a better spatial resolution (1,020 genes detected simultaneously at millimeter-scale volumes encompassing ~30,000 cells). To overcome the limitation of short reads in ISS, a new method utilizing bidirectional sequencing chemistry and an imaging transcript-specific

barcode was developed. With this bidirectional sequencing method, the generated reads were efficiently extracted and assembled into longer reads using NGS technology (16). One of the most important aspects of sequencing methodology is maximizing throughput for reliable gene detection. Hybridization-based *in situ* sequencing (HybISS) achieved this goal by replacing the SBL technique of ISS with sequencing-by-hybridization (SBH) to detect hundreds of target genes and successfully applied in mouse and human brains. The SBH enabled detecting genes with increased signal-to-noise using autofluorescence quenching (17). When classifying cell types, the low throughput of ISS technology might not be sufficient to have reliable outcomes. To address this issue, probabilistic cell typing by *in situ* sequencing (pciSeq) introduced an approach that uses ample scRNA-seq classification to identify cell types using multiplexed *in situ* RNA detection (18). Specifically, the probability of assigning each read to a given cell type was calculated using a Bayesian algorithm derived from scRNA-seq cluster data.

Another attempt has been made with ExSeq where cDNA amplicons are eluted from the sample and re-sequenced *ex situ* with NGS for long-read untargeted and targeted *in situ* RNA sequencing (19). ExSeq yielded the readout of thousands of genes, including splice variants, and enabled highly multiplexed mapping of RNAs when applied in mouse hippocampal neuron cells and tissues.

Capture-based spatially resolved transcriptomics Tissues preserved in FFPE blocks can be stored from 3-10 years. Using laser capture microdissection (LCM), Romanens *et al.* extracted tissues related to breast cancer that were preserved in FFPE blocks using an optimized protocol for hematoxylin and eosin (H&E) staining with the best parameters for dyes, incubation time, the thickness of tissue, and surface area for microdissection (20). In the study, Romanens and colleagues spatially characterized tumor and immune cells in triple-negative breast cancer. Using the immunofluorescence-guided laser capture microdissection (immuno-LCM-RNAseq) technique, it became easier to dissect tissues and study particular genes of cells that have

functions that are highly dependent on the spatial organization in the tissue (21). In their study, the quality of RNA was immensely improved by a modified high-salt protocol and RNase inhibitor, which enabled the investigation of full-length transcripts and isoforms.

Compared with the LCM-seq method, Photo-isolation chemistry has the advantage in spatial resolution because its photochemical isolation technique uses photo-caged oligo-deoxynucleotides for *in situ* reverse transcription (22). This approach has allowed to analyze cells in complicated and fine structures, detecting 8,000 genes with $7x10^4$ unique read counts per single cell in mouse embryos. These advanced LCM can profile more than 10,000 genes from dozens of cells with spatial resolution up to subcellular or subnuclear levels. While the sequencing depth is high enough, the analysis is restricted to a limited number of regions.

Another strategy to obtain spatially resolved transcriptomic data is to use pre-arranged set of barcoded reverse transcription (RT)-primers on glass slides. Spatial Transcriptomics (ST) analyzes the transcriptome in individual tissue sections while maintaining two-dimensional positional information of tissues. Specifically, slide sections are delivered to a glass slide that bears RT-primers and sets of DNA barcodes *in situ*. After generating cDNA libraries, NGS was carried out with the sequencing libraries. Then, the spatially resolved whole-transcriptome information is provided by aligning tissue images with the sorted RNA-seq data that include spatial information carried by the barcodes. However, the spatial resolution of ST is limited to 100 μm with a center-to-center distance of 200 μm which yields around 50 to 100 cells per single spot (23, 24). The higher resolution of ST was achieved with the 10X Genomics Visium platform that features a physical spot size of 55 μm diameter and a center-to-center distance of 100 μm. As a result, the 10X Genomics Visium platform has been one of the most popular choices for mapping spatial profiles of tissues in interest in various fields, including neuroscience, cancer biology, and developmental study (25-27). Spatial cellular resolution and the quality of analysis can be further improved by combining two or more different spatially

resolved transcriptomic techniques. Chen *et al.* combined ISS and capture-based ST methods to identify a plaque-induced gene network in micro-and astroglia cells and oligodendrocytes in the Alzheimer's disease model of mouse and human. Then, the oligodendrocyte gene responses were confirmed at the single-cell level using ISS with target-specific probes with known cell type markers. Combining those two techniques elucidated how amyloid plaques are involved in the neurodegenerative process (28). In another study, Ji *et al.* carried out three techniques of single-cell, spatially resolved transcriptomics, and multiplexed ion beam imaging (MIBI) to investigate the tumor microenvironment of cutaneous squamous cell carcinoma (29). Spatially resolved transcriptomes from 17,064 spots were obtained utilizing both conventional ST and 10X Genomics Visium kits. The combination of three methods revealed ligand-receptor networks of specific cell types, suggesting tumor-specific keratinocytes as a signaling hub for intercellular communication.

The techniques for spatial transcriptomics with barcoded oligonucleotide capture array described, however, have limitation in the spatial resolution of up to 55~100 μm due to the physical size of capturing spots. To resolve the issue of low spatial resolution in capture-based sequencing methodology, bead-based capturing sequencing was developed. In 2019, Rodriques *et al.* developed Slide-seq for higher cellular resolution genome-wide analysis using DNA-barcoded 10 μm beads onto a rubber-coated glass slide (30). In Slide-seq, each bead's distinct barcoded sequence is determined via SBL methods. Using the beads, the resolution comparable to the size of an individual cell was achieved by NGS of barcoded RNA-seq libraries and successfully applied to the hippocampus area in mice. Slide-seq defined finer cellular subpopulations of cerebellar cell types, which had not been identified in previous spatially resolved single-cell sequencing studies. Additionally, high-definition spatial transcriptomics (HDST) was developed as another effort to accomplish higher resolution for barcoded transcripts (31). HDST randomly deposits barcoded poly(d)T oligonucleotides into a 2 μm well

with bead array-based fabrication. Mouse brain and primary breast cancers were profiled with 25x higher resolution than that of Slide-seq using this method, resulting in a reconstruction of the spatial architecture of mouse brain and primary breast cancer tissues. Although Slide-seq has a superior spatial resolution compared to other techniques, it has a lower transcript detection sensitivity that limits biological applications. To circumvent these restrictive factors of the initial version, Slide-seqV2 was developed to offer higher capturing efficiency for transcripts by implementing a readily available monobase-sequencing strategy instead of dibase-sequencing, leading to a dramatic increase the capture efficiency (32). Another advancement from Slide-seq was the application of an error robust computational approach using NGS. With the higher number of unique molecular identifier (UMI) per bead (494 UMIs per 10 μm), which is higher than HDST, Slide-seqV2 demonstrated that it can detect a rare transcript, such as dendritic mRNA in CA1 neurons. Cho *et al.* developed a high-resolution spatial barcoding technology, Seq-Scope, with a 0.5~0.8 μm center-to-center resolution (33). This technique was developed based on the NGS technology that utilizes randomly barcoded single-molecule oligonucleotides with an outstanding transcriptome capture capacity (~4,700 UMIs/cell on average), comparable to the conventional scRNA-seq methods. This technical improvement enabled visualizing the histological organization of the transcriptome architecture in liver tissues at subcellular level. More recently, another high-resolution spatial barcoding technology, SpaTial Enhanced Resolution Omics-sequencing (Stereo-seq) was introduced with a center-to-center resolution of 500~715 nm (34). This method applies microfluidic channels perpendicularly for unique pairwise barcoding on each spot of the tissue section. Stereo-seq can capture up to 133,775 UMI per 100 μm diameter bin, superior to other available technologies, including Seq-Scope compared to the same bin resolution. Stereo-seq was capable of profiling whole mouse embryo, demonstrating that it can be applied to a much larger tissue such as whole human brain sections.

Integration of spatially resolved transcriptomic data with other methods

While scRNA-seq and spatially resolved transcriptomic data have limitations, a computational integration of two or more data modalities can better characterize spatial cell type compositions and local cell states in the tissues. Here, we describe the improvements benefiting from the novel *in silico* computational algorithms, which integrate spatially resolved transcriptomic data with additional data modalities, such as scRNA-seq, and several other Omics approaches to better understand spatial architecture of tissues and to derive biological insights.

Combination with scRNA-seq probabilistic cell typing by *in situ* sequencing (pciSeq) was used to integrate scRNA-seq data with ISS for cluster analysis (18). To design a gene panel, genes were grouped leveraging prior scRNA-seq cell type classification. Using pciSeq, Qian *et al.* were able to map the inhibitory neurons of hippocampal area CA1, making it feasible to confirm results for ISS. Before this study, the ISS RNA detection had not been proven to distinguish fine cortical cell types identified from previous reports. Defining how cell states correlate with spatial cell distribution and investigating how the local signaling environment impacts molecular signatures and, ultimately, the fate of cells is inevitably important to understand mouse embryo changes during early development. Lohoff and colleagues delineated the precise location of distinct cell types within a single reference scaffold by combining a high-resolution seqFISH map with scRNA-seq (35). From this study, integration of single-cell transcriptome with spatial context enabled not only identification of 387 target genes from seqFISH but also the imputation of the spatially resolved map at cellular resolution in a mouse organogenesis model.

To improve the low spatial resolution of capture-based spatial methods, the two data modalities from each scRNA-seq and spatially resolved transcriptomic methods were integrated into other

studies. Cell type signature defined by scRNA-seq was integrated with Slide-seq data, facilitating the discovery of spatially defined gene expression patterns from mouse hippocampus. For reconstructing expression of each Slide-seq data, non-negative matrix factorization regression (NMFreg) was utilized as a combination of cell-type signatures defined in the reference scRNA-seq data (30). In addition, several computational techniques for combining single-cell and MIBI or multiple spatial transcriptomics were reported to achieve a higher cellular resolution (28, 29). Moncada *et al.* integrated scRNA-seq with the ST method to study the microenvironment of primary pancreatic tumors by introducing multimodal intersection analysis (MIA) and reported that fibroblast-specific genes in the tissue of pancreatic ductal adenocarcinoma overlap significantly with the set of genes specific to the cancer region from the ST data (36). The study also identified the colocalization of inflammatory fibroblasts and cancer cells expressing a stress-response gene module.

Deep learning-based spatial information Machine learning algorithms have been extensively used as an imputation method for predicting cell types based on the context of relevant datasets. For example, DEEPsc, an artificial neural network model for categorizing cell types, was trained to predict a specific cell type with sets of genes from the scRNA-seq reference atlas dataset (37) and achieved an accuracy comparable to several existing methods (2-norm, infinity norm, mean percent difference, and large margin nearest neighbor) for applications utilizing 3,000 highly variable genes. Additionally, BayesSpace applies a Bayesian statistical method that takes the spatial gene expression profile from neighborhoods as a prior and achieves a super-resolution image (38). It is noteworthy that BayesSpace does not require separate scRNA-seq gene expression signature or preselected marker genes. SPICEMIX, another approach using non-negative matrix factorization (NMF) based on probabilistic latent variable modeling, calculates spatial affinity between the metagenes and their proportions of

neighboring cells (39). Data acquired from seqFISH+ and STARmap were used to demonstrate its ability to refine the identification of cell types in mouse primary visual cortex. SpaOTsc was developed to better understand cell-cell communications in the spatially resolved transcriptomic datasets and inferred the spatial distance between every pair of cell types by computing the optimal transport distance measured by utilizing spatial measurements of a relatively small number of genes without the scRNA-seq dataset (4). The results from SpaOTsc indicate that signal sender cells exhibit more spatial localization patterns, while the locations of the signal receivers are more scattered over throughout the tissues of zebrafish and the mouse visual cortex. SpaOTsc can be used both to integrate non-spatial single-cell measurements with the spatial data and to reconstruct spatial cellular dynamics in tissues.

Deconvolution of spatially resolved transcriptomics Owing to the lower cellular resolution of spatial barcoding capture-based methods, the proportions of discrete cell types for a given spot have been inferred by various deconvolution algorithms. Robust cell type decomposition (RCTD) is based on statistical model maximum-likelihood estimation to approximate the proportions of spatially localized cellular subtypes in spatially resolved transcriptomic data such as Slide-seq or 10X Genomics Visium datasets (40). RCTD accurately recapitulated known cell type spatial distribution in both Slide-seq and 10X Genomics Visium in the mouse brain tissues. Alternatively, SPOTlight uses NMF alongside non-negative least squares for accurate and sensitive cell-type detection, seamlessly applied in mouse brain and 22 immune subpopulations from pancreatic adenocarcinoma samples with a curated annotation (41). Recently, SpatialDWLS was shown to have a higher degree of sensitivity and accuracy than RCTD and SPOTlight using the dampened weighted least squares (DWLS) algorithm in which the weight is selected to minimize the overall relative error rate to infer cell-type composition (42). By applying SpatialDWLS to spatial transcriptomics dataset of mouse brain and human

embryonic heart, increased abundance of different cell types was observed during development. Finally, DestVI was developed to alleviate the problem of the complicated deconvolution, especially when there is a continuum of cell states that cannot be clearly distinguished (43). To address this problem, deconvolution of spatially resolved transcriptomic profiles used Variational Inference (DestVI), a Bayesian model-based multi-resolution cell-type deconvolution algorithm, was developed. DestVI learns cell type-specific profiles and sub-cell type variations using a conditional deep generative model. By combining scRNA-seq and spatially resolved transcriptomic data from the same tissue, DestVI was used to study the immune interplay within lymph nodes and explore the spatial tissue architecture of the mouse tumor microenvironment. Notably, DestVI outperforms existing discrete deconvolution approaches such as RCTD, SPOTlight, and Seurat.

Cell type inference via image-based machine learning H&E-stained histology images are easy and cheap to obtain and routinely generated in clinics. Several studies integrated spatially resolved transcriptomics and histopathology images to extract feature information and make predictions by various machine learning algorithms. ST-Net implements a convolutional neural network model and was trained on 68 breast tissue sections from 23 patients with breast cancer to predict the gene expression based on histopathology images (44). In addition, HisToGene adopts Vision Transformer for image recognition and predicts super-resolution gene expression. Using the same dataset from ST-Net as a training set, HisToGene outperformed ST-Net in both gene expression prediction and clustering tissue regions accuracy (45). stLearn also integrates three types of datasets, including spatial dimensionality, tissue morphology, and genome-wide transcriptional profile using a deep learning network model (46) and predict cell type clustering, intercellular interaction, and reconstruction of spatial transition gradients, which were all successfully conducted with brain and breast cancer datasets. SpaCell and CoSTA also utilize

a convolutional neural network model to predict malignant cells in prostate cancer and quantify the level of spatial expression relationships between each pair of genes from mouse brain samples, respectively (47, 48). Alternatively, STUtility uses raw 10X Genomics Visium RNA-seq and image data as input and processes the images, aligns consecutive stacked tissue images, and finally visualize them all together in 3D. Another useful functionality of STUtility includes NMF to decompose 10X Genomics Visium data into a lower dimension and a method to identify neighboring capture spots in a spatial network (49). This approach was applied in human breast cancer tissues to define the leading edge of the tumor region, eventually leading to the delineation of tumor heterogeneity between the tumor core and the tumor front. Bergenstråhle *et al.* also constructed a deep generative model for spatial data fusion by combining low-sensitivity, low-resolution *in situ* RNA capturing expression data with high-resolution histological image data (50). This method inferred de-noised full-transcriptome spatial gene expression at the same resolution as the image data in mouse olfactory bulb by jointly training a recognition neural network that maps the image data to the variational parameters of the latent state to optimize model parameters.

Cellular protein information Single-cell profiling via proteomics approach has become progressively comprehensive, and unbiased profiling of protein expression has had a broad impact in biomedical research (51). Recent advances in techniques such as epitope-based imaging, mass cytometry, and mass spectrometry enable protein expression to be mapped across tissue with high resolution. Stoeckius *et al.* introduced a method named cellular indexing of transcriptomes and epitopes by sequencing (CITE-seq) that integrates cellular protein and transcriptome measurements in an effort to provide phenotypic information such as cell-surface protein expression levels (52). This multimodal analysis on cellular protein, using oligonucleotide-labeled antibodies and high-throughput scRNA-seq, revealed phenotypes including a detailed multimodal characterization of immune subpopulations that had not been

ascertained by using scRNA-seq alone. The study carried out by Schulz *et al.* used extended imaging mass cytometry (IMC), based on the simultaneous staining followed by multiplexed detection of mRNA and proteins in tissue (53). This approach that couples the RNAscope-based ISH with simultaneous antibody staining was used in 70 breast cancer samples and demonstrated a moderate correlation between HER2 and CK19 mRNA and protein expression levels. However, IMC has technical challenges for probing spatial distribution due to the limited number of labels and/or antibodies for labeling proteins. To get around this, Piehowski *et al.* developed an automated imaging approach, Nanodroplet Processing in One plot for Trace Samples (nanoPOTS), utilizing label-free nanoproteomics to analyze tissue voxels (54). nanoPOTS increased the number of target proteins to more than 2,000 with spatial resolution of 100 μm. Combined with LCM and ultrasensitive nanoLC-MS/MS, large sample sets were collected and processed with high sensitivity capturing the unique proteome mapping in a mouse uterus model.

Spatial ATAC-seq To capture spatial epigenetic information in tissue at the single-cell level and genome scale, single-cell combinatorial indexing on microbiopsies that is assigned to positions for the assay for transposase accessible chromatin (sciMAP-ATAC) was introduced (55). sciMAP-ATAC produced data of similar quality to non-spatial sci-ATAC of cells within a 214-micron cubic area and was submitted to the adult mouse primary somatosensory cortex and human primary visual cortex to successfully characterize the spatial progressive nature of cerebral ischemic infarction. Integration of sciMAP-ATAC with single-nucleus RNA-seq and single-cell chromatin accessibility datasets demonstrated high concordance for most cell types. Deng *et .al.* also reported spatially resolved chromatin accessibility profiling method (spatial-ATAC-seq) (56). The basic concept of ATAC-seq utilizing NGS with Tn5 transposition chemistry was improved with a microfluidic barcoding system, introducing two barcode solutions to the tissue surface using an array of microchannels for *in situ* ligation

perpendicularly. By integrating the spatial-ATAC-seq data with the scRNA-seq data, each cell type and organ-specific cell type was assigned in mouse embryos. Furthermore, differences between the epigenetic state and protein expression of specific marker genes enabled distinguishing cell types in mouse brain and human tonsil.

DISCUSSION

While spatially resolved transcriptomic technology offers tremendous opportunity to discover spatial heterogeneity in the disease state, characterize spatial expression blueprints during development, and elucidate spatial architecture at the molecular level, its potential lies beyond that as it is still in the early days of development. Unfortunately, none of the currently available spatially resolved transcriptomic technologies are perfect, and the choice of methods depends on study design, the biological question, and often a balance between the cell and/or transcript throughput and spatial resolution. A few overlooked caveats include the requirement for a large number of pseudocolors and barcodes in image-based techniques and a low spatial resolution in capture-based methods. This has brought new technological challenges. The integrative computational algorithms that combine the spatially resolved transcriptomic data and other data modalities have significantly contributed to not only overcoming the key challenges faced by current spatially resolved transcriptomic technologies but gaining fundamental biological insights. The development of novel computational tools will continue to play a significant role in exploring large-scale spatially resolved transcriptomic datasets, translating the consequences of newly acquired spatial patterns, and elucidating principles of the underlying biology. The advent of such integrative approaches will ultimately shed light on the mechanisms that explain the essential differences in spatial architectures of healthy and diseased tissues.

ACKNOWLEDGMENTS

The authors are grateful to Junho Song for critical reading of the manuscript. This research was supported by the National Research Foundation of Korea (NRF) grants funded by the South Korean government (2020R1F1A1076705).

CONFLICTS OF INTEREST

The authors declare no conflict of interest.

FIGURE LEGENDS

Figure 1. Overview of two main categories of spatially resolved transcriptomic techniques: Image-based (Left) and capture-based (Right) methods. Imaging-based methods detect target genes by sequencing directly into a fixed tissue section (ISS) or hybridizing a complementary fluorescent probe (ISH). Both image-based methods are suitable for analyzing subcellular transcript patterns of a cell. Capture-based methods are divided into three sub-categories: directly obtaining a specific region of interest from a tissue section using laser capture microdissection (LCM), customized slides, or bead arrays to capture mRNAs by oligonucleotide-based spatial barcodes followed by NGS. Combining computational strategies enables the comprehensive mapping of cell types in spatially resolved transcriptomic data with a specific spatial resolution by each method.

Table 1. Spatially resolved transcriptomic techniques

Table 2. Integration of spatially resolved transcriptomic data with other methods

REFERENCES

1. Hwang B, Lee JH and Bang D (2018) Single-cell RNA sequencing technologies and bioinformatics pipelines. Exp Mol Med 50, 1-14
2. Stark R, Grzelak M and Hadfield J (2019) RNA sequencing: the teenage years. Nat Rev Genet 20, 631-656
3. Roth R, Kim S, Kim J and Rhee S (2020) Single-cell and spatial transcriptomics approaches of cardiovascular development and disease. BMB Rep 53, 393-399
4. Cang Z and Nie Q (2020) Inferring spatial and signaling relationships between cells from single cell transcriptomic data. Nat Commun 11, 2084
5. Asp M, Bergenstråhle J and Lundeberg J (2020) Spatially Resolved Transcriptomes-Next Generation Tools for Tissue Exploration. Bioessays 42, e1900221
6. Dries R, Chen J, Del Rossi N, Khan MM, Sistig A and Yuan GC (2021) Advances in spatial transcriptomic data analysis. Genome Res 31, 1706-1718
7. Shah S, Lubeck E, Schwarzkopf M et al (2016) Single-molecule RNA detection at depth by hybridization chain reaction and tissue hydrogel embedding and clearing. Development 143, 2862-2867
8. Codeluppi S, Borm LE, Zeisel A et al (2018) Spatial organization of the somatosensory cortex revealed by osmFISH. Nat Methods 15, 932-935
9. Moffitt JR, Hao J, Wang G, Chen KH, Babcock HP and Zhuang X (2016) High-throughput single-cell gene-expression profiling with multiplexed error-robust fluorescence in situ hybridization. Proc Natl Acad Sci U S A 113, 11046-11051
10. Moffitt JR, Bambah-Mukku D, Eichhorn SW et al (2018) Molecular, spatial, and functional single-cell profiling of the hypothalamic preoptic region. Science 362
11. Eng CL, Lawson M, Zhu Q et al (2019) Transcriptome-scale super-resolved imaging in tissues by RNA seqFISH. Nature 568, 235-239
12. Kishi JY, Lapan SW, Beliveau BJ et al (2019) SABER amplifies FISH: enhanced multiplexed imaging of RNA and DNA in cells and tissues. Nat Methods 16, 533-544
13. Goh JJL, Chou N, Seow WY et al (2020) Highly specific multiplexed RNA imaging in tissues with split-FISH. Nat Methods 17, 689-693
14. Borovec J, Kybic J, Arganda-Carreras I et al (2020) ANHIR: Automatic Non-Rigid Histological Image Registration Challenge. IEEE Trans Med Imaging 39, 3042-3052
15. Wang X, Allen WE, Wright MA et al (2018) Three-dimensional intact-tissue sequencing of single-cell transcriptional states. Science 361
16. Fürth D, Hatini V and Lee JH (2019) In Situ Transcriptome Accessibility Sequencing (INSTA-seq). bioRxiv, 722819
17. Gyllborg D, Langseth CM, Qian X et al (2020) Hybridization-based in situ sequencing (HybISS) for spatially resolved transcriptomics in human and mouse brain tissue. Nucleic Acids Res 48, e112

18. Qian X, Harris KD, Hauling T et al (2020) Probabilistic cell typing enables fine mapping of closely related cell types in situ. Nat Methods 17, 101-106
19. Alon S, Goodwin DR, Sinha A et al (2021) Expansion sequencing: Spatially precise in situ transcriptomics in intact biological systems. Science 371
20. Romanens L, Chaskar P, Tille J-C et al (2020) Spatial transcriptomics of tumor microenvironment in formalin-fixed paraffin-embedded breast cancer. bioRxiv, 2020.2001.2031.928143
21. Zhang X, Hu C, Huang C et al (2021) Robust Acquisition of High Resolution Spatial Transcriptomes from Preserved Tissues with Immunofluorescence Based Laser Capture Microdissection. bioRxiv, 2021.2007.2013.452123
22. Honda M, Oki S, Kimura R et al (2021) High-depth spatial transcriptome analysis by photo-isolation chemistry. Nat Commun 12, 4416
23. Salmén F, Ståhl PL, Mollbrink A et al (2018) Barcoded solid-phase RNA capture for Spatial Transcriptomics profiling in mammalian tissue sections. Nat Protoc 13, 2501-2534
24. Ståhl PL, Salmén F, Vickovic S et al (2016) Visualization and analysis of gene expression in tissue sections by spatial transcriptomics. Science 353, 78-82
25. Yao Z, van Velthoven CTJ, Nguyen TN et al (2021) A taxonomy of transcriptomic cell types across the isocortex and hippocampal formation. Cell 184, 3222-3241.e3226
26. Wu SZ, Al-Eryani G, Roden DL et al (2021) A single-cell and spatially resolved atlas of human breast cancers. Nat Genet 53, 1334-1347
27. Fawkner-Corbett D, Antanaviciute A, Parikh K et al (2021) Spatiotemporal analysis of human intestinal development at single-cell resolution. Cell 184, 810-826.e823
28. Chen WT, Lu A, Craessaerts K et al (2020) Spatial Transcriptomics and In Situ Sequencing to Study Alzheimer's Disease. Cell 182, 976-991.e919
29. Ji AL, Rubin AJ, Thrane K et al (2020) Multimodal Analysis of Composition and Spatial Architecture in Human Squamous Cell Carcinoma. Cell 182, 497-514.e422
30. Rodriques SG, Stickels RR, Goeva A et al (2019) Slide-seq: A scalable technology for measuring genome-wide expression at high spatial resolution. Science 363, 1463-1467
31. Vickovic S, Eraslan G, Salmén F et al (2019) High-definition spatial transcriptomics for in situ tissue profiling. Nat Methods 16, 987-990
32. Stickels RR, Murray E, Kumar P et al (2021) Highly sensitive spatial transcriptomics at near-cellular resolution with Slide-seqV2. Nat Biotechnol 39, 313-319
33. Cho CS, Xi J, Si Y et al (2021) Microscopic examination of spatial transcriptome using Seq-Scope. Cell 184, 3559-3572.e3522
34. Chen A, Liao S, Cheng M et al (2021) Spatiotemporal transcriptomic atlas of mouse organogenesis using DNA nanoball patterned arrays. bioRxiv, 2021.2001.2017.427004
35. Lohoff T, Ghazanfar S, Missarova A et al (2021) Integration of spatial and single-cell transcriptomic data elucidates mouse organogenesis. Nat Biotechnol
36. Moncada R, Barkley D, Wagner F et al (2020) Integrating microarray-based spatial

transcriptomics and single-cell RNA-seq reveals tissue architecture in pancreatic ductal adenocarcinomas. Nat Biotechnol 38, 333-342
37. Maseda F, Cang Z and Nie Q (2021) DEEPsc: A Deep Learning-Based Map Connecting Single-Cell Transcriptomics and Spatial Imaging Data. Front Genet 12, 636743
38. Zhao E, Stone MR, Ren X et al (2021) Spatial transcriptomics at subspot resolution with BayesSpace. Nat Biotechnol 39, 1375-1384
39. Chidester B, Zhou T and Ma J (2021) SPICEMIX: Integrative single-cell spatial modeling for inferring cell identity. bioRxiv, 2020.2011.2029.383067
40. Cable DM, Murray E, Zou LS et al (2021) Robust decomposition of cell type mixtures in spatial transcriptomics. Nat Biotechnol
41. Elosua-Bayes M, Nieto P, Mereu E, Gut I and Heyn H (2021) SPOTlight: seeded NMF regression to deconvolute spatial transcriptomics spots with single-cell transcriptomes. Nucleic Acids Res 49, e50
42. Dong R and Yuan GC (2021) SpatialDWLS: accurate deconvolution of spatial transcriptomic data. Genome Biol 22, 145
43. Lopez R, Li B, Keren-Shaul H et al (2021) Multi-resolution deconvolution of spatial transcriptomics data reveals continuous patterns of inflammation. bioRxiv, 2021.2005.2010.443517
44. He B, Bergenstråhle L, Stenbeck L et al (2020) Integrating spatial gene expression and breast tumour morphology via deep learning. Nat Biomed Eng 4, 827-834
45. Pang M, Su K and Li M (2021) Leveraging information in spatial transcriptomics to predict super-resolution gene expression from histology images in tumors. bioRxiv, 2021.2011.2028.470212
46. Pham D, Tan X, Xu J et al (2020) stLearn: integrating spatial location, tissue morphology and gene expression to find cell types, cell-cell interactions and spatial trajectories within undissociated tissues. bioRxiv, 2020.2005.2031.125658
47. Tan X, Su A, Tran M and Nguyen Q (2020) SpaCell: integrating tissue morphology and spatial gene expression to predict disease cells. Bioinformatics 36, 2293-2294
48. Xu W, Gao Y, Wang Y and Guan J (2021) Protein-protein interaction prediction based on ordinal regression and recurrent convolutional neural networks. BMC Bioinformatics 22, 485
49. Bergenstråhle J, Larsson L and Lundeberg J (2020) Seamless integration of image and molecular analysis for spatial transcriptomics workflows. BMC Genomics 21, 482
50. Bergenstråhle L, He B, Bergenstråhle J et al (2020) Super-resolved spatial transcriptomics by deep data fusion. bioRxiv, 2020.2002.2028.963413
51. Kelly RT (2020) Single-cell Proteomics: Progress and Prospects. Mol Cell Proteomics 19, 1739-1748
52. Stoeckius M, Hafemeister C, Stephenson W et al (2017) Simultaneous epitope and transcriptome measurement in single cells. Nat Methods 14, 865-868
53. Schulz D, Zanotelli VRT, Fischer JR et al (2018) Simultaneous Multiplexed Imaging of mRNA and Proteins with Subcellular Resolution in Breast Cancer Tissue Samples by Mass Cytometry. Cell Syst 6, 25-36.e25

54. Piehowski PD, Zhu Y, Bramer LM et al (2020) Automated mass spectrometry imaging of over 2000 proteins from tissue sections at 100-μm spatial resolution. Nat Commun 11, 8

55. Thornton CA, Mulqueen RM, Torkenczy KA et al (2021) Spatially mapped single-cell chromatin accessibility. Nat Commun 12, 1274

56. Deng Y, Bartosovic M, Ma S et al (2021) Spatial-ATAC-seq: spatially resolved chromatin accessibility profiling of tissues at genome scale and cellular level. bioRxiv, 2021.2006.2006.447244

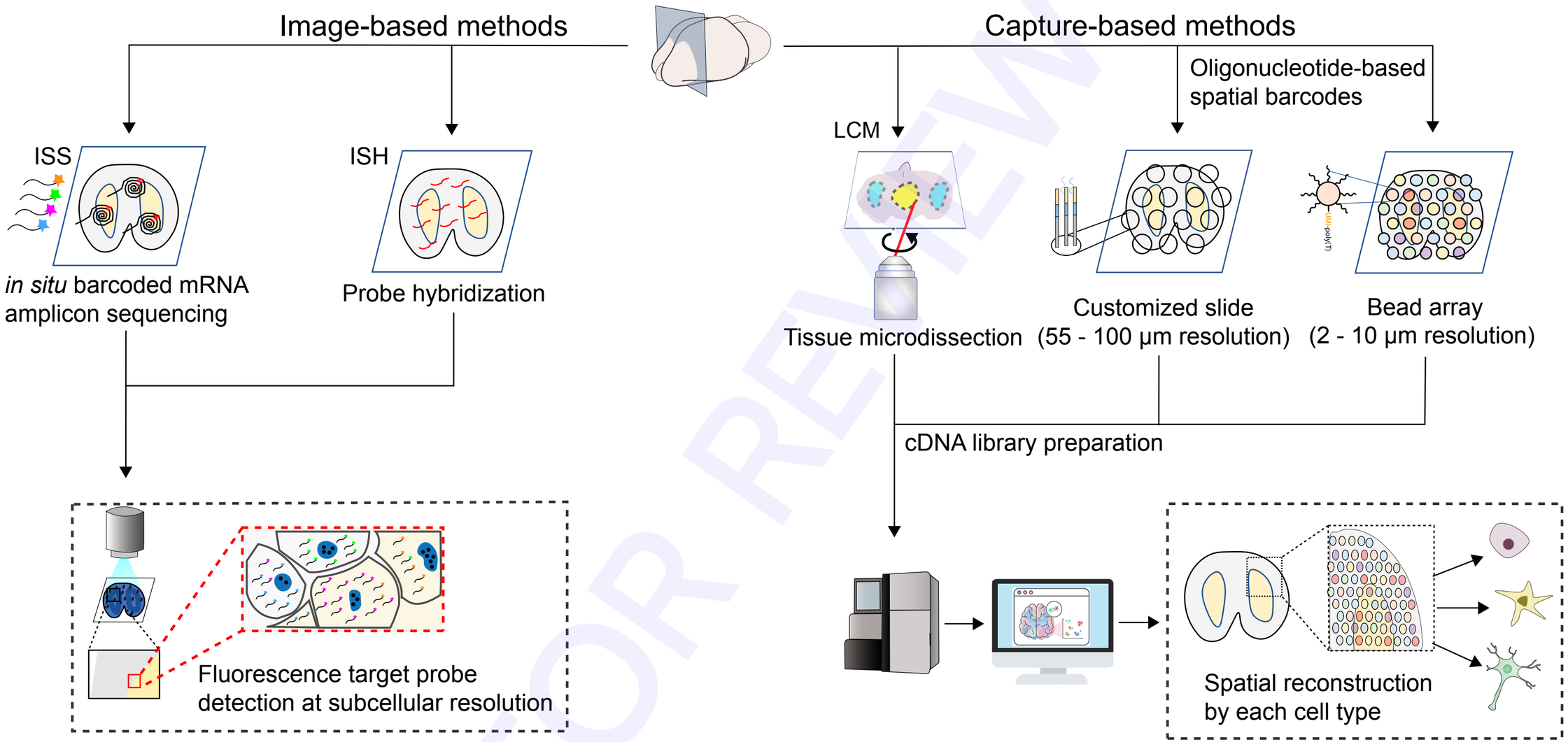
Image-based methods
Capture-based methods
Oligonucleotide-based spatial barcodes
ISS
ISH
LCM
in situ barcoded mRNA amplicon sequencing
Probe hybridization
Tissue microdissection
Customized slide (55 - 100 μm resolution)
Bead array (2 - 10 μm resolution)
cDNA library preparation
Fluorescence target probe detection at subcellular resolution
Spatial reconstruction by each cell type

Table 1. Spatially resolved transcriptomic techniques

Techniques	Features	Target genes	Application	Programming language	Reference
Image-based spatially resolved transcriptomics: ISH					
smFISH smHCR	Short DNA probes complementary to mRNA targets trigger chain reactions	39 probes (smHCR)	Zebra fish, mouse brain	MATLAB	Shah *et al.*, 2016 Development (7)
osmFISH	Binding of 20 nucleotide-long fluorescently labeled DNA probes	33 marker genes	mouse brain	Python	Codeluppi *et al.*, 2018 Nat Methods (8)
MERFISH	Using chemical cleavage instead of photobleaching to remove fluorescent signals	130 genes in up to 100,000 cells	cultured U-2 OS cells	MATLAB	Moffitt *et al.*, 2016 PNAS (9)
MERFISH-based analysis platform	In situ cell-type identification and mapping in combination with scRNA-seq	Targeting a set of 155 genes	mouse hypothalamic preoptic region	MATLAB	Moffitt *et al.*, 2018 Science (10)
seqFISH+	Enables visualization of the subcellular localization	10,000 genes in single cells	NIH/3T3, mouse brain	MATLAB	Eng *et al.*, 2019 Nature (11)
SABER	Additional signal amplification or applying serial imaging with DNA-Exchange	18,000 probes targeting a 3.9-Mb region	mouse retinal tissue	MATLAB & Python	Kishi *et al.*, 2019 Nat Methods (12)
Split-FISH	Alternative approach to reduce off-target background fluorescence by integrating split-probe strategy with multiplexed FISH	317 genes in single cell	mouse brain, liver, kidney, ovary	Python	Goh *et al.*, 2020 Nat Methods (13)
Image-based spatially resolved transcriptomics: ISS					
STARmap	Integrated with hydrogel-tissue chemistry and targeted signal amplification	160 to 1,020 genes simultaneously	mouse brain	Python	Wang *et al.*, 2018 Science (15)
INSTA-seq	Sequences two bases simultaneously from both ends of the cDNA fragments	up to 443,304 UMIs in total	drosophila retina	R	Fürth *et al.*, 2019 BioRxiv (16)
HybISS	New barcoding system via sequence-by-hybridization chemistry	119 genes for PLP design	mouse visual cortex, human brain	MATLAB	Gyllborg *et al.*, 2020 Nucleic Acids Res (17)
pciSeq	Bayesian algorithm derived from scRNA-seq clusters data	designed 755 probes for 99 genes	mouse CA1 interneuron	Python	Qian *et al.*, 2020 Nat Methods (18)
ExSeq	cDNA amplicons are eluted from the sample and re-sequenced	up to 3,039 genes with untargeted approach	mouse brain, mouse visual cortex, human breast cancer	MATLAB	Alon *et al.*, 2021 Science (19)
Capture-based spatially resolved transcriptomics: LCM					
exome-capture RNA-sequencing	Optimized standard protocol for hematoxylin and eosin (H&E) staining	Whole exome	7 tumor samples of TNBC	MATLAB	Romanens *et al.*, 2020 BioRxiv (20)
immuno-LCM-RNAseq	RNA quality was significantly improved using modified protocol	Up to 60 cells were demonstrated to be sufficient quality	mouse small intestine	Python	Zhang *et al.*, 2021 BioRxiv (21)
PIC	Photo-irradiated cells were suppressed cDNA amplification	8,000 genes were detected with 7 × 10^4 unique read counts	mouse embryo	R	Honda *et al.*, 2021 Nat Commun (22)
Capture-based spatially resolved transcriptomics: Oligonucleotide-based spatial barcode on slide					
ST	Arrayed reverse transcription primers with unique positional barcodes	Up to 200 million oligonucleotides in each of 1007 features	mouse brain, human breast cancer	R	Ståhl *et al.*, 2016 Science (24); Salmén *et al.*, 2018 Nat Protoc (23)
Multimodal analysis	Combined single-cell RNA sequencing with ST	Median depth of 1,629 UMIs/spot and 967 genes/spot	human cSCC	MATLAB & R	Ji *et al.*, 2020 Cell (29)
ST	Combined ST and ISS	Mean 31,283 UMIs and 6,578 unique genes per TD	AD mouse model, mouse and human brain	Python	Chen *et al.*, 2020 Cell (28)
Capture-based spatially resolved transcriptomics: Oligonucleotide-based spatial barcode on bead array					
Slide-seq	DNA-barcoded beads with known positions (in situ indexing)	1.5 million beads, of which 770,000 can be analyzed	mouse cerebellum and hippocampus	MATLAB & R	Rodriques *et al.*, 2019 Science (30)
HDST	Barcoded poly(d)T oligonucleotides into 2-μm wells with a randomly ordered bead array-based	2,893,865 individual barcoded beads	mouse brain, primary breast cancer	Python	Vickovic *et al.*, 2019 Nat Methods (31)
Slide-seq V2	Improvements in library generation, bead synthesis and array indexing	Mean 45,772 UMIs in 110 μm diameter area	mouse hippocampus	MATLAB & R & Python	Stickels *et al.*, 2021 Nat Biotechnol (32)
Seq-Scope	Based on a solid-phase amplification using an Illumina sequencing platform	Up to 5.88 ~ 19.7 genes were identified per HDMI pixel	human liver and colon	Python	Cho *et al.*, 2021 Cell (33)
Stereo-seq	Combined DNA nanoball pattern arrays and tissue RNA capture	Up to 133,776 UMIs per 100 μm diameter	mouse brain	Not identified yet	Chen *et al.*, 2021 BioRxiv (34)

Table 2. Integration of spatially resolved transcriptomic data with other methods

Techniques		Features	Input data	Application	Programming language	Reference
Combination with scRNA-seq						
	pciSeq	Bayesian algorithm derived from scRNA-seq clusters data with ISS	Designed 755 probes for 99 genes	mouse CA1 interneuron	Python	Qian *et al.*, 2020 Nat Methods (18)
	seqFISH	Computing the ratio of the performance and prediction scores with scRNA-seq data	Each cell contained avg 196 mRNA from 93.2 genes	embryo development in brain and gut	R	Lohoff *et al.*, 2021 Nat Biotechnol (35)
	Multiple spatial transcriptomics	Unbiased approach with additional in situ hybridization using RNAscope and multi-molecule ISS	Mean 31,283 UMIs and 6,578 unique genes per tissue domain	AD mouse model, mouse and human brain	Python	Chen *et al.*, 2020 Cell (25)
	Slide-seq	NMFreg that reconstructs expression of each cell type signatures defined by scRNA-seq	1.5 million beads, of which 770,000 could be analyzed	mouse cerebellum and hippocampus	MATLAB & R	Rodriques *et al.*, 2019 Science (30)
	Integrating microarray-based ST and MIA	Enrichment analysis that two-tailed Student's t-test were used to compare expression of those marker genes	2,500~3,300 UMIs and 1,400~1,700 unique expressed genes per single cell	pancreatic ductal adenocarcinoma	R	Moncada *et al.*, 2020 Nat Biotechnol (36)
Deep learning-based spatial information						
	DEEPsc	Deep-learning network was trained with spatial position feature vectors as simulated scRNA-seq data	Started with top 3,000 highly variable genes	drosophila embryo, Zebrafish embryo, murine frontal cortex	MATLAB	Maseda *et al.*, 2021 Front Genet (37)
	BayesSpace	Bayesian statistical method that uses the information from spatial neighborhoods to achieve super-resolution images	10X Genomics Visium data, does not require independent single-cell data or marker gene preselection	brain, melanoma, invasive ductal carcinoma, ovarian adenocarcinoma	R	Zhao *et al.*, 2021 Nat Biotechnol (38)
	SPICEMIX	Enhances the NMF of gene expression with a graphical representation of the spatial relationship of cells	2,470 genes in 523 cells (seqFISH+), 930 cells and 1,020 genes (STARmap)	mouse primary visual cortex	Python	Chidester *et al.*, 2020 BioRxiv (39)
	SpaOTsc	Infer the spatial distance between every pair of cells by computing the optimal transport distance	851~15,413 cells and 10,495~45,789 genes (scRNA-seq), 64~1,549 spatial positions and 47~1,020 genes	Zebrafish embryo, Drosophila embryo, mouse visual cortex	Python	Cang *et al.*, 2020 Nat Commun (4)
Deconvolution of spatially resolved transcriptomics						
	RCTD	Statistical model assumed to be Poisson distributed and maximum-likelihood estimation (MLE) used to infer the cell types	Slide-seq and 10X Genomics Visium data	mouse brain	R	Cable *et al.*, 2021 Nat Biotechnol (40)
	SPOTlight	NMF along with non-negative least squares (NNLS) model with both the basis and coefficient matrices with cell type marker genes	41,986 cells were merged to identify a total of 10,623 immune cells	mouse brain, pancreatic adenocarcinoma	R	Elosua-Bayes *et al.*, 2021 Nucleic Acids Res (41)
	SpatialDWLS	Dampened weighted least squares (DWLS) model with cell-type specific gene signatures from a public scRNA-seq dataset as a reference	10,000 genes in 523 cells (seqFISH+)	mouse brain, human heart	R	Dong *et al.*, 2021 Genome Biol (42)
	DestVI	Bayesian model for multi-resolution deconvolution of cell types using Variational Inference	Pair of ST and scRNA-seq from same tissue	murine lymph node, mouse tumor model	Python	Lopez *et al.*, 2021 BioRxiv (43)
Cell type Inference via image-based machine learning						
	ST-Net	Deep learning algorithm that combines ST and histology images to predict the target gene expression of each spot	30,612 spots in 68 breast tissue sections	breast cancer	Python	He *et al.*, 2020 Nat Biomed Eng (44)
	HisToGene	employs a modified Vision Transformer model for gene expression prediction from histology images	9,612 spots and 785 genes in breast cancer tissue	breast cancer	PyTorch	Pang *et al.*, 2021 BioRxiv (45)
	stLearn	Deep neural network model to predict hotspots where cell-cell interactions are more likely to occur	Feature vectors from H&E images of the tissue section	mouse brain, human brain, breast cancer	Python	Pham *et al.*, 2020 BioRxiv (46)
	SpaCell	Normalized count data and H&E staining images were trained with convolutional neural network	Tissue morphology and spatial gene expression data	prostate cancer, amyotrophic lateral sclerosis	Python	Tan *et al.*, 2020 Bioinformatics (47)
	CoSTA	Clustering by Gaussian mixture model (GMM) and weight updating as commonly performed in training neural networks	Image-type matrix of MERFISH and Slide-seq data	mouse brain, brain Injury	Python	Xu *et al.*, 2021 BMC Bioinformatics (48)
	STUtility	NMF to decompose ST data and identification and extraction of neighbouring capture-spots	10x Genomics Visium data	mouse brain, breast cancer tissue, lymph node, rheumatoid arthritis	R	Bergenstråhle *et al.*, 2020 BMC Genomics (49)
	ISST	Image-based *in silico* spatial transcriptomics without hybridization or sequencing	12 sections from the mouse olfactory bulb	mouse olfactory bulb, human breast cancer	Python	Bergenstråhle *et al.*, 2020 BioRxiv (50)
Cellular protein information						
	CITE-seq	Oligonucleotide-labeled antibodies are used to integrate cellular protein and	Common immune subpopulation	human HeLa, mouse 4T1 cell, immune	R	Stoeckius *et al.*, 2017 Nat

		transcriptome measurements	markers (CD8a, CD3e, CD19, CD56, CD16, CD11c and CD14)	subpopulation		Methods (52)
	IMC	Epitope-based imaging methods that employ a mass spectrometer for readout to infer RNA-to-protein correlations	Detected three mRNA simultaneously (HER2, CK19 and CXCL10)	breast cancer	MATLAB & R & Python	Schulz *et al.*, 2018 Cell Syst (53)
	nanoPOTS	Unique proteins were identified via combination of LCM and ultrasensitive nanoLC-MS/MS	> 2,000 proteins with 100 μm spatial resolution	mouse luminal epithelial cell, stromal cell, glandular epithelial cell	R	Piehowski *et al.*, 2020 Nat Commun (54)
Spatial ATAC-seq						
	sciMAP-ATAC	Spatially resolved, single-cell profiling of chromatin states from a single tissue punch	Mean 12,052 - 30,212 passing reads per cell	mouse and human brain, cerebral ischemia model system	R	Thornton *et al.*, 2021 Nat Commun (55)
	Spatial-ATAC-seq	DNA barcode solutions were introduced to the tissue surface using an array of microchannels	36,303 - 100,786 unique fragments per pixel	mouse embryos, human tonsil tissue	R	Deng *et al.*, 2021 BioRxiv (56)