Biophys Rep 2022, 8(3):119–135
https://doi.org/10.52601/bpr.2021.210037

Biophysics Reports

REVIEW

Special Topic - SINGLE CELL

Spatial transcriptomics: new dimension of understanding biological complexity

Zhuxia Li[1,4], **Guangdun Peng**[1,2,3] ✉

[1] Ceter for Cell Lineage and Development, CAS Key Laboratory of Regenerative Biology, Guangdong Provincial Key Laboratory of Stem Cell and Regenerative Medicine, Guangzhou Institutes of Biomedicine and Health, Chinese Academy of Sciences, Guangzhou 510530, China
[2] Center for Cell Lineage and Atlas, Bioland Laboratory (Guangzhou Regenerative Medicine and Health Guangdong Laboratory), Guangzhou 510005, China
[3] Institute for Stem Cell and Regeneration, Chinese Academy of Sciences, Beijing 100101, China
[4] University of Chinese Academy of Sciences, Beijing 100049, China

Received: 1 August 2021 / Accepted: 18 October 2021

Abstract Cells and tissues are exquisitely organized in a complex but ordered manner to form organs and bodies so that individuals can function properly. The spatial organization and tissue architecture represent a keynote property underneath all living organisms. Molecular architecture and cellular composition within intact tissues play a vital role in a variety of biological processes, such as forming the complicated tissue functionality, precise regulation of cell transition in all living activities, consolidation of central nervous system, cellular responses to immunological and pathological cues. To explore these biological events at a large scale and fine resolution, a genome-wide understanding of spatial cellular changes is essential. However, previous bulk RNA sequencing and single-cell RNA sequencing technologies could not obtain the important spatial information of tissues and cells, despite their ability to detect high content transcriptional changes. These limitations have prompted the development of numerous spatially resolved technologies which provide a new dimension to interrogate the regional gene expression, cellular microenvironment, anatomical heterogeneity and cell-cell interactions. Since the advent of spatial transcriptomics, related works that use these technologies have increased rapidly, and new methods with higher throughput and resolution have grown quickly, all of which hold great promise to accelerate new discoveries in understanding the biological complexity. In this review, we briefly discussed the historical evolution of spatially resolved transcriptome. We broadly surveyed the representative methods. Furthermore, we summarized the general computational analysis pipeline for the spatial gene expression data. Finally, we proposed perspectives for technological development of spatial multi-omics.

Keywords Spatial transcriptomics, Spatial data analysis, Spatial multi-omics, Histology, Single-cell sequencing

INTRODUCTION

Multicellular organisms rely on a variety of cellular events such as cell movement, communication between interacted cells and position-dependent gene expression at distinct microenvironments. Studying these cellular events on the anatomical structure is essential to the understanding of many important biological and pathological questions. In essence, the orchestrated composition of cells and tissues at predefined locations is a prerequisite for the complex functionality of biological system. For example, the

✉ Correspondence: peng_guangdun@gibh.ac.cn

© The Author(s) 2022

steady-state lymph nodes have intrinsic location patterns including cortex, paracortex and medulla, composed by distinct cell types (Stoltzfus *et al*. 2020). The correct position and interaction between immune cell subpopulations are critical for lymph nodes to function well and maintain immune homeostasis. The approaches to systematically survey a cell's transcriptome in its tissue architecture at genome-wide level become a biological imperative.

Recently, plenty of spatial gene profiling technologies have been developed, aiming to reveal the molecular changes on 2D and 3D tissue samples at given positions and even at single cell or subcellular resolution (Chen *et al*. 2021; Cho *et al*. 2021). The modern age technologies are originated from historical tools that can quantify gene expression and preserve spatial information. Some essential spatial gene atlas and databases during numerous dynamic processes have been accomplished by these historical methods. For example, histological staining using chemical dye such as hematoxylin and eosin, has been performed on tissue samples containing cells of distinct types and positions, and the tissue organization and anatomical structure are visualized under electron or light microscopes (Alturkistani *et al.* 2015). Furthermore, an improvement on histological staining has been achieved by targeting more specific tissue regions or cell subgroups with the application of molecular techniques, for example, genes or gene promoters are fused to fluorescent signals to allow for gene trapping or promotor trapping, which are popular tools to investigate the influence of specific genes during embryonic development (Skarnes *et al*. 1995). Nevertheless, gene trapping and promotor trapping are both random insertions, which means the vector is probably biasedly inserted into some large gene or promoter sequences. Moreover, only one or a few selected genes could be studied at one time. Of note, *in situ* hybridization (ISH) or immunohistochemistry (IHC) is an accurate and effective method, which detects multiple spatial patterns of gene or protein with the aid of nucleic acid probes or antibodies. With the development of sequencing technology in the late 20th century, genomic information of model species was completed, which enhanced the design and synthesis of *in situ* reporters. For instance, fused cosmids with reporter and upstream regulators based on whole genome sequence were constructed to reveal the expression of many development-dependent genes in nematode *Caenorhabditis elegans* (Lynch *et al*. 1995).

While traditional spatial techniques have common unfulfillment-samples and genes that can be simultaneously examined are limited. In addition, these methods with a few predefined genes and a small sampling size may result in the biased descriptions. It is also unclear in which position the whole genome is transcribed on the natural tissue structure, which is indispensable for exploring molecular distributions, discovering novel spatial expression patterns and querying the function of candidate genes in development, diseases and cancer. Therefore, expanding our knowledge on the expression of transcripts meanwhile preserving the morphology of tissues will facilitate the understanding of cell type heterogeneity, cell-cell interaction, cell fate dynamics in normal and abnormal biological contexts. Hence, unbiased and systematic gene profiling of tissues with spatial coordinates requires the implement of spatial transcriptome technologies.

Even though basic principles and capacities of spatial transcriptomic technologies vary widely, methods of transcriptome-level profiling can be roughly classified into three groups according to their space-procuring protocols: technologies based on FISH, technologies based on isolation through microdissection or photoreaction, and array-based technologies.

TECHNOLOGIES BASED ON FISH

In situ hybridization (ISH) was first reported in 1969 by Pardue and Gall to detect the hybrid of DNA and RNA. The main principle of ISH for visualizing target molecules is the detection of hybridization signals of labeled molecule probes including isotope, fluorescence and enzyme by autoradiography or microscopy (Francoz *et al*. 2016). FISH has been the standard approach and the routine procedure in various experimental contexts (Shaffer *et al*. 2017). Powered by ISH, a number of imaging-based RNA sequencing techniques have been developed to generate high-resolution spatial molecular atlas of different organ at different developmental systems at single-cell and subcellular levels (Lohoff 2021).

Nevertheless, there are still drawbacks of these technologies, such as the challenge of complicated image processing and the difficulties to keep the balance between detection efficiency, sensitivity and the detected number of transcripts. Importantly, a large amount of customized costly probes must be designed and synthesized with *a priori* knowledge, which may limit the application and obstruct new findings.

smFISH

Compared to the long probes targeting one transcript in ISH, single molecule FISH (smFISH) utilizes a series of

 © The Author(s) 2022

shorter fluorescent probes (often 40–50 nucleotides) hybridizing to different sequences of one nucleic acid molecule (Femino *et al*. 1998). Individual small probes are coupled with five fluorophores (such as CY3 and FITC). In this way, the targeted DNA or RNA can be marked with numerous fluorescence, and then imaged by fluorescent microscopy to obtain the information of its high intensity light spots (Fig. 1A). The image information of each gene or mRNA molecule is constrained by deconvolution algorithms into voxels. Other interference voxels introduced by autofluorescent signal and background noise are

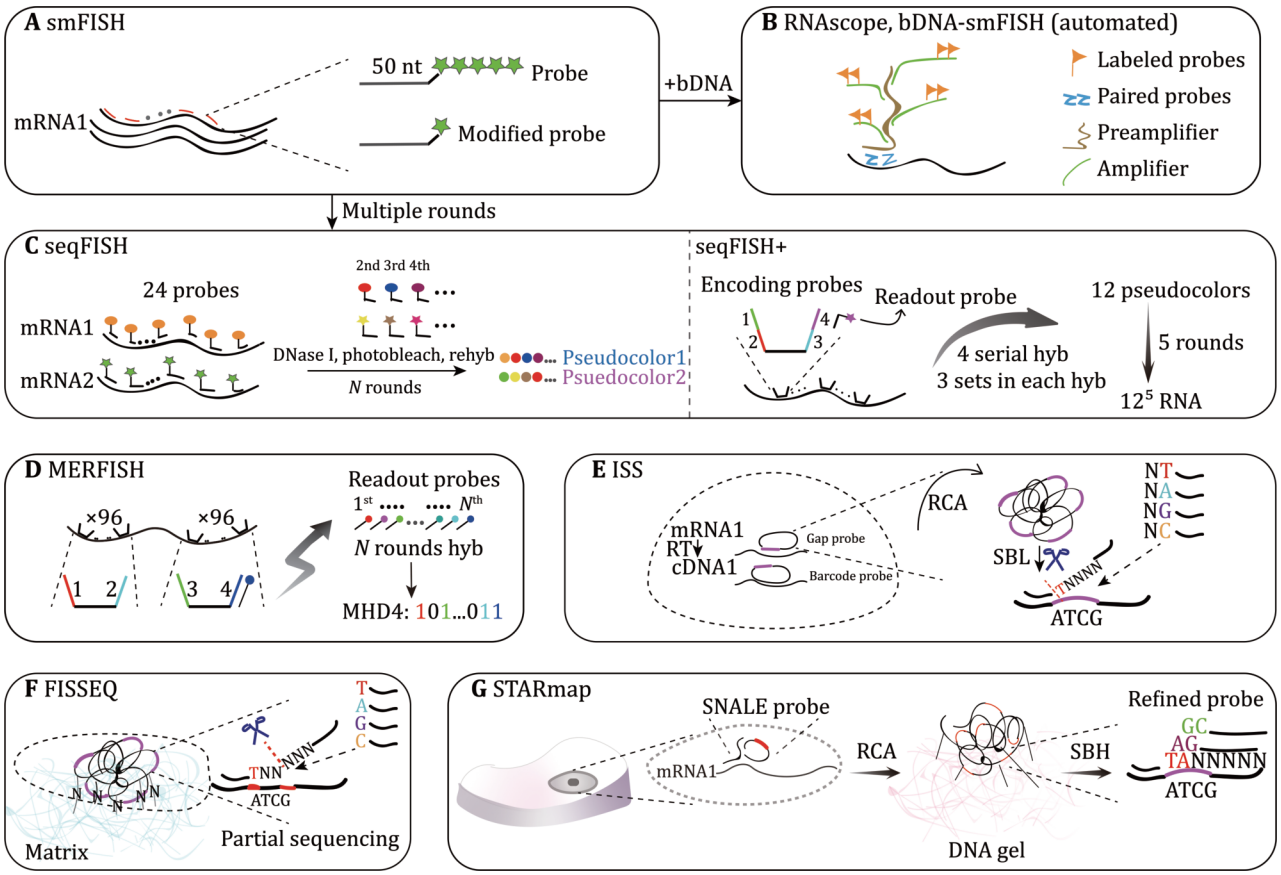

Fig. 1 Technologies based on FISH. **A** smFISH: 50 nt probes labeled with the same five fluorescence are bound to targeted mRNA. Improved probes are merely labeled with one fluorescence. **B** RNAscope and bDNA-smFISH: based on smFISH, branched DNA technology is merged to amplify mRNA signals in the background of noise. Probes pair firstly hybridize to specific sequences on mRNA, which provide a site that preamplifier binds. Several amplifiers with multiple labeled probes are bound with preamplifiers. bDNA-smFISH can be completed by automated liquid handling. **C** seqFISH: a series of 24 probes with respectively distinct fluorophores are loaded into individual target mRNA. In the first round, 24 probes with specific color are introduced, then probes are removed by DNase I and photobleached. Identical probes labeled with different fluorophore are rehybridized at the next successive rounds. Finally, each mRNA can be detected as unique composite colors of its probes. seqFISH+ applies encoding probes instead of common probes, which are composed of target complementary sequence and four readout sequences that can bind to distinct colored readout probes. In each four hybridizing rounds, 12 sets of readout probes with just three colors are used to generate 12 pseudocolors representing targeted mRNA. Therefore, after five pseudocolor rounds, each mRNA is marked by a composite color of five pseudocolors. **D** MERFISH: two sets of 96 encoding probes with two readout parts are linked to a particular target. Different fluorescence conjugated readout probes are added respectively in N rounds, decoding mRNA by MHD4 which "1" means hybridizing signal and "0" means no signal. **E** ISS: cDNA reversely transcribed from mRNA is identified by padlock probes. Here two strategies can be used: (1) Probe with a gap at the complementary site is circularized to form target sequence; (2) Probe perfectly binding to mRNA has a barcode on the remaining site. After circularization, multiple target sequences or barcode sequences can be recognized through SBL. **F** FISSEQ: cDNA nanoball can be fixed on the extracellular matrix by ligation of amine-reaction and partially sequenced. **G** STARmap: mRNA of cells residing in thick tissue directly links to SNAIL probes. Primer pairing to target mRNA and padlock probe initiates RCA. Barcodes of nanoball conjugated with DNA gel are then encoded by SBH in which the base pair of sequencing probes are labeled with correspondent color for error correction (SEDAL). MHD4: modified Hamming distance; SBL: sequencing by ligation; SBH: sequencing by hybridization

© The Author(s) 2022

filtered by setting a threshold. Therefore, smFISH provides the advantages of high-resolution and sensitivity in visualizing and detecting single target molecules. For example, smFISH was described to successfully identify the dynamic snapshots of the transcription initiation, termination, processing and transporting of β-actin mRNA in normal rat kidney cells using oligonucleotide probes targeting to 3-UTR(s), 5-UTR(s), intron and splice conjugation regions on the β-actin transcript (Femino *et al*. 1998).

smFISH allows us to study interested coding and noncoding genes and RNA at the single-cell level and visualize their subcellular locations robustly and quantitatively. However, it is difficult to synthesize and purify multiple fluorescence labeled oligonucleotide probes, and the hybridization properties might be altered when the target has a high concentration. An improved method was published in 2008. In this procedure, larger amounts of shorter probes (20 nucleotides) that are labeled with single fluorophores at the 3' end were used (Raj *et al.* 2008) (Fig. 1A). This updated smFISH protocol mitigates the potential off-target effects and false positives from a few probes in the detection of FISH signals. However, only a few nucleic acid molecules can be studied simultaneously owing to the general spectral overlap defect of fluorescent microscopy in single cells.

smFISH combined with branched DNA technology such as RNAscope (Wang *et al.* 2012) and bDNA-smFISH (Battich *et al.* 2013) significantly enhances the intensity of signal, in which genetic target probe, preamplifier, amplifier and fluorescent probes are used to make cascades targeting a particular gene (Fig. 1B). Thousands of human single cells were imaged by bDNA-smFISH with a high ratio of target signal to background noise. For example, smFISH and related methods have been exploited to study the subcellular localizations and cell–cell variability of human lncRNA (Cabili *et al.* 2015), examine the relationship between cellular volume and transcription abundance (Padovan-Merhar *et al.* 2015) and quantitate the 3D mRNA content on tissue of *Drosophila* embryos (Trcek *et al.* 2017).

Sequential fluorescence *in situ* hybridization (seqFISH)

smFISH has superior resolution but is still limited by target transcripts that can be simultaneously detected. To address this issue, seqFISH (Lubeck *et al.* 2014) improves the multiplexity of target molecules detection by hybridization of multiple rounds of probes. 24 oligonucleotide probes coupled with identical fluorescence are firstly hybridized to each mRNA of fixed cells and imaged *in situ*. Next, the probes are raveled out by DNase I and photobleached to prepare for later rounds. The same steps are repeated using the same group of probes but with different fluorophores in the following hybridization rounds. The sequential color barcode at a fixed locus of the same mRNA finally aligns together. Therefore, after N rounds of FISH, each mRNA possesses a unique sequential barcode composition (Fig. 1C). In principle, seqFISH has the capability of distinguishing N^m target molecules using m species of probes with N rounds of hybridization. seqFISH has been validated by studying 12 mRNA by employing four sets of probes in two rounds (Lubeck *et al.* 2014).

In 2018, the same research group further improved seqFISH to profile the transcription of 10,421 genes. In this study, the author applied seqFISH to intron regions of nascent RNA (Shah *et al.* 2018). 25 primary probes complementary with different sites of target RNA share the same bilateral overhangs, and each overhang consists of two distinct readout sequences which provides the binding places of readout probes. Readout probes labeled with fluorescence are introduced into the reaction and bound with the corresponding overhang sequence on primary probes. In the first four rounds of hybridization, readout probes are removed at each round while primary probes are still hybridized with introns. Three colored readout probes that separately denote three distinct RNA molecules are imaged each time. Therefore 12 targets at different locations with just three colors labeled readout probes are collected after four rounds, which are transformed into 12-pseudocolors due to different combinations of position and color. When five series of four rounds of hybridization are completed, the sequential barcodes of five 12-pseudocolors can identify 12^5 RNA species in principle (Fig. 1C). It shows the potential to profile the whole transcriptome in single fixed cells.

It is worthy to note that as the hybridization rounds increase, drop out events and imaging crowding occur more frequently which may confound the precise identification of targets.

Multiplexed error-robustness fluorescence *in situ* hybridization (MERFISH)

Zhuang *et al*. established a highly multiplexed spatial technology called MERFISH with the power to detect and correct errors during hybridization rounds (Chen *et al.* 2015). They skillfully incorporated the superiority of Oligopaints (Beliveau *et al.* 2012) and seqFISH, which markedly increased the efficiency of hybridization and the throughput of gene expression profiles. Another

 © The Author(s) 2022

new strategy called binary code of modified Hamming distance (MHD4) was adopted in MERFISH to make this procedure equipped with the robustness of error detection and correction. Two sets of 96 encoding probes are synthesized for a single target gene. Every encoding probe contains a gene target sequence in the center and two readout sequences in the flanks which will bind to diverse complementary readout probes conjugated with a particular fluorophore. Individual gene has a barcode consisting of four types of color. In each round of successive application of *N* readout probes on multiple RNA species, only one kind of readout probe with one color is added and subsequently photobleached. Ultimately, targeted transcripts will produce *N* bites of binary code where "1" marks the successful detection of the signal from hybridized readout probe while "0" means no signal (Fig. 1D). The performance of MERFISH has been confirmed by imaging hundreds and thousands of genes residing in hundreds of single cells (Chen *et al.* 2015). Similar to seqFISH, detection accuracy and efficiency will be influenced as binding rounds go up. Besides, the high cost of probes synthesis and the speed of imaging restrict its wide implementation.

To improve the adaptability of this method, Zhuang *et al*. optimized MERFISH and further expanded the multiplexity in two ways (Moffitt *et al.* 2016). Instead of photobleaching, the readout probes were clipped by a chemical method during successive rounds of smFISH to reduce the scale-dependent imaging time. On the other hand, they extended the microscopic field of view (FOV) to prevent the complex and time-consuming alteration of small FOVs. They implemented this advanced MERFISH on 100,000 cultured cells by imaging the spatial distribution of 140 genes.

In 2018, the same group further merged expansion microscopy into MERFISH, which facilitated the measurement of these molecules with high density compared to MERFISH (Wang *et al.* 2018a). High abundance cDNA is stretched by binding on the expandable gel. The efficiency and accuracy of imaging high concentration molecules boost largely by expansion microscopy. Additionally, this technology is versatile for its compatibility with other ISH and imaging technologies. For instance, branched DNA technology is fused into MERFISH, making it stronger for visualizing shorter RNA species with fewer encoding probes in less exposure time even using lower power laser illumination sources. Remarkably, the size of fluorescent spots and inter-spot brightness variance is not influenced when increasing FISH signal (Xia *et al.* 2019).

Ouroboros smFISH (osmFISH)

For the purpose of resolving the optical overlap between different fluorescence, one option is to reduce the throughput of genes each time. A similar technology named ouroboros smFISH that discards multiplexed barcoding and amplification identifies a small number of genes in each probe bound round (Codeluppi *et al.* 2018). Therefore, spots of the individual transcripts can be detected accurately without optical distraction from others. After several rounds of continuous hybridization and removal of unique fluorescently probes, a moderate quantity of marker genes' expression profiles is collected without interference on comparatively low abundance ones. osmFISH can perform on thin larger scale tissue such as the brain section to resolve spatial organization of cell types within the anatomical structure. Although it is designed as a semiautomated pipeline, it is still time-consuming.

In situ sequencing (ISS)

ISS makes directly visualizing and *in situ* sequencing the base pair of target genes a reality. Two strategies are adopted in ISS, including gap-targeted sequencing and barcode-targeted sequencing with padlock probes (Ke *et al.* 2013). In gap-targeted sequencing, padlock probes bind to cDNA that are reversely transcribed from target mRNA with a gap, and cyclized by polymerization and ligation reactions. While in barcode-targeted sequencing, padlock probes perfectly hybridize to cDNA with a barcode sequence on the remaining non-hybridization region, which are circularized only by ligation step. Through rolling circle amplification (RCA) using target primer, rolling-circle products (RCPs) with plenty of replicated target sequences are synthesized finally. Target sequences on the RCP are then sequenced by ligation with the aid of anchor probes which binds to RCP near the target sequences. Subsequently, 9-mer interrogation probe with a fixed 1/4 color representing four natural bases (A/T/C/G) on one random site and eight degenerated bases on the other sites, ligates to anchor probe when its fixed base matches well with the one on the target sequence. Interrogation probes are then imaged and removed for SOLiD (sequencing by oligonucleotide ligation and detection) sequencing of remaining bases on the target sequence (Fig. 1E). Images of the base after several ligations should reflect the real sequence of the target region.

Barcoded anatomy resolved by sequencing (BARseq) combines multiplexed analysis of projections by sequencing (MAPseq) with ISS to achieve the mapping

© The Author(s) 2022

of long-range neuronal projection, which possesses the power to illustrate the spatial anatomy organization of the nerve system by gene expression features *in situ* (Chen *et al.* 2019).

ISS has the ability to read short target sequences. Mutation of nucleotides can be also recognized by gap-targeted sequencing. But RCA has the disadvantages of increasing the size of fluorescent spots which will impede the discrimination of close molecules. It is also inefficient because RNA needs to be reversely transcribed, which may introduce additional technical errors. Recently, some emerging techniques have circumvented this limitation to improve the efficiency of target gene detection, such as BOLORAMIS and SCRINSHOT (Liu *et al*. 2021; Sountoulidis *et al*. 2020), which elegantly utilize barcoded probes that are directly hybridized to RNA and ligated by SplintR ligase to generate RCA copies. However, the number of genes that can be studied in each experiment awaits significant amelioration.

Fluorescent *in situ* RNA sequencing (FISSEQ)

Researchers have directly sequenced 8102 genes on human fibroblast cells using FISSEQ (Lee *et al.* 2014), which also employs RCA to magnify fluorescent signals of target single molecules. In this method, RNA is reversely transcribed into cDNA using random hexamer primers. cDNA with aminoallyl dUTP cross-links to the matrix through amine reactive linking reaction mediated by PEG. Fixed cDNA is amplified by RCA after circularization to form a single strand nanoball with the size of 200–400 nm in diameter, which similarly cross-links by PEG. Based on sequencing by ligation, partial sequencing is used to sample sequence the RCP in order to minimize the rounds of sequencing but not impair the recognition of each cDNA's location (Fig. 1F). Nevertheless, FISSEQ needs frequent transition between the small fields of view, and the detection sensitivity is influenced by the increase of detected genes.

Spatially-resolved transcript amplicon readout mapping (STARmap)

Inherited from ISS, STARmap realizes the 3-dimension (3D) sequencing for intact thick tissue in single-cell resolution. The hydrogel-tissue chemistry methodology is employed in STARmap to maintain the accuracy of multiple detections by stabilizing the location and integrity of amplicons that are produced through enzyme catalyzed RCA (Wang *et al.* 2018b). The pair of probes are introduced to increase the specificity of detected mRNA by lowering the background noise resulting from the mismatches of probes to targets. Padlock probe that can bind to both RNA and modified primers will initiate RCA only when there is correct hybridization of primers with mRNA and padlock probes. Here the approach using padlock probes and modified primers is defined as SNAIL (specific amplification of nucleic acids via intramolecular ligation). The amplicon with numerous repeated 5 nt gene-specific identifier fragments, modified with amine during RCA, is embedded into the hydrogel framework via hydrogel-tissue chemistry. Subsequently, the unlinked proteins and lipids are cleared for enhancing optical transparency and diffusion rate. Notably, the sequencing by ligation is applied to hydrogel-tissue with refined sequencing probes which identifies two bases on the identifier segment in one of five rounds, offering the capacity of sequencing with two-base encoding for error correction (SEDAL). STARmap successfully resolved the distribution and anatomical pattern of different genotypes by sequencing over a thousand mRNA in single-cell on 3D brain tissues.

TECHNOLOGIES BASED ON ISOLATION THROUGH MICRODISSECTION OR PHOTOREACTION

FISH based technology aims at counting and recognizing partial transcriptomes *in situ* within intact tissues. Alternatively, specifically selecting the regions of interest (ROI) or cells from original tissue also allows us to analyze the spatial profiles at the genome-wide level. Moreover, the advantage of high throughput and sequencing depth enhances the illumination of molecular dynamics of precisely controlled biological processes (Brown *et al*. 2002; van den Brink *et al*. 2020).

Laser capture microdissection (LCM)

LCM is one of the location-directed capturing strategies, which puts the tissues between laser beam and thermoplastic film for accurate isolation of pre-defined cells (Emmert-Buck *et al*. 1996). It begins with mounting frozen tissue on the glass slice followed by regular histological staining. A thin, sterile transparent film with an adhesion layer laid above the tissue slice will be cut by laser beams at predefined locations. Cells in these positions can be separated from the remaining tissues via the focal adhesion force of activated film or adhesive caps, which are then subjected to downstream molecular biological analysis (Fig. 2A). LCM has the capability to interrogate heterogeneous tissues such as

 © The Author(s) 2022

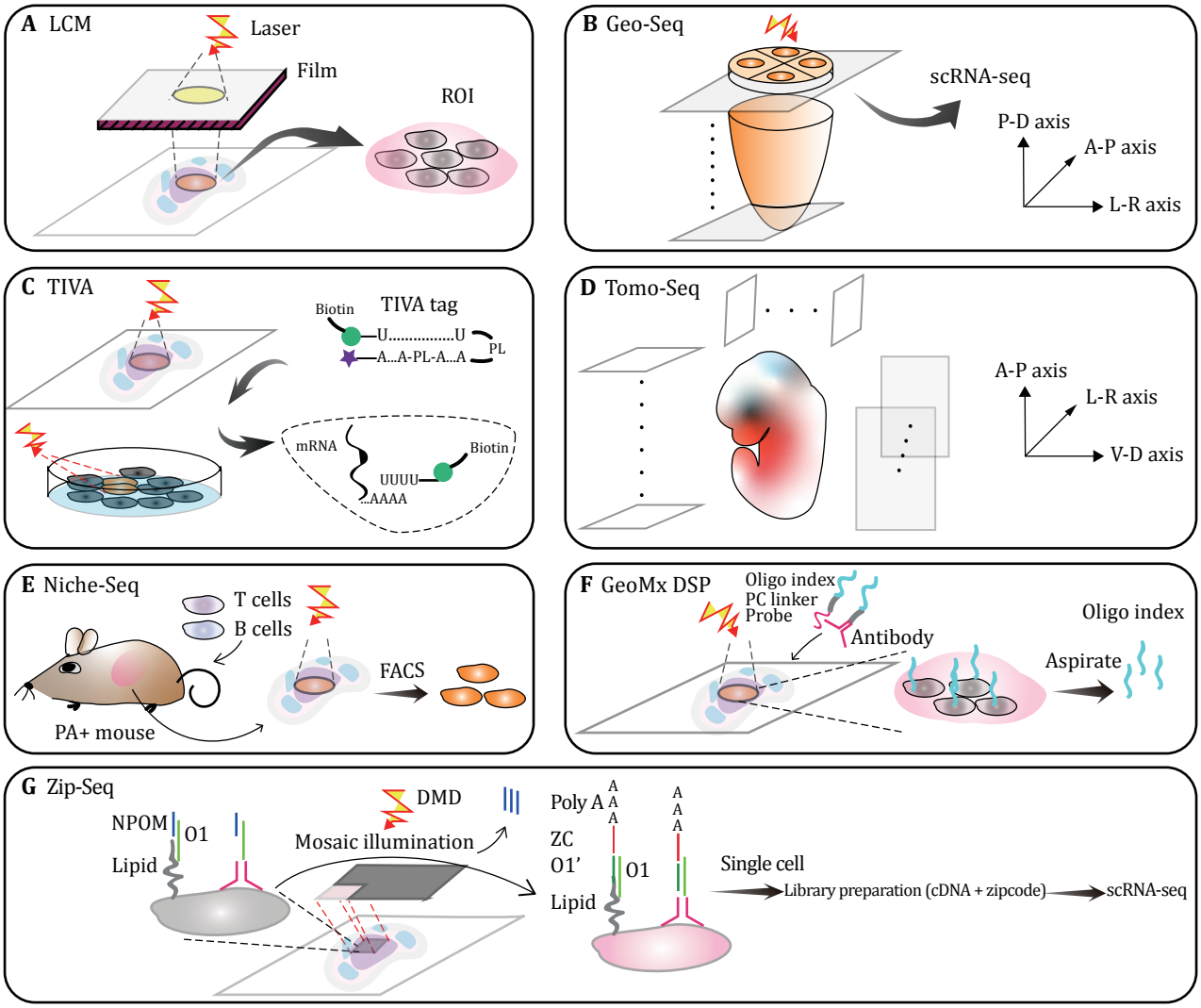

Fig. 2 Technologies based on isolation through microdissection or photoreaction. **A** LCM: thermoplastic films covered on frozen tissue sections are activated at interesting regions by laser beam irradiation. The tissue region of interest is isolated along with the film. **B** Geo-seq: embryonic slices along the proximal and distal axis are dissected into ROIs at four quadrants of the anterior-posterior and left-right axis. Each region of interest (ROI) is then analyzed by scRNA-seq. **C** TIVA: TIVA tags with mRNA capture sequence which is sealed by complementary sequence enter cells and are then activated at defined loci of tissue section or culture cells via illumination incises photobleached linkers (PL). Activated tags capture mRNA by released biotin labeled poly(U) sequence. Captured mRNA can be selectively sequenced after biotin affinity purification. **D** Tomo-Seq: identical embryos are sliced into 50–100 sections on the main body axis, respectively. Individual slices are sequenced to reconstruct 3D spatial transcriptome. **E** Niche-Seq: engrafted cells such as T and B cells are injected into photoactivatable reporter transgenic mice to identify histological structures. Immunological cells on the tissue section are photoactivated at specific histological regions. Activated cells are then isolated by FASC. **F** GeoMx DSP: barcoded antibodies and probes containing photocleavable (PC) linker are incubated with targeted mRNA and proteins of cells on the tissue slices. Spatially resolved indexed barcodes are aspirated into microcapillary after PC linkers break via ultraviolet lighting at the target locus of tissues. Oligo barcodes are counted to reflect the abundance of target mRNA and proteins. **G** ZipSeq: cells on the tissue section are labeled with 6-nitropiperonyloxylmethyl (NPOM) caged O1 segment which is introduced through the combination of general cellular membrane antigen with barcoded antibody or lipid insertion into the membrane. Illumination cleaves NPOM, exposing O1 to catch complementary O1' of zipcoded oligo at selected tissue regions. Target cells are isolated from tissue dissociated single cell suspension and then libraries of zipcode and cDNA from aimed cells are undergone scRNA-seq

tumors and brain, and to scrutinize the transcriptional activity in specific cells in the specific locations of the tissue. LCM also suits for formalin- or alcohol-fixed paraffin-embedded tissues and cellular preparations.

© The Author(s) 2022

But there are several disadvantages of LCM, *e.g*., specific cell types are often hardly discerned merely from morphological staining tissue and potential harm to DNA or RNA molecules may happen during laser microdissection.

Immunofluorescent LCM (IF-LCM) can select cells of specific types with the aid of antibody recognition (Murakami *et al.* 2000). To minimize the damage to mRNA when labeling cells with fluorescence in an aqueous solution, IF-LCM labels cells at a rapid speed of just one minute. Labeled cells are microscopely selected and collected by LCM. IF-LCM elucidated the genotype-based gene transcription in a few specific locations of tissues.

Examples of research using LCM grow rapidly. Bandyopadhyay *et al*. constructed the expression profiling of motor neurons in mouse spinal cord by LCM and microarray. This study prompts our understanding of transcriptional changes in spinal cord motor neurons during neurogenic and neurodegenerative diseases (Bandyopadhyay *et al.* 2014). More examples of down-stream utilization of LCM are overviewed elsewhere (Datta *et al.* 2015).

Geographical position sequencing (Geo-seq)

To bridge LCM with RNA sequencing, Geo-seq and LCM-seq adapted and optimized scRNA-seq on LCM-isolated defined low-input cells (Nichterwitz *et al*. 2016; Chen *et al*. 2017). As a showcase of Geo-seq, Peng *et al*. systematically examined the epiblast cell populations in mid-gastrulation mouse embryos (Fig. 2B). A 3D digital transcriptome data was built up to investigate the continuous transition of gene expression and pathways that models the cell fate segregation and regionalization of germ layer-related stem cells (Peng *et al.* 2016). The merit of Geo-seq is its capability to build a high confidence 3D spatial atlas with serial sections at a resolution of as few as ten cells (Chen *et al.* 2017).

In 2019, Geo-seq was applied to the exploration of the temporal and spatial patterns of code and non-code transcripts in post-implantation mouse embryos (E5.5–E7.5). It successfully revealed the lineage trajectory between germ layers at consecutive time points. This rich data resource of spatiotemporal transcriptomic pattern accelerated our recognition of the regulation of lineage propensity of pluripotent stem cells and identification of core regulatory mechanisms during early embryo development (Peng *et al*. 2019).

However, all these methods based on LCM consume intensive labor and time and the throughput becomes a limiting factor when a large number of samples are surveyed.

Transcriptome *in vivo* analysis (TIVA)

LCM crudely dissects fixed tissue sections mounted on glass slides so that the natural interaction of cells may be lost. TIVA initiates the non-destructive spatial profiling of gene expression of selected local cells (Lovatt *et al.* 2014). Live cells in culture and tissue section can intake TIVA tag via cell-penetrating peptide (CPP) which will dissociate from TIVA in cytoplasm. Capture oligos labeled with Cy5 and biotin are released from TIVA tag by irradiation. The pair fluorophores can also excite the fluorescence resonance energy transfer (FRET) for confirming the uncaging events. Activated capture oligo catches mRNA in the target cells with photocleavable linkers that are labeled with Cy3. Captured mRNA via biotin affinity purification of lysates are sequenced by RNA-seq (Fig. 2C). As a proof-of-concept, spatially diverse distributions of single neuronal cells were investigated by TIVA.

Although being advantageous as an *in vivo* exploration method, the complexity of TIVA tag and efficiency of conjugation and hybridization still awaits further optimization.

Tomography sequencing (Tomo-seq)

In addition to laser microdissection, cryosection along interested directions can preserve spatial ordering along latitude or longitude axes. Inspired by this, Tomo-seq in zebrafish embryos acquired 3D developmental regulators and genes with definite spatial patterns in different phases of embryonic morphogenesis (Junker *et al.* 2014). Each of the three embryos was sliced into 50–100 sections according to coordinates of anterior-posterior, left-right and ventral-dorsal body axis respectively, which were eventually reconstructed into an entire 3D embryo image by mathematical algorithms (Fig. 2D). Extracted RNA from different embryonic slices in one axis was reversely transcribed into cDNA using uniquely barcoded poly(T) primer in individual tubes so that all cDNA can be linearly amplified in mixed pools to reduce PCR bias between samples. Finally, a spatial coordinate was approximated through the iterative proportional fitting algorithm. Although the spatial analysis has been improved into 3-dimension with Tomo-seq, the accuracy of genes deduced from inadequate computational fitting is not satisfying. Importantly, the reconstruction requires several morphologically identical samples. All these

 © The Author(s) 2022

limitations restrict its application in some rare biological tissues.

Niche-seq

In order to figure out the rare cell types and their dynamic changes in cellular and molecular levels within the immunological niche, Niche-seq was established in 2017 to apply in live animal or *ex vivo* tissues via the combination of photoactivation (PA) and scRNA-seq technologies (Medaglia *et al.* 2017). A photoactivatable reporter transgenic mouse was intravenously injected with labeled T and B cells to mark the main histological texture of the lymph node. Tissues in different states including normal and infection were photoactivated to express GFP fluorescence within the target regions. FACS were used to enrich specific fluorescent cells followed by single-cell sequencing (Fig. 2E). Various cell types containing scarce populations like NK cells in individual niches were analyzed to illuminate the heterogeneity between immune cell subsets and the dynamic view of the composition of cell types as well as molecular pathways that were induced by treatment. However, the spatial resolution is coarse and the method is limited by available photoactivated colors and transgenic manipulation.

GeoMx digital spatial profiling (GeoMx Dsp)

Simultaneous profiling of RNA and proteins offers more comprehensive insights than single omics. GeoMx Dsp captures the information of gene expression and protein concentration in particular regions or cell types by utilizing photocleaved oligo barcodes (Merritt *et al.* 2019). RNA probes or protein antibodies coupled with a photocleavable linker that connects to an indexing oligo are incubated with targeted mRNA or proteins on the fixed FFPE tissue slice. At the same time, up to four fluorescence labeled biomarkers can identify the histological structure or cell populations. Regions of interest (ROIs) are selectively illuminated by a digital-micromirror device (DMD) or dual-DMD directed UV light. Photocleaved oligos in targeted areas undergo releasing, gathering by capillary aspiration and dispensing into microplates for final quantitative detection through single-molecule counting nCounter System or RNA-seq (Fig. 2F). Hundreds of RNA and proteins can be simultaneously profiled in enormous ROIs via this powerful method. However, its resolution can hardly reach single cells, and it is not economical when expanding to vast numbers of samples and targets.

ZipSeq

Instead of obtaining position by isolating specific tissue regions, ZipSeq directly allocates spatial tags (referred to as “zipcode”) on the cell surface at interesting locations through the binding of the high-affinity antibody or the stable lipid insertion into the membrane. The anchor strand contains an overhang sequence (O1) that is caged by thymidine-conjugated 6-nitropiperonyloxylmethyl (NPOM) (Hu *et al*. 2020). Controlled mosaic digital micromirror device (DMD) illumination releases NPOM at selected regions so that the O1 anchor strands will be exposed. Activated local cells are incorporated with special zipcodes by the complementary O1’ strand which hybridizes to the uncaged overhanging sequence of anchor DNA double strands. Besides O1’ sequence, the zipcode strand has an 8 bp barcode for identifying spatial information and a 32 bp poly(A) fragment for amplification in the later step of cDNA library construction. Meanwhile, anchor strands can be labeled with fluorescence in order to filter cells conjugated with the specific antibody. After several rounds of adding distinct zipcodes at specific locations on the tissue section, collected cells are subjected to scRNA-seq analysis to retrieve the zipcode and transcriptome from selected or all cells in target regions (Fig. 2G). The ability of spatial recognization lacks precision due to the area size of defined mosaic pattern.

ARRAY BASED TECHNOLOGIES

Technologies based on FISH and microdissection usually identify insufficient genes or samples at each run. Even though FISH based spatial transcriptome technologies can directly visualize and even sequence RNA *in situ*, the fluorescent channels can’t satisfy rapid and high-throughput profiling. The down-stream image analysis is complicated, which sometimes requires costly equipment and professional experts. Microdissection related methods have the potential of isolating interested regions and even single cells from tissue slice, however, it is expensive and laborious despite some semiautomated or automated devices that can be integrated. Recently, high-throughput methods based on the spatially barcoded array have been developed to obtain the whole transcriptome of the tissue section and the positional information simultaneously. With the commercialized platform like 10× Genomics Visium, an increasing number of researchers have utilized these convenient platforms to answer many biological and pathological questions. For

© The Author(s) 2022

example, Hou *et al*. probed the transcriptional profiles of the human liver organization via integration of scRNA-seq and Visium (Hou *et al*. 2021). The molecular atlas of specific brain layers was constructed by 10× Genomics Visium to obtain an exquisite reference for position-dependent single cells (Maynard *et al*. 2021). Recently, spatial variances of tumor microenvironment were interrogated during tumor formation, development, metastasis, and a poor curative efficacy was found to correlate with special tumor subclones by 10× Genomics Visium (Nagasawa *et al*. 2021; Zhang *et al*. 2021).

Spatial transcriptomics (ST)

As a pioneering technology to combine *in situ* RNA amplification and spatially barcoded array, ST prints 1007 pre-defined spots with a size of 100 μm diameter and an inter-spots distance of 200 μm on the 6.2 mm × 6.6 mm array. The spatially barcoded oligo probes consist of cleavage part, sequencing fragment handle, positional barcode, randomized unique molecular identifier (UMI) and mRNA capture poly(T) sequence. Frozen tissue slices are posted on the array and then fixed, HE stained and imaged to obtain the histological images that are subsequently utilized to align the spatially barcoded mRNA in each spot to their corresponding tissue regions. mRNA over the whole tissue section is captured by the poly(T) sequence and polymerizes into RNA-DNA hybrid after mounted tissues are enzymatically permeabilized. The second cDNA strands are synthesized by reverse transcription, followed by library construction and sequencing. Through spatial barcoding and in-array amplification, histological position and gene expression profiling are parallelly registered (Ståhl *et al*. 2016) (Fig. 3A). ST has the strength to unbiasedly delineate gene expression from diverse sample locations for processing simultaneously on the same printed array slides. The manual and automated protocols of ST are both available and the latter is undoubtedly hands-on time saving (Jemt *et al.* 2016; Ståhl *et al*. 2016).

In 2019, 10× Genomics company commercialized ST and further developed it into 10× Genomics Visium technology with increased sensitivity (55 μm diameter spot) and throughput (5000 spots per capture area) (Fig. 3A). Depending on tissue types and sampling thickness, up to 5000 genes from 1–10 cells on average per spatial spot can be obtained. The array-based ST, therefore has broad applications in characterizing different species, organs, developmental stages, biological processes and diseases (Maynard *et al.* 2021; Nagasawa *et al.* 2021). Nevertheless, the cellular resolution has not been readily acquired before more precise printing technology is available.

Slide-seq

Slide-seq improves the spatial resolution by adapting densely packed bead arrays to capture more subtle tissue features. A monolayer of 10 μm microbeads with specific spatial barcodes are loaded on the rubber-coated glass coverslip. The "puck" surface captures released mRNA from the frozen tissue sections that are mounted onto the bead array. The positional index of each bead can be determined by SOLiD sequencing. The captured mRNA then undergoes library preparation for 3'-end RNA-seq (Fig. 3B). Computational algorithms are also developed for data mining such as non-negative matrix factorization regression (NMFreg) for assigning scRNA-seq identified cell types to beads and finding out spatially non-random genes (Chen *et al.* 2020; Rodriques *et al*. 2019). Most beads are annotated as single cells even though a part of beads have two or more cells. Slide-seq and later updated Slide-seq2 (Chen *et al*. 2020) made enormous progress on the cellular resolution of the spatial transcriptome.

High-definition spatial transcriptomics (HDST)

HDST makes a step forward to the higher resolution of spatial gene profiling. In HDST, spatially indexed oligos are randomly distributed into the 2 μm beads of hexagonal well on the array. The spatial barcodes are generated by a split-and-pool approach. Similarly, the positional index of barcodes is pre-decoded through sequential fluorophore hybridization and mRNA of the tissue section is sequenced by RNA-seq (Vickovic *et al.* 2019) (Fig. 3C).

With higher spatial resolution obtained, the mRNA capture and detection efficiency on single beads with Slide-seq or HDST remain a bottleneck.

Deterministic barcoding in tissue for spatial omics sequencing (DBiT-seq)

A simple way to allocate address index to tissue is through deploying individual coordinate parameters on the 2D axis of the tissue section. DBiT-seq labels fixed tissues on the slide with two sets of 50 distinct barcodes which flow through tissue via a microfluidic chip for position recording. Each barcode flowing channel can be set as 10, 25 or 50 μm width (Liu *et al.* 2020) (Fig. 3D), which means that near single-cell resolution can be obtained at the most barcoded areas. However, as the resolution increases, smaller regions of

 © The Author(s) 2022

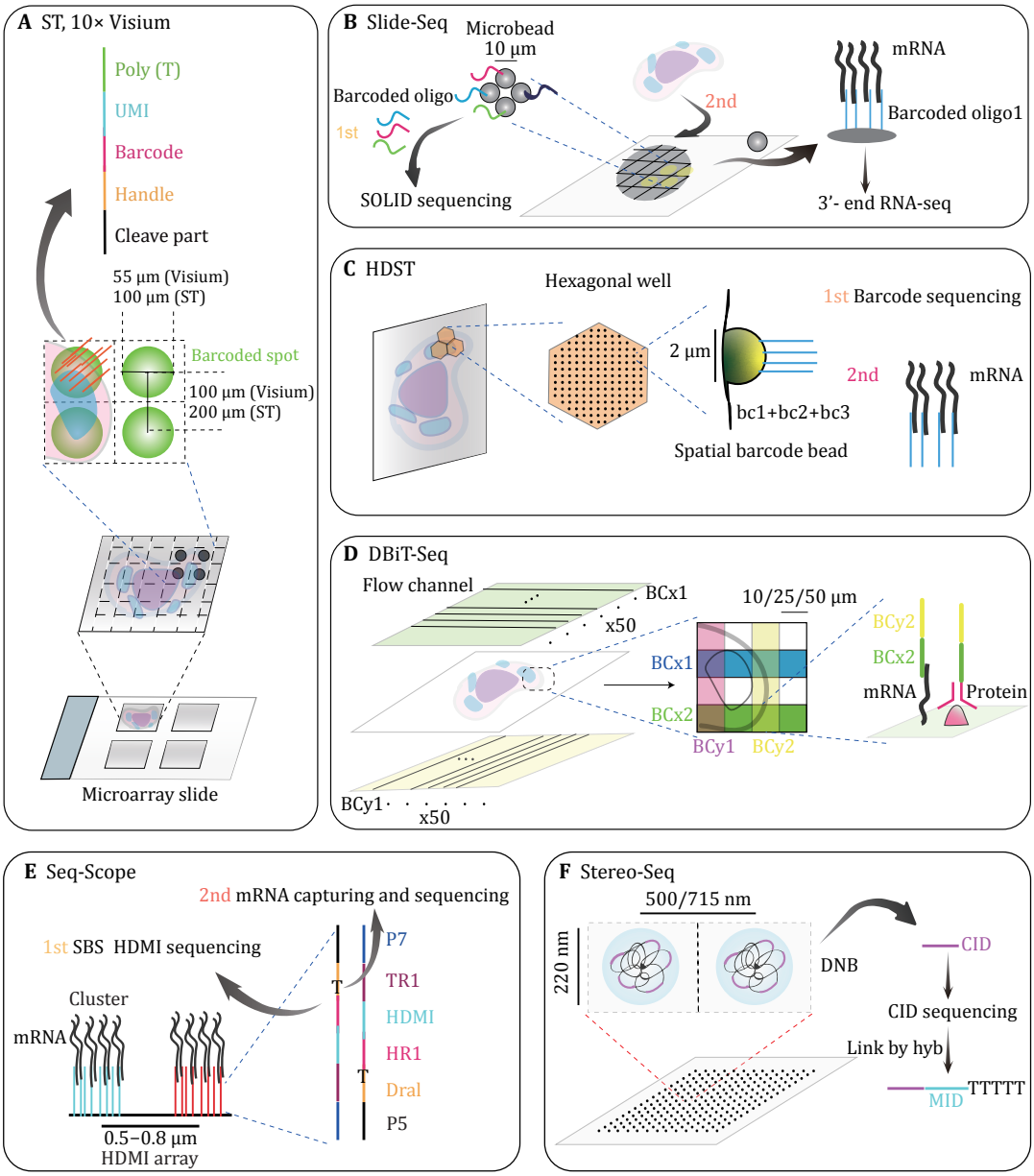

Fig. 3 Technologies based on the array. **A** ST and 10× Visium: tissue slice samples are simultaneously sequenced on the same array chip. Thousands of spatially barcoded spots are evenly distributed on the slide. The diameter of the spot is 100 μm (ST) or 55 μm (Visium) and the distance of spot center between two adjacent spots is 200 μm (ST) or 100 μm (Visium) respectively. Each spot has unique spatially barcoded oligo clusters whose poly (T) captures mRNA for later RNA-seq. **B** Slide-Seq: indexed microbeads are randomly organized into monolayer mRNA capture surfaces. The distance between two microbeads is 10–20 μm. Firstly, the spatial index oligo is identified by SOLiD sequencing. Then mRNA can be captured and analyzed spatially through 3'- end sequencing. **C** HDST: 2 μm beads are randomly placed into hexagonal wells on glass slides. Each bead is marked with combinatorial barcodes through split-and-pool with three sets of barcodes. The spatial barcodes are also decoded via seqFISH. mRNA is consequently spatially demultiplexed. **D** DBiT-Seq: two sets of 50 different barcodes flow through microfluidic chips on two perpendicular axes. The tissue section is successively posted on the two chips and labeled with barcode *x* and barcode *y*. The barcode can be coupled with mRNA and proteins when it is conjugated to dT and antibodies. **E** Seq-Scope: molecular clusters are synthesized by the Illumina platform. The distance between the two clusters is 0.5–0.8 μm. Each cluster is composed of plentiful indexed oligos. Single indexed oligo contains P5, Dra1, Oligo-dT, TR1 (TruSeq Read 1), HDMI (high-definition map coordinate identifier), HR1 (HDMI Read 1) and P7. The positions of clusters are confirmed by SBS decoding HDMI and followed by mRNA sequencing. **F** Stereo-Seq: DNA nanoballs (DNB) are stably docked into grid pattern array slide. The diameter of the nanoball is 220 nm and the center-to-center distance is 500 or 715 nm. Each nanoball has multiple identical barcodes which are resolved by sequencing to generate coordinate identifiers (CID). Next, molecular identifiers (MID) and oligo dT are incorporated into CID through hybridization

© The Author(s) 2022

tissue can be recovered with a fixed number of channels. Moreover, the possible risk of lateral spreading of barcodes will cause positioning confusion.

Seq-scope

Based on the Illumina sequencing platform, Seq-Scope firstly generates HDMI (high-definition map coordinate identifier) clusters of 0.5–0.8 μm spaced apart. Each cluster is composed of plentiful spatially indexed probes. Single indexed oligo contains P5, Dra1, Oligo-dT, TR1 (TruSeq Read 1), HDMI, HR1 (HDMI Read 1) and P7 sequences. The spatial index will be decoded by sequencing by synthesis (SBS). Afterwards, tissue spatial expression profiles at the cellular and subcellular levels can be revealed by the second RNA sequencing step (Cho *et al.* 2021) (Fig. 3E). The efficiency of mRNA capturing has been further improved in comparison with HDST. Therefore, Seq-Scope opens the door of subcellular expression visualization via spatially high throughput sequencing, which holds great promises for implementing high-definition spatial transcriptomic studies in biological and clinical contexts.

Seq-Scope also reputably solved problems that existed in previous array-based methods such as mixed cell types in each dot unit and unclear histological boundary when plotting clustering results onto the spatial position. However, there are some challenges like scratch-related data loss. Besides, more advanced algorithms are needed to extract valuable information at the subcellular level because currently good performance was only obtained with 10 μm cluster grids for reliable cell type classification.

Spatio-temporal enhanced REsolution omics-sequencing (Stereo-seq)

The BGI company recently developed a method with nanometer resolution and a remarkable capture rate on large frozen tissue sections (Chen *et al.* 2021). Repeated random spatial barcodes DNA nanoball (DNB) of 220 nm of diameter and 500 or 715 nm of inter-distance produced by RCA are docked into the grid-pattern array slide. After imaging and sequencing, the spatial barcode of each DNB, *i.e.*, the coordinate identifier (CID) can be determined. Molecular identifiers (MID) and oligo-dT are hybridized and ligated to CID containing DNBs which follows mRNA capturing, library preparation and sequencing (Fig. 3F). With its superior capture rate and density, up to 200 mm^2 tissue sections have been spatially sequenced. As centimeter-scale tissue samples can be analyzed, Stereo-seq greatly facilities spatial transcriptome with a large field of view. Of note, the multiplexing is significantly increased by Stereo-seq so that one can deal with hundreds of samples on continuous slides simultaneously. Technical influences caused by batch operations could also be minimized. However, Stereo-seq can not capture enough mRNA features in each DNB. Besides, it is difficult to distinguish cellular boundary similar to other subcellular spatial approaches. DNB barcode generation and mRNA sequencing on the platform other than BGI sequencer have not been assessed.

Table 1 summarizes the technical metrics of mentioned spatial transcriptomic technologies.

SPATIAL MULTI-OMICS

From RNA-seq to scRNA-seq, the heterogeneity of single cells is resolved which gives us the opportunity to gain a profound insight into complicated tissues such as the organization of cell subtypes and dynamic states during biological processes. Comparably, spatial transcriptomics presents a map of molecular architecture in particular tissue regions. Some existing spatial methodologies have shown their strength to grab not only RNA but also proteins and epigenetic information. DBiT-seq profiles whole genome gene expression and proteins of interest at the same time on the frozen fixed tissue (Su *et al.* 2021). Recently, it further demonstrated that epigenome can be analyzed with DBiT-seq pipeline (Deng *et al*. 2021a, b). ExM-MERFISH also could simultaneously detect RNA abundance and the level of protein of interest through hybridization of oligo-conjugated antibodies (Wang *et al.* 2018a). GeoMx Dsp was initially designed to realize the multi-modal information of photocleaved oligo-barcoded RNA and protein content in particular regions or cell types (Merritt *et al.* 2019). The combination of array based spatial transcriptome technology and antibody based protein measurement was developed into spatial multi-omics (SM-Omics) to ensure automatic detection of the transcripts and proteins (Vickovic *et al.* 2020).

Likewise, chromosome organization and interaction with *cis*-elements as well as the dynamic location of active chromatin topology domains can be visualized by the advanced multiplexed FISH platforms such as MERFISH and OligoFISSEQ (Nguyen *et al.* 2020; Su *et al.* 2020) along with gene expression profiles. Now spatial multi-omics approaches have exceeded the incomplete description of cell states merely by transcriptomics and offered multi-modality information for us to explore the

 © The Author(s) 2022

Table 1 overview of technical metrics of mentioned spatial transcriptomic technologies

Method	Resolution	Multi-plexity	Gene coverage	Sample	Strength	Limitation
smFISH	Subcellular	Low	Low	Fresh frozen/ FFPE	High detection sensitivity	Low throughput
RNAscope	Subcellular	Low	Low	Fresh frozen/ FFPE	High resolution and sensitivity	Optical overlap
seqFISH	Subcellular	Low	~100	Fresh frozen	Multiplexity	Predefined probes; drop out error
seqFISH+	Subcellular	Low	10,000	Fresh frozen	Higher multiplexity	Predefined probes; drop out error
MERFISH	Subcellular	Low	Hundreds to thousands	Cultured cells/ fresh frozen	Super high resolution; error detection and correction	Predefined probes; detect efficiency and accuracy
osmFISH	Subcellular	Low	Low	Fresh frozen	Low optical overlap	Lower throughput and multiplexity
ISS	Subcellular	Low	Low	Fresh frozen/ FFPE	Mutation and isoform detection; reduced background	Optical overlap
FISSEQ	Subcellular	Low	Whole transcriptome	Fresh frozen/ FFPE	Genome-wide detection	Impaired sensitivity; time consuming; inefficient
STARmap	Subcellular	Low	Thousands	Thick fresh frozen/ FFPE	No RT step; high specificity; error correction	Complexity of image processing; small FOV
LCM-seq/ Geo-seq	Single to a few cells	Low	Whole transcriptome	Fresh frozen/ FFPE	High depth; tissue compatibility; full-length transcript	Time-consuming; low throughput
TIVA	Single cell	Low	Whole transcriptome	Live cells	*In vivo*; natural state; isolate rare cell types	Complexity of tag; efficiency of conjugation and hybridization
Tomo-seq	Tissue slice	Medium	Whole transcriptome	Fresh frozen	Not requiring specialized equipment	Require similar samples; lack positional details; reconstruction artefact
Niche-seq	Single cell	Low	Whole transcriptome	Live cells/*in vitro* tissue	Isolation of rare immune cells; *in vivo*; natural state	Need transgenic animal
GeoM Dsp	A few cells	Low	Hundreds	Fresh frozen/ FFPE	Ability to analyze protein; compatible with fluorescent staining	Throughput; costly
ZipSeq	Single cell	Low	Whole transcriptome	Fresh frozen/ FFPE	Ability to analyze protein; combinatorial zipcode	Complexity and hybridization efficiency of combinatorial zipcode; positional details lack
ST/Visium	50/100 μm spot	High	Whole transcriptome	Fresh frozen	High throughput	Not single-cell resolution
Slide-seq	10 μm bead/ subcellular	High	Whole transcriptome	Fresh frozen	Improved spatial resolution	Low sensitivity;
HDST	2 μm bead/ subcellular	High	Whole transcriptome	Fresh frozen	Subcellular resolution	Low sensitivity

© The Author(s) 2022

Continued

Method	Resolution	Multi-plexity	Gene coverage	Sample	Strength	Limitation
DBiT-seq	10/25/50 μm mosaic grid	High	Whole transcriptome	Fresh frozen/ FFPE	Flexibility	Risk of lateral expansion of barcode
Seq-Scope	0.5–0.8 μm distant cluster	High	Whole transcriptome	Fresh frozen	high resolution	Require spatial barcode sequencing
Stereo-seq	220 nm	High	Whole transcriptome	Fresh frozen	Ability for larger tissue section; super high resolution	Require spatial barcode sequencies; detection sensitivity

accurate identity and characteristics of cell types in various biological contexts. However, the integration of multi-modal measurements requires suitable computational methods. There still needs a lot of work to combine other layers of omics into transcriptomics in the same tissue section, *e.g*., metabolomics and epitranscriptome, to enable a holistic understanding of biological processes.

DATA ANALYSIS

Spatial transcriptomic data should be preprocessed and mined through distinct analytic pipelines dependent on the platform of the sequencing technique. Today there is a mass of available packages and online references for convenient and quick data analysis. A concise summary of spatial analytical methods has been reviewed elsewhere (Longo *et al*. 2021). In most cases, a general single-cell RNA-seq data analysis pipeline can be followed. However, spatial transcriptome data often have different properties, for example, multiple cells could be mixed in one spatial unit leading to potential confounding heterogeneity and segmentation issues, and joint merging of histology images information to enhance spatial gene expression analysis is necessary (Hu *et al*. 2021).

Compared to single cell sequencing, spatially resolved data requires additional spatial barcode demultiplexing to identify the position of transcripts when aligning reads to the genome reference (Navarro *et al*. 2017). Spatial variable genes that are active at specific domains of tissue sections can be computed by a number of open resource packages, such as SpatialDE, SPARK and SPATA (Kueckelhaus *et al.* 2020; Sun *et al.* 2020; Svensson *et al.* 2018). Cell communications within intact tissue texture are explored by ligand–receptor interaction which helps scientists have a better understanding of the interaction of the cells and their microenvironment. Reconstruction of spatial trajectory along undissociated tissue slices can be established at a shared space depending on the gradient changes of expression of distinct cell types (Dries *et al.* 2020; Pham *et al.* 2020; Ren *et al.* 2020; Tran *et al.* 2021), which helps the molecular navigation of space-dependent cell transitions.

To efficiently integrate single-cell and spatial transcriptome data, many existing tools such as bindSC (Dou *et al*. 2020), SPOTlight (Elosua-Bayes *et al*. 2021), RCTD (Cable *et al*. 2021) and Seurat (Stuart *et al*. 2019) are developed through either deconvolution or spatial mapping (Longo *et al*. 2021). Most importantly, analytic algorithms that fully harness histologic images and other tissue architecture modalities will deeply transform the current computational spatial data workflow. Emerging methods such as stLearn and SEDR (Pham *et al*. 2020; Chen *et al*. 2021) that incorporate morphological features and machine learning models have already empowered the exploration of the rich information in spatial transcriptome data.

SUMMARY AND PERSPECTIVES

Anatomists help us to perceive the composition of our organs and tissues. Histologists clarify the cellular organization and characteristics of tissues. Molecular technologists take a deeper probe into biological substances in cytoplasm and nucleus. While RNA-seq serves as the means to depict a rough picture of mRNA levels of culture cells or tissues on average, the non-negligible cell heterogeneity is then conquered by scRNA-seq. However, cell suspension from intact tissue made it impossible to infer the original anatomical regions and detailed locations. To circumvent this limitation, a variety of spatial transcriptomics technologies have been developed to explore the spatial dimension for comprehensive understanding of the

 © The Author(s) 2022

relationship between cellular function and location, the interaction of cell populations and dynamic trajectory of specific subpopulations reacting to stimulus or treatment. These technologies have enabled spatially resolved molecular profiling on almost all types of tissue in model organisms. Studies on pathological samples including diseases and tumors with these spatial omics technologies will undoubtedly uncover the alteration of molecules at important histological sites, leading to the development of novel discoveries for clinical diagnosis, drug design, disease prevention and prognosis evaluation.

Although new methods become available from time to time, no singular spatial transcriptomics technique is currently best for all situations. The limits to resolution and sensitivity, as well as throughput, accessibility, and affordability shall be considered in designing the experimental pipeline. We envision that the integration of multimodal information from single-cell RNA-seq data, spatial multi-omics, and histological features will fully unleash the power of spatial transcriptome to disentangle the complex interplay in both physiological and pathological conditions.

Acknowledgements We apologize to all authors whose work has not been cited in this review due to space constrains. We thank the critical reading from all the Peng lab members. This work was supported in part by National Key R&D Program of China (2018YFA0801402, 2018YFA0107200), the "Strategic Priority Research Program" of the Chinese Academy of Sciences (XDA16020404), National Natural Science Foundation of China (31871456), Guangdong Basic and Applied Basic Research Foundation (2019B151502054), Frontier Research Program of Bioland Laboratory (Guangzhou Regenerative Medicine and Health Guangdong Laboratory, 2018GZR110105013), Jiazi Research Innovative Project of Bioland Laboratory (2019GZR110108001), Science and Technology Planning Project of Guangdong Province (2020B1212060052).

Compliance with Ethical Standards

Conflict of interest Guangdun Peng and Zhuxia Li declare that they have no conflict of interest.

Human and animal rights and informed consent This article does not contain any studies with human or animal subjects performed by any of the authors.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

Alturkistani HA, Tashkandi FM, Mohammedsaleh ZM (2015) Histological stains: a literature review and case study. Glob J Health Sci 8(3): 72–79

Bandyopadhyay U, Fenton WA, Horwich AL, Nagy M (2014) Production of RNA for transcriptomic analysis from mouse spinal cord motor neuron cell bodies by laser capture microdissection. J Vis Exp 83: e51168. https://doi.org/10.3791/51168

Battich N, Stoeger T, Pelkmans L (2013) Image-based transcriptomics in thousands of single human cells at single-molecule resolution. Nat Methods 10(11): 1127–1133

Beliveau BJ, Joyce EF, Apostolopoulos N, Yilmaz F, Fonseka CY, McCole RB, Chang Y, Li JB, Senaratne TN, Williams BR, Rouillard JM, Wu CT (2012) Versatile design and synthesis platform for visualizing genomes with Oligopaint FISH probes. Proc Natl Acad Sci USA 109(52): 21301–21306

Brown VM, Ossadtchi A, Khan AH, Cherry SR, Leahy RM, Smith DJ (2002) High-throughput imaging of brain gene expression. Genome Res 12: 244-254

Cabili MN, Dunagin MC, McClanahan PD, Biaesch A, Padovan-Merhar O, Regev A, Rinn JL, Raj A (2015) Localization and abundance analysis of human lncRNAs at single-cell and single-molecule resolution. Genome Biol 16(1): 20. https://doi.org/10.1186/s13059-015-0586-4

Cable DM, Murray E, Zou LS, Goeva A, Macosko EZ, Chen F, Irizarry RA (2021) Robust decomposition of cell type mixtures in spatial transcriptomics. Nat Biotechnol. https://doi.org/10.1038/s41587-021-00830-w

Chen A, Liao S, Cheng M, Ma K, Wu L, Lai Y, Yang J, Li W, Xu J, Hao S, Lu H, Chen X, Liu X, Huang X, Lin F, Tang X, Li Z, Hong Y, Fu D, Jiang Y, Peng J, Liu S, Shen M, Liu C, Li Q, Wang Z, Wang Z, Huang X, Yuan Y, Volpe G, Ward C, Muñoz-Cánoves P, Thiery JP, Zhao F, Li M, Kuang H, Wang O, Lu H, Wang B, Ni M, Zhang W, Mu F, Yin Y, Yang H, Lisby M, Cornall RJ, Uhlen M, Esteban MA, Li Y, Liu L, Xu X, Wang J (2021) Large field of view-spatially resolved transcriptomics at nanoscale resolution. bioRxiv. https://doi.org/10.1101/2021.01.17.427004

Chen H, Murray E, Laumas A, Li J, Nie X, Hotaling J, Guo J, Cairns BR, Macosko EZ, Cheng CY, Chen F (2020) Dissecting mammalian spermatogenesis using spatial transcriptomics. bioRxiv. https://doi.org/10.1101/2020.10.17.343335

Chen J, Suo S, Tam PP, Han JJ, Peng G, Jing N (2017) Spatial transcriptomic analysis of cryosectioned tissue samples with Geo-seq. Nat Protoc 12(3): 566–580

Chen KH, Boettiger AN, Moffitt JR, Wang S, Zhuang X (2015) RNA imaging. Spatially resolved, highly multiplexed RNA profiling in single cells. Science 348(6233): aaa6090. https://doi.org/10.1126/science.aaa6090

Chen X, Sun YC, Zhan H, Kebschull JM, Fischer S, Matho K, Huang ZJ, Gillis J, Zador AM (2019) High-Throughput mapping of long-range neuronal projection using *in situ* sequencing. Cell 179(3): 772–786

Cho CS, Xi J, Si Y, Park SR, Hsu JE, Kim M, Jun G, Kang HM, Lee JH (2021) Microscopic examination of spatial transcriptome using Seq-Scope. Cell 184(13): 3559–3572

Codeluppi S, Borm LE, Zeisel A, La Manno G, van Lunteren JA, Svensson CI, Linnarsson S (2018) Spatial organization of the somatosensory cortex revealed by osmFISH. Nat Methods

15(11): 932–935
Datta S, Malhotra L, Dickerson R, Chaffee S, Sen CK, Roy S (2015) Laser capture microdissection: big data from small samples. Histol Histopathol 30(11): 1255–1269
Deng Y, Bartosovic M, Ma S, Zhang D, Liu Y, Qin X, Su G, Xu ML, Halene S, Craft JE, Castelo-Branco G, Fan R (2021a) Spatial-ATAC-seq: spatially resolved chromatin accessibility profiling of tissues at genome scale and cellular level. bioRxiv. https://doi.org/10.1101/2021.06.06.447244
Deng Y, Zhang D, Liu Y, Su G, Enninful A, Bai Z, Fan R (2021b) Spatial epigenome sequencing at tissue scale and cellular level. bioRxiv. https://doi.org/10.1101/2021.03.11.434985
Dou J, Liang S, Mohanty V, Cheng X, Kim S, Choi J, Li Y, Rezvani K, Chen R, Chen K (2020) Unbiased integration of single cell multi-omics data. bioRxiv. https://doi.org/10.1101/2020.12.11.422014
Dries R, Zhu Q, Dong R, Linus Eng C-H, Li H, Liu K, Fu Y, Zhao T, Sarkar A, Bao F, George RE, Pierson N, Cai L, Yuan G-C (2021) Giotto, a toolbox for integrative analysis and visualization of spatial expression data. Genome Biol : 22(1): 78. https://doi.org/10.1186/s13059-021-02286-2
Elosua-Bayes M, Nieto P, Mereu E, Gut I, Heyn H (2021) SPOTlight: seeded NMF regression to deconvolute spatial transcriptomics spots with single-cell transcriptomes. Nucleic Acids Res 49: e50. https://doi.org/10.1093/nar/gkab043
Emmert-Buck MR, Bonner RF, Smith PD, Chuaqui RF, Zhuang Z, Goldstein SR, Weiss RA, Liotta LA (1996) Laser capture microdissection. science 274(5289): 998–1001
Femino AM, Fay SF, Fogarty K, Singer RH (1998) Visualization of single RNA transcripts *in situ*. science 280(5363): 585–590
Francoz E, Ranocha P, Pernot C, Le Ru A, Pacquit V, Dunand C, Burlat V (2016) Complementarity of medium-throughput *in situ* RNA hybridization and tissue-specific transcriptomics: case study of Arabidopsis seed development kinetics. Sci Rep 6: 24644. https://doi.org/10.1038/srep24644
Hou X, Yang Y, Li P, Zeng Z, Hu W, Zhe R, Liu X, Tang D, Ou M, Dai Y (2021) Integrating spatial transcriptomics and single-cell RNA-seq reveals the gene expression profling of the human embryonic liver. Front Cell Dev Biol 9: 652408. https://doi.org/10.3389/fcell.2021.652408
Hu J, Schroeder A, Coleman K, Chen C, Auerbach BJ, Li M (2021) Statistical and machine learning methods for spatially resolved transcriptomics with histology. Comput Struct Biotechnol J 19: 3829–3841
Hu KH, Eichorst JP, McGinnis CS, Patterson DM, Chow ED, Kersten K, Jameson SC, Gartner ZJ, Rao AA, Krummel MF (2020) ZipSeq: barcoding for real-time mapping of single cell transcriptomes. Nat Methods 17: 833–843
Jemt A, Salmen F, Lundmark A, Mollbrink A, Fernandez Navarro J, Stahl PL, Yucel-Lindberg T, Lundeberg J (2016) An automated approach to prepare tissue-derived spatially barcoded RNA-sequencing libraries. Sci Rep 6: 37137. https://doi.org/10.1038/srep37137
Junker JP, Noel ES, Guryev V, Peterson KA, Shah G, Huisken J, McMahon AP, Berezikov E, Bakkers J, van Oudenaarden A (2014) Genome-wide RNA tomography in the zebrafish embryo. Cell 159(3): 662–675
Ke R, Mignardi M, Pacureanu A, Svedlund J, Botling J, Wahlby C, Nilsson M (2013) *In situ* sequencing for RNA analysis in preserved tissue and cells. Nat Methods 10(9): 857–860
Kueckelhaus J, von Ehr J, Ravi VM, Will P, Joseph K, Beck J, Hofmann UG, Delev D, Schnell O, Heiland DH (2020) Inferring spatially transient gene expression pattern from spatial transcriptomic studies. bioRxiv. https://doi.org/10.1101/2020.10.20.346544
Lee JH, Daugharthy ER, Scheiman J, Kalhor R, Yang JL, Ferrante TC, Terry R, Jeanty SS, Li C, Amamoto R, Peters DT, Turczyk BM, Marblestone AH, Inverso SA, Bernard A, Mali P, Rios X, Aach J, Church GM (2014) Highly multiplexed subcellular RNA sequencing *in situ*. Science 343(6177): 1360–1363
Liu Y, Yang M, Deng Y, Su G, Enninful A, Guo CC, Tebaldi T, Zhang D, Kim D, Bai Z, Norris E, Pan A, Li J, Xiao Y, Halene S, Fan R (2020) High-spatial-resolution multi-omics sequencing via deterministic barcoding in tissue. Cell 183(6): 1665–1681
Liu S, Punthambaker S, Iyer EPR, Ferrante T, Goodwin D, Fürth D, Pawlowski AC, Jindal K, Tam JM, Mifflin L, Alon S, Sinha A, Wassie AT, Chen F, Cheng A, Willocq V, Meyer K, Ling K-H, Camplisson CK, Kohman RE, Aach J, Lee JH, Yankner BA, Boyden ES, Church GM (2021) Barcoded oligonucleotides ligated on RNA amplified for multiplexed and parallel *in situ* analyses. Nucleic Acids Res 49: e58. https://doi.org/10.1093/nar/gkab120
Lohoff T, Ghazanfar S, Missarova A, Koulena N, Pierson N, Griffiths JA, Bardot ES, Eng CL, Tyser R, Argelaguet R, Guibentif C, Srinivas S, Briscoe J, Simons BD, Hadjantonakis AK, Göttgens B, Reik W, Nichols J, Cai L, Marioni JC (2021) Integration of spatial and single-cell transcriptomic data elucidates mouse organogenesis. *Nat Biotechnol*. https://doi.org/10.1038/s41587-021-01006-2
Longo SK, Guo MG, Ji AL, Khavari PA (2021) Integrating single-cell and spatial transcriptomics to elucidate intercellular tissue dynamics. Nat Rev Genet 22(10): 627–644
Lovatt D, Ruble BK, Lee J, Dueck H, Kim TK, Fisher S, Francis C, Spaethling JM, Wolf JA, Grady MS, Ulyanova AV, Yeldell SB, Griepenburg JC, Buckley PT, Kim J, Sul JY, Dmochowski IJ, Eberwine J (2014) Transcriptome *in vivo* analysis (TIVA) of spatially defined single cells in live tissue. Nat Methods 11(2): 190–196
Lubeck E, Coskun AF, Zhiyentayev T, Ahmad M, Cai L (2014) Single-cell *in situ* RNA profiling by sequential hybridization. Nat Methods 11(4): 360–361
Lynch AS, Briggs D, Hope IA (1995) Developmental expression pattern screen for genes predicted in the *C. elegans* genome sequencing project. . Nat Genet 11(3): 309–313
Maynard KR, Collado-Torres L, Weber LM, Uytingco C, Barry BK, Williams SR, Catallini JL, 2nd, Tran MN, Besich Z, Tippani M, Chew J, Yin Y, Kleinman JE, Hyde TM, Rao N, Hicks SC, Martinowich K, Jaffe AE (2021) Transcriptome-scale spatial gene expression in the human dorsolateral prefrontal cortex. Nat Neurosci 24(3): 425–436
Medaglia C, Giladi A, Stoler-Barak L, De Giovanni M, Salame TM, Biram A, David E, Li H, Iannacone M, Shulman Z, Amit I (2017) Spatial reconstruction of immune niches by combining photoactivatable reporters and scRNA-seq. Science 358(6370): 1622–1626
Merritt CR, Ong GT, Church S, Barker K, Geiss G, Hoang M, Jung J, Liang Y, McKay-Fleisch J, Nguyen K, Sorg K, Sprague I, Warren C, Warren S, Zhou Z, Zollinger DR, Dunaway DL, Mills GB, Beechem JM (2019) High multiplex, digital spatial profiling of proteins and RNA in fixed tissue using genomic detection methods. bioRxiv. https://doi.org/10.1101/559021
Moffitt JR, Hao J, Wang G, Chen KH, Babcock HP, Zhuang X (2016) High-throughput single-cell gene-expression profiling with multiplexed error-robust fluorescence *in situ* hybridization. Proc Natl Acad Sci USA 113(39): 11046–11051
Murakami H, Liotta L, Star RA (2000) IF-LCM: laser capture microdissection of immunofluorescently defined cells for mRNA analysis rapid communication. Kidney Int 58(3): 1346–1353
Nagasawa S, Kuze Y, Maeda I, Kojima Y, Motoyoshi A, Onishi T, Iwatani T, Yokoe T, Koike J, Chosokabe M, Kubota M, Seino H, Suzuki A, Seki M, Tsuchihara K, Inoue E, Tsugawa K, Ohta T, Suzuki Y (2021) Genomic profiling reveals heterogeneous

 © The Author(s) 2022

populations of ductal carcinoma *in situ* of the breast. Commun Biol 4(1): 438. https://doi.org/10.1038/s42003-021-01959-9

Navarro JF, Sjostrand J, Salmen F, Lundeberg J, Stahl PL (2017) ST Pipeline: an automated pipeline for spatial mapping of unique transcripts. Bioinformatics 33: 2591–2593

Nguyen HQ, Chattoraj S, Castillo D, Nguyen SC, Nir G, Lioutas A, Hershberg EA, Martins NMC, Reginato PL, Hannan M, Beliveau BJ, Church GM, Daugharthy ER, Marti-Renom MA, Wu CT (2020) 3D mapping and accelerated super-resolution imaging of the human genome using *in situ* sequencing. Nat Methods 17(8): 822–832

Padovan-Merhar O, Nair GP, Biaesch AG, Mayer A, Scarfone S, Foley SW, Wu AR, Churchman LS, Singh A, Raj A (2015) Single mammalian cells compensate for differences in cellular volume and DNA copy number through independent global transcriptional mechanisms. Mol Cell 58(2): 339–352

Peng G, Suo S, Chen J, Chen W, Liu C, Yu F, Wang R, Chen S, Sun N, Cui G, Song L, Tam PP, Han JD, Jing N (2016) Spatial transcriptome for the molecular annotation of lineage fates and cell identity in mid-gastrula mouse embryo. Dev Cell 36(6): 681–697

Peng G, Suo S, Cui G, Yu F, Wang R, Chen J, Chen S, Liu Z, Chen G, Qian Y, Tam P, Han JJ, Jing N (2019) Molecular architecture of lineage allocation and tissue organization in early mouse embryo. Nature 572(7770): 528–532

Pham D, Tan X, Xu J, Grice LF, Lam PY, Raghubar A, Vukovic J, Ruitenberg MJ, Nguyen Q (2020) stLearn: integrating spatial location, tissue morphology and gene expression to find cell types, cell-cell interactions and spatial trajectories within undissociated tissues. bioRxiv. https://doi.org/10.1101/2020.05.31.125658

Raj A, van den Bogaard P, Rifkin SA, van Oudenaarden A, Tyagi S (2008) Imaging individual mRNA molecules using multiple singly labeled probes. Nat Methods 5(10): 877–879

Ren X, Zhong G, Zhang Q, Zhang L, Sun Y, Zhang Z (2020) Reconstruction of cell spatial organization based on ligand-receptor mediated self-assembly. bioRxiv. https://doi.org/10.1101/2020.02.13.948521

Rodriques SG, Stickels RR, Goeva A, Martin CA, Murray E, Vanderburg CR, Welch J, Chen LM, Chen F, Macosko EZ (2019) Slide-seq: a scalable technology for measuring genome-wide expression at high spatial resolution. Science 363(6434): 1463–1467

Shah S, Takei Y, Zhou W, Lubeck E, Yun J, Eng CL, Koulena N, Cronin C, Karp C, Liaw EJ, Amin M, Cai L (2018) Dynamics and spatial genomics of the nascent transcriptome by intron seqFISH. Cell 174(2): 363–376

Shaffer SM, Dunagin MC, Torborg SR, Torre EA, Emert B, Krepler C, Beqiri M, Sproesser K, Brafford PA, Xiao M, Eggan E, Anastopoulos IN, Vargas-Garcia CA, Singh A, Nathanson KL, Herlyn M, Raj A (2017) Rare cell variability and drug-induced reprogramming as a mode of cancer drug resistance. Nature 546: 431–435

Skarnes WC, Moss JE, Hurtley SM, Beddington RS (1995) Capturing genes encoding membrane and secreted proteins important for mouse development. Proc Natl Acad Sci USA 92(14): 6592–6596

Ståhl PL, Salmén F, Vickovic S, Lundmark A, Navarro J F, Magnusson J, Giacomello S, Asp M, Westholm JO, Huss M, Mollbrink A, Linnarsson S, Codeluppi S, Borg Å, Pontén F, Costea P, Sahlén P, Mulder J, Bergmann O, Lundeberg J, Frisén J (2016) visualization and analysis of gene expression in tissue sections by spatial transcriptomics. Science 353(6294): 78–82

Stoltzfus CR, Filipek J, Gern BH, Olin BE, Leal JM, Wu Y, Lyons-Cohen MR, Huang JY, Paz-Stoltzfus CL, Plumlee CR, Pöschinger T, Urdahl KB, Perro M, Gerner MY (2020) CytoMAP: a spatial analysis toolbox reveals features of myeloid cell organization in lymphoid tissues. Cell Rep 31(3): 107523. https://doi.org/10.1016/j.celrep.2020.107523

Sountoulidis A, Liontos A, Nguyen HP, Firsova AB, Fysikopoulos A, Qian X, Seeger W, Sundström E, Nilsson M, Samakovlis C (2020) SCRINSHOT enables spatial mapping of cell states in tissue sections with single-cell resolution. PLoS Biol 18(11): e3000675. https://doi.org/10.1371/journal.pbio.3000675

Su G, Qin X, Enninful A, Bai Z, Deng Y, Liu Y, Fan R (2021) Spatial multi-omics sequencing for fixed tissue via DBiT-seq. STAR Protoc 2(2): 100532. https://doi.org/10.1016/j.xpro.2021.100532

Su JH, Zheng P, Kinrot SS, Bintu B, Zhuang X (2020) Genome-scale imaging of the 3D organization and transcriptional activity of chromatin. Cell 182(6): 1641–1659

Sun S, Zhu J, Zhou X (2020) Statistical analysis of spatial expression patterns for spatially resolved transcriptomic studies. Nat Methods 17(2): 193–200

Svensson V, Teichmann SA, Stegle O (2018) SpatialDE: identification of spatially variable genes. Nat Methods 15(5): 343–346

Tran M, Yoon S, Teoh M, Andersen S, Lam PY, Purdue P, Raghubar A, Hanson SJ, Devitt K, Jones K, Walters S, Tuong ZK, Kulasinghe A, Monkman J, Soyer HP, Frazer I, Nguyen Q (2021) Spatial analysis of ligand-receptor interaction in skin cancer at genome-wide and single-cell resolution. bioRxiv. https:// doi.org/10.1101/2020.09.10.290833

Trcek T, Lionnet T, Shroff H, Lehmann R (2017) mRNA quantification using single-molecule FISH in *Drosophila* embryos. Nat Protoc 12(7): 1326–1348

van den Brink SC, Alemany A, van Batenburg V, Moris N, Blotenburg M, Vivié J, Baillie-Johnson P, Nichols J, Sonnen KF, Martinez Arias A, van Oudenaarden A (2020) Single-cell and spatial transcriptomics reveal somitogenesis in gastruloids. Nature 582: 405–409

Vickovic S, Eraslan G, Salmen F, Klughammer J, Stenbeck L, Schapiro D, Aijo T, Bonneau R, Bergenstrahle L, Navarro JF, Gould J, Griffin GK, Borg A, Ronaghi M, Frisen J, Lundeberg J, Regev A, Stahl PL (2019) High-definition spatial transcriptomics for *in situ* tissue profiling. Nat Methods 16(10): 987–990

Vickovic S, Lötstedt B, Klughammer J, Segerstolpe Å, Rozenblatt-Rosen O, Regev A (2020) SM-Omics: an automated platform for high-throughput spatial multi-omics. bioRxiv. https://doi.org/10.1101/2020.10.14.338418

Wang F, Flanagan J, Su N, Wang LC, Bui S, Nielson A, Wu X, Vo HT, Ma XJ, Luo Y (2012) RNAscope: a novel *in situ* RNA analysis platform for formalin-fixed, paraffin-embedded tissues. J Mol Diagn 14(1): 22–29

Wang G, Moffitt JR, Zhuang X (2018a) Multiplexed imaging of high-density libraries of RNAs with MERFISH and expansion microscopy. Sci Rep 8(1): 4847. https://doi.org/10.1038/s41598-018-22297-7

Wang X, Allen WE, Wright MA, Sylwestrak EL, Samusik N, Vesuna S, Evans K, Liu C, Ramakrishnan C, Liu J, Nolan GP, Bava FA, Deisseroth K (2018b) Three-dimensional intact-tissue sequencing of single-cell transcriptional states. Science 361(6400):eaat5691.https://doi.org/10.1126/science.aat5691

Xia C, Babcock HP, Moffitt JR, Zhuang X (2019) Multiplexed detection of RNA using MERFISH and branched DNA amplification. Sci Rep 9(1): 7721. https://doi.org/10.1038/s41598-019-43943-8

Zhang L, Mao SQ, Yao ML, Chao N, Yang Y, Ni Y, Song T, Liu Z, Yang Y, Li W (2021) Spatial transcriptome sequencing revealed spatial trajectory in the non-small cell lung cancer. bioRxiv. https://doi.org/10.1101/2021.04.26.441394

© The Author(s) 2022