SpaceSequest: A unified pipeline for spatial transcriptomics data analysis

Yu H. Sun[1,#,*], Sarbottam Piya[1,#], Zhengyu Ouyang[2, #], Yirui Chen[1], Jake Gagnon[3], Shaolong Cao[1], Haotian Zhang[1], Baobao Song[1], Jing Zhu[1], Khyati Chandratre[1], Houlin Yu[1], Wenxing Hu[1], Matthew Ryals[4], Fergal Casey[1], Dann Huh[1], Baohong Zhang[1]

[1] Research, Biogen Inc., Cambridge, MA, 02142, USA

[2] Data Science, BioInfoRx Inc., Madison, WI, 53719, USA

[3] Biostatistics, Biogen Inc., Cambridge, MA, 02142, USA

[4] PharmaLex, Conshohocken, PA, 19428, USA

Contribute equally

* Correspondence: Yu H. Sun, yuhenry.sun@biogen.com

ABSTRACT (limit of 250 words)

Background

Spatial transcriptomics has emerged as one of the most powerful tools for gaining biological insights, enabling researchers to uncover intricate relationships between gene expression patterns and tissue architecture. Recent advances in the field have resulted in a variety of new platforms, including Visium, Visium HD, and Xenium from 10x Genomics, as well as GeoMx and CosMx from NanoString Technologies, which has now been acquired by the Bruker Corporation. However, the existence of diverse spatial transcriptomics platforms and various data formats poses challenges in standardizing data analysis. Thus, there remains a critical gap in the availability of a comprehensive pipeline capable of conducting end-to-end analysis that is necessary to extract biological insights from multiple spatial transcriptomics platforms.

Results

Here, we present SpaceSequest, a tailored pipeline that utilizes cutting-edge computational methodologies to conduct a thorough analysis, enabling the extraction of crucial biological insights from five major spatial transcriptomics technologies. SpaceSequest performs (1) standardized quality control and general data processing, (2) key analyses customized for each spatial platform, (3) automated cell type annotation and deconvolution, and (4) high-quality figure and analysis result generation. In addition, SpaceSequest allows for smooth integration with cellxgene VIP and Quickomics for user-friendly data access and interactive visualization.

Conclusions:

SpaceSequest is a unified and comprehensive pipeline designed for the analysis, visualization, and publication of spatial transcriptomics data from various platforms. The source code is available at https://github.com/interactivereport/SpaceSequest. To facilitate seamless installation and usage, we have also created a detailed Bookdown tutorial that can be accessed through https://interactivereport.github.io/SpaceSequest/tutorial/docs/index.html.

Keywords:
Spatial transcriptomics, Visium, Visium HD, Xenium, GeoMx, CosMx, single-cell RNA-seq, deconvolution

BACKGROUND

The advent of single-cell RNA sequencing (scRNA-seq) has revolutionized our understanding of cellular heterogeneity and gene expression dynamics. It represents a significant advancement from bulk RNA-seq, which only provides average expression profiles from mixed cell populations, thereby obscuring the underlying cellular diversity [1–3]. Although scRNA-seq generates invaluable insights into cellular heterogeneity, it lacks spatial information, as it dissociates cells from their original tissue context [4, 5]. Understanding the intricate spatial organization of gene expression within complex tissues and organs is crucial for unraveling the mysteries of biology [6–8]. Spatial transcriptomics (ST) has emerged as a transformative technology that quantifies RNA within intact tissue sections, enabling researchers to explore gene expression patterns within their native spatial context [6, 9–12].

ST can provide precise insights into disease progression by discerning heterogeneity within diseased tissue, which is powerful for studying complex diseases such as tumor and neurological disorders[13–17]. For example, using ST data, Berglund et al. not only distinguished healthy and diseased areas but were able to identify re-stratified tumor heterogeneity by analyzing gene expression gradients within the tumor microenvironment, specifically in the stroma adjacent to the tumor area [18]. These advancements have unlocked new avenues for addressing critical biological inquiries. First, ST plays a pivotal role in characterizing the cell-type composition of tissues. While definitions are often derived from large-scale single-cell RNA-seq or epigenetics datasets and projected onto ST data, the technology also empowers the direct generation of cell-type definitions from the spatial transcriptomics data itself[19, 20]. This application has been widely utilized, resulting in the creation of compositional atlases for a diverse range of tissues, including the nervous system [21–24], human kidney [25–27], heart [28, 29], testes [30], and lung [31, 32]. Second, ST facilitates the exploration of cellular interactions, revealing the rules and patterns governing spatial covariation among distinct cell types. For instance, studies examining the mouse visual cortex identified a preferential spatial proximity of inhibitory neuron

subtypes, shedding light on their organizational principles[33]. Similar proximity analyses have been instrumental in elucidating gene expression patterns linked to Alzheimer's disease and in characterizing histopathological responses to traumatic brain injury[34–37]. Finally, ST's transcriptome-wide data on in situ gene expression empowers the elucidation of molecular interactions between tissue components[7]. Through the identification of ligand-receptor pairs among cell types at varying spatial proximities, we can discern the intricate communication networks that underlie phenomena such as tumor interactions with the microenvironment, immune infiltrates in tissues, and the establishment of developmental gradients[38, 39].

There are two distinct ST methods: (i) sequencing-based ST, and (ii) imaging-based ST. In the sequencing-based ST method, a DNA-barcoded surface is used to capture RNA molecules from the intact tissue section. These captured RNA molecules are then subjected to next-generation sequencing to determine their respective gene identities[6]. By contrast, in imaging-based methods, multiple mRNA transcripts are simultaneously detected within intact tissue using fluorescence in situ hybridization (FISH) or direct in-situ sequencing[6, 12]. The imaging-based methods, such as 10x Genomics Xenium and NanoString CosMx, are suitable for hypothesis testing as these methods allow transcriptomic profiling at single cell resolution albeit for a limited number of target genes. Conversely, sequencing-based methods which produce an unbiased, whole-transcriptome profile, are suitable for hypothesis generation. Typical methods include 10x Genomics Visium, Visium HD, and NanoString GeoMx with sequencing readout. However, transcriptomics profiles obtained from these methods are not directly resolved at single cell resolution.

One critical drawback of current grid-based ST profiling methods (Visium, Slide-seq [38, 40], and HDST [41] is their inability to achieve single-cell resolution. These techniques either combine multiple cells or fractions of multiple cells at each tissue location, limiting the ability to discern individual cell contributions. Nevertheless, this limitation can be addressed through computational integration, leveraging the reference transcriptome signatures of cell types obtained from coupled single-cell RNA sequencing (scRNA-seq) profiles. By integrating ST data with scRNA-seq data, researchers can deconvolute cell type composition and retrieve further information from the spatially resolved gene expression patterns[42, 43].

The landscape of spatial transcriptomics data analysis has seen the development of several data analysis pipelines, including ST Pipeline [44] and Spacemake [45]. ST Pipeline primarily focuses on the essential tools and scripts required to process raw files derived from Visium data in FASTQ format, providing datasets primed for downstream analysis [44]. In contrast, Spacemake represents a more comprehensive solution, offering a wider range of functionalities compared to ST Pipeline. This tool is adept at handling spatial transcriptomic data from diverse platforms, such as Visium, Slide-seq, Seq-scope, etc [45]. Another distinctive feature of Spacemake is its utilization of NovoSpaRc [46] that allows mapping snRNASeq data to spatial coordinates [45]. Recently developed pipelines are focused on specific tasks, for example, PIPEFISH[47] was designed for FISH spatial transcriptomics, ENACT was specifically designed for Visium HD data analysis[48], while Panpipes[49] was created for spatial multi-omics analysis. Although these existing pipelines are useful, their capabilities fall short in extracting comprehensive and meaningful biological insights from diverse spatial transcriptomics datasets.

Here we introduce SpaceSequest, our comprehensive spatial transcriptomics data analysis pipeline tailored for the analysis of several major platforms, spanning 10X Genomics Visium, Visium HD, Xenium, as well as NanoString GeoMx and CosMx platforms. By utilizing cutting-edge computational methodologies, SpaceSequest streamlines a series of essential analyses to extract meaningful biological insights. Compatible with diverse input formats, the pipeline seamlessly navigates through critical steps including clustering, deconvolution, cell type annotation, and differential expression analysis. This integrated approach equips researchers with a powerful toolkit to dissect the spatially defined molecular landscape, facilitating a deeper understanding of complex biological systems and their microenvironmental interactions.

IMPLEMENTATION

General data processing

SpaceSequest supports various input formats and processes the data through a series of automated steps, leveraged by multiple well-designed packages including Seurat 5 [50] and anndata[51]. For sequencing-based strategies (10x Genomics Visium, Visium HD, and NanoString GeoMx with sequencing readouts), FASTQ processing is required before running

SpaceSequest. The input files for running Visium and Visium HD workflows, including fundamental quality metrics and the unique molecular identifiers (UMI) count matrix, can be generated through 10x Genomics Space Ranger software. Since our pipeline does not rely on gene annotation or species-specific details, users are required to accurately select the appropriate species during Space Ranger execution. For GeoMx, the GeoMx NGS Pipeline developed by NanoString is required to convert FASTQ to digital count conversion (DCC) files, which can be further processed by SpaceSequest. In the case of imaging-based technologies such as Xenium and CosMx, files generated by the platform can be readily used for SpaceSequest data processing.

Clustering

Clustering is a powerful tool for understanding the structure of the data, as well as the spatial tissue domains. For general clustering of single cell expression in CosMx and Xenium, or bins in Visium HD, we utilized the UMAP dimension reduction followed by the FindNeighbors and FindClusters functions. For Visium technology, we have incorporated two more clustering tools, SpaGCN [52] and BayesSpace [53]. SpaGCN leverages graph convolutional networks to integrate gene expression, spatial location, and histology data, allowing for the identification of spatial domains with distinct expression patterns [52]. In contrast, BayesSpace is a Bayesian statistical tool that leverages prior spatial knowledge to cluster transcriptomically similar spots in proximity [53]. Additionally, BayesSpace offers the capability to enhance resolution by reallocating gene counts from larger spots in array-based data, such as Visium, to more refined sub-spots based on spatial information from nearby spots [53].

Cell type annotation

Automated cell type annotation can be achieved through transferring labels from a reference single-cell RNA-seq dataset. For single-cell-level assays, including CosMx and Xenium, we incorporated Azimuth to predict the most likely cell type for each cell[54]. A UMAP plot will be generated after running the pipeline for users to examine the cell type annotation results.

Deconvolution

For deconvolution purposes, we have incorporated two state-of-the-art deconvolution tools, Tangram [55] and Cell2location [42]. We selected these two packages for their outstanding

performance in a recent benchmarking study[43]. Tangram, a deep-learning framework, excelled in creating comprehensive single-cell-resolution spatial gene-expression maps and integrating histological and anatomical data from the same specimens. It adeptly handled low-quality genes in high-resolution methods, offered single-cell resolution for lower-resolution techniques, and ensured broad genome-wide coverage for targeted approaches[55]. Meanwhile, Cell2location, a Bayesian model, showcased its ability to discern fine-grained cell types within spatial transcriptomic data and constructs detailed cellular maps of diverse tissues. In the first step, Cell2location estimated reference cell type signatures from single-cell and single-nucleus RNA-seq profiles, aligned with gene expression profiles for user-specified cell types and subpopulations. In the subsequent stage, Cell2location utilized these reference signatures in conjunction with one or more spatial transcriptomic datasets to deconstruct mRNA counts at individual spatial locations into their respective reference cell types[42].

Ligand receptor interactions

Here, we employ SpaTalk to investigate spatially resolved cell-cell interactions that will enable us to get insights into tissue homeostasis, development, and disease[56]. This tool leverages a graph network along with a knowledge graph to model and evaluate the ligand-receptor-target signaling network operating between neighboring cells[56]. SpaTalk serves as an optional step in the Visium data processing workflow, providing researchers with a powerful means to unravel the complex dynamics of inter-cellular interactions within tissues[56].

Standardized output

SpaceSequest generates a series of output files that are ready for result query and further analysis. The R objects are saved as RData files, and the python anndata objects are saved as h5ad files. These files contain expression matrices, dimension reduction results, and intermediate data generated during pipeline processing.

RESULTS

Highlights of SpaceSequest

The pipeline incorporates several key workflows to perform various analyses corresponding to each platform. Key strengths of this pipeline include:

- 1) Self-explainable configuration files that ensure easy set up.
- 2) Standardized and reproducible data processing powered by R and Python.
- 3) Customized tools are integrated in the pipeline for each spatial platform.
- 4) Comprehensive result generation, including R/Python objects and key figures.
- 5) Streamlined data visualization enabled by cellxgene VIP and Quickomics.

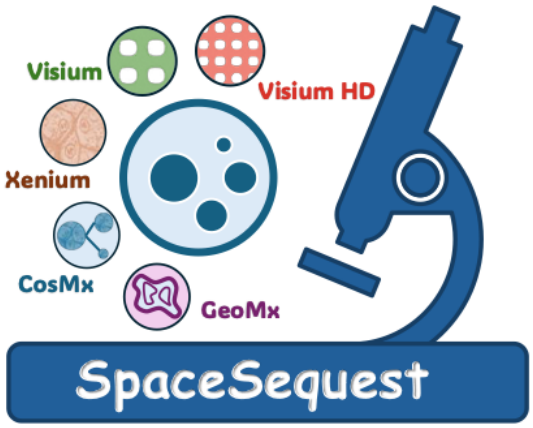

Figure 1. Components of SpaceSequest. The pipeline integrates five major workflows tailored to perform a series of analysis for each platform, including Visium, Visium HD, Xenium, CosMx, and GeoMx.

Overall pipeline design

SpaceSequest is a unified toolkit which contains a set of workflows. Each workflow can be activated by calling the workflow script, and necessary configuration files will be created accordingly. The main configuration files are in YAML format, ensuring standardized formatting and flexible modification. To execute the pipeline, in most of the times, users are typically required to fill in a sample metadata file, which documents the details of each data, such as sample name, file path, and associated sample-level information. The overall pipeline components are illustrated in Figure 1.

Visium data analysis incorporates multiple tools

The 10x Genomics Visium platform was released in 2019 and continues to serve as a powerful commercial spatial transcriptomics technology since then. The Visium slides contain an array of barcoded 55 µm spots, with two major tissue capture sizes: 6.5x6.5 mm or 11x11 mm. Sequencing data in FASTQ format are processed through Space Ranger, and SpaceSequest continues the analysis by utilizing the binned output files. Firstly, SpaceSequest performs standard quality control on the data by removing spots with low gene detection rates, low UMI counts, etc. The filtered data are then integrated into a new h5ad file, which is piped to downstream tools for further analysis.

We tested our pipeline on a public dataset generated from the brain tissue of an Alzheimer's disease mouse model, which has mutations in APP and PSEN1 genes[57]. We also downloaded the Azimuth mouse motor cortex single-nucleus RNA-seq data as a reference for cell type annotation[58]. The data analysis is achieved by calling the ***visium*** command in SpaceSequest, which reads all input data and performs a series of quality control steps (Figure 2A). The cleaned data are then integrated for downstream processing, which includes spatial clustering and cell type deconvolution results. To leverage the power of parallel computing, SpaceSequest is able to submit sbatch jobs to execute the following four tools simultaneously: BayesSpace[53], Cell2location[42], SpaGCN[52], and Tangram[55]. The jobs will be monitored until results have been generated and saved in separate directories. To provide a few examples, BayesSpace clustering step will cluster the spots across samples and identify comparable spatial clusters (Figure 2B). While Cell2location results can be visualized by cell abundance heatmaps overlayed with the tissue image (Figure 2C). Additionally, the pipeline utilizes SpaTalk for cell-cell communication analysis, which requires Cell2location output so it is only executed once Cell2location results are available (Figure S1)[56]. Overall, our Visium data analysis workflow ensures automated data processing and analysis with a streamlined, one-stop-shop design (Figure 2A).

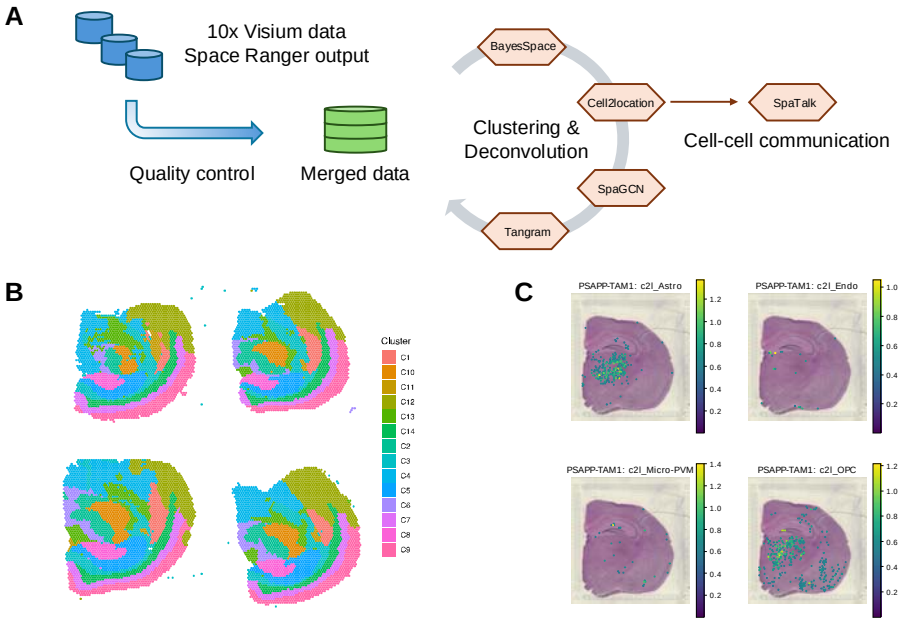

Figure 2. 10x Visium data analysis. (A) Workflow for Visium data processing. (B) Spatial clustering results generated by BayesSpace. This run contains four samples, and BayesSpace identified shared clusters across them. (C) Predicted cell type abundance by Cell2location in one sample (PSAPP-TAM1), highlighting astrocytes, OPCs, microglia, and endothelial cells.

Streamlined Xenium and CosMx data analysis

The 10x Genomics Xenium and NanoString (now acquired by the Bruker Corporation) CosMx platforms are both imaging-based, high-plex in-situ technologies capable of capturing up to thousands of RNAs simultaneously at single-cell resolution. They require a set of probes, usually from pre-designed panels, with additional space for customization as well. Xenium data consists of a set of well-defined files in various formats, such as the cell_feature_matrix files (a folder with three associated files, or a single h5 file), transcript files (transcripts.csv.gz, transcripts.parquet, transcripts.zarr.zip), and cell summary files (cells.csv.gz, cells.parquet), etc. The LoadXenium function from the Seurat package reads the necessary files from the directory and converts them into a Seurat object for downstream analysis.

CosMx data are stored as a collection of plain text files in the comma-separated values (csv) format. These files include a metadata file (metadata_file), a file storing field of view (fov) positions (fov_positions_file), an expression matrix file (exprMat_file), a polygon file (polygons), and a transcript data file (tx_file). Data loading is achieved by the LoadNanostring function, which converts these csv files to a Seurat object.

SpaceSequest contains two workflows—***xenium*** and ***cosmx*** —to streamline the process from data loading to downstream analysis. Once Xenium or CosMx data are successfully loaded, these workflows incorporate standard procedures that perform quality control, clustering, cell type annotation, and differential expression (DE) analysis, etc (Figure 3A). Moreover, the workflows generate UMAP and tissue visualizations, enabling users to examine the dimension reduction and cell type annotation results. Example plots generated using public datasets are shown in Figure 3B-D. The Xenium data were generated from a healthy human brain FFPE sample profiled with the Xenium Human Brain Gene Expression Panel containing 266 targets. On the other hand, the CosMx data were derived from human frontal cortex FFPE tissue using >6,000 target genes. After running the workflow, various analysis results and the processed Seurat object will be saved for future queries and additional analyses, including further exploration by Cellxgene VIP[59].

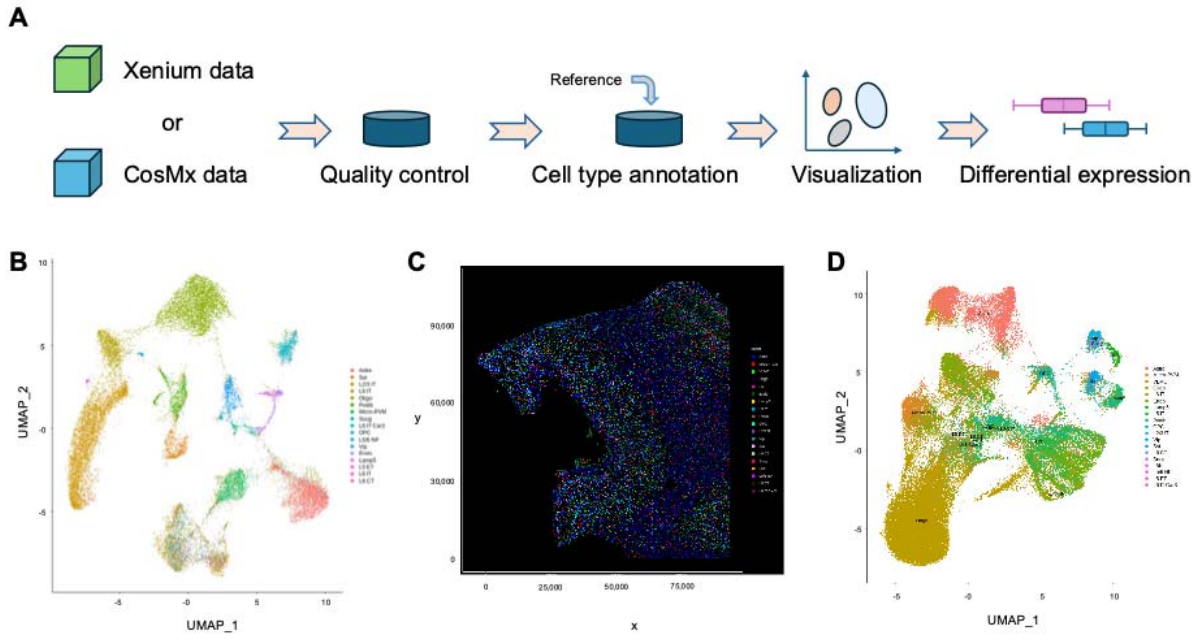

Figure 3. Xenium and CosMx data analysis. (A) Streamlined workflows for analyzing Xenium and CosMx data. (B) UMAP visualization of Xenium data, labelled by predicted cell types. (C) Tissue and cell visualization of CosMx data. (D) UMAP visualization of CosMx data with predict cell types labelled.

Enhanced GeoMx data processing and visualization

The NanoString GeoMx platform is a useful technique to compare gene expression differences in defined regions on the tissue slides. It uses probes to detect the whole transcriptome in the selected region of interests (ROIs). However, GeoMx is not a single-cell resolution technique; instead, it requires a sufficient number of cells within the region to generate robust signals. Thus, GeoMx is considered as a mini-bulk RNA detection platform, capable of profiling >18,000 major genes for human and mouse samples. In our pipeline, we developed ***geomx***, a workflow that takes DCC files, sample metadata, and a set of comparison groups as input. It performs a series of analysis, and outputs differential expression analysis results (Figure 4A). This workflow generates four csv files that are fully compatible with other tools including some Shiny applications[60, 61], for example, Quickomics, an open-source R Shiny application for flexible data exploration and visualization[60]. After uploading the data, users can examine PCA plots (Figure 4B), volcano plots (Figure 4C), heatmaps (Figure 4D), and a number of on-the-fly visualizations through Quickomics. This strategy not only automates GeoMx data analysis, especially for studies with many samples, but bridges the gap between raw data and insight generation through Quickomics, which creates a link to facilitate data storage and future analysis.

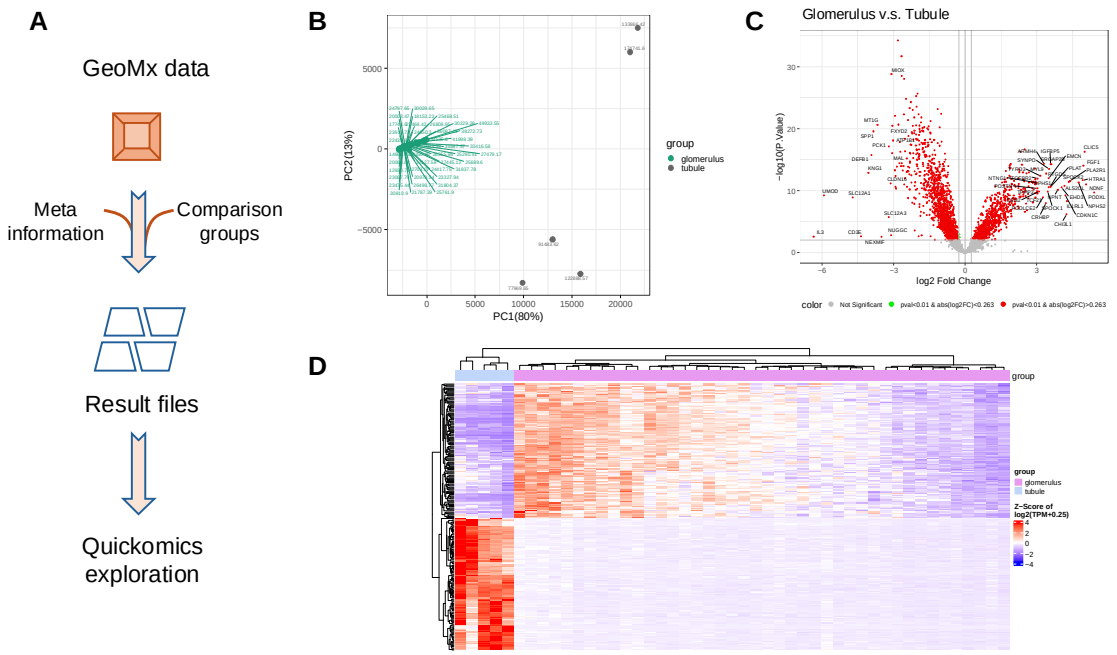

Figure 4. GeoMx data analysis. (A) Overview of GeoMx data analysis workflow. (B) PCA plot showing samples from two kidney regions, glomerulus and tubule. (C) Volcano plot highlighting differentially expressed genes between glomerulus and tubule. (D) Heatmap displaying the top differentially expressed genes between glomerulus and tubule, with samples and genes clustered by dendrograms.

Standardized data analysis of the cutting-edge Visium HD platform

In 2024, 10x Genomics released a new platform called Visium HD, featuring high-resolution, gap-free detection of tissue RNAs. This platform is now compatible with FFPE, fixed frozen, and fresh frozen samples. To keep pace with the rapid evolution of spatial transcriptomics technologies, we incorporated ***visiumhd*** into SpaceSequest for Visium HD data processing. The workflow takes Space Ranger outputs, performs bin-level quality control, dimension reduction, and cell type annotation when an Azimuth reference dataset is available (Figure 5A). By executing the pipeline, users receive plots to examine the gene counts (Figure 5B) and cell type annotation results (Figure 5C). Furthermore, we developed a new function, getVisiumHDspotCol, to extract the mean RGB intensities when a paired high-resolution immunofluorescence image is available (Figure 5D). This function facilitates accurate bin

selection using nuclei staining, cell type marker staining, or other biomarker staining, for example, pathological protein markers.

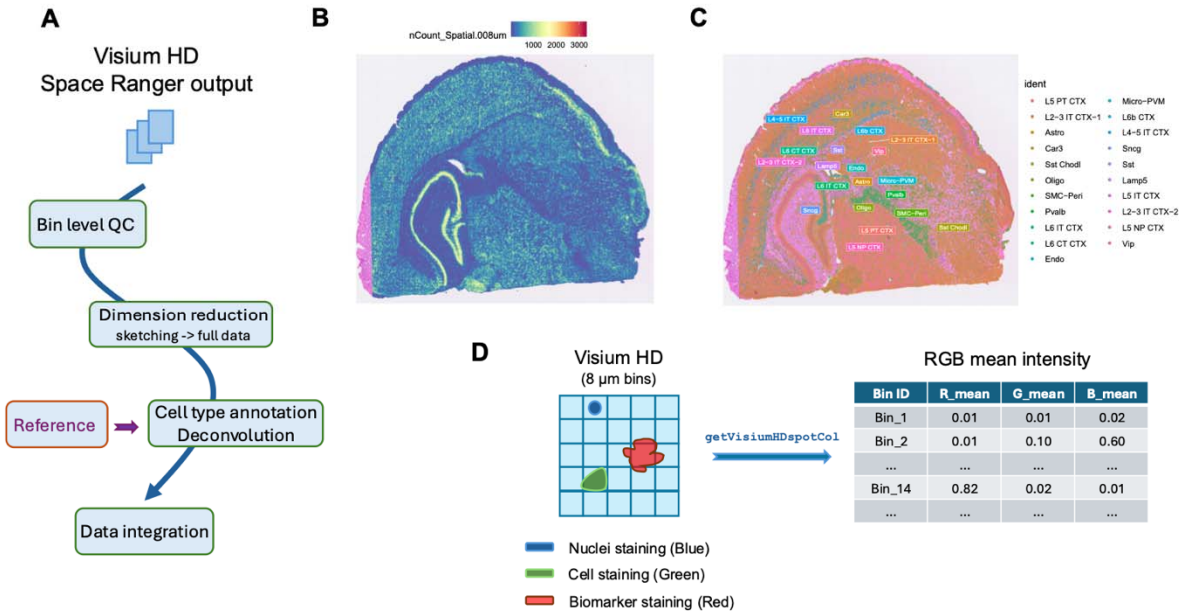

Figure 5. Visium HD data analysis. (A) Overview of the general processing workflow for Visium HD data. (B) Heatmap showing the number of UMI counts (nCount) detected across the tissue section at 8 µm bin resolution. (C) Visualization of annotated cell types on the tissue section. (D) Illustration of a new function that calculates the average RGB intensities across bins.

DISCUSSION

In this study, we developed SpaceSequest, a unified pipeline for spatial transcriptomics data processing. Currently, it supports five major platforms and incorporates customized software and packages to maximize the information that can be extracted from the data. To make the pipeline accessible, we provided a detailed tutorial aimed at lowering the barrier for researchers to apply this SpaceSequest to their own studies. Guided by the FAIR principles (Findable, Accessible, Interoperable, and Reusable), SpaceSequest enhances data access, interpretation, and facilitates data sharing by generating standardized results. In summary, SpaceSequest is a user-friendly, end-to-end solution for spatial transcriptomics data analysis and interpretation.

Abbreviations

scRNA-seq: Single-cell RNA-seq

NGS: Next-generation sequencing

csv: comma-separated values

fov: field of view

ST: spatial transcriptomics

UMI: unique molecular identifiers

FISH: fluorescence in situ hybridization

DE: Differential expression

DCC: digital count conversion

PCA: Principal component analysis

OPC: Oligodendrocyte progenitor cells

UMAP: Uniform Manifold Approximation and Projection

Cellxgene VIP: Cellxgene visualization in plug-in

LRI: ligand-receptor interaction

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

Y.H.S., S.P., Y.C., J.G., S.C., B.S., J.Z., W.H., F.C., D.H., and B.Z. are employees of Biogen and hold stocks from the company. H.Z., K.C. and H.Y. were co-ops at Biogen and do not hold Biogen's stocks. Z.O. and M.R. are employees of BioInfoRx Inc. and PharmaLex, respectively.

Funding

This study is funded by Biogen Inc.

Author Contributions Statement

Y.H.S., S.P., and Z.O. developed the pipeline and wrote the manuscript, with the help from Y.C., J.G., S.C., H.Z., B.S., J.Z., K.C., H.Y., W.H., M.R., F.C., D.H., and B.Z. All authors reviewed the manuscript.

Acknowledgement

We acknowledge the scientists at Biogen for their valuable feedback on the pipeline.

Availability and Requirements

No new data were generated from this study. Several public datasets were used to demonstrate SpaceSequest. The 10x Visium data were from a public study using mouse brain tissues with NCBI accession number GSE203424 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE203424). Public data were downloaded from 10x Genomics dataset webpage (https://www.10xgenomics.com/datasets) to test Xenium (https://www.10xgenomics.com/datasets/xenium-human-brain-preview-data-1-standard) and Visium HD (FFPE: https://www.10xgenomics.com/datasets/visium-hd-cytassist-gene-expression-libraries-of-mouse-brain-he, Fixed frozen: https://www.10xgenomics.com/datasets/visium-hd-cytassist-gene-expression-mouse-brain-fixed-frozen) workflows, respectively. Moreover, we used a public human kidney GeoMx dataset (http://nanostring-public-share.s3-website-us-west-2.amazonaws.com/GeoScriptHub/Kidney_Dataset_for_GeomxTools.zip) and a human brain CosMx dataset (https://nanostring.com/products/cosmx-spatial-molecular-imager/ffpe-dataset/human-frontal-cortex-ffpe-dataset/) from NanoString websites to test GeoMx and CosMx pipelines respectively. More details related to data downloading and processing can be found at

the tutorial (https://interactivereport.github.io/SpaceSequest/tutorial/docs/index.html) associated with this pipeline.

Project name: SpaceSequest

Project home page: https://github.com/interactivereport/spacesequest

Operating system(s): Linux, Mac OS

Programming language: Bash (Unix Shell)

Other requirements: Conda

License: MIT license

Any restrictions to use by non-academics: No

Supplementary figure legends

Figure S1: Examples of SpaTalk results. (A) Highlighted pairs between astrocytes (Astro) and GABAergic neurons. (B) Significantly enriched ligand-receptor interaction (LRI) pairs between astrocytes and GABAergic neurons**.** (C) Highlighted pairs between microglia and macrophage cells (Micro_PVM) and Glutamatergic neurons. (D) Significantly enriched ligand-receptor interaction (LRI) pairs between microglia/macrophages and Glutamatergic neurons.

REFERENCES

1. Jovic D, Liang X, Zeng H, Lin L, Xu F, Luo Y. Single-cell RNA sequencing technologies and applications: A brief overview. Clinical & Translational Med. 2022;12:e694. https://doi.org/10.1002/ctm2.694.

2. Wang S, Sun S-T, Zhang X-Y, Ding H-R, Yuan Y, He J-J, et al. The Evolution of Single-Cell RNA Sequencing Technology and Application: Progress and Perspectives. IJMS. 2023;24:2943. https://doi.org/10.3390/ijms24032943.

3. Potter SS. Single-cell RNA sequencing for the study of development, physiology and disease. Nat Rev Nephrol. 2018;14:479–92. https://doi.org/10.1038/s41581-018-0021-7.

4. Du J, Yang Y-C, An Z-J, Zhang M-H, Fu X-H, Huang Z-F, et al. Advances in spatial transcriptomics and related data analysis strategies. J Transl Med. 2023;21:330. https://doi.org/10.1186/s12967-023-04150-2.

5. Adema K, Schon MA, Nodine MD, Kohlen W. Lost in space: what single-cell RNA sequencing cannot tell you. Trends in Plant Science. 2024;29:1018–28. https://doi.org/10.1016/j.tplants.2024.03.010.

6. Williams CG, Lee HJ, Asatsuma T, Vento-Tormo R, Haque A. An introduction to spatial transcriptomics for biomedical research. Genome Med. 2022;14:68. https://doi.org/10.1186/s13073-022-01075-1.

7. Longo SK, Guo MG, Ji AL, Khavari PA. Integrating single-cell and spatial transcriptomics to elucidate intercellular tissue dynamics. Nat Rev Genet. 2021;22:627–44. https://doi.org/10.1038/s41576-021-00370-8.

8. Saviano A, Henderson NC, Baumert TF. Single-cell genomics and spatial transcriptomics: Discovery of novel cell states and cellular interactions in liver physiology and disease biology. Journal of Hepatology. 2020;73:1219–30. https://doi.org/10.1016/j.jhep.2020.06.004.

9. Li X, Wang C-Y. From bulk, single-cell to spatial RNA sequencing. Int J Oral Sci. 2021;13:36. https://doi.org/10.1038/s41368-021-00146-0.

10. Moses L, Pachter L. Museum of spatial transcriptomics. Nat Methods. 2022;19:534–46. https://doi.org/10.1038/s41592-022-01409-2.

11. Tian L, Chen F, Macosko EZ. The expanding vistas of spatial transcriptomics. Nat Biotechnol. 2023;41:773–82. https://doi.org/10.1038/s41587-022-01448-2.

12. Wang Y, Liu B, Zhao G, Lee Y, Buzdin A, Mu X, et al. Spatial transcriptomics: Technologies, applications and experimental considerations. Genomics. 2023;115:110671. https://doi.org/10.1016/j.ygeno.2023.110671.

13. Jain S, Eadon MT. Spatial transcriptomics in health and disease. Nat Rev Nephrol. 2024;20:659–71. https://doi.org/10.1038/s41581-024-00841-1.

14. Piwecka M, Rajewsky N, Rybak-Wolf A. Single-cell and spatial transcriptomics: deciphering brain complexity in health and disease. Nat Rev Neurol. 2023;19:346–62. https://doi.org/10.1038/s41582-023-00809-y.

435 15. Zhang L, Xiong Z, Xiao M. A Review of the Application of Spatial Transcriptomics in Neuroscience.
436 Interdiscip Sci Comput Life Sci. 2024;16:243–60. https://doi.org/10.1007/s12539-024-00603-4.

437 16. Jung N, Kim T-K. Spatial transcriptomics in neuroscience. Exp Mol Med. 2023;55:2105–15.
438 https://doi.org/10.1038/s12276-023-01093-y.

439 17. Tsai H-H, Piya S, Wang J, Zhu J, Hu W, Gehrke AR, et al. Spatial transcriptomics reveals
440 heterogeneous cell–cell interactions among brain regions in a cuprizone model consistent with multiple
441 sclerosis lesions. bioRxiv. 2025;:2025.02.07.637125. https://doi.org/10.1101/2025.02.07.637125.

442 18. Berglund E, Maaskola J, Schultz N, Friedrich S, Marklund M, Bergenstråhle J, et al. Spatial maps of
443 prostate cancer transcriptomes reveal an unexplored landscape of heterogeneity. Nat Commun.
444 2018;9:2419. https://doi.org/10.1038/s41467-018-04724-5.

445 19. Vandereyken K, Sifrim A, Thienpont B, Voet T. Methods and applications for single-cell and spatial
446 multi-omics. Nat Rev Genet. 2023;24:494–515. https://doi.org/10.1038/s41576-023-00580-2.

447 20. Ma Y, Zhou X. Spatially informed cell-type deconvolution for spatial transcriptomics. Nat Biotechnol.
448 2022;40:1349–59. https://doi.org/10.1038/s41587-022-01273-7.

449 21. Shi H, He Y, Zhou Y, Huang J, Maher K, Wang B, et al. Spatial atlas of the mouse central nervous
450 system at molecular resolution. Nature. 2023;622:552–61. https://doi.org/10.1038/s41586-023-06569-
451 5.

452 22. Huuki-Myers LA, Spangler A, Eagles NJ, Montgomery KD, Kwon SH, Guo B, et al. A data-driven single-
453 cell and spatial transcriptomic map of the human prefrontal cortex. Science. 2024;384:eadh1938.
454 https://doi.org/10.1126/science.adh1938.

455 23. Qian X, Coleman K, Jiang S, Kriz AJ, Marciano JH, Luo C, et al. Spatial transcriptomics reveals human
456 cortical layer and area specification. Nature. 2025. https://doi.org/10.1038/s41586-025-09010-1.

457 24. Han L, Liu Z, Jing Z, Liu Y, Peng Y, Chang H, et al. Single-cell spatial transcriptomic atlas of the whole
458 mouse brain. Neuron. 2025;113:2141-2160.e9. https://doi.org/10.1016/j.neuron.2025.02.015.

459 25. Lake BB, Menon R, Winfree S, Hu Q, Melo Ferreira R, Kalhor K, et al. An atlas of healthy and injured
460 cell states and niches in the human kidney. Nature. 2023;619:585–94. https://doi.org/10.1038/s41586-
461 023-05769-3.

462 26. Wu H, Liu F, Shangguan Y, Yang Y, Shi W, Hu W, et al. Integrating spatial transcriptomics with single-
463 cell transcriptomics reveals a spatiotemporal gene landscape of the human developing kidney. Cell
464 Biosci. 2022;12:80. https://doi.org/10.1186/s13578-022-00801-x.

465 27. Chen D, Shao M, Song Y, Ren G, Guo F, Fan X, et al. Single-cell RNA -seq with spatial transcriptomics
466 to create an atlas of human diabetic kidney disease. The FASEB Journal. 2023;37:e22938.
467 https://doi.org/10.1096/fj.202202013RR.

468 28. Mohenska M, Tan NM, Tokolyi A, Furtado MB, Costa MW, Perry AJ, et al. 3D-cardiomics: A spatial
469 transcriptional atlas of the mammalian heart. Journal of Molecular and Cellular Cardiology.
470 2022;163:20–32. https://doi.org/10.1016/j.yjmcc.2021.09.011.

471 29. Lee SE, Joo JH, Hwang HS, Chen S-F, Evans D, Lee KY, et al. Spatial transcriptional landscape of human
472 heart failure. European Heart Journal. 2025;:ehaf272. https://doi.org/10.1093/eurheartj/ehaf272.

473 30. Chen H, Murray E, Sinha A, Laumas A, Li J, Lesman D, et al. Dissecting mammalian spermatogenesis
474 using spatial transcriptomics. Cell Reports. 2021;37:109915.
475 https://doi.org/10.1016/j.celrep.2021.109915.

476 31. Madissoon E, Oliver AJ, Kleshchevnikov V, Wilbrey-Clark A, Polanski K, Richoz N, et al. A spatially
477 resolved atlas of the human lung characterizes a gland-associated immune niche. Nat Genet.
478 2023;55:66–77. https://doi.org/10.1038/s41588-022-01243-4.

479 32. Vannan A, Lyu R, Williams AL, Negretti NM, Mee ED, Hirsh J, et al. Spatial transcriptomics identifies
480 molecular niche dysregulation associated with distal lung remodeling in pulmonary fibrosis. Nat Genet.
481 2025;57:647–58. https://doi.org/10.1038/s41588-025-02080-x.

482 33. Zhang Y, Miller JA, Park J, Lelieveldt BP, Long B, Abdelaal T, et al. Reference-based cell type matching
483 of in situ image-based spatial transcriptomics data on primary visual cortex of mouse brain. Sci Rep.
484 2023;13:9567. https://doi.org/10.1038/s41598-023-36638-8.

485 34. Chen W-T, Lu A, Craessaerts K, Pavie B, Sala Frigerio C, Corthout N, et al. Spatial Transcriptomics and
486 In Situ Sequencing to Study Alzheimer's Disease. Cell. 2020;182:976-991.e19.
487 https://doi.org/10.1016/j.cell.2020.06.038.

488 35. Yu M, Risacher SL, Nho KT, Wen Q, Oblak AL, Unverzagt FW, et al. Spatial transcriptomic patterns
489 underlying amyloid-β and tau pathology are associated with cognitive dysfunction in Alzheimer's
490 disease. Cell Reports. 2024;43:113691. https://doi.org/10.1016/j.celrep.2024.113691.

491 36. Navarro JF, Croteau DL, Jurek A, Andrusivova Z, Yang B, Wang Y, et al. Spatial Transcriptomics
492 Reveals Genes Associated with Dysregulated Mitochondrial Functions and Stress Signaling in Alzheimer
493 Disease. iScience. 2020;23:101556. https://doi.org/10.1016/j.isci.2020.101556.

494 37. Zheng P, Zhang N, Ren D, Yu C, Zhao B, Zhang Y. Integrated spatial transcriptome and metabolism
495 study reveals metabolic heterogeneity in human injured brain. Cell Reports Medicine. 2023;4:101057.
496 https://doi.org/10.1016/j.xcrm.2023.101057.

497 38. Ferri-Borgogno S, Zhu Y, Sheng J, Burks JK, Gomez JA, Wong KK, et al. Spatial Transcriptomics Depict
498 Ligand–Receptor Cross-talk Heterogeneity at the Tumor-Stroma Interface in Long-Term Ovarian Cancer
499 Survivors. Cancer Research. 2023;83:1503–16. https://doi.org/10.1158/0008-5472.CAN-22-1821.

500 39. Moreno J, Gluud LL, Galsgaard ED, Hvid H, Mazzoni G, Das V. Identification of ligand and receptor
501 interactions in CKD and MASH through the integration of single cell and spatial transcriptomics. PLoS
502 ONE. 2024;19:e0302853. https://doi.org/10.1371/journal.pone.0302853.

503 40. Rodriques SG, Stickels RR, Goeva A, Martin CA, Murray E, Vanderburg CR, et al. Slide-seq: A scalable
504 technology for measuring genome-wide expression at high spatial resolution. Science. 2019;363:1463–7.
505 https://doi.org/10.1126/science.aaw1219.

506 41. Vickovic S, Eraslan G, Salmén F, Klughammer J, Stenbeck L, Schapiro D, et al. High-definition spatial
507 transcriptomics for in situ tissue profiling. Nat Methods. 2019;16:987–90.
508 https://doi.org/10.1038/s41592-019-0548-y.

509 42. Kleshchevnikov V, Shmatko A, Dann E, Aivazidis A, King HW, Li T, et al. Cell2location maps fine-
510 grained cell types in spatial transcriptomics. Nat Biotechnol. 2022;40:661–71.
511 https://doi.org/10.1038/s41587-021-01139-4.

512 43. Li H, Zhou J, Li Z, Chen S, Liao X, Zhang B, et al. A comprehensive benchmarking with practical
513 guidelines for cellular deconvolution of spatial transcriptomics. Nat Commun. 2023;14:1548.
514 https://doi.org/10.1038/s41467-023-37168-7.

515 44. Navarro JF, Sjöstrand J, Salmén F, Lundeberg J, Ståhl PL. ST Pipeline: an automated pipeline for
516 spatial mapping of unique transcripts. Bioinformatics. 2017;33:2591–3.
517 https://doi.org/10.1093/bioinformatics/btx211.

518 45. Sztanka-Toth TR, Jens M, Karaiskos N, Rajewsky N. Spacemake: processing and analysis of large-scale
519 spatial transcriptomics data. GigaScience. 2022;11:giac064.
520 https://doi.org/10.1093/gigascience/giac064.

521 46. Moriel N, Senel E, Friedman N, Rajewsky N, Karaiskos N, Nitzan M. NovoSpaRc: flexible spatial
522 reconstruction of single-cell gene expression with optimal transport. Nat Protoc. 2021;16:4177–200.
523 https://doi.org/10.1038/s41596-021-00573-7.

524 47. Cisar C, Keener N, Ruffalo M, Paten B. A unified pipeline for FISH spatial transcriptomics. Cell
525 Genomics. 2023;3:100384. https://doi.org/10.1016/j.xgen.2023.100384.

526 48. Kamel M, Song Y, Solbas A, Villordo S, Sarangi A, Senin P, et al. ENACT: End-to-End Analysis of Visium
527 High Definition (HD) Data. Bioinformatics. 2025;41:btaf094.
528 https://doi.org/10.1093/bioinformatics/btaf094.

529 49. Curion F, Rich-Griffin C, Agarwal D, Ouologuem S, Rue-Albrecht K, May L, et al. Panpipes: a pipeline
530 for multiomic single-cell and spatial transcriptomic data analysis. Genome Biol. 2024;25:181.
531 https://doi.org/10.1186/s13059-024-03322-7.

532 50. Hao Y, Stuart T, Kowalski MH, Choudhary S, Hoffman P, Hartman A, et al. Dictionary learning for
533 integrative, multimodal and scalable single-cell analysis. Nat Biotechnol. 2024;42:293–304.
534 https://doi.org/10.1038/s41587-023-01767-y.

535 51. Virshup I, Rybakov S, Theis FJ, Angerer P, Wolf FA. anndata: Annotated data. bioRxiv.
536 2021;:2021.12.16.473007. https://doi.org/10.1101/2021.12.16.473007.

537 52. Hu J, Li X, Coleman K, Schroeder A, Ma N, Irwin DJ, et al. SpaGCN: Integrating gene expression,
538 spatial location and histology to identify spatial domains and spatially variable genes by graph
539 convolutional network. Nat Methods. 2021;18:1342–51. https://doi.org/10.1038/s41592-021-01255-8.

540 53. Zhao E, Stone MR, Ren X, Guenthoer J, Smythe KS, Pulliam T, et al. Spatial transcriptomics at subspot
541 resolution with BayesSpace. Nat Biotechnol. 2021;39:1375–84. https://doi.org/10.1038/s41587-021-
542 00935-2.

543 54. Hao Y, Hao S, Andersen-Nissen E, Mauck WM, Zheng S, Butler A, et al. Integrated analysis of
544 multimodal single-cell data. Cell. 2021;184:3573-3587.e29. https://doi.org/10.1016/j.cell.2021.04.048.

545 55. Biancalani T, Scalia G, Buffoni L, Avasthi R, Lu Z, Sanger A, et al. Deep learning and alignment of
546 spatially resolved single-cell transcriptomes with Tangram. Nat Methods. 2021;18:1352–62.
547 https://doi.org/10.1038/s41592-021-01264-7.

548 56. Shao X, Li C, Yang H, Lu X, Liao J, Qian J, et al. Knowledge-graph-based cell-cell communication
549 inference for spatially resolved transcriptomic data with SpaTalk. Nat Commun. 2022;13:4429.
550 https://doi.org/10.1038/s41467-022-32111-8.

551 57. Castranio EL, Hasel P, Haure-Mirande J, Ramirez Jimenez AV, Hamilton BW, Kim RD, et al. Microglial
552 *INPP5D* limits plaque formation and glial reactivity in the PSAPP mouse model of Alzheimer's disease.
553 Alzheimer's & Dementia. 2023;19:2239–52. https://doi.org/10.1002/alz.12821.

554 58. Yao Z, Van Velthoven CTJ, Nguyen TN, Goldy J, Sedeno-Cortes AE, Baftizadeh F, et al. A taxonomy of
555 transcriptomic cell types across the isocortex and hippocampal formation. Cell. 2021;184:3222-
556 3241.e26. https://doi.org/10.1016/j.cell.2021.04.021.

557 59. Li K, Ouyang Z, Chen Y, Gagnon J, Lin D, Mingueneau M, et al. Cellxgene VIP unleashes full power of
558 interactive visualization and integrative analysis of scRNA-seq, spatial transcriptomics, and multiome
559 data. bioRxiv. 2022;:2020.08.28.270652. https://doi.org/10.1101/2020.08.28.270652.

560 60. Gao B, Zhu J, Negi S, Zhang X, Gyoneva S, Casey F, et al. Quickomics: exploring omics data in an
561 intuitive, interactive and informative manner. Bioinformatics. 2021;37:3670–2.
562 https://doi.org/10.1093/bioinformatics/btab255.

563 61. Gao B, Sun YH, Zhang X, Lin T, Li W, Admanit R, et al. xOmicsShiny: an R Shiny application for cross-
564 omics data analysis and pathway mapping. Bioinformatics Advances. 2024;5:vbaf097.
565 https://doi.org/10.1093/bioadv/vbaf097.

566