StereoCell enables highly accurate single-cell segmentation for spatial transcriptomics

Mei Li[1,2,#], Huanlin Liu[1,#], Min Li[1,#], Shuangsang Fang[1,3,#], Qiang Kang[1], Jiajun Zhang[1,4], Fei Teng[1,4], Dan Wang[5,1], Weixuan Cen[1], Zepeng Li[1], Ning Feng[1], Jing Guo[1], Qiqi He[1], Leying Wang[1], Tiantong Zheng[1], Shengkang Li[1,6], Yinqi Bai[1], Min Xie[4], Yong Bai[1], Sha Liao[1,4], Ao Chen[1,4], Susanne Brix[2,*], Xun Xu[1,7,*], Yong Zhang[1,6,7,*] , Yuxiang Li[1,6,7,*]

[1] BGI-Shenzhen, Shenzhen 518103, China

[2] Department of Biotechnology and Biomedicine, Technical University of Denmark, 2800 Kgs. Lyngby, Denmark

[3] BGI-Beijing, Beijing 102601, China

[4] BGI Research-Southwest, BGI, Chongqing 401329, China

[5] BNU-HKBU United International College, Zhuhai 519087, China

[6] BGI Research-Wuhan, BGI, Wuhan 430074, China

[7] Guangdong Bigdata Engineering Technology Research Center for Life Sciences, Shenzhen 518083, China

[#] These authors contributed equally to this work.

*Correspondence: sbrix@dtu.dk, xuxun@genomics.cn, zhangyong2@genomics.cn, liyuxiang@genomics.cn

Abstract

Owing to recent advances in resolution and field-of-view, spatially resolved sequencing has emerged as a cutting-edge technology that provides a technical foundation for the interpretation of large tissues at the single-cell level. To generate accurate single-cell spatial gene expression profiles from high-resolution spatial omics data and associated images, a powerful one-stop toolbox is required. Here, we present StereoCell, an image-facilitated cell segmentation framework for high-resolution and large field-of-view spatial transcriptomics. StereoCell provides a comprehensive and systematic platform for the generation of high-confidence single-cell spatial data, which includes image stitching, registration, nuclei segmentation, and molecule labeling. During image stitching and molecule labeling, StereoCell delivers better-performing algorithms to reduce stitching error and time, in addition to improving the signal-to-noise ratio of single-cell gene expression data, in comparison with existing methods. Additionally, StereoCell has been shown to obtain highly accurate single-cell spatial data using mouse brain tissue, which facilitated clustering and annotation.

Introduction

Spatially resolved technology generates comprehensive data regarding the distribution of molecules that can be used to identify the location and function of cells within tissues, which helps to broaden our understanding of organ development[1], tumor heterogeneity[2] and cancer evolution[3, 4]. This rapidly developing field is focused on obtaining detailed molecular information at single-cell resolution, as well as spatial information per molecule of tissue in a large field-of-view[5], the latter of which allows the generation of 3D maps representing biological function at the organ level. Technology using Stereo-seq[6] in a large field-of-view can realize whole-organ sequencing and analysis of most tissues in model organisms (e.g., mouse brain (1 cm × 1 cm) and mouse embryo (1 cm × 2 cm)). Single-cell resolution technology, such as Stereo-seq[6], MERFISH[7], and Slide-seq[8], make it possible to explore spatial omics data at the single-cell, or even subcellular level. Solutions that merge the two to obtain accurate spatial single-cell gene expression profiles in large field-of view settings will be an important steppingstone and advance spatial resolved technology by enabling the construction of organ-wide 3D maps of cellular organization and tissue function.

Although spatially resolved technologies vary in resolution and field-of-view, one common feature is their ability to produce traditional high-content tissue images using dyes such as hematoxylin-eosin (H&E) or 4,6-diamidino-2-phenylindole (DAPI) for cellular or nucleus staining. The generated images provide important information regarding tissue and cell morphology but are not currently fully disposable in spatially resolved pipelines; nevertheless, they generate auxiliary information that can be used for subsequent biological data mining. Two main difficulties exist: insufficient precision in image mosaics and inaccuracy of molecule labeling. The first of these difficulties is limited by the small imaging area, mechanical tolerance, and imperfect calibration of microscope technology (e.g., 10 × imaging lens, ~1 mm × 1 mm/tile), which are only sufficient for visual review and not accurate enough for quantitative single-cell analysis. For example, a 1 cm × 1 cm object generates approximately 10 × 10 tile images, while a 3 cm × 3 cm object generates thousands of tiles. Existing image stitching methods, such as ASHLAR[9] and MIST[10], can meet most of the stitching accuracy requirements; however, stitching misalignment is a major issue and it is difficult to collect evaluation datasets to improve the precision and accuracy of image mosaics in a large field-of-view, which is an obvious obstacle in the expansion of the scope of application of these image mosaic methods. The second of these difficulties, which is the key to achieving high-precision single-cell resolution, is classifying the transcripts and other molecules into the owner cell, a procedure termed “molecule labeling”; however, molecule labeling is not an easy task to perform due to the low efficiency of molecule capture and diffusion in wet-lab experiments. Certain techniques, such as pciSeq[11], perform molecule labeling according to nuclear location; however, many transcripts (messenger RNA) are present outside the nucleus, including the cytoplasm and extracellular environment (due to molecule movement and diffusion), which greatly affects the correct annotation of transcripts to the corresponding cell. Accordingly, more recent technology has considered this issue, such as 10X Xenium, which includes more transcripts by extending the distance outside the nucleus (15 μm)[12]. Since the diameter of immune, stromal, and tumor cells varies, it is challenging to obtain accurate single-cell data using such a generalized model that introduces differing levels of noise. Many available methods, including Baysor[13] and Pixel-seq[14], require high-quality data, such as the number of captured molecules and their density distribution, which may not be applicable to most sequencing-based spatial omics data. Furthermore, existing

spatial data analysis frameworks, such as Seurat[15], RCTD[16], Cell2location[17], Squidpy[18], and Spateo[19], only focus on the downstream analysis tasks or one specific aspect of the spatial data analysis, or consider only cell segmentation. A fully functional solution to performing single-cell molecule labeling of spatial omics data from a large field-of-view at high resolution is not yet available.

In the StereoCell technology and software analysis suite presented herein, we combine large field-of-view tissue images and high-resolution gene expression data at the single-cell level to obtain high-confidence spatially resolved single-cell transcriptomics data. StereoCell provides a systematic platform for obtaining accurate single-cell spatial data, which includes image stitching, registration, nuclei segmentation, and molecule labeling. For image stitching, StereoCell employs the best-performing algorithm, MFWS, to reduce stitching errors in large-field-of-view datasets. For molecule labeling, StereoCell includes a novel algorithm that increases the signal-to-noise ratio of the single-cell gene expression matrix in comparison with existing approaches. Here, we demonstrate the performance of StereoCell when applied to a mouse brain dataset, from which it obtains high-confidence spatially resolved single-cell data enabling cell clustering and annotation.

Results

Overview of the StereoCell pipeline

StereoCell provides a complete processing workflow that includes image stitching, registration, tissue segmentation, nuclei segmentation, and molecule labeling, to generate a spatially resolved single-cell gene expression profile with high confidence (Fig. 1A). Firstly, we developed the multiple FFT-weighted stitching (MFWS) algorithm, which is based on frequency domain information and allows the accurate stitching of high-resolution images in a wide field-of-view, in addition to improvements in cell dislocation errors, caused by inaccurate image stitching, and low efficiency caused by the huge number of tiles (Fig. 1B). Secondly, the pipeline enables molecule labeling using the cell morphology and Gaussian mixture model (GMM) algorithm MLCG, which applies a probability-based statistical method to the nuclei segmentation mask for accurate assignment of molecules to the most probable cell, obtaining a gene expression profile at the single-cell level (Fig. 1C). The high-confidence single-cell spatial transcriptomics data obtained by StereoCell was able to achieve more reliable downstream analysis, which is useful for the rapid exploration of spatial omics datasets and the advancement of methods for spatial omics data acquisition at single-cell resolution.

StereoCell provides a high-precision stitching method for morphological tissue images

The first procedure in StereoCell involves accurate stitching of the morphological image. Processing of a morphological image from a tissue slide requires the stitching of an array of multiple image tiles generated by microscopy. Microscopes can automatically capture the tiles of a tissue one

by one and stitch the tiles together using the built-in stitching method; however, overlapping areas between adjacent tiles generated during microscope movement may be imprecise due to mechanical tolerance and imperfect calibration. Such stitching errors are therefore common in mosaic images and need to be removed if the goal is to achieve single-cell resolution in spatial omics applications. Fig. 2A displays a mouse brain mosaic image on a chip with a size of ~2 cm × 2 cm and a tile size of 2424 × 2031 pixels, corresponding to a physical size of ~1.2 mm × 1.0 mm. To visualize seams and stitching errors, we selected two adjacent tiles from the image (Fig. 2B). For accurate stitching results, the lower part of tile 1 (dotted area) and the upper part of tile 2 (dotted area) should be accurately overlaid. Shadows that represent inaccurate overlap of cells can be easily seen in the stitching results produced by the microscope (Fig. 2C, left); however, MFWS is able to stitch the tiles accurately, resulting in the absence of shadows (Fig. 2C, right). The stitching results of two tiles within non-tissue areas (full-colored part) where individual track lines can be clearly seen using auto contrast adjustment are also shown (Fig. 2C). The applied Stereo-seq chip contains periodic track lines with a size of 1500 nm (~3 pixels), which is apparent from the image (Fig. 2C, F). We used a collection of data to evaluate the accuracy and efficiency of available stitching algorithms as compared with MFWS.

We first applied MFWS to a public stitching dataset (Phase Image Tiles[20], bottom of Supplementary Table 1), and the results show that the relative offset errors of MFWS are comparable with those of MIST, both being concentrated within 5 pixels, while ASHLAR has larger offsets (>10-pixels) (Fig. 2D). Moreover, the offset error distribution is much more concentrated by MFWS. Next, we generated a ground-truth dataset based on the track lines in the chip and performed an analysis of nuclei-stained images of mouse brain tissues using different chip sizes from 1 cm × 1 cm to 2 cm × 3 cm, which correspond to tile numbers ranging from 99 to 667 (Supplementary Table 1). The relative offset errors were calculated by comparison with the ground truth. MFWS was shown to perform significantly better than ASHLAR and MIST with respect to the relative offset errors for all image size combinations (Fig. 2E). Further, we fixed a calibration point and calculated the absolute offset errors for each method. An example of a 10-pixel (~5 μm) dislocation, roughly corresponding to half a cell, is shown in Fig. 2F, right. In all the datasets, the errors generated by MFWS remain within a maximum of 10-pixels in the tissue area, while the other methods produce a shift greater than 10-pixels, which may misplace the cell (Fig. 2G). As the number of tiles increases, the run time becomes a significant factor for stitching algorithms. The run time of MFWS is significantly shorter than that of the other methods, mostly due to the embedded spectral calculation based on FFT (Fig. 2H).

StereoCell generates highly accurate single-cell gene expression profiles for spatial omics dataset

Following stitching of the morphological images, the combined morphological tissue image is computationally overlaid onto the tissue gene expression map (Supplementary Fig. 2), and StereoCell identifies single cells within the merged image. StereoCell performs better on nuclei-stained grayscale images and H&E images generated by Stereo-seq than does Deepcell[21] and Cellpose[22] (Supplementary Fig. 3).

We applied StereoCell to the spatial transcriptomics data of the mouse olfactory bulb. After obtaining the nuclei segmentation masks, the molecules located inside the nucleus are assigned to the corresponding cell. Subsequently, StereoCell fits the molecular density distribution to each cell nucleus and relabels the molecules outside the nucleus using the molecule labeling correction method to obtain spatial single-cell gene expression profiles (Fig. 3A). After molecule labeling, the average value of uniquely expressed genes and total gene counts in a single cell increased by approximately 2.34 and 2.56 times, respectively (Fig. 3B). StereoCell provides better overall clustering using the "Leiden" algorithm[23] (Fig. 3C), and StereoCell's silhouette coefficient (an index for cluster evaluation)[24] is 0.05 as compared with 0.02 for the nuclei mask. We then manually annotated the cell types by comparing the differentially expressed genes in each cluster with the marker genes of cell types inferred by a reference dataset[25]. Of note, there are fewer scattered points and a more concentrated distribution in the spatial position of the astrocyte layer (Moran's I: 0.40 vs. 0.30) and dopaminergic neuron layer (Moran's I: 0.52 vs. 0.37) after molecule labeling using StereoCell (Fig. 3C). We further explored the marker gene distribution and expression in the annotated cell types, and all the marker genes provided by the reference dataset[25] had high expression using the StereoCell mask as compared with a nuclei mask (Fig. 3D). Subsequently, we compared subtypes of granule cells discovered using a nuclei mask vs. the StereoCell mask. The nuclei mask was able to identify two subtypes of granule cells: granule cell 0 and granule cell 3, while the StereoCell mask enabled the identification of three granule cell subtypes: granule cell 0, granule cell 1, and granule cell 3. We explored the expression of granule cell 1 subtype-marker genes, Syt1, Scg2, and Cplx1, among all granule cells but there was no obvious difference in expression when using a nuclei mask; however, higher expression was found in granule cell 1 as compared with the other two subtypes when using the StereoCell mask (Fig. 3E). This indicates that StereoCell displays enhanced capture of transcript signals from rare cells, facilitating better annotation and enhancing single-cell spatial resolution. Finally, we estimated and compared the similarity of spatial gene expression between a reference single-cell dataset of the mouse olfactory bulb and our Stereo-seq-generated mouse olfactory bulb dataset using a nuclei mask vs. the StereoCell mask. Tissue cells were auto annotated by Tangram[26] and correlation analysis was performed between the expression matrixes of single cells using each mask and the single-cell sequencing expression profile of the reference[25]. The correlation coefficient using annotation based on the StereoCell mask is higher (0.4% to 8.9%) than that based on the nuclei mask. Overall, the cell matrix generated by StereoCell shows a higher correlation with the single-cell reference than that generated by a nuclei mask (Fig. 3F), indicating that StereoCell provides a spatial single cell-level gene expression profile closer to that generated by single-cell sequencing. Taken together, the results suggest that StereoCell provides a more accurate spatial single-cell transcriptomics expression matrix.

StereoCell dissects the structural composition of mouse brain cortex data at single-cell resolution

For identification of single cells in tissues, approaches that directly divide adjacent tissue regions into bins are both convenient and fast, and are a common way to perform spatial omics data analysis. Accordingly, we compared the ability of StereoCell with an approach based on different

sized bins (BinX means X × X DNBs, DNA nanoballs)[6] to reconstruct known cellular regions in the mouse brain. Firstly, we compared the clustering of data generated by StereoCell vs. Bin100, Bin50, and Bin20 (Fig. 4A). Then, we evaluated the clustering results of several methods using the silhouette coefficient (Fig. 4B), which provides information on how similar a given cell or bin is to other similar cells/bins (cohesion) in comparison with non-similar cells/bins (separation). StereoCell obtained the highest silhouette coefficient, followed by Bin20, Bin100, and Bin50. We calculated the total expression (total gene counts) and number of uniquely expressed genes identified using StereoCell vs. the different bin sizes (Fig. 4C).

Bin100 divided the tissue image into 20 clusters, and some important areas in the tissue (such as the brain hub and hippocampus) could not be clearly identified. In comparison with Bin100, Bin50 yielded 28 clusters and identified the important areas of the tissue more clearly; however, the tissue cortex and blood cells were poorly identified. Bin20 classified the cells into 30 clusters, while StereoCell generated 25 clusters. Both these approaches were able to identify important areas in the tissue, division of the cortex and blood cells was obvious, and differences in clustering were noticeable. The clustering results of Bin20 and StereoCell highly reflect the real situation, both being close to single-cell accuracy (the size of the cells in the mouse brain is approximately 10 μm, which corresponds to Bin20, i.e., 20 nm × 500 nm)[6]. The visualization results of cell boundaries obtained by the two methods, however, show that Bin20 more often divided a single cell into two or more cells, while StereoCell more accurately divided the cell area and is consistent with the actual cell distribution in the tissue (Fig. 4D). Moreover, Bin20 resulted in the slicing of approximately 90% of the cells, with only ~10% of the cell nuclei being completely covered in the nuclei-stained image, while only ~2% of the cells were sliced using StereoCell and most of the nuclei are included (Fig. 4E). The results of Bin20 and StereoCell were then annotated using Spatial-ID[27] (Fig. 4F, top) with the Adolescent mouse brain as a reference[28]. Bin20 annotated 29 different cell types, while StereoCell was able to annotate 37 different cell types. In the ACTE series, both Bin20 and StereoCell annotated 2 subtypes. In the MEINH series, Bin20 annotated 2 isoforms, while StereoCell annotated 3 isoforms. In the TEGLU series, Bin20 annotated 10 subtypes, while StereoCell annotated 13 subtypes. In the TEINH series, Bin20 annotated 5 subtypes and StereoCell annotated 4 subtypes, which is the only case where StereoCell annotated fewer subtypes. In addition, in comparison with Bin20, StereoCell annotated 2 subtypes of the ACNT series and 4 subtypes of the TEINH series. Overall, the annotation results of StereoCell are more accurate than those of Bin20. When zooming in on the cortical area (Fig.4F, bottom), the staining position of StereoCell and the nucleus is better aligned, and the cell state and tissue structure are more in line with the actual situation. We found that in the cortical area of the mouse brain, StereoCell can annotate each layer more accurately than Bin20, including the gear gyrus and cortical regions, in comparison with the Allen mouse brain dataset (Fig. 4G). The annotation of cells using StereoCell resulted in consistency with respect to the position of the cells in the Allen mouse brain dataset. Although Bin20 was also capable of annotating the gear gyrus and cortex to align with the location in the Allen mouse brain dataset, the division between layers was not as accurate. Expression of marker genes in the gear gyrus (DGGRC2) and cortex (TEGLU3, TEGLU4, TEGLU7, and TEGLU8) was also fully captured using the StereoCell algorithm, which was not the case when using the Bin20 approach (Fig. 4H).

Overall, StereoCell provides single-cell spatial transcriptomics maps with a higher signal-to-noise ratio in comparison with BinX. Moreover, the single-cell maps generated using StereoCell

align well with cellular annotation within different brain regions and can provide a valuable reference for studies of cellular interaction networks in health and disease.

Discussion

There is still a long way to go before the potential of spatially resolved technology can be fully utilized, and continued improvement of both protocols and analytical methods for processing spatial omics data is required. Here, we present StereoCell, an image-facilitated cell segmentation framework for high-resolution and large-field-of-view spatial omics that provides a complete and systematic platform for obtaining high-confidence single-cell spatial data. The implemented morphological tissue image stitching procedure proved to be accurate and reliable for single-cell identification and was flexible and convenient in terms of run time. The high-precision stitching method based on FFT and a benchmark dataset based on chip track lines were found to be useful for the correction of stitching to almost subcellular precision. Additionally, we implemented a new molecule labeling method based on GMM and nuclei segmentation to generate spatial single-cell gene expression data with a higher signal-to-noise ratio, which yielded more reliable analysis of cell clustering and end-to-end annotation. Moreover, we optimized the processing modules used for image registration to include track lines on the chip, nuclei segmentation, and tissue cutting. StereoCell is equipped with an extensive testing suite, implemented in a robust continuous integration pipeline that contains a graphic user interface for manual image stitching, registration, and labeling. StereoCell's rich documentation in the form of a functional application programming interface, examples, and tutorial workflows, is easy to navigate and accessible to both experienced developers and beginner analysts.

In the present study, we demonstrate that auxiliary staining can be extremely valuable to resolve difficult cases. The molecule labeling or cell segmentation results will likely be improved using better staining, such as labeling of outer membranes, and better techniques for segmenting such images. Additionally, better photo capturing using enhanced AI recognition models will advance our insight into biological traits and speed up the identification of features that enable us to distinguish between healthy and diseased tissue. It is noteworthy that the entire StereoCell framework can be applied to generate single-cell gene expression for Stereo-seq, and most of its modules are also suitable for other spatial transcriptomics technology.

The goal here was to develop a general tool for combined data generation and single-cell level analysis of spatial omics data. The data presented herein are based on spatial transcriptomics alone. In future studies, we will combine spatial information with additional biological characteristics, such as proteomics, to integrate multi-omics information at the single-cell spatial level, improve the signal-to-noise ratio, and develop interfaces and extensions that support user-defined data and methods. StereoCell has the potential to serve as a bridge between the fields of image analysis and molecular omics and provides a foundation for the development of next generation computational methods for spatially resolved technology.

References

1. Chen, A. et al. Spatiotemporal transcriptomic atlas of mouse organogenesis using DNA

nanoball-patterned arrays. Cell **185**(10), 1777-1792 (2022).
2. Lu, T. et al. Spatially resolved epigenomic profiling of single cells in complex tissues. Cell **185**(23), 4448-4464 (2022).
3. Erickson, A. et al. Spatially resolved clonal copy number alterations in benign and malignant tissue. Nature **608**(7922), 360-367 (2022).
4. Grünwald, B. T. et al. Spatially confined sub-tumor microenvironments in pancreatic cancer. Cell **184**(22), 5577-5592 (2021).
5. Cui Zhou, D. et al. Spatially restricted drivers and transitional cell populations cooperate with the microenvironment in untreated and chemo-resistant pancreatic cancer. Nat. Genet. **54**(9), 1390-1405 (2022).
6. Fang, S. et al. Computational Approaches and Challenges in Spatial Transcriptomics. Genom. Proteom. Bioinf. (2022).
7. Chen, K. H. et al. Spatially resolved, highly multiplexed RNA profiling in single cells. Science **348**(6233), aaa6090 (2015).
8. Rodriques, S. G. et al. Slide-seq: a scalable technology for measuring genome-wide expression at high spatial resolution. Science **363**(6434), 1463-1467 (2019).
9. Muhlich, J. L. et al. Stitching and registering highly multiplexed whole-slide images of tissues and tumors using ASHLAR. Bioinformatics **38**(19), 4613-4621 (2022).
10. Chalfoun, J. et al. MIST: Accurate and Scalable Microscopy Image Stitching Tool with Stage Modeling and Error Minimization. Sci. Rep. **7**(1), 4988 (2017).
11. Qian, X. et al. Probabilistic cell typing enables fine mapping of closely related cell types in situ. Nat. Methods **17**(1), 101-106 (2020).
12. Janesick, A. et al. High resolution mapping of the breast cancer tumor microenvironment using integrated single cell, spatial and in situ analysis of FFPE tissue. bioRxiv (2022).
13. Petukhov, V. et al. Cell segmentation in imaging-based spatial transcriptomics. Nat. Biotechnol. **40**(3), 345-354 (2022).
14. Fu, X. et al. Polony gels enable amplifiable DNA stamping and spatial transcriptomics of chronic pain. Cell **185**(24), 4621-4633 e17 (2022).
15. Hao, Y. et al. Integrated analysis of multimodal single-cell data. Cell **184**(13), 3573-3587 (2021).
16. Cable, D. M. et al. Robust decomposition of cell type mixtures in spatial transcriptomics. Nat. Biotechnol. **40**(4), 517-526 (2022).
17. Kleshchevnikov, V. et al. Cell2location maps fine-grained cell types in spatial transcriptomics. Nat. Biotechnol. **40**(5), 661-671 (2022).
18. Palla, G. et al. Squidpy: a scalable framework for spatial omics analysis. Nat. Methods **19**(2), 171-178 (2022).
19. Qiu, X. et al. Spateo: multidimensional spatiotemporal modeling of single-cell spatial transcriptomics. BioRxiv (2022).
20. Phase_Image_Tiles. https://isg.nist.gov/BII_2015/Stitching/Phase_Image_Tiles.zip
21. Greenwald, N. F. et al. Whole-cell segmentation of tissue images with human-level performance using large-scale data annotation and deep learning. Nat. Biotechnol **40**(4), 555-565 (2022).
22. Stringer, C. et al. Cellpose: a generalist algorithm for cellular segmentation. Nat. Methods **18**(1), 100-106 (2021).
23. Traag, V. A. et al. From Louvain to Leiden: guaranteeing well-connected communities. Sci. Rep. **9**(5233 (2019).

24. Rousseeuw, P. J. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. J. Comput. Appl. Math. **20**(53-65 (1987).
25. Brann, D. H. et al. Non-neuronal expression of SARS-CoV-2 entry genes in the olfactory system suggests mechanisms underlying COVID-19-associated anosmia. Sci. Adv. **6**(31), eabc5801 (2020).
26. Biancalani, T. et al. Deep learning and alignment of spatially resolved single-cell transcriptomes with Tangram. Nat. Methods **18**(11), 1352-1362 (2021).
27. Shen, R. et al. Spatial-ID: a cell typing method for spatially resolved transcriptomics via transfer learning and spatial embedding. Nat. Commun. **13**(1), 7640 (2022).
28. Zeisel, A. et al. Molecular Architecture of the Mouse Nervous System. Cell **174**(4), 999-1014 e22 (2018).

Methods

StereoCell relies on common Python packages, such as libvips, pillow, numpy, and PyQt, for image processing, data analysis, and the image interface. For key modules of StereoCell, we provide self-developed algorithms and tools for evaluation and data adjustment. We also introduce some validated methods in our modules; for example, the watershed and DeepCell methods are integrated into the cell segmentation module. StereoCell provides a set of visual solutions, and its GUI was developed using PyQt. Detailed instructions and operational documentation, including tutorials and sample workflows, are also provided.

Image stitching

We developed an image stitching algorithm named Multiple FFT-weighted Stitching (MFWS). The algorithm procedures are as follows:

1. Calculate the overlapping matrixes of adjacent tiles in the horizontal and vertical direction s. We reconstructed the input data to improve the robustness and accuracy of the algorithm. For each tile, we calculate its actual overlap value with its neighbors in the horizontal and vertical directions. The calculation method applied in both directions is the same; a pair of tiles in the vertical direction is taken as an example for further introduction.

i) Multiple divisions of the overlapping regions of adjacent tiles: the overlapping regions of adjacent tiles are marked as A_f and A_m, respectively. We divided A_f and A_m by N-1 to obtain N subregion images, marked $\{f_i\}$ and $\{m_i\}$.

ii) Calculation of the subregion spectra: we applied the fast Fourier transform algorithm to obtain $2 \times N$ frequency spectra based on $\{f_i\}$ and $\{m_i\}$.

iii) Calculation of the cross-power spectra of the paired subregions: the subregion images in A_f and A_m are paired, such as f_i and m_i. N cross-power spectra can be obtained using the formula for image cross-power calculation. Theoretically, any one of these spectra can reflect the true offset information.

iv) Weighted superposition of the mutual power spectrum: due to the low signal-to-noise ratio of some subregions in the previous step, the calculated value differs greatly from the actual value. The overall accuracy of the algorithm can be significantly improved by weighted enhancement of

partial blocks and reduction of residual subblocks. Our weighting coefficient is based on the peak value of the mutual power spectrum because experiments show that the greater the peak value, the higher the precision and reliability of the overlap value. A unique cross-power spectrum is obtained after the weighted superposition of N spectra. Then, the corresponding spatial domain correlation graph can be obtained through Inverse Fast Fourier Transform (IFFT). The coordinate of the peak value represents the required offset value.

v) Generation of the offset matrixes: carrying out the above steps for each pair of adjacent tiles generates the offset matrices O_H and O_V.

2. Obtain the global coordinates of each tile. Transformation of local coordinates to stitching coordinates is required, since image stitching demands the unique coordinates of each tile in the reference system, i.e., the global coordinate matrix L. In the O_H and O_V matrixes, offsets corresponding to low-confidence values are not credible and need to be eliminated; however, this leads to the existence of single or multiple connected domains. Therefore, when obtaining the coordinates, we first find the center of each connected domain, complete the splicing in the connected domain according to the relative relationship among the neighbors, and then use the experience value to fill multiple connected domains to complete the splicing of the entire image.

3. Complete the image mosaic and fusion generation. According to the tiles and the coordinate matrix L, the size of the mosaic image is obtained, the value is traversed, and the seam fusion is completed synchronously during the stitching process.

Image registration

The gene expression matrix was transformed into an image (map), in which the intensity value of each pixel was proportional to the total count of molecules of all genes expressed at position DNB (DNA nanoball, x, y). To spatially assign the expressed genes to the corresponding cells, the nuclei-stained microscopic image was registered to the transformed gene count image. Sequencing chips with sample tissue attached, imaging, and sequencing were designed with periodic track lines (horizontal and vertical) to assist base calling and image registration. The track lines were displayed as dark straight lines in both the nuclei-stained and gene count images. Image registration was achieved using the gene count image as a template and performing an affine transformation on the nuclei-stained image. In particular, the track lines were detected in both images by line searching algorithms, and the intersections of the horizontal and vertical track lines were located. The scaling parameter was calculated by comparing the length of the line segments between intersections within the two images. The rotation angle was defined as the difference between the horizontal track lines and the horizontal direction of the image coordinate system. For the offset between them, since the track line patterns were periodic, we first calculated the center of gravity of the tissue regions within the two images, roughly matched them together, and then matched the track line intersections to fine-tune the image registration. Details are shown in Supplementary Fig. 1B.

Tissue segmentation

The tissue segmentation process consists of two steps: preprocessing and model inference. In addition, a self-scoring mechanism was added to optimize the processing time and eliminate the effort of manually checking the results.

During the preprocessing step, the histogram enhancement method was used to improve the separation of tissue and background. Subsequently, during the model inference step, Bi-Directional ConvLSTM U-Net[29] based on U-Net[30] was used to predict tissue masks. In comparison with the original U-Net, the following improvements are incorporated: (i) BCDU-Net replaces the convolution sequence with a densely connected convolution in the final step of the encoding path to encourage feature reuse; (ii) BConvLSTM is used to combine the feature maps extracted from the corresponding encoding path with the previous decoding up-convolutional layer in a non-linear manner; and (iii) the batch normalization is added to accelerate the convergence of the network.

When dealing with a large amount of unstable data, the model could not achieve ideal results with some complex or poor-quality tissue images, and the self-evaluation mechanism helped to filter these potentially unsatisfactory results. When the nuclear image possessed a high gray value, we used the threshold value obtained in the histogram enhancement step as a reference value and assumed that the pixel points higher than this value were valid pixels. The effective pixel densities of the tissue inside the mask and in the area surrounding the mask were calculated in sections, and the ratio of the two values was used as the basis for evaluation. The process of tissue segmentation using the algorithm and specific segmentation examples are shown in Supplementary Fig. 2.

Molecule labeling

We developed molecule labeling based on cell morphology and the GMM(MLCG) algorithm, which has two main steps: (i) cell morphological segmentation (nuclei-stained images) and (ii) molecule labeling correction. In comparison with other methods, MLCG combined the cell morphological information and the molecule matrix to obtain a more accurate spatial single-cell expression profile.

Nuclei segmentation

Our cell segmentation pipeline consists of three major sequential steps (Supplementary Fig. 3A): (i) image preprocessing to transform the image with noise reduction and ensure that the cells are relatively obvious against the background; (ii) segmentation model inference using the proposed model to segment on a grayscale image; and (iii) mask postprocessing to separate large areas of adhesion.

Preprocessing method

Prior to full cell segmentation, median filtering was employed to smooth the noise that may be present in the input image. To both enhance the cell and homogenize the image background, the output of the median filtering operation was input into the pixel-wise subtraction process.

Cell segmentation

A deep neural network based on U-Net[30] with a suitable backbone modified from resnet was

used to segment the cells. The thresholds were applied to the output inferred by the model to create the mask. The details of the model architecture are given in the next section and shown in Supplementary Fig. 3B.

Model architecture

Deep Residual U-Net (ResUnet)[31] underpinned the psaUnet architecture, which combined Deep Residual net[32], U-Net[30], and EPSANet[33]. The proposed psaUnet contained 5 encoder blocks and 5 decoder blocks. Based on U-Net, we used resnet as a feature-extraction method. The first block in the encoder was a combination of 7 × 7 convolution, batch normalization, Rectified Linear Unit (ReLU) activation, maxpool, and identity mapping. The rest of the block was made up of 2–5 psablocks, each of which was a residual block embedded by the Pyramid Squeeze Attention (PSA) module[33]. The PSA module was employed as an attention unit to obtain detailed information from the feature maps.

PSA module

As illustrated by the PSA module in Supplementary Fig. 3B, the main operation of PSA implementation contained four steps. Firstly, the input tensors were sliced into a multi-scale feature map channel-wise using the Squeeze and Contact (SPC) module. Secondly, the SEWeight module was applied to extract the channel-wise attention vector from the different scale feature maps. Thirdly, the weighted channel-wise attention vector was generated by applying the "Softmax" activation function to the channel-wise attention vector. Fourthly, a refined feature map was obtained by taking the element-wise product of a multi-scale feature map with the corresponding weighted channel-wise attention vector as the output. In the proposed PSAUnet, we used the PSA module instead of the 3 × 3 convolution in the bottleneck, which was the basic feature-extracting unit of resnet[31] to ensure that the model pays more attention to the highlighted feature representation. The module was defined as follows:

$$Output = Softmax\left(se\left(SPC(X)\right)\right) \otimes SPC(X)$$

Where, X is the input tensor, SPC and se represent the SPC and SEWeight module, respectively, $Softmax$ is the "softmax" function, and $\otimes$ is the operation of the element-wise product.

Psablock

A deeper network can bring a higher accuracy; however, as the network becomes deeper, gradient vanishes and renders the network difficult to train. To solve the problem of gradient vanishing, resnet was introduced. With residual units and skipping of connections in the dropout layer, the training of a deep network becomes faster and the feature learning ability becomes stable. For this reason, we use the architecture of resnet as the backbone of the encoder in our proposed model.

Postprocessing method

To obtain a more reliable segmentation, the mask generated by the segmentation model is input to the area filtering operation. Further, the opening operation (erode and dilatate) is applied to revise the cells shape and boundaries. In addition, the optional watershed approach is provided to adjust the size and boundaries of previously revised segmented cells.

To better identify different forms and types of cells, we also provided other cell segmentation methods in addition to the self-developed model, such as the traditional image segmentation method (watershed algorithm) and the industry's widely acclaimed deep learning models (DeepCell[21], Cellpose[22], and Cellpose 2.0[34]). Some segmentation examples are shown in Supplementary Fig. 3D.

Nuclei mask correction

To obtain spatially resolved high-precision single-cell omics data, we need to be able to accurately label transcripts to cells. Firstly, the morphological tissue image is used to identify nuclear boundaries as described under nuclei segmentation. The specific transcripts that are contained within each nucleus are then assigned. A GMM[35] algorithm is used to estimate the probability of each transcript belonging to a given cell based on the initial nuclei segmentation map. This is performed by modeling each cell as a GMM distribution, combining its spatial position, transcript count, and density. Simultaneously, we also estimate the probability scores of transcripts in the acellular region. In our model, the probability of assignment of a transcript to a given cell declines progressively with distance from the cell center, but increases with transcript count and density.

A mixture model is a probabilistic model that can be used to represent a population distribution with K sub-distributions; in other words, it is a mixture of *K* sub-distributions. GMM can be regarded as a model composed of *K* single Gaussian models. To formalize this mathematically, we denote:

$$P(x|\theta) = \sum_{k=1}^{K} \alpha_k \varphi(x|\theta_k),\ \ \theta = (\alpha_k, \theta_k),$$

where,

$$\varphi(x|\theta_k) = \frac{1}{(2\pi)^{\frac{D}{2}} |\Sigma_k|^{\frac{1}{2}}} \exp\left(-\frac{(x-\mu_k)^T {\Sigma_k}^{-1} (x-\mu_k)}{2}\right),\ \ \theta_k = (\mu_k, \Sigma_k).$$

$$\alpha_k \geq 0,\ \ \sum_{k=1}^{K} \alpha_k = 1.$$

We focused on a single cell as an example to describe the entire correction process. Based on the results of nuclei segmentation, we could locate the spatial position of the nucleus. Firstly, GMM was used to fit the molecular distribution of the current cells. Here, we expanded the cell boundary and captured as much as possible of the true distribution of data in cells rather than nuclei. According to the empirical value of the cell area, we finally determined the fitting range of 100 pixels × 100 pixels. Subsequently, GMM was used to calculate the probability score of extracellular molecules within the fitting range. Finally, using the maximum and minimum probability scores within the fitting range, the quartile was used as the adaptive threshold of the current cell. When the probability score was greater than the threshold, the molecule was re-divided into cells. In this way, we could obtain all the molecules that should belong to cells with high confidence.

Mouse olfactory bulb data analysis

Spatial single-cell gene expression matrixes of mouse olfactory bulb data were generated by StereoCell, and the stitched image and gene expression map were exported. The nuclei mask and StereoCell mask images were drawn using the Python (3.8) script. The matrixes were input into Stereopy (v6.0)[36] to calculate the total gene counts and number of uniquely expressed genes through the quality control function. During the filtering process, the cells with fewer than 150 expressed genes and more than 5% mitochondrial genes were removed, and genes present in less than 3 cells were also removed. The gene expression matrixes were then normalized using the “SCTransform”

function. The differentially expressed genes were summarized by Principal Component Analysis (PCA) to reduce the data dimensionality. With these settings, we ran the uniform manifold approximation and projection (UMAP) algorithm to obtain 2D data projections, followed by "Leiden" clustering to identify all clusters within the dataset. The silhouette coefficient[24] was calculated using the "metrics.silhouette_samples" and "metrics.silhouette_score" functions in the "sklearn" package. The related formula is as follows:

$$S(i) = \frac{b(i) - a(i)}{\max\{a(i), b(i)\}}$$

Where, $S(i)$ is the silhouette coefficient, $a(i)$ indicates the average distance between the i-th sample and other samples in its cluster, and $b(i)$ is the average distance between the i-th sample and the samples in other clusters. The moran's I was calculated using the "metrics.morans_i" function in the "Scanpy" package[37]. The related formula is as follows:

$$I = \frac{n}{S_0} \frac{\sum_{i=1}^{n} \sum_{j=1}^{n} \omega_{i,j} z_i z_j)}{\sum_{i=1}^{n} z_i^2}$$

Where, I is the moran's I, z_i is the deviation of the attribute of factor i from its mean value, $w_{i,j}$ is the spatial weight between factors i and j, n is equal to the factor integration, and S_0 is the aggregation of all spatial weights. The marker genes of each cell type were obtained from a reference[25], and the dot plot of expression was drawn. Tangram[26] was used to annotate the cell types according to the reference[25]. We combined the single-cell expression matrix of each annotation, which was used to calculate the Spearman correlation coefficient.

Mouse brain data analysis

Spatial single-cell gene expression matrixes of mouse brain data were generated by StereoCell and subsequently analyzed using Stereopy (v0.6.0)[36]. The cells with fewer than 200 expressed genes and more than 5% mitochondrial genes were removed, and the genes present in less than 3 cells were also removed. After filtering, the gene expression matrixes were normalized using the "SCTransform" function. To reduce the data dimensionality, we summarized the differentially expressed genes using PCA and then ran the UMAP algorithm to obtain 2D data projections, followed by "Leiden" clustering to identify all clusters within the dataset. The silhouette coefficient[24] was calculated using the "metrics.silhouette_samples" and "metrics.silhouette_score" functions in the "sklearn" package. The quality control function of Stereopy was used to compute the total genes and number of uniquely expressed genes. We exported the gene map, nuclei-stained image, outlines of molecule labeling, and bin20 to visually demonstrate the segmentation effect of different methods. Spatial-ID[27], with the Adolescent mouse brain[28] dataset as a reference, was used to annotate cell types. The marker genes of each cell type were obtained from the "L5_All.agg.loom" file, which was downloaded from the Adolescent mouse brain website[28], and the matrix plot of expression was drawn.

References

29. Azad, R. et al. Bi-Directional ConvLSTM U-Net with Densley Connected Convolutions. 406-415 (2019).

30. Ronneberger, O. et al. U-Net: Convolutional Networks for Biomedical Image Segmentation. 234-241 (2015).
31. Zhang, Z. et al. Road Extraction by Deep Residual U-Net. IEEE Geosci. Remote S. **15**(5), 749-753 (2018).
32. He, K. et al. Deep Residual Learning for Image Recognition. 770-778 (2016).
33. Zhang, H. et al. EPSANet: An Efficient Pyramid Squeeze Attention Block on Convolutional Neural Network. arXiv (2021).
34. Pachitariu, M. and Stringer, C. Cellpose 2.0: how to train your own model. Nat. Methods **19**(12), 1634-1641 (2022).
35. Douglas, R. Gaussian Mixture Models. Encyclopedia of Biometrics, Spring, (2009).
36. Stereopy. https://stereopy.readthedocs.io/en/latest/index.html
37. Wolf, F. A. et al. SCANPY: large-scale single-cell gene expression data analysis. Genome Biol. **19**(1), 15 (2018).

Data and code availability

StereoCell was used to perform the workflow analysis in this paper. Data, code and graphical user interface (GUI) software are available at the following GitHub repository with the detailed manual documentation, including tutorials and example workflows: https://github.com/BGIResearch/StereoCell.

Acknowledgements

This work is part of the "SpatioTemporal Omics Consortium" (STOC) paper package. A list of STOC members is available at http://sto-consortium.org. We would like to thank the MOTIC China Group, in addition to Bichao Chen, Canqiang Xu, Xiangxiang Liao, Kui Su, Zhonghan Deng, Yiwen Wu, Bohan Zhang, Chao Liu, Mengnan Cheng, Qing Zhou, Marija Ivanovic, Vladimir Kovacevic, Lei Han, Xing Guo, and Wei Peng for their help. This work was supported by Wuhan Supercomputing Center for the data training.

Author contributions

Conceptualization: Mei Li, Huanlin Liu, Min Li, and Shuangsang Fang.
Project administration and supervision: Susanne Brix, Xun Xu, Yong Zhang, and Yuxiang Li.
Algorithm development and implementation: Huanlin Liu and Min Li.
Data collection, processing, and application: Mei Li, Huanlin Liu, Min Li, Jiajun Zhang, Qiang Kang, Fei Teng, Weixuan Cen, Zepeng Li, Ning Feng, Jing Guo, Qiqi He, and Leying Wang.
Method comparisons: Huanlin Liu, Min Li, and Qiang Kang.
Biological interpretation: Jiajun Zhang, Shuangsang Fang, Ying Zhang, Qiang Kang, and Fei Teng.
Data visualization: Tiantong Zhen and Shengkang Li.
Project coordination: Mei Li, Shuangsang Fang, Sha Liao, and Ao Chen.
Manuscript writing and figure generation: Mei Li, Huanlin Liu, Min Li, Jiajun Zhang, Shuangsang

Fang, Qiang Kang, Fei Teng, and Dan Wang.
Manuscript review: Susanne Brix, Yinqi Bai, Min Xie, and Yong Bai.

Funding

National Key R&D Program of China (2022YFC3400400)

Figure legends

Fig. 1 | Overview of StereoCell: an integrated technology and software analysis suite for spatial single-cell expression profiling.

A. Schematic overview of the StereoCell pipeline. The spatial gene expression matrix and the morphological (nuclei-stained) image of the tissue are obtained using the developed spatial transcriptomics technology Stereo-seq and microscopy, respectively. The tiles obtained by microscopy are stitched together to generate a large mosaic image of the whole tissue to be merged with the gene expression matrix. Subsequently, tissue and nuclei segmentation are performed simultaneously on the overlaid image to obtain the tissue area and cell nucleus mask. Finally, molecule assignment is applied to obtain spatial single-cell gene expression profiles.

B. Multiple FFT-weighted stitching (MFWS). Overlapping regions between adjacent tiles are cut, and spectral information is calculated based on the Fast Fourier Transform (FFT) algorithm. The spectral information from the cut pairs is then weighted to obtain the cross-power spectrum from which the maximum spectral position is determined as the offset of the adjacent tiles. Subsequently, tiles are stitched using the offset value. See Supplementary Fig. 1A for details.

C. Flowchart of molecule labeling based on cell morphology and the Gaussian mixture model (GMM) algorithm MLCG. Molecules within the nuclear boundary are assigned to a given cell, while molecules outside the nucleus are labeled to the cell that gains the largest probability from GMM fitting.

Fig. 2 | StereoCell includes a high-precision mosaic method for image stitching (MFWS), enabling large-field morphological image analysis at single-cell spatial accuracy.

A. Mosaic image of mouse brain tissue; chip size 2 cm × 2 cm.

B. The red box indicates the edge of the tissue (containing cells and non-tissue); two neighboring tiles (in the vertical direction) are shown in different colors. The part that needs to be overlaid by stitching is the lower part of tile 1 (within the dotted area) and the upper part of tile 2 (within the dotted area).

C. The stitching results for the two tiles (yellow box in B) using the stitching method built in to the microscope (left) and MFWS (right). In each sub-figure, the left part shows the overlap of cells after stitching and the right part shows the track lines in the non-tissue area. The contrast of the background was increased to clearly show the track lines.

D. Comparison of the relative errors produced by MIST, ASHLAR, and MFWS using the public dataset “Phase_Image_Tiles.”

E. Comparison of the relative errors obtained using MIST, ASHLAR, and MFWS on mouse brain datasets sequenced by Stereo-seq, where different chip sizes from 1 cm × 1 cm to 2 cm × 3 cm were applied, which corresponds to a number of tiles from 11 × 9 to 23 × 29. The evaluation is based on the ground truth calculated using the track lines on the Stereo-seq chip.
F. An example of the stitching error in pixels in the zoomed-in image; the 10-pixels roughly correspond to half a cell.
G. Bar graph illustrating the maximal accumulation of stitching errors using MIST, ASHLAR, and MFWS on the datasets from **E**.
H. Line graph displaying the run time of MIST, ASHLAR, and MFWS on the datasets.

Fig. 3 | The StereoCell mask provides single-cell spatial data with a higher signal-to-noise ratio that facilitates finer cell clustering and annotation of the mouse olfactory bulb.
A. Flowchart of MLCG processing of mouse olfactory bulb data.
B. Comparison of the number of uniquely expressed genes and the total gene count per cell using a nuclei mask vs. the StereoCell mask on mouse olfactory bulb data. Top: density plot of the total gene counts per cell. Bottom: density plot of the number of uniquely expressed genes per cell.
C. Comparison of the clustering results generated using the "Leiden" algorithm on a nuclei mask vs. the StereoCell mask. Top row, left: spatial clustering results of mouse olfactory bulb data generated using a nuclei mask vs. the StereoCell mask; right: comparison of unique cells estimated by the Silhouette coefficient. Middle and bottom rows, left: spatial distribution of cells in the inferred clusters of astrocytes and dopaminergic neurons generated using a nuclei mask vs. the StereoCell mask; right: comparison of the spatial autocorrelation estimated by Moran's I. AONM/T cell, anterior olfactory nucleus mitral/tufted cell; VIPP neuron, vasoactive intestinal peptide positive neuron.
D. Comparison of marker gene expression in the main cellular clusters inferred by the "Leiden" algorithm generated using a nuclei mask vs. the StereoCell mask. AONM/T cell, anterior olfactory nucleus mitral/tufted cell; VIPP neuron, vasoactive intestinal peptide positive neuron.
E. Comparison of cellular subtype identification using a nuclei mask vs. the StereoCell mask. First column: subtypes of granule cells identified using a nuclei mask vs. the StereoCell mask. Other columns: respective gene expression heatmaps of Syt1, Scg2, and Cplx1, which have been reported to be marker genes for granule cell 1, using a nuclei mask vs. the StereoCell mask.
F. Spearman correlation between gene expression in similar cells in tissue slides of the mouse olfactory bulb and a single-cell reference of the mouse olfactory bulb using a nuclei mask vs. the StereoCell mask. AONM/T cell, anterior olfactory nucleus mitral/tufted cell; VIPP neuron, vasoactive intestinal peptide positive neuron.

Fig. 4 | Clustering and annotation of mouse brain tissue by StereoCell vs. different BinX approaches.
A. Identification of gene expression clusters in mouse brain tissue using different bin sizes vs. StereoCell.
B. A silhouette coefficient-based evaluation of clustering performance.
C. Violin plots displaying the total gene counts and number of uniquely expressed genes per cell captured using different bin sizes vs. StereoCell.
D. Comparison of the cell boundary generated using Bin20 and StereoCell. The first image shows

the mouse brain gene expression matrix map and the nuclei-stained image, which were merged in green and blue, respectively. The second image displays an enlarged map of the local area from the first image. The third image is a comparison of the cell boundary obtained using Bin20 vs. StereoCell.

E. Comparison of the proportion of cell nuclei covered or intersected using Bin20 vs. StereoCell.

F. Comparison of end-to-end annotation using Bin20 vs. StereoCell. Colors represent the same cell type in the Bin20 and StereoCell annotation charts. The enlarged local comparison map of the cortex shows the spatial distribution and boundary information of different cortical cells (bottom sub-figure).

G. Comparison of cell annotation based on Bin20 vs. StereoCell using the Allen mouse brain atlas as a reference. The spatial distribution of the gear gyrus (DGGRC2, black) and the cortex (TEGLU3, TEGLU4, TEGLU7, and TEGLU8) is shown.

H. Heatmap displaying the consistency of marker gene expression within the different substructures in G (gear gyrus (DGGRC2) and cortex (TEGLU3, TEGLU4, TEGLU7, and TEGLU8) layers)). The upper part, Bin20 results; lower part, StereoCell results.

Fig. 1

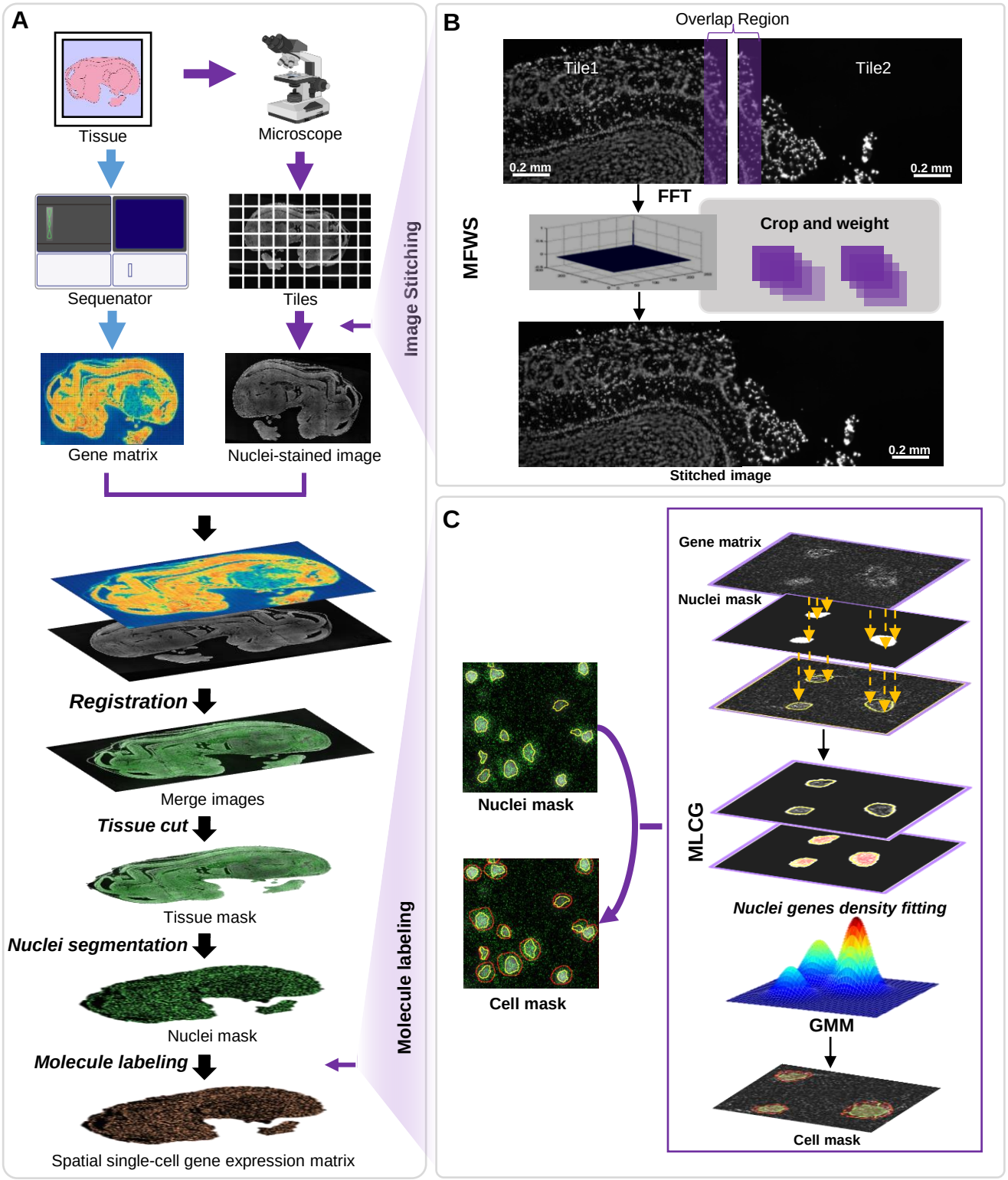

A
Tissue
Microscope
Sequenator
Tiles
Gene matrix
Nuclei-stained image
Registration
Merge images
Tissue cut
Tissue mask
Nuclei segmentation
Nuclei mask
Molecule labeling
Spatial single-cell gene expression matrix
Image Stitching
Molecule labeling
B
Overlap Region
Tile1
Tile2
0.2 mm
0.2 mm
FFT
MFWS
Crop and weight
0.2 mm
Stitched image
C
Nuclei mask
Cell mask
Gene matrix
Nuclei mask
MLCG
Nuclei genes density fitting
GMM
Cell mask

Fig. 2

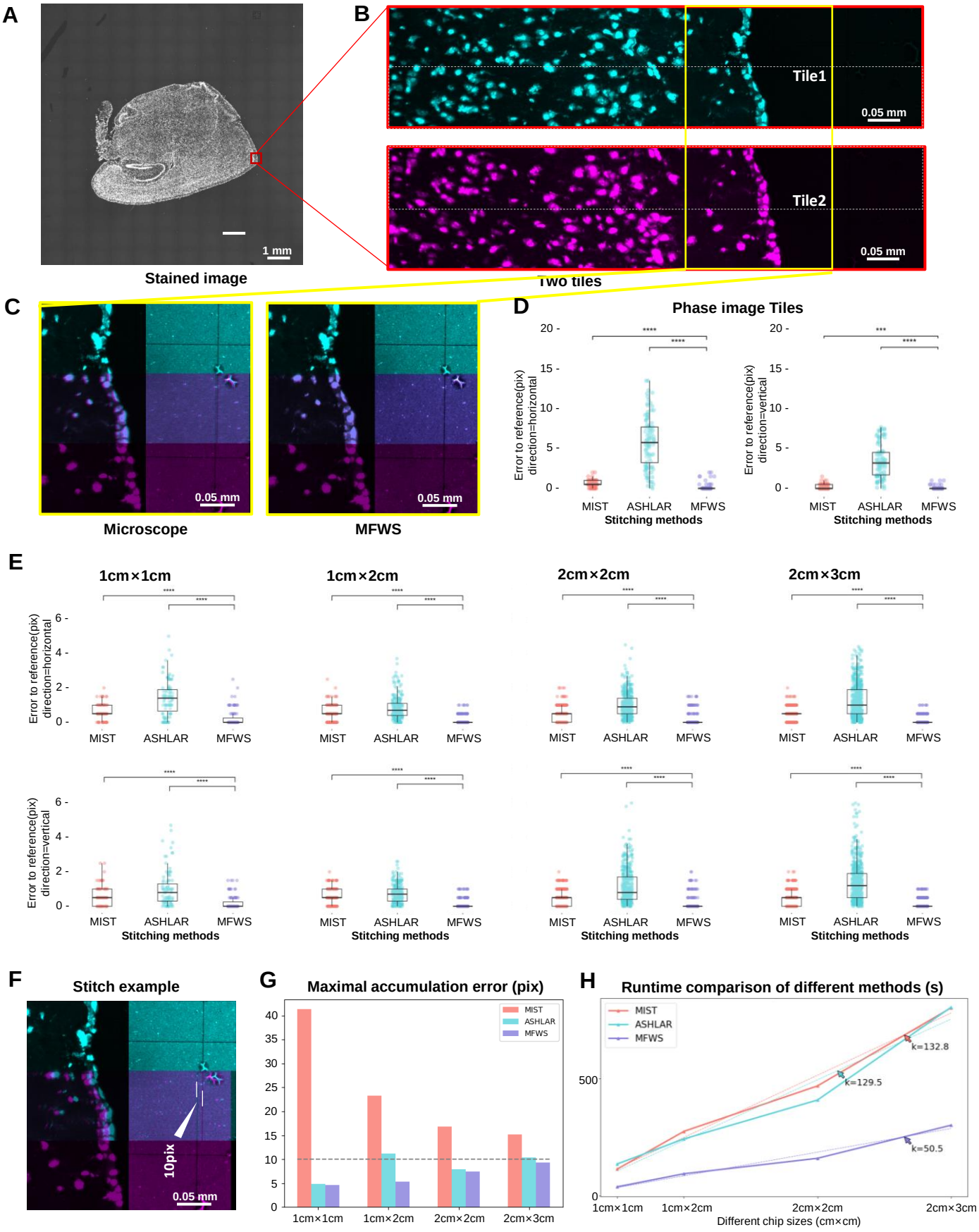

A
1 mm
Stained image
B
Tile1
0.05 mm
Tile2
0.05 mm
Two tiles
C
0.05 mm
0.05 mm
Microscope
MFWS
D
Phase image Tiles
Error to reference(pix) direction=horizontal
Error to reference(pix) direction=vertical
MIST
ASHLAR
MFWS
Stitching methods
E
1cm×1cm
1cm×2cm
2cm×2cm
2cm×3cm
Error to reference(pix) direction=horizontal
Error to reference(pix) direction=vertical
MIST
ASHLAR
MFWS
Stitching methods
F
Stitch example
10pix
0.05 mm
G
Maximal accumulation error (pix)
MIST
ASHLAR
MFWS
1cm×1cm
1cm×2cm
2cm×2cm
2cm×3cm
H
Runtime comparison of different methods (s)
MIST
ASHLAR
MFWS
k=132.8
k=129.5
k=50.5
1cm×1cm
1cm×2cm
2cm×2cm
2cm×3cm
Different chip sizes (cm×cm)

Fig. 3

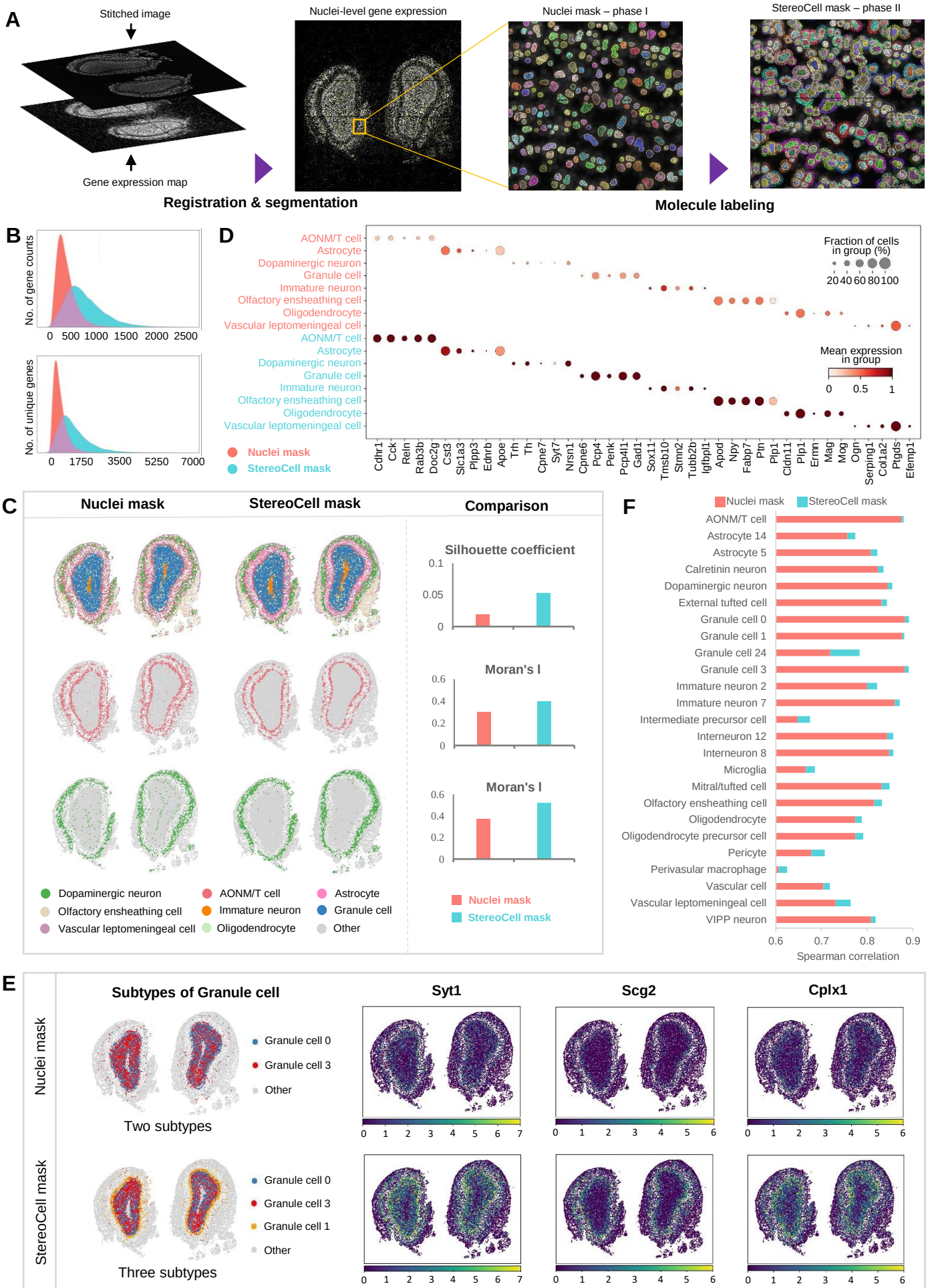
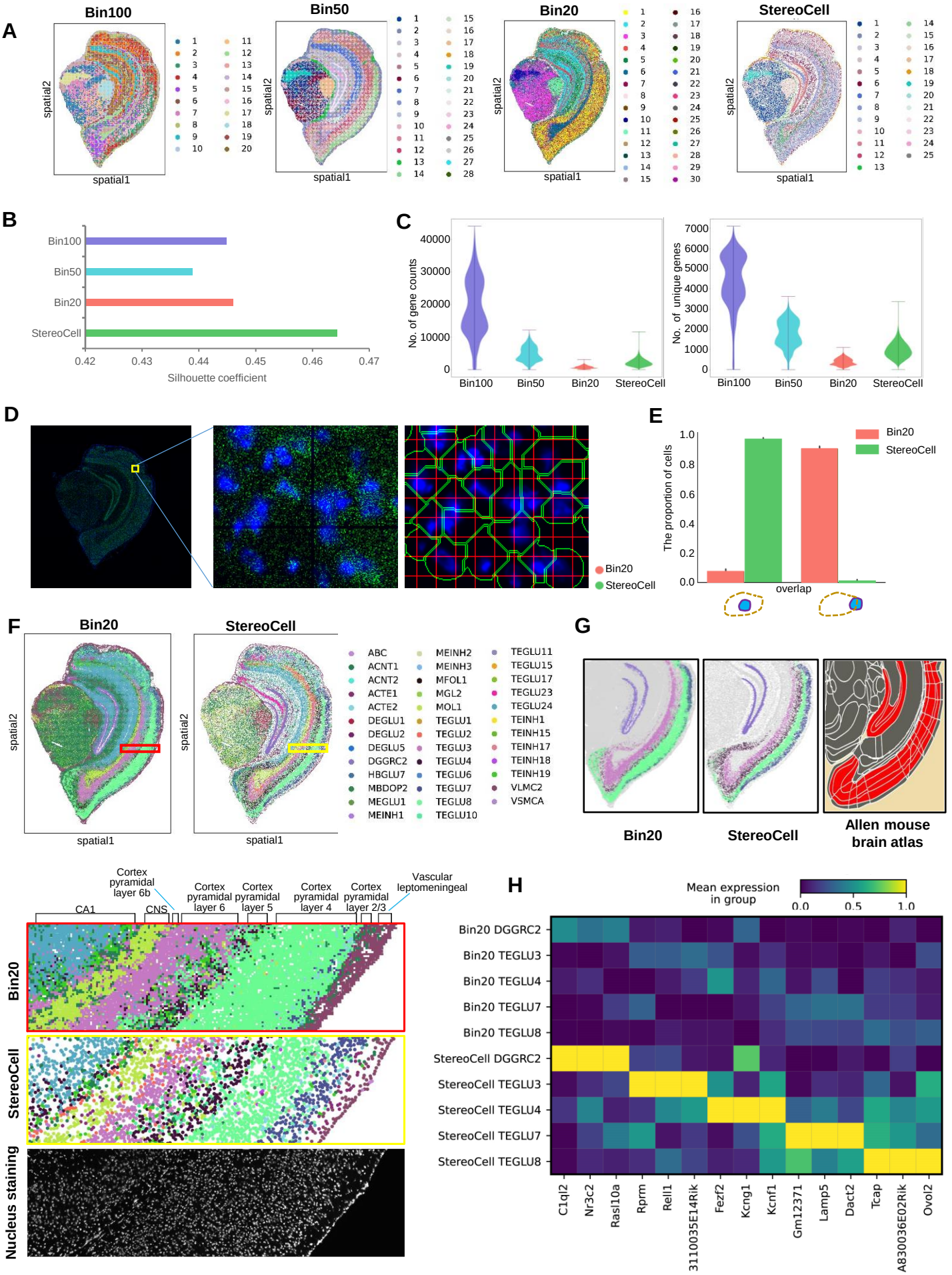

A
Stitched image
Gene expression map
Nuclei-level gene expression
Nuclei mask – phase I
StereoCell mask – phase II
Registration & segmentation
Molecule labeling
B
No. of gene counts
0 500 1000 1500 2000 2500
No. of unique genes
0 1750 3500 5250 7000
D
AONM/T cell
Astrocyte
Dopaminergic neuron
Granule cell
Immature neuron
Olfactory ensheathing cell
Oligodendrocyte
Vascular leptomeningeal cell
Fraction of cells in group (%)
20 40 60 80 100
Mean expression in group
0 0.5 1
Nuclei mask
StereoCell mask
Cdhr1 Cck Reln Rab3b Doc2g Cst3 Slc1a3 Plpp3 Ednrb Apoe Trh Th Cpne7 Syt7 Nrsn1 Cpne6 Pcp4 Penk Pcp4l1 Gad1 Sox11 Tmsb10 Stmn2 Tubb2b Igfbpl1 Apod Npy Fabp7 Ptn Plp1 Cldn11 Plp1 Ermn Mag Mog Ogn Serping1 Col1a2 Ptgds Efemp1
C
Nuclei mask
StereoCell mask
Comparison
Silhouette coefficient
0.1
0.05
0
Moran's I
0.6
0.4
0.2
0
Moran's I
0.6
0.4
0.2
0
Dopaminergic neuron
AONM/T cell
Astrocyte
Olfactory ensheathing cell
Immature neuron
Granule cell
Vascular leptomeningeal cell
Oligodendrocyte
Other
Nuclei mask
StereoCell mask
F
Nuclei mask
StereoCell mask
AONM/T cell
Astrocyte 14
Astrocyte 5
Calretinin neuron
Dopaminergic neuron
External tufted cell
Granule cell 0
Granule cell 1
Granule cell 24
Granule cell 3
Immature neuron 2
Immature neuron 7
Intermediate precursor cell
Interneuron 12
Interneuron 8
Microglia
Mitral/tufted cell
Olfactory ensheathing cell
Oligodendrocyte
Oligodendrocyte precursor cell
Pericyte
Perivasular macrophage
Vascular cell
Vascular leptomeningeal cell
VIPP neuron
0.6 0.7 0.8 0.9
Spearman correlation
E
Subtypes of Granule cell
Syt1
Scg2
Cplx1
Nuclei mask
Granule cell 0
Granule cell 3
Other
Two subtypes
StereoCell mask
Granule cell 0
Granule cell 3
Granule cell 1
Other
Three subtypes
0 1 2 3 4 5 6 7
0 1 2 3 4 5 6

Fig. 4


A
Bin100
Bin50
Bin20
StereoCell
spatial1
spatial2
B
Bin100
Bin50
Bin20
StereoCell
0.42
0.43
0.44
0.45
0.46
0.47
Silhouette coefficient
C
No. of gene counts
No. of unique genes
Bin100
Bin50
Bin20
StereoCell
D
Bin20
StereoCell
E
The proportion of cells
overlap
Bin20
StereoCell
F
Bin20
StereoCell
spatial1
spatial2
ABC
ACNT1
ACNT2
ACTE1
ACTE2
DEGLU1
DEGLU2
DEGLU5
DGGRC2
HBGLU7
MBDOP2
MEGLU1
MEINH1
MEINH2
MEINH3
MFOL1
MGL2
MOL1
TEGLU1
TEGLU2
TEGLU3
TEGLU4
TEGLU6
TEGLU7
TEGLU8
TEGLU10
TEGLU11
TEGLU15
TEGLU17
TEGLU23
TEGLU24
TEINH1
TEINH15
TEINH17
TEINH18
TEINH19
VLMC2
VSMCA
G
Bin20
StereoCell
Allen mouse brain atlas
Cortex pyramidal layer 6b
CA1
CNS
Cortex pyramidal layer 6
Cortex pyramidal layer 5
Cortex pyramidal layer 4
Cortex pyramidal layer 2/3
Vascular leptomeningeal
Bin20
StereoCell
Nucleus staining
H
Mean expression in group
0.0
0.5
1.0
Bin20 DGGRC2
Bin20 TEGLU3
Bin20 TEGLU4
Bin20 TEGLU7
Bin20 TEGLU8
StereoCell DGGRC2
StereoCell TEGLU3
StereoCell TEGLU4
StereoCell TEGLU7
StereoCell TEGLU8
C1ql2
Nr3c2
Rasl10a
Rprm
Rell1
3110035E14Rik
Fezf2
Kcng1
Kcnf1
Gm12371
Lamp5
Dact2
Tcap
A830036E02Rik
Ovol2