StereoCell enables high accuracy single cell segmentation for spatial transcriptomic dataset

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

[1] BGI-Shenzhen, Shenzhen 518103, China

[2] BGI-Beijing, Beijing 100101, China

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

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

[#] These authors contributed equally as the first authors.

*Correspondence: xuxun@genomics.cn, zhangyong2@genomics.cn, liyuxiang@genomics.cn

Abstract

With recent advances in resolution and field-of-view, spatially resolved sequencing has emerged as a cutting-edge technology that provides a technical foundation for interpreting large tissues at the spatial single-cell level. To handle the high-resolution spatial omics dataset with associated images and generate spatial single-cell level gene expression, a powerful one-stop toolbox is required. Here, we propose StereoCell, an image-facilitated cell segmentation framework for high-resolution and large field-of-view spatial omics. StereoCell offers a comprehensive and systematic solution to generating high-confidence spatial single-cell data, including image stitching, registration, nuclei segmentation, and molecule labeling. In image stitching and molecule labeling, StereoCell delivers the best-performing algorithms to reduce stitching error and improve the signal-to-noise ratio of single-cell gene expression compared to existing methods. Meanwhile, as demonstrated using mouse brain, StereoCell has been shown to obtain high-accuracy spatial single-cell data, which facilitates clustering and annotation.

Introduction

Spatially-resolved technologies generate spatially resolved, comprehensive data about molecules that define the identification and function of cells in tissue[1], broaden its application in organ development[2], tumor heterogeneity[1, 3], broaden its application in organ development[2], tumor heterogeneity[3], cancer microenvironment[4, 5], cancer evolution[4, 5]. The rapid development of spatially-resolved technologies tends to move toward obtaining detailed molecules information at single-cell resolution, as well as spatial location information per molecule of tissue at a larger field of view[6], the latter of which is beneficial to reveal biological function at 3D organ level. Currently, some spatial technologies have already provided the possibility for single-cell resolution and a large-field of views. At large field-of-views, technologies containing Stereo-seq[1], can realize the whole view sequencing and analysis of most tissues in model organisms (e.g., mouse brain (1cm × 1cm) and mouse embryo (1cm × 2cm). The single-cell resolution technologies, containing Stereo-seq[1], MERFISH[7], Slide-seq[8], make it possible to explore spatial omics data at the single-cell or even subcellular level. However, how to obtain accurate spatial single-cell gene expression profiles generated by the aforementioned spatially resolved technologies remains elusive.

Although spatially resolved technologies vary in resolution and field of view, one common feature is the ability to produce traditional high-content tissue images using tissue staining dyes such as Hematoxylin-eosin (H&E), or 4,6-Diamidino-2-phenylindole, dihydrochloride (DAPI) for cellular or nucleus staining. Images generated using tissue stains provide important basic information about tissue and cell morphology, but are currently not fully leveraged in spatially-resolved pipelines, although they generate auxiliary information that can be used for subsequent biological data mining. There are two main difficulties: insufficient precision in image mosaic and accuracy of molecule labeling. The first one is limited by the small imaging area, mechanical tolerance, and imperfect calibration of microscope technology (e.g., 10X imaging lens, ~1mm × 1mm), which are only sufficient for visual review and not accurate enough for quantitative single-cell analysis. For example, the object 1cm × 1cm generates about 10 × 10 tile images, and the object 3cm × 3cm generates nearly thousands of tile images. The existing image stitching methods, such as ASHLAR[9] and MIST[10], can meet most of the stitching accuracy requirements. But they all have the problem of stitching misalignment, and meanwhile, it is difficult to collect evaluation datasets to improve the precision and accuracy of image mosaic at large field of view, which has been an obvious obstacle in expanding the scope of application of these image mosaic methods. Secondly, the key to achieve high-precision, single-cell resolution is to classify the transcriptomic molecules into their owner cells, this procedure is termed “molecule labeling”. However, molecule labeling is not easy task due to the low efficiency of molecule capture and molecule diffusion in the wet-lab experiments. Some techniques, such as pciSeq[11], perform molecule labeling according to the nuclear location. However, many transcripts (messenger RNA) are present outside the nucleus, including messenger RNAs in the cytoplasm and outside the cell (due to molecular movement and diffusion), which greatly affects the correct annotation of transcripts to the corresponding cells. Therefore, more recent technologies have taken note of this issue. 10X Xenium technology includes transcripts that extend outside the nucleus for a certain distance (15μm) from the nucleus[12]. Since the diameter of immune cell, stromal cell and tumor cell varies, however, this crude method is difficult to obtain accurate single-cell data and is prone to introduce a lot of noise. Many available methods like Baysor[13], Pixel-seq[14], require high data quality, such as the number of captured molecules, their

density distribution, which may not applicable to most sequencing-based spatial omics data. Furthermore, existing spatial data analysis frameworks, such as Seurat[15], RCTD[16], Cell2location[17], Squidpy[18], Spateo[19], etc., either focus on the downstream analysis tasks alone, or focus on one specific aspect of the spatial data analysis, or only take cell segmentation into consideration. The fully functional solution to perform single-cell molecule labeling in spatial omics data with a large-field of view and high resolution is not yet available.

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 spatial single-cell transcriptomic data. StereoCell provides the systematic solution for obtaining high-confidence spatial single-cell data, including image stitching, registration, nuclei segmentation, and molecule labeling. For image stitching, StereoCell contributes the best-performance algorithm named MFWS to reduce the stitching error on large field-of-view datasets. And for molecule labeling, StereoCell proposed a novel algorithm to increase the signal-to-noise ratio of single-cell gene expression compared to existing approaches. At the end, StereoCell is applied to mouse brain dataset and proven to obtain a high-confidence spatial single-cell data for cell clustering and annotation.

Result

Overview of the StereoCell pipeline

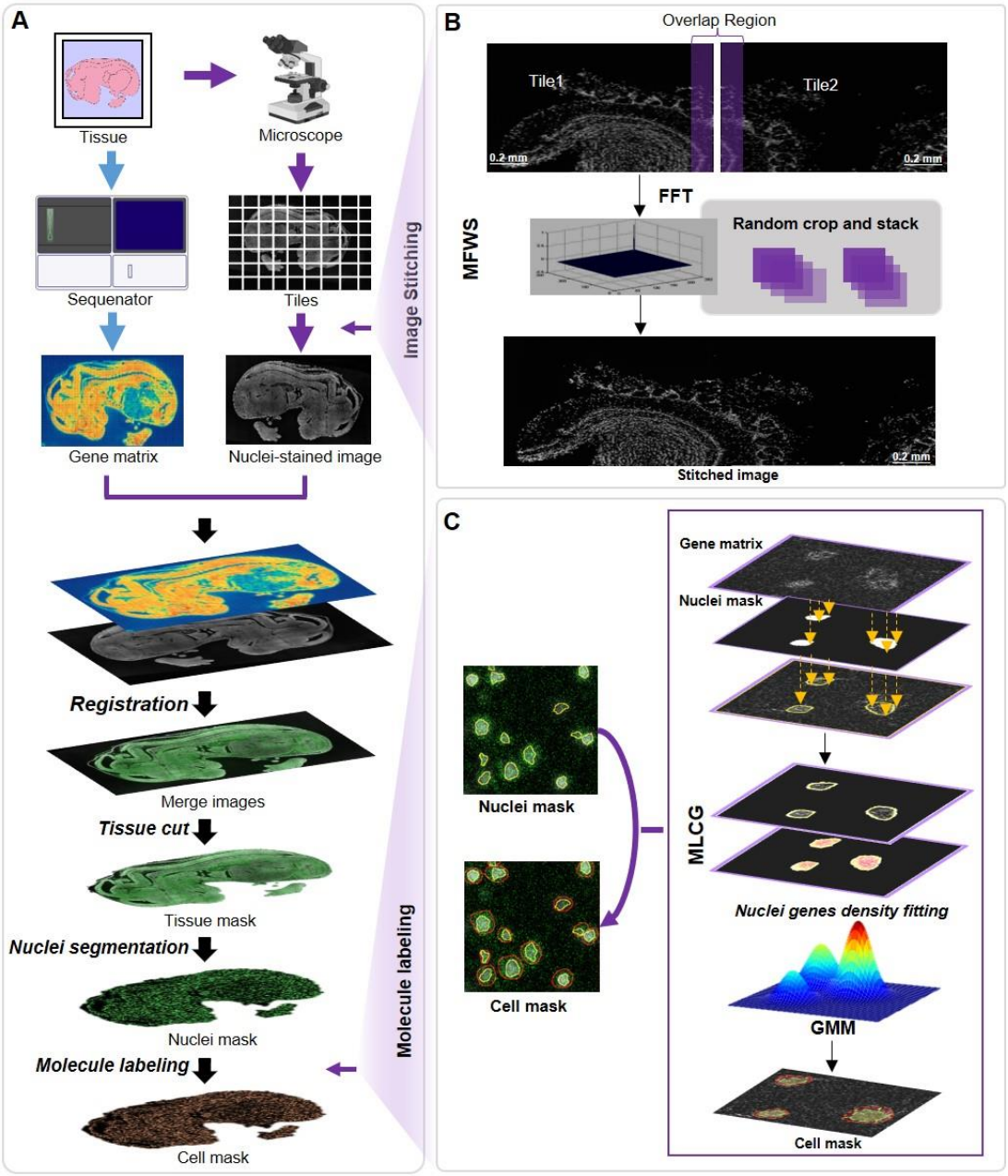


Fig. 1 | Overview of StereoCell - an integrated technology and software analysis suite for single-cell expression profiling of spatial data.

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 microscope, respectively. The tiles obtained by microscope are stitched together to make up a large mosaic image of the whole tissue to be merged with the gene expression matrix. Next, tissue segmentation 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 stacked, and spectral information is calculated based on the Fast Fourier Transform (FFT) algorithm to take the maximum spectral position as the offset of the adjacent tiles. Then, tiles are stitched with the offset value. See also in Supplementary Fig. 1A.
C. The flow chart of molecule labeling based on cell morphology and gaussian mixture model (GMM) (MLCG) algorithm. Molecules within the nucleus boundary are assigned to the given cell, and molecules outside the nucleus are labeled to the cell with the largest probability using GMM fitting.

StereoCell provides a complete processing workflow, including image stitching, registration, tissue segmentation, nuclei segmentation and molecular labeling, and finally generates a spatial single-cell gene expression profile with high confidence (Fig. 1A). Firstly, we develop multiple FFT weighted stitching (MFWS) algorithm, which is based on the frequency domain information and allows accurate stitching of high-resolution images with a wide field of view, along with improvements in the cell dislocation caused by inaccurate image stitching and low efficiency caused by a huge number of tiles (Fig. 1B). Secondly, it provides molecule labeling based on cell morphology and gaussian mixture model (GMM) (MLCG) algorithm, a probability statistical method based on the nuclei segmentation mask to accurately assign molecules to the most probable cell to obtain a single-cell level gene expression profile (Fig. 1C). The high-confidence spatial single-cell transcriptomic data obtained by StereoCell was able to achieve more reliable downstream analysis, which is useful for quickly exploring spatial omics datasets and developing the methods for spatial omics data at single-cell resolution.

StereoCell provides high precision stitching method for morphological tissue images

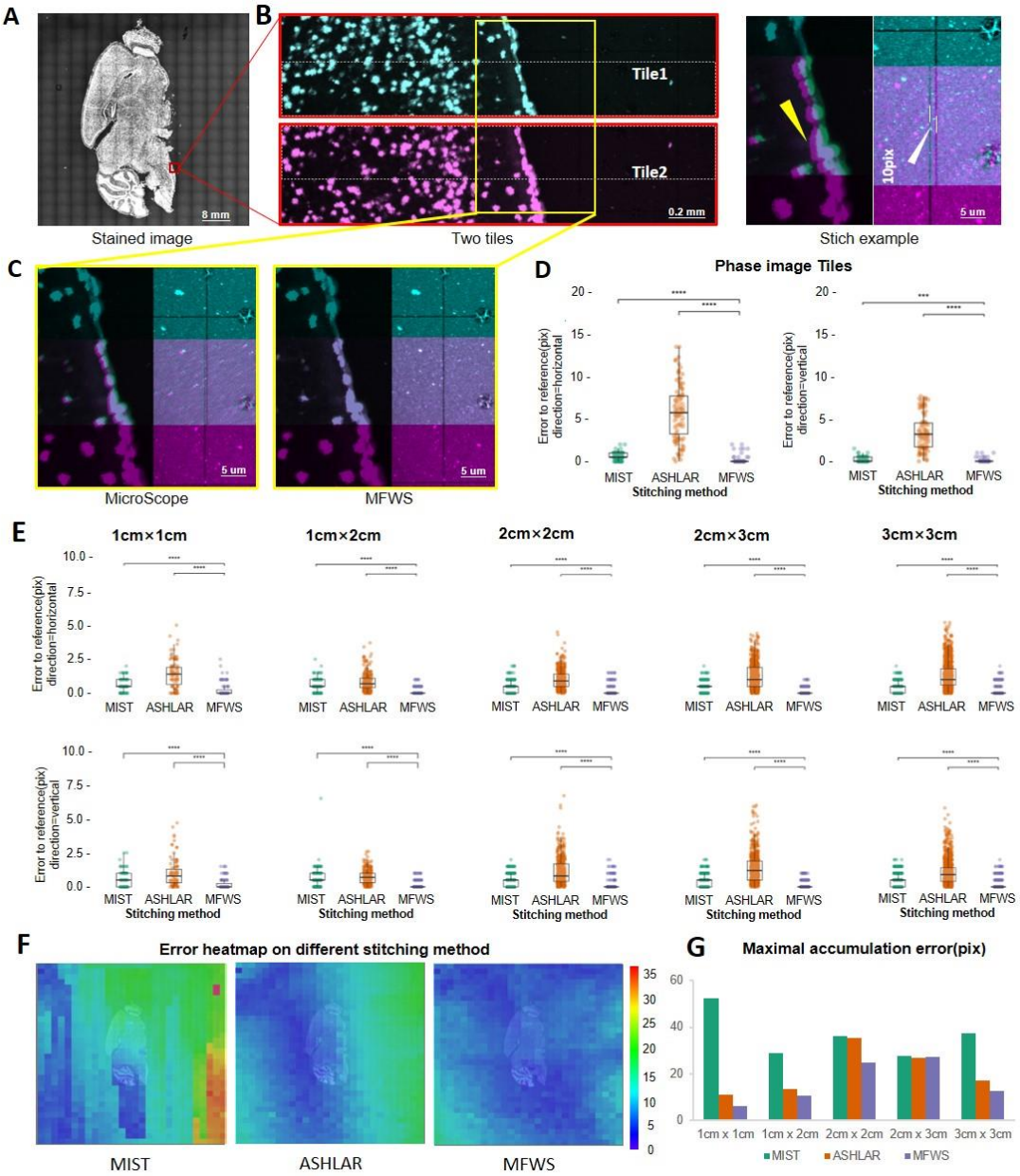


Fig. 2 | StereoCell includes a high-precision mosaic method for image stitching (MFWS), enabling large-field morphological image analysis to spatial single-cell accuracy.

A. Part of mosaic image of mouse brain, chip size in 3cm × 3cm.

B. Left: The red box is placed at the edge area of the tissue (containing cells and non-tissue area), the two neighboring tiles (in vertical direction, part of the tiles) are shown in different colors. The part that needs to be overlaid by stitching is the lower part of tile 1 (in the dotted area) and the upper part of tile 2 (in the dotted area). Right: An example of the stitching error in pixels in the stitched zoom-in image. The 10 pixels correspond to about half a cell.

C. The stitching result of the two tiles (yellow box in B) with the built-in stitching method of the microscope and with the use of MFWS. 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. And the contrast of the background area was increased to show the track line clearly.
D. The comparison of the relative errors of MIST, ASHLAR and MFWS on public data set "Phase_Image_Tiles".
E. The comparison of the relative errors of MIST, ASHLAR and MFWS on the datasets. We select different chip sizes from 1cm × 1 cm to 3cm × 3 cm, and the number of tiles is from 12 × 9 to 34 × 31, All the samples are mouse brain and sequenced by Stereo-seq. The evaluation is based on the ground truth guided by the track line on Stereo-seq data.
F. The heatmap displays the local stitching errors of each tile calculated by MIST, ASHLAR and MFWS based on a 3cm × 3cm slide.
G. The bar graph displays maximal accumulation stitching errors of MIST, ASHLAR and MFWS on the datasets.

The first procedure in StereoCell involves accurate stitching of the morphological image. Processing a morphological image from a tissue slide involves stitching an array of multiple image tiles generated by a microscope. Microscopes can automatically capture the tiles of the tissue one by one and stitch the tiles by the built-in stitching method of the microscope. The overlaps between adjacent tiles that are generated during microscope movement may however be imprecise due to mechanical tolerances and imperfect calibration. Such stitching errors are therefore commonly present in mosaic images and need to be removed with the goal to achieve single-cell resolution in spatial omics applications. Fig. 2A displays a mouse brain mosaic image on a chip with a size of about 3cm × 3cm and a tile size of 2428 × 2035 pixels, corresponding to a physical size of ~1.2mm × 1.0mm. To visualize seams and stitching errors, we selected two adjacent tiles (Fig. 2B left) from the image. With the intention of getting an accurate stitching result, the lower part of tile 1 (dotted area), and the upper part of tile 2 (dotted area) should be accurately overlaid. Fig. 2B right shows an example of 10 pixels (5μm) dislocation which corresponds to about half a cell. Shadows that represent inaccurate overlapping of cells can be easily seen in the stitching result handled by the microscope (Fig. 2C, left). MFWS is able to stitch the tiles accurately, resulting in no shadows (Fig. 2C, right). We also present the stitching result of two tiles with non-tissue areas, where individual track lines can be clearly seen using auto contrast adjustment, and we can easily view the stitching result simultaneously (Fig. 2C). According to the track line, the stitching errors are measured to calculate the cell dislocation.

We applied MFWS to a public stitching data set (Supplementary Table 1)[20] and compared with existing stitching methods. The result shows that the relative offset errors are more concentrated and with the mean shifts of 0.12 (horizontal direction) and 0.06 (vertical direction) pixels, followed by MIST with the mean shifts of 0.59 (horizontal direction) and 0.18 (vertical direction) pixels and 8.8 pixels and ASHLAR with the mean shifts of 5.81 (horizontal direction) and 3.44 (vertical direction) pixels (Fig. 2D). The applied Stereo-seq chip contains periodic track lines with a 1500nm (~3 pixel) size, which is apparent from the image (Fig. 2B, C). We used a collection of data to evaluate the accuracy and efficiency of available stitching algorithms vs. MFWS. We made a ground truth data set according to these track lines, and included nuclei-stained images of the mouse brain with different chip sizes from 1cm × 1cm to 3cm × 3cm, corresponding to the number of tiles ranging

from 108 to 1054 (Supplementary Table 1). The relative offset errors are calculated by comparing them with the ground truth. The results show that MFWS causes lower relative offset errors than ASHLAR and MIST (Fig. 2E). Further, we fixed a calibration point and calculated the absolute offset errors of each method. On the 3cm × 3cm dataset, the errors of MFWS stay within a maximum of 10 pixels in the tissue area, while other methods, such as ASHLAR and MIST, have a more than 15-pixel shift, which may misplace the cell (Fig. 2F). For more datasets, MFWS shows a less maximal accumulation error than ASHLAR and MIST (Fig. 2G).

StereoCell generates high accuracy single-cell gene expression profile for spatial omics data

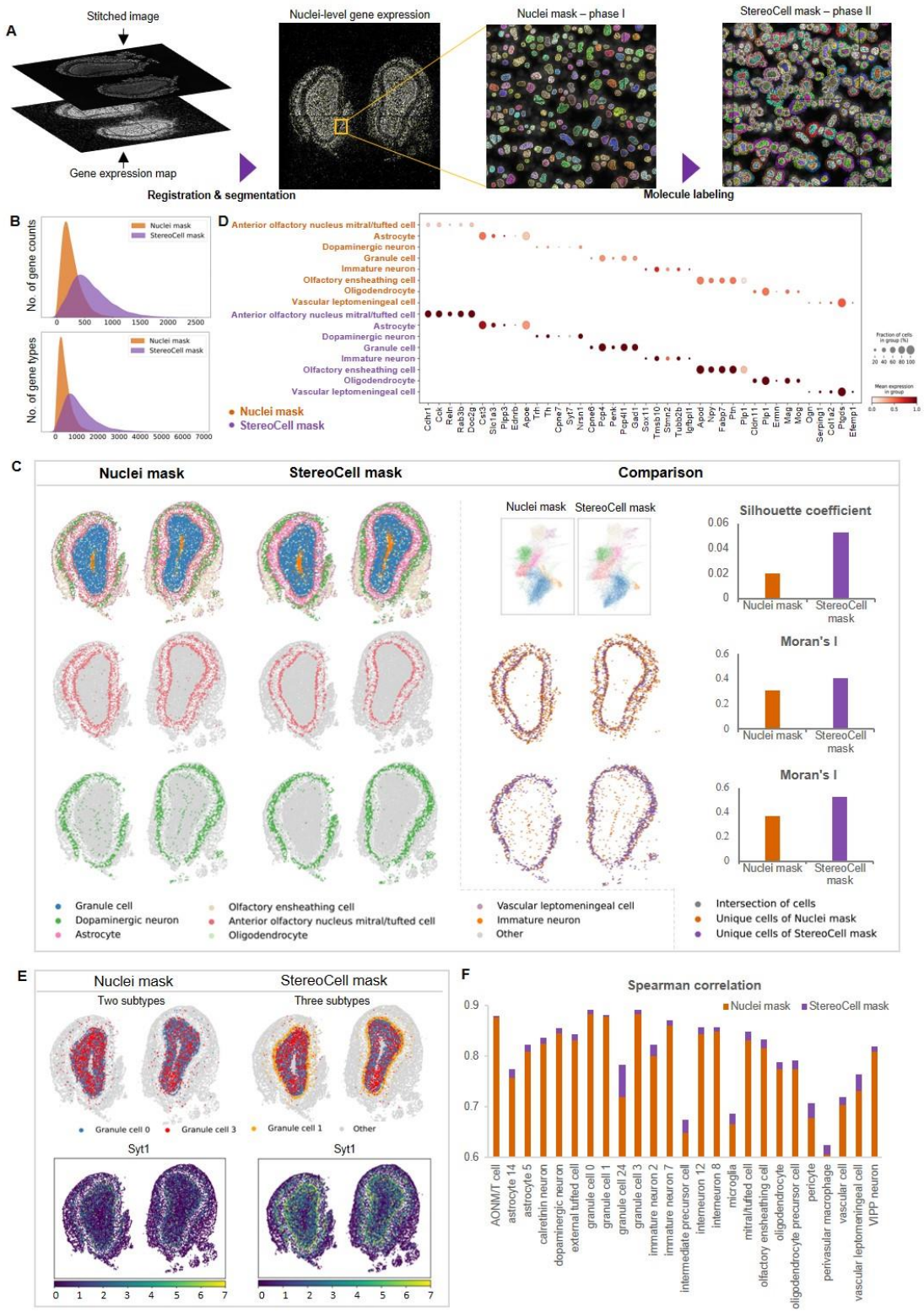


Fig. 3 | The molecular labeling method can provide spatial single-cell data with a higher signal-to-noise ratio and facilitate finer cell clustering and annotation on mouse olfactory bulb data.

A. Flowchart of MLCG processing on mouse olfactory bulb data.
B. Comparison of the number of expressed gene types and the total gene counts per cell under nuclei mask and StereoCell mask on mouse olfactory bulb data. Up: the density plot of total gene counts per cell. Down: the density plot of the number of gene types per cell.
C. Comparison of clustering results by the "Leiden" algorithm using the nuclei mask and StereoCell mask. First line, left part: spatial clustering result of mouse olfactory bulb data under nuclei mask and StereoCell mask; right part: umap visualization of clustering result under nuclei mask and StereoCell mask, and the comparison for clustering result estimated by silhouette coefficient. Second and third line, left part: the spatial distribution of cells in the inferred clusters of Astrocyte and Dopaminergic neuron under nuclei mask and StereoCell mask; right part: the comparison of cells in the inferred clusters of Astrocyte and Dopaminergic neuron under nuclei mask and StereoCell mask, and the comparison for spatial autocorrelation estimated by moran's I.
D. Marker gene expression comparison of the main clusters inferred by "Leiden" algorithm using a nucleus mask and the StereoCell mask.
E. Comparison of subtype identification under nuclei mask and StereoCell mask. First line, the subtypes identified for Granule cell with nuclei mask and StereoCell mask. Second line, the gene expression of Syt1 which is reported as a marker gene for Granule cell 1 with nuclei mask and StereoCell mask.
F. Spearman correlation between the cell-level gene expression using a nucleus mask and the StereoCell mask in tissue slides of the mouse olfactory bulb versus a single-cell reference of the mouse olfactory bulb. AONM/T cell: anterior olfactory nucleus mitral/tufted cell, VIPP neuron: vasoactive intestinal peptide positive neuron.

Upon stitching of the morphological image, the morphological tissue image is computationally overlaid onto the tissue gene expression map (Supplementary Fig. 2). Next, StereoCell identifies single cells within the merged image. StereoCell performs better on nuclei-stained grayscale images and H&E images generated by Stereo-seq (Supplementary Fig. 3).

We applied StereoCell to the spatial transcriptome data of the mouse olfactory bulb. After obtaining the nuclei segmentation masks, the molecules falling inside the nucleus were assigned to the corresponding cell. Then, StereoCell fit the molecular density distribution in each cell nucleus and relabels the molecules outside the nucleus based on the molecule labeling correction method to obtain spatial single-cell gene expression profiles (Fig. 3A). After the molecule labeling, the average values of expressed gene type number and total gene counts in a single cell increased about 2.34 and 2.56 times, respectively (Fig. 3B). StereoCell provides a better overall clustering result by the "Leiden" algorithm[21] (Fig. 3C) and its silhouette coefficient (an index for cluster evaluation)[22] is 0.05 compared to the nucleus segmentation mask with a silhouette coefficient of 0.02. After that, we manually annotated the cell types by comparing the differential expressed genes of each cluster with the marker genes of cell types inferred by the reference[23]. To be noted, there are fewer scattered points and a more concentrated distribution in the spatial position at 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 (Fig. 3C). We further explored the marker gene distribution and expression in all annotated cell types. All of the marker genes provided by the reference[23] shows high expression in dataset under the StereoCell mask compared to the nuclei mask (Fig. 3D). We next compared the subtypes of Granule cell discovered under the nuclei mask and StereoCell mask. The dataset under the nuclei mask can

support identifying two subtypes of Granule cell: Granule cell 0 and Granule cell 3, while three under the StereoCell mask: Granule cell 0, Granule cell 3 and Granule cell 1. We explored the expression of marker genes of Granule cell 1. Syt1 and other maker genes of Granule cell 1 can be clearly seen that their expression have no obvious difference in Granule cells under the nuclei mask, while has a higher expression in Granule cell 1 compared to other two subtypes under StereoCell mask (Fig. 3E, Supplementary Fig. 4). It indicates that StereoCell is able to retain more signal for rare cells and thus can be better annotated, which facilitates in-depth scientific research. Finally, we estimated and compared the similarity of spatial gene expression with single cell datasets under nuclei mask and StereoCell mask. Cells are auto annotated by Tangram[24] and the correlation analysis is performed between the expression matrices of the spatial single cell under each mask and the single cell sequencing expression profile of the reference[23]. The correlation of StereoCell mask on each annotation is higher (0.4% to 8.9%) than that of nuclei mask. The cell matrix generated by StereoCell overall has a higher correlation with the single cell reference than the gene matrix generated by nuclei mask (Fig. 3F). It indicates that StereoCell provides a spatial cell-level gene expression profile closer to single-cell sequencing gene expression. All these results suggested that StereoCell provides a higher accuracy single-cell spatial transcriptomics expression matrix.

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

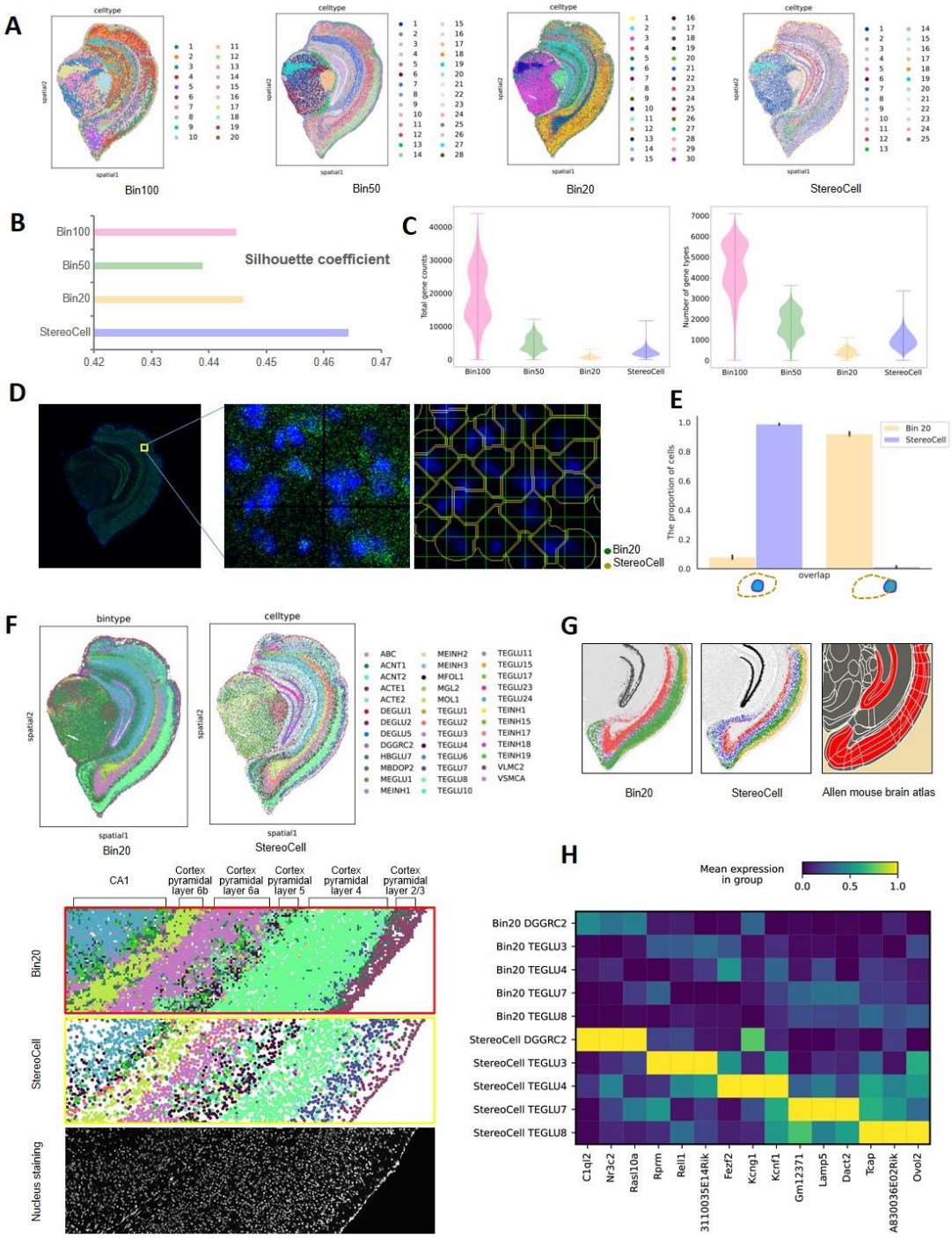


Fig. 4 | Clustering and annotation of mouse brain tissue by StereoCell vs. different BinX approaches.

A. Identification of gene expression clusters in mouse brain tissue using different bin sizes and StereoCell.

B. A silhouette coefficient-based evaluation of clustering performance.

C. Violin plots display total gene counts and gene type number per cell processed by different bin

sizes and StereoCell.

D. Cell boundary comparison using Bin20 and StereoCell. The first image shows the mouse brain gene expression matrix map and the nuclei-stained image which were merged in two colors, green and blue respectively. The second image displays the enlarged map of the local area from the first image. The third image is a comparison diagram of the cell boundary result of the Bin20 and StereoCell in this area.

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

F. Comparison of end-to-end annotation results using Bin20 and StereoCell. The colors that represent the same cell type in the Bin20 and StereoCell annotation charts have been unified. The local enlarged comparison map of the cortex shows the spatial distribution and boundary information of different cortical cells.

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

H. Heatmap of marker gene expression in annotation results in G. It displays the marker gene expression in gear gyrus (DGGRC2) and cortex (TEGLU3, TEGLU4, TEGLU7 and TEGLU8) layer. The upper part is results based on Bin20, and the results in the lower part were based on StereoCell.

Compared with the method of obtaining cell boundaries based on images, the approach that directly divides adjacent regions as downstream analysis units according to the matrix is more convenient and faster, and it is also a common way to perform spatial omics data analysis. We compared StereoCell with bins of different sizes (BinX means X × X DNBs, DNA nano balls)[1] on mouse brain tissue. Firstly, we compared the clustering results on data generated by StereoCell, Bin100, Bin50, and Bin20 (Fig. 4A). We evaluated the clustering results of several methods by silhouette coefficient (Fig. 4B), StereoCell obtained the highest silhouette coefficient, followed by Bin20, then Bin100, and Bin50. We calculated the total expression (total gene counts) and the number of gene types in cells on data processed by different bin sizes and StereoCell (Fig. 4C). Bin100 was divided into 20 clusters and some important areas in the tissue (such as the brain hub, hippocampus, etc.) couldn't be clearly identified. Compared with Bin100, Bin50 yielded 28 clusters and identified the important areas of the tissue more clearly, but still had a poor identification of tissue cortex and blood cells. Bin20 classified the cells into 30 clusters while StereoCell generated 25 clusters. Both could identify important areas in the tissue and the division of cortex and blood cells was obvious, but the differences in clustering were noticeable.

The clustering results of Bin20 and StereoCell reflected the reality at a very high level and both are close to single-cell accuracy (the cell size of a mouse brain is about 10μm, which corresponds to Bin20, i.e., 20nm × 500nm)[1]. The visualization results of cell boundaries obtained by the two methods (Fig. 4D) showed that Bin20 divided a single cell into two or more cells in most cases, while StereoCell more accurately divided the cell area and was consistent with the real cell distribution in the tissue. Fig. 4E showed Bin20 slicing of approximately 90% of the cells, and only about 10% of the cell nuclei were completely covered in the nuclei-stained image, while using StereoCell, only ~2% of the cells and most of the nuclei were included. The results of Bin20 and StereoCell were annotated by Spatial-ID[25] (Fig. 4F, right) with the Adolescent mouse brain as a reference[26]. 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 was the only case where StereoCell annotated fewer subtypes. In addition, compared with Bin20, StereoCell annotated 2 subtypes of the ACNT series and 4 subtypes of the TEINH series. The annotation results of StereoCell were overall more accurate. When zooming in on the cortical area (Fig.4 F left), the staining position of StereoCell and the nucleus were more aligned, and the cell state and tissue structure were more in line with reality. We found that in the cortical area of the mouse brain, StereoCell could annotate each layer more clearly than Bin20, such as the gear gyrus and cortical regions of the mouse brain and compared with the Allen mouse brain (Fig.4 G). The annotation of the StereoCell results was consistent with the position of the Allen mouse brain. Although Bin20 could also annotate the cogwheel and cortex and match the location of the Allen mouse brain, the division between layers was not obvious. Statistically observing the expression of marker genes in the gear wheel (DGGRC2) and cortex (TEGLU3, TEGLU4, TEGLU7 and TEGLU8) regions, and the results of StereoCell showed that the marker genes had significantly higher expression levels than Bin20.

Altogether, compared with BinX, StereoCell provided single-cell spatial transcriptomics maps with a higher signal-to-noise ratio. Moreover, the single-cell maps generated using StereoCell aligned well with cellular annotation within different brain regions, and could 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 technologies can be fully utilized. This will require continued improvement of both protocols and analytical methods for processing the spatial omics data. Here we present StereoCell, the first systematic method which generates more accurate spatial gene expression profiles for downstream bioinformatic analysis of spatial omics data at a large field-of-view and single-cell resolution. The implemented morphological tissue image stitching proved to be accurate and reliable for single cell identification. The high-precision stitching method based on FFT and a benchmark dataset based on chip track lines are also useful for correcting the stitching to almost subcellular precision. Also, we proposed a new molecule labeling method based on GMM and nuclei segmentation to generate spatial single-cell gene expression data with higher signal-to-noise ratio, yielding more reliable analysis of cell clustering and end-to-end annotation. We also optimized the processing modules on image registration tools based on chip track lines, nuclei segmentation, and tissue cutting. StereoCell is equipped with an extensive testing suite, implemented in a robust continuous integration pipeline. We provide a graphic user interface for manual image stitching, registration, labeling. StereoCell's rich documentation in the form of functional application programming interface documentation, examples, and tutorial workflows, is easy to navigate and accessible to both experienced developers and beginner analysts.

In our study, we demonstrate that auxiliary staining can be extremely valuable when resolving difficult cases. The molecular labeling or cell segmentation results will probably be improved overall using better staining, like the labeling of outer membranes and better techniques for segmenting such images. Additionally, better staining photos with enhanced AI recognition models will provide us more insight into biological traits, such as identifying features with the ability to distinguish between healthy and diseased tissue. Besides, it needs to mention that the whole StereoCell framework can be applied to generate single-cell gene expression for Stereo-seq and most of the its modules are also suitable for other spatial transcriptomics technologies.

Our goal here was to develop a generalist tool for combined data generation and single-cell level analysis of spatial omics data. The herein presented data are based on spatial transcriptomics alone. Next, we will combine spatial information and additional biological characteristics, such as proteomics to integrate multi-omics information at the spatial single-cell 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 technologies.

Online Methods

StereoCell relies on common python packages for image processing, data analysis and image interface, such as libvips, pillow, numpy and PyQt. 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 in the cell segmentation module. StereoCell provides a set of visual solutions, and its GUI is developed using PyQt. The detailed instructions and operational documentation, including tutorials and sample workflows, are also provided.

Image stitching

We developed the image stitching algorithm named Multiple FFT Weighted Stitching (MFWS). The algorithm procedures are as follows:

1. Calculate the overlap matrices of adjacent tiles in horizontal and vertical direction. 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 horizontal and vertical direction. The calculation method applied for horizontal and vertical direction is the same, a pair of two tiles in vertical direction is taken as an example for further introduction.

i) Multiple division in the overlapping regions of adjacent tiles: The overlap regions of adjacent tiles are marked as A_f and A_m for two tiles, respectively. We divided A_f and A_m for N-1 times and get N subregion images each, marked as $\{f_i\}$ and $\{m_i\}$.

ii) Calculation of subregion spectrum: We applied the fast Fourier transform algorithm and obtained $2*N$ frequency spectrums based on $\{f_i\}$ and $\{m_i\}$.

iii) Calculation of cross-power spectrum of pairing 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 by using the formula of image cross-power calculation. Theoretically, any one of these spectra can reflect the true offset information.

iv) Weighted superposition of 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 the Inverse Fast Fourier Transform (IFFT). The coordinate of the peak value represents the required offset value.

v) Generation the offset matrices: Carrying out the above steps for each pair of adjacent tiles, we can generate the offset matrices O_H and O_V.

2. Obtain the global coordinates of each tile. It requires transformation from local coordinates to stitching coordinates since image stitching requires the unique coordinates of each tile in the reference system, i.e., the global coordinate matrix L. In the O_H and O_V matrices, offsets corresponding to low confidence values are not credible and need to be eliminated. But 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 between 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 was obtained, the value was traversed, and the seam fusion was completed synchronously during the stitching process.

Image registration

The gene expression matrix was transformed into an image (map), where the intensity value at each pixel was proportional to the total molecule counts of all genes expressed at position DNB (DNA nano ball, x, y). To assign the expressed genes to the spatially corresponding cells, the nuclei-stained microscopic image was registered to the transformed gene count image. Sequencing chips attached with sample tissue, imaging and sequencing were designed with periodic track lines (horizontal and vertical track lines) to assist base calling and image registration. The track lines were displayed as dark straight lines in both the nuclei-stained image and the gene count image. Image registration was achieved by 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 crosses of the horizontal and vertical track lines were located. The scaling parameter was calculated by comparing the length of the line segments between crosses 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 of the two images and roughly match them together, then matched the track line crosses to perform fine-tune the image registration. See details 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.

In the preprocessing step, the histogram enhancement method is used to improve the separation of tissue and background. Then, in the model inference step, we use Bi-Directional ConvLSTM U-Net[27] based on U-Net[28] to predict tissue masks. Compared with the original U-Net, it incorporates the following improvements: (i) BCDU-Net replaces the convolution sequence with a densely connected convolution in the last step of the encoding path to encourage feature reuse; (ii) BConvLSTM was 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 couldn't achieve ideal results on some complex or poor imaging quality tissue images, and the self-evaluation mechanism could help to filter these potentially unsatisfactory results. According to the characteristic that the tissue

part of the nucleus image had a higher 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 in the image were valid pixels. The effective pixel density outside and inside the tissue of the surrounding area of the mask were calculated in sections, and the ratio of the two values was used as the basis for the evaluation. The process of tissue segmentation algorithm and specific segmentation examples were shown respectively in Supplementary Fig. 2.

Molecule labeling

We developed the molecule labeling based on cell morphology and GMM(MLCG) algorithm. It had two main steps: (i)cell morphological segmentation (nuclei-stained images); and (ii) molecule labeling correction. Compared with other methods, MLCG jointed the cell morphological information and the molecule matrix to get 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 to make cells relatively obvious on 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.

Segmentation of cell

A deep neural network based on U-Net[28] with a suitable backbone modified from resnet was used to segment 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.

Model architecture

Deep Residual U-Net (ResUnet)[29] underpinned the psaUnet architecture, which combined Deep Residual net[30], U-Net[28] and EPSANet[31]. The diagram of the proposed psaUnet architecture was shown in Supplementary Fig. 3B. 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; For the rest of the block was made up of 2~5 psablocks, each of them was a residual block embedded by Pyramid Squeeze Attention (PSA) module[31]. The Pyramid Squeeze Attention（PSA）module that was employed as an attention unit to obtain detailed information from feature maps was described in Supplementary Fig. 3B. The diagram of the proposed psaUnet architecture was shown in Supplementary Fig. 3B. 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; For the rest of the block was made up of 2~5 psablocks, each of them was a residual block embedded

by Pyramid Squeeze Attention (PSA) module[31]. The Pyramid Squeeze Attention（PSA）module that was employed as an attention unit to obtain detailed information from feature maps was described in Supplementary Fig. 3B. Deep Residual U-Net (ResUnet)[29] underpinned the psaUnet architecture, which combined Deep Residual net[30], U-Net[28] and EPSANet[31]. The diagram of the proposed psaUnet architecture was shown in Supplementary Fig. 3B. 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; For the rest of the block was made up of 2~5 psablocks, each of them was a residual block embedded by Pyramid Squeeze Attention (PSA) module[31]. The Pyramid Squeeze Attention（PSA）module that was employed as an attention unit to obtain detailed information from feature maps was described in Supplementary Fig. 3B. The diagram of the proposed psaUnet architecture was shown in Supplementary Fig. 3B. 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; For the rest of the block was made up of 2~5 psablocks, each of them was a residual block embedded by Pyramid Squeeze Attention (PSA) module[31]. The Pyramid Squeeze Attention（PSA）module that was employed as an attention unit to obtain detailed information from feature maps was described in Supplementary Fig. 3B.

PSA Module

As illustrated by PSA module in Supplementary Fig. 3B, the main operation of implementing PSA took four steps. First, the input tensors were sliced into a multi-scale feature map on the channel-wise using the Squeeze and Contact (SPC) module. Second, the SEWeight module was applied to extract the channel-wise attention vector from the different scale feature maps. Third, the weighted channel-wise attention vector was generated by applying the "Softmax" activation function to the channel-wise attention vector. Fourth, a refined feature map was obtained by taking element-wise product of a multi-scale feature map with the corresponding weight 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[29] to let the model pay 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 the X was the input tensor, SPC and se represented the SPC and SEWeight module respectively, the $Softmax$ was the "softmax" function, and $\otimes$ was the operation of element-wise product.

Psablock

Thought deeper network can bring a higher accuracy, the network gets deeper, gradient will vanish as the meantime making the network difficult to train. To solve the problem of gradient vanishing, resnet is introduced. With residual unit and skipping of connections in dropout layer, the training of deep network becomes faster and the ability of feature learning becomes stable. For this reason, we use architecture of resnet as the backbone architecture of encoder in our proposed model.

Postprocessing method

To get a more reliable segmentation result, 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 shape and boundaries of cell. In addition, the optional watershed approach was 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 besides the self-developed model, such as the traditional image segmentation method (watershed algorithm), the industry's widely acclaimed deep learning models (DeepCell[32], Cellpose[33] and Cellpose 2.0[34]). Some segmentation examples are shown in Supplementary Fig. 3D. To better identify different forms and types of cells, we also provided other cell segmentation methods besides the self-developed model, such as the traditional image segmentation method (watershed algorithm), the industry's widely acclaimed deep learning models (DeepCell[32], Cellpose[33] and Cellpose 2.0[34]). Some segmentation examples are shown in Supplementary Fig. 3D.

Molecule labeling correction

In order to obtain spatial high-precision single-cell omics data, we need to be able to accurately label transcripts to cells. First, the morphological tissue image was used to identify nuclei boundaries as described under nuclei segmentation. The specific transcripts that are contained within each nucleus are then assigned. A GMM[35] algorithm was next 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 the,

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

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 took a single cell as an example to describe the entire correction process. Based on the results of nucleus segmentation, we could locate the spatial position of the nucleus. First of all, GMM was used to fit the molecular distribution of current cells. Here we expanded the cell boundary and captured as much as possible the true distribution of data in cells rather than nuclei. According to the empirical value of cell area, we finally determined the fitting range of 100 pixel * 100 pixel. Then, 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 molecules that should belong to cells with high confidence.

Mouse olfactory bulb data analysis

Spatial single-cell gene expression matrices of mouse olfactory bulb data were generated by StereoCell, and the stitched image and gene expression map were exported in this process. The nuclei mask and StereoCell mask images were drawn by Python (3.8) script. The matrices were input into Stereopy (v6.0)[36] to calculate the total gene counts and number of gene types through the quality control function. In the filtering process, the cells with fewer than 150 expressed genes and more than 5% of mitochondrial genes were removed, and the genes present in less than 3 cells were removed. Then the gene expression matrices were normalized by the "SCTransform" function. The variably 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 two-dimensional data projections, followed by "Leiden" clustering to identify all clusters within the dataset. The silhouette coefficient[22] was calculated by the "metrics.silhouette_samples" and "metrics.silhouette_score" functions in the "sklearn" package. The related formula as shown below:

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

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

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

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[23], and then the dot plot of expression was drawn. The Tangram[24] was used to annotate the cell types with the reference[23]. We made the single cell expression matrix of each annotation into a bulk. The bulk was used to calculate Spearman correlation coefficient with the bulk made of the same annotated single cell expression matrix in the reference.

Mouse brain data analysis

Spatial single-cell gene expression matrices of mouse brain data were generated by StereoCell, and then use Stereopy (v0.6.0)[36] for downstream analysis. The cells with fewer than 200 expressed genes and more than 5% of mitochondrial genes were removed, and the genes present in less than 3 cells were removed. After filtering, the gene expression matrices were normalized by the "SCTransform" function. To reduce the data dimensionality, we summarized the variably expressed genes by PCA. Then we ran the UMAP algorithm to obtain two-dimensional data projections, followed by "Leiden" clustering to identify all clusters within the dataset. The silhouette coefficient[22] was calculated by "metrics.silhouette_samples" and "metrics.silhouette_score" functions in "sklearn" package. The quality control function of Stereopy was used to compute the total genes and number of gene types. We exported the gene map, nuclei-stained image, outlines of

molecule labeling and bin20 to visually show the segmentation effect of different methods. The Spatial-ID[25] with the Adolescent mouse brain[26] as a reference, was used to annotate the cell type. 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[26] and then the matrix plot of expression was drawn.

Data and Code availability

StereoCell was used to perform the workflow analyses in the paper, the 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, and 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, Wei Peng for their help.

Author contributions

Conceptualization: Mei Li, Huanlin Liu, Min Li, Shuangsang Fang.
Project administration and supervision: Susanne Brix, Xun Xu, Yong Zhang, Yuxiang Li.
Algorithm development and implementation: Huanlin Liu, Min Li.
Data datasets 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, Leying Wang;
Methods comparisons: Huanlin Liu, Min Li, Qiang Kang.
Biological interpretation: Jiajun Zhang, Shuangsang Fang, Ying Zhang, Qiang Kang, Fei Teng.
Data visualization: Tiantong Zhen, Shengkang Li.
Project coordination: Mei Li, Shuangsang Fang, Sha Liao, Ao Chen.
Manuscript writing and figure generation: Mei Li, Huanlin Liu, Min Li, Jiajun Zhang, Shuangsang Fang, Qiang Kang, Fei Teng, Dan Wang.
Manuscript reviewing: Susanne Brix, Yinqi Bai, Min Xie, Yong Bai.
Major revision: Shuangsang Fang, Jiajun Zhang, Dan Wang, Susanne Brix.
Final revision: Mei Li. All authors approved the manuscript.

Funding

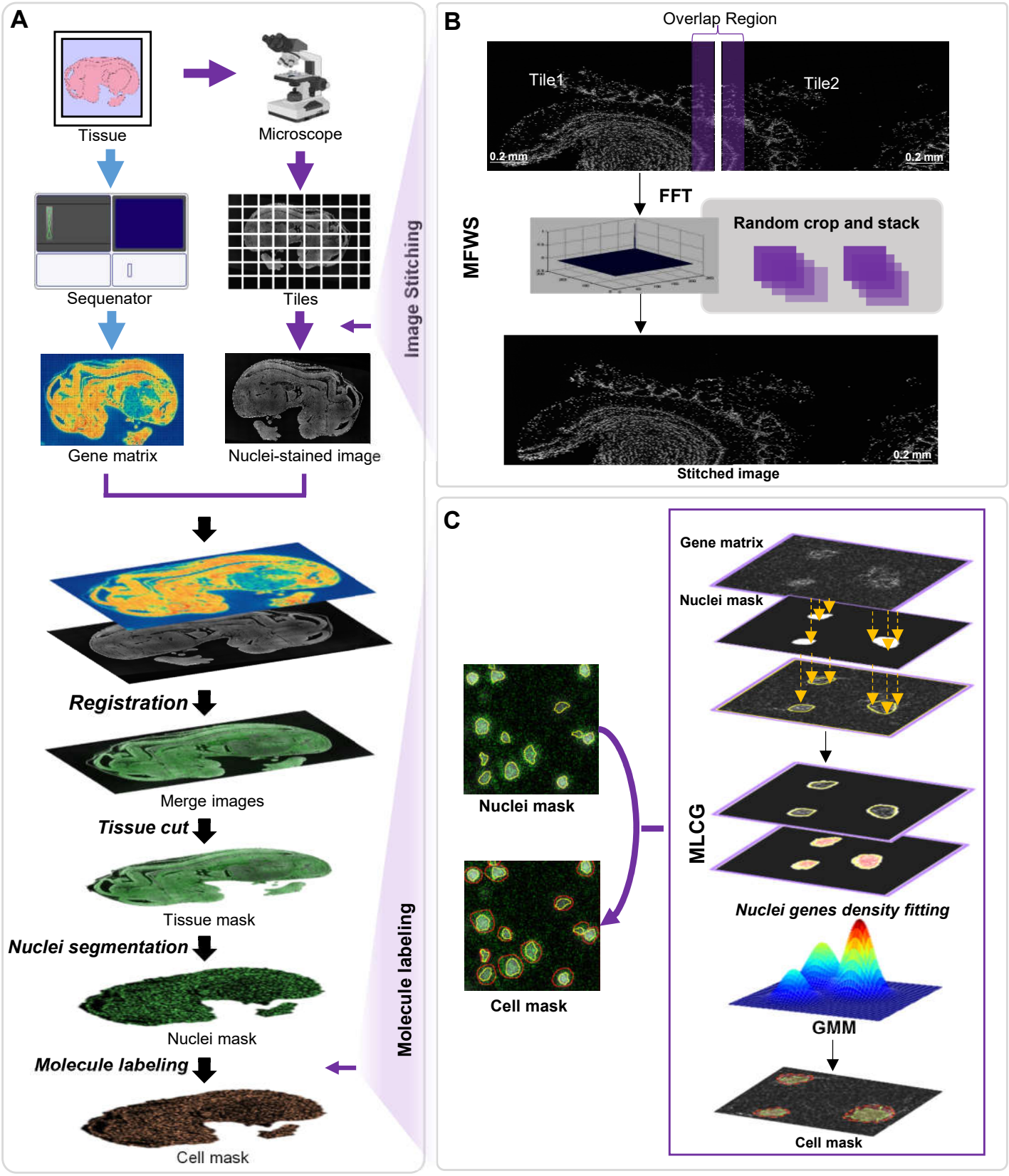
National Key Research and Development Program of China No. 2022YFC3400400

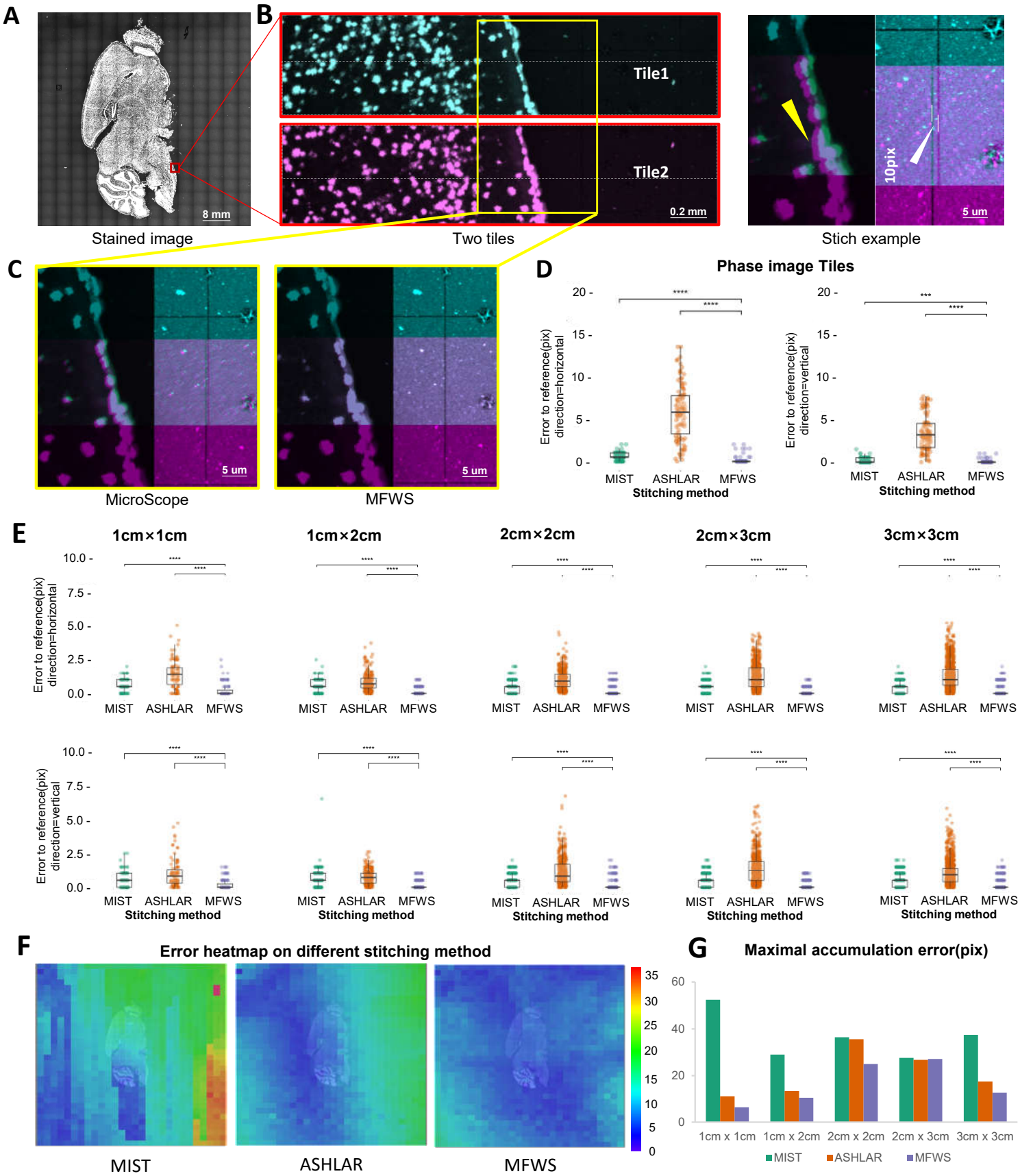
Reference

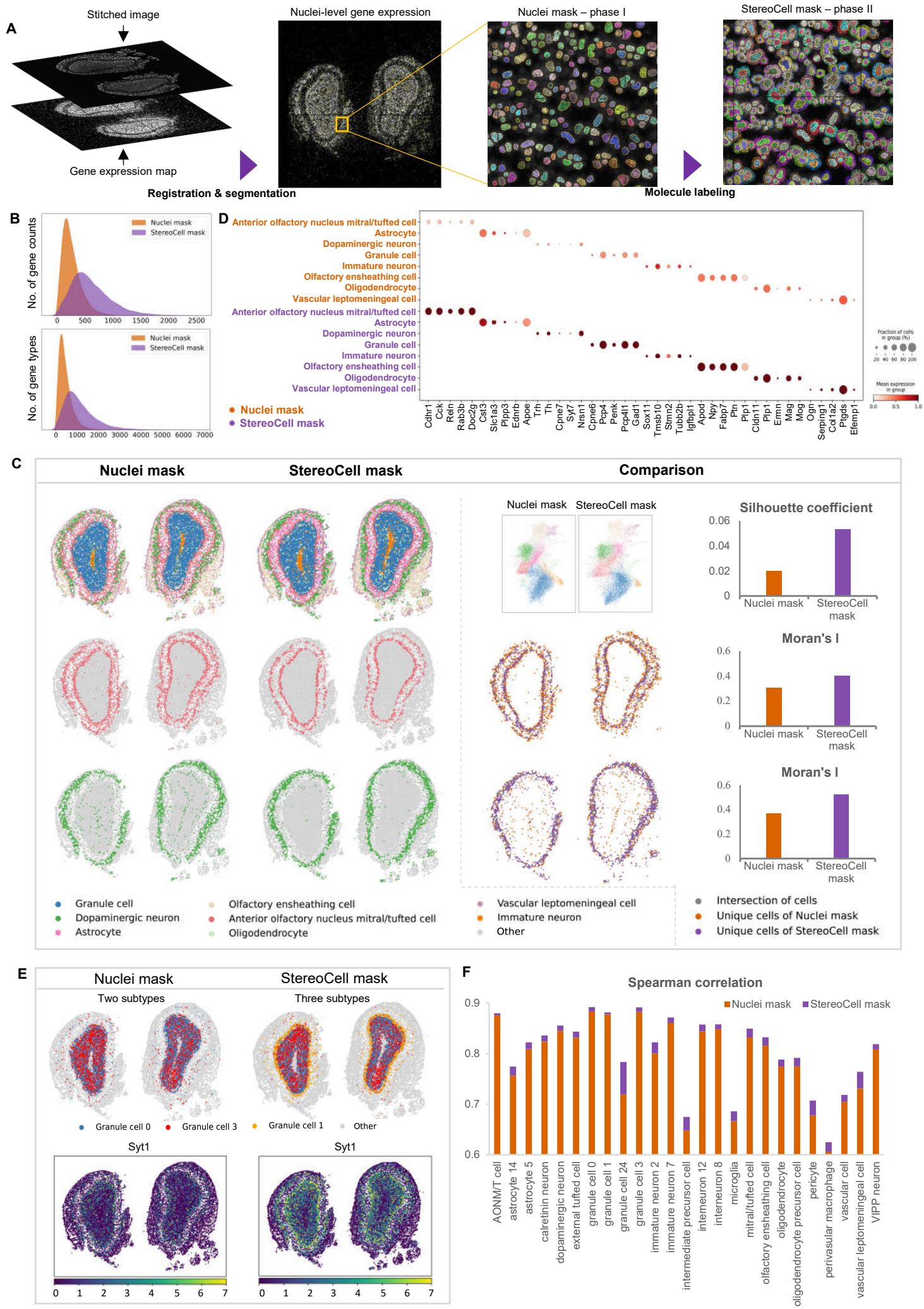
1. Chen, A. et al. Spatiotemporal transcriptomic atlas of mouse organogenesis using DNA nanoball-patterned arrays. Cell **185**(10), 1777-1792 (2022).
2. Lu, T. et al. Spatially resolved epigenomic profiling of single cells in complex tissues. Cell

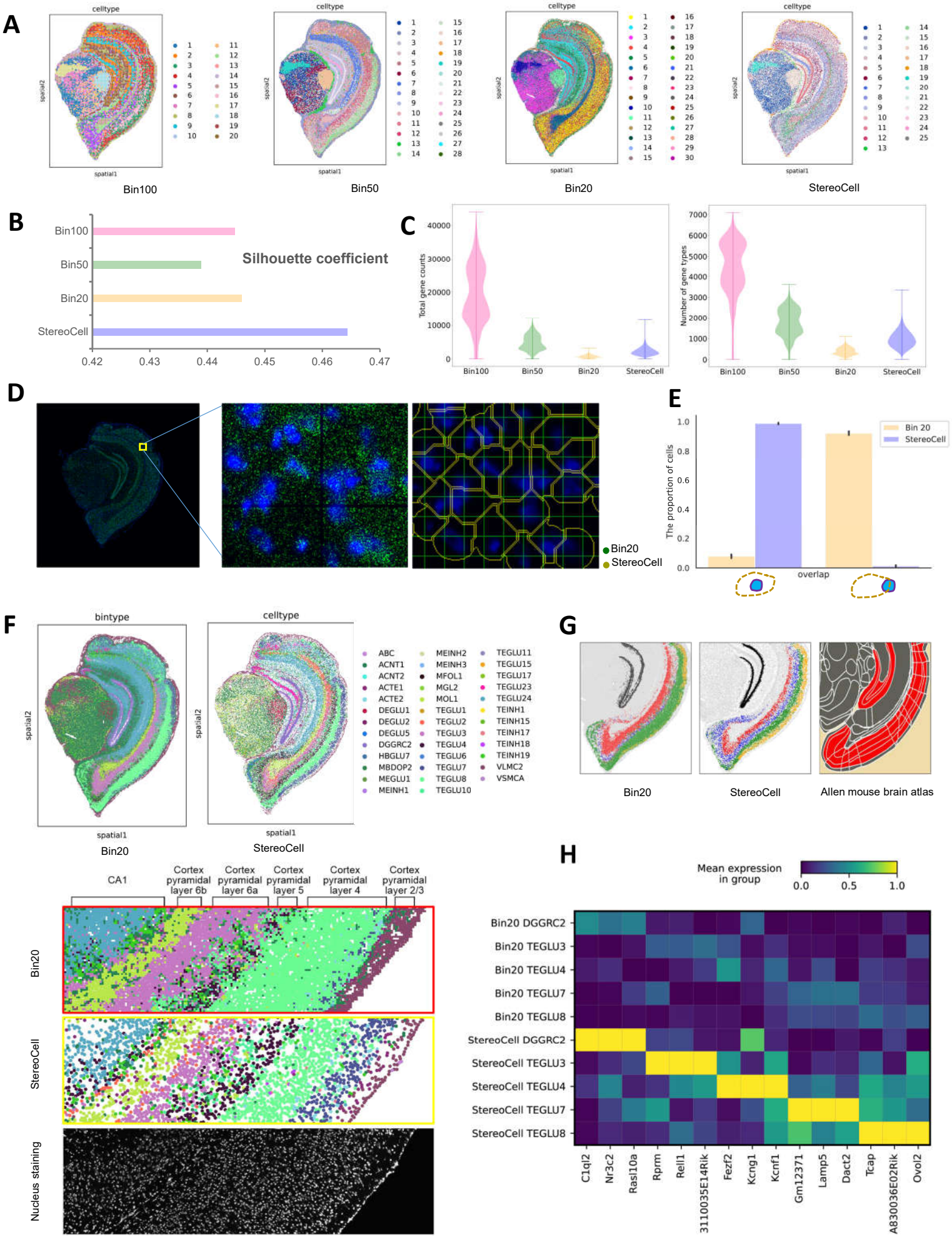
185(23), 4448-4464 (2022).
3. Erickson, A. et al. Spatially resolved clonal copy number alterations in benign and malignant tissue. Nature **608**(7922), 360-367 (2022).
4. Grünwald, B. T. et al. Spatially confined sub-tumor microenvironments in pancreatic cancer. Cell **184**(22), 5577-5592 (2021).
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. Shah, S. et al. In Situ Transcription Profiling of Single Cells Reveals Spatial Organization of Cells in the Mouse Hippocampus. Neuron **92**(2), 342-357 (2016).
21. Traag, V. A. et al. From Louvain to Leiden: guaranteeing well-connected communities. Sci. Rep. **9**(5233 (2019).
22. Rousseeuw, P. J. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. J. Comput. Appl. Math. **20**(53-65 (1987).
23. 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).

24. Biancalani, T. et al. Deep learning and alignment of spatially resolved single-cell transcriptomes with Tangram. Nat. Methods **18**(11), 1352-1362 (2021).
25. 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).
26. Zeisel, A. et al. Molecular Architecture of the Mouse Nervous System. Cell **174**(4), 999-1014 e22 (2018).
27. Azad, R. et al. Bi-Directional ConvLSTM U-Net with Densley Connected Convolutions. 406-415 (2019).
28. Ronneberger, O. et al. U-Net: Convolutional Networks for Biomedical Image Segmentation. 234-241 (2015).
29. Zhang, Z. et al. Road Extraction by Deep Residual U-Net. IEEE Geosci. Remote S. **15**(5), 749-753 (2018).
30. He, K. et al. Deep Residual Learning for Image Recognition. 770-778 (2016).
31. Zhang, H. et al. EPSANet: An Efficient Pyramid Squeeze Attention Block on Convolutional Neural Network. arXiv (2021).
32. 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).
33. Stringer, C. et al. Cellpose: a generalist algorithm for cellular segmentation. Nat. Methods **18**(1), 100-106 (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).


A
Tissue
Microscope
Sequenator
Tiles
Gene matrix
Nuclei-stained image
Registration
Merge images
Tissue cut
Tissue mask
Nuclei segmentation
Nuclei mask
Molecule labeling
Cell mask
Image Stitching
Molecule labeling
B
Overlap Region
Tile1
Tile2
0.2 mm
0.2 mm
FFT
MFWS
Random crop and stack
0.2 mm
Stitched image
C
Nuclei mask
Cell mask
Gene matrix
Nuclei mask
MLCG
Nuclei genes density fitting
GMM
Cell mask


A
8 mm
Stained image
B
Tile1
Tile2
0.2 mm
Two tiles
10pix
5 um
Stich example
C
5 um
MicroScope
5 um
MFWS
D
Phase image Tiles
Error to reference(pix) direction=horizontal
Error to reference(pix) direction=vertical
MIST
ASHLAR
MFWS
Stitching method
E
1cm×1cm
1cm×2cm
2cm×2cm
2cm×3cm
3cm×3cm
Error to reference(pix) direction=horizontal
Error to reference(pix) direction=vertical
MIST
ASHLAR
MFWS
Stitching method
F
Error heatmap on different stitching method
MIST
ASHLAR
MFWS
G
Maximal accumulation error(pix)
1cm x 1cm
1cm x 2cm
2cm x 2cm
2cm x 3cm
3cm x 3cm
MIST
ASHLAR
MFWS


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
No. of gene types
Nuclei mask
StereoCell mask
D
Anterior olfactory nucleus mitral/tufted cell
Astrocyte
Dopaminergic neuron
Granule cell
Immature neuron
Olfactory ensheathing cell
Oligodendrocyte
Vascular leptomeningeal cell
Nuclei mask
StereoCell mask
C
Nuclei mask
StereoCell mask
Comparison
Silhouette coefficient
Moran's I
Granule cell
Dopaminergic neuron
Astrocyte
Olfactory ensheathing cell
Anterior olfactory nucleus mitral/tufted cell
Oligodendrocyte
Vascular leptomeningeal cell
Immature neuron
Other
Intersection of cells
Unique cells of Nuclei mask
Unique cells of StereoCell mask
E
Nuclei mask
StereoCell mask
Two subtypes
Three subtypes
Granule cell 0
Granule cell 3
Granule cell 1
Other
Syt1
F
Spearman correlation
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


A
celltype
spatial1
spatial2
Bin100
Bin50
Bin20
StereoCell
B
Silhouette coefficient
Bin100
Bin50
Bin20
StereoCell
0.42
0.43
0.44
0.45
0.46
0.47
C
Total gene counts
Number of gene types
Bin100
Bin50
Bin20
StereoCell
D
Bin20
StereoCell
E
The proportion of cells
overlap
Bin 20
StereoCell
F
bintype
celltype
Bin20
StereoCell
G
Bin20
StereoCell
Allen mouse brain atlas
H
CA1
Cortex pyramidal layer 6b
Cortex pyramidal layer 6a
Cortex pyramidal layer 5
Cortex pyramidal layer 4
Cortex pyramidal layer 2/3
Bin20
StereoCell
Nucleus staining
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