Wu *et al. Genome Biology* (2024) 25:278
https://doi.org/10.1186/s13059-024-03421-5

Genome Biology

METHOD

Open Access

STASCAN deciphers fine-resolution cell distribution maps in spatial transcriptomics by deep learning

Ying Wu[1,2,3], Jia-Yi Zhou[1,2,3,4], Bofei Yao[1,2,3], Guanshen Cui[1,2], Yong-Liang Zhao[1,2,3], Chun-Chun Gao[1,2], Ying Yang[1,2], Shihua Zhang[5,6,7*†] and Yun-Gui Yang[1,2,3,4,8,9*†]

†Shihua Zhang and Yun-Gui Yang jointly supervised this work.

*Correspondence: zsh@amss.ac.cn; ygyang@big.ac.cn

[1] China National Center for Bioinformation, Beijing 100101, China
[5] NCMIS, CEMS, RCSDS, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China
Full list of author information is available at the end of the article

Abstract

Spatial transcriptomics technologies have been widely applied to decode cellular distribution by resolving gene expression profiles in tissue. However, sequencing techniques still limit the ability to create a fine-resolved spatial cell-type map. To this end, we develop a novel deep-learning-based approach, STASCAN, to predict the spatial cellular distribution of captured or uncharted areas where only histology images are available by cell feature learning integrating gene expression profiles and histology images. STASCAN is successfully applied across diverse datasets from different spatial transcriptomics technologies and displays significant advantages in deciphering higher-resolution cellular distribution and resolving enhanced organizational structures.

Keywords: Spatial transcriptomics, Cell annotation, Deep learning, Imputation, Multimodal data integration

Background

The spatial distribution and function of cells are intimately related, and their characterization can provide valuable insights into their function and potential impact on biological processes, including development and disease [1]. The emerging spatial transcriptomics (ST) technologies allow to capture gene expression while preserving spatial context in tissue, which improves our understanding of the structure and the cellular composition of different organs across different species [2–4].

Current ST technologies are typically divided into two categories: (1) imaging-based approaches that capture the probe-targeted genes through in situ sequencing or in situ hybridization, which can achieve single-cell resolution but are limited by low throughput and transcriptome coverage. Although some emerging imaging-based approaches have been extended to the whole transcriptome level with high throughput [5, 6], they

still suffer from the bias of pre-designed gene panels and are limited by the expensive specialized equipment and labor-intensive workflow; and (2) next-generation sequencing (NGS)-based approaches that capture transcripts from tissues combined with encoded spatial positional information before sequencing [1, 2]. By contrast, NGS-based approaches allow high throughput and unbiased coverage for measuring gene expression at the entire transcriptome level. More accessible commercial products of NGS-based approaches also promote their broad adoption [7].

However, the NGS-based approaches still suffer from several limitations. The significant one is the spatial resolution restricted by the area and sparsity of captured domains (defined as spots) [8]. For example, the spots of 10 × Visium are 55 μm in diameter which is too large to capture the cells at single-cell resolution [9]. Besides, there is a 100-μm center-to-center distance between spots, which leaves approximately 54–80% of the entire tissue uncharted [10], leading to a substantial reduction of spatial resolution of transcripts. In addition, DBiT-seq [4] was designed to produce smaller and more dense spots (10, 25, or 50 μm in width). Even though it provides higher spatial resolution, the transcriptional information in uncharted areas is still inevitably lost. Other approaches, such as Slide-seq [11] and Stereo-seq [12], also aim to achieve single-cell or subcellular resolution by using small and densely packed spots. However, they are also affected by the fact of multiple fractions of cells within a single spot.

Moreover, the uncaptured cells in uncharted areas lead to a limited spatial cellular resolution in not only 2D but also 3D levels. Currently, ST technologies enable us to depict cellular maps of planar tissue sections. By continuously stacking ST data from planar tissue sections in order, a 3D cellular map can be constructed to decipher the natural morphology of organs or organisms and provide a better interpretation of the heterogeneity among diverse tissue structures [13–15]. However, the high cost only allows a small portion of consecutive tissue sections to be sequenced, incurring the problem of uncharted areas along the *z*-axis and ultimate low 3D resolution [7, 14, 15].

To improve the cellular resolution of ST technologies, current computational methods typically resolve gene profiles to perform cell-type deconvolution by integrating ST data with signatures of single-cell references and annotating cell types for captured domains [16], such as Cell2location [17], Seurat [18], and RCTD [19]. However, the deconvolution methods suffer from the potential "dropout" resulting in only partial overlap between marker genes in the ST data and single-cell reference [20–22]. Besides, they are easily affected by the inaccurate cell-type annotations of the single-cell reference data, introducing bias during the selection of marker genes whose expression may be undetectable or fluctuant arising from cross-platform and batch effects, or is inconsistent in different literature [23, 24]. More importantly, these methods only aim to improve computationally cellular resolution within captured domains by inferring the proportion or abundance of cell type. However, how to enhance the spatial cellular resolution by predicting cell distribution in uncharted areas and imputing cell distribution between tissue sections in the *z*-axis remains to be solved.

In addition to the gene expression information derived from sequencing data, morphological information is also usually used to identify and characterize cell types of medical images [25–27]. Current deconvolution methods tend to focus on gene expression data, often overlooking the morphological information carried by images of ST datasets,

resulting in a waste of image resources. More importantly, morphological information can also contribute to increasing the accuracy of cell annotation. Some emerging computational methods integrating gene expression and morphological information have been developed for ST data. For instance, Tangram [21] synthesizes histological images to estimate the proportion of cells in each spot during single-cell deconvolution, and MUSE [28] combines transcriptional profiles and morphological features to characterize the cells and tissue regions. However, none of them involve the improvement of spatial cellular resolution in uncharted areas. Furthermore, some developed methods leverage ST data and histology images to enhance spatial gene profiles in captured domains and even uncharted areas [10, 29–32]. Nevertheless, due to the limitation of cell-type deconvolution methods mentioned above and the need for improving the accuracy of the gene expression-imputing models especially when performing predictions in complex tissue types [10, 29, 30, 33], there are still obstacles to directly using the enhanced gene expression profile to predict cell types.

To this end, we introduce STASCAN, a Spatial TrAnscriptomics-driven Spatially Cellular ANnotation tool. STASCAN enables cell-type predictions in uncharted areas across tissue sections and subdivided-resolved annotation of cells within captured areas, thereby greatly enhancing spatial cellular resolution. In addition, STASCAN succeeds in generating cell distribution maps solely from histological images of adjacent sections, enabling the construction of a more detailed 3D cellular atlas of organs with reduced experimental costs. Furthermore, we evaluated the applicability of STASCAN in diverse datasets from different ST technologies. We consistently observed significant improvements of STASCAN in cell granularity and comprehensive characterization of resolved cell patterns. For instance, STASCAN identified a micrometer-scale oval structure, which was confirmed as a smooth muscle bundle near the tracheal wall and has not been identified by other methods. Moreover, STASCAN provided a refined distribution of cell-type niches in human cardiac and mouse embryonic tissue, facilitating our understanding of the states of disease and development.

Results

Overview of STASCAN

STASCAN adopts a deep learning model to utilize both the spatial gene expression profiles from the ST technology and the corresponding histological images. Leveraging these multi-modal data, STASCAN depicts a fine-resolution cell distribution map in tissues by generating annotations of cell types for the spots or subdivided spots among captured and uncharted areas (Fig. 1a). Firstly, STASCAN extracts the spot images from slide images based on location information and infers highly reliable cell labels for each spot based on spatial gene expression using deconvolution during the pre-labeling process. Secondly, STASCAN constructs a base convolutional neural network (CNN) model (VGG16 architecture) [34] and trains the base CNN model using the cell-type labeled spot images as input. Additionally, STASCAN provides optional section-specific training, which can fine-tune the base CNN model through transfer learning to improve the prediction accuracy for a specific single section. Finally, through ample training,

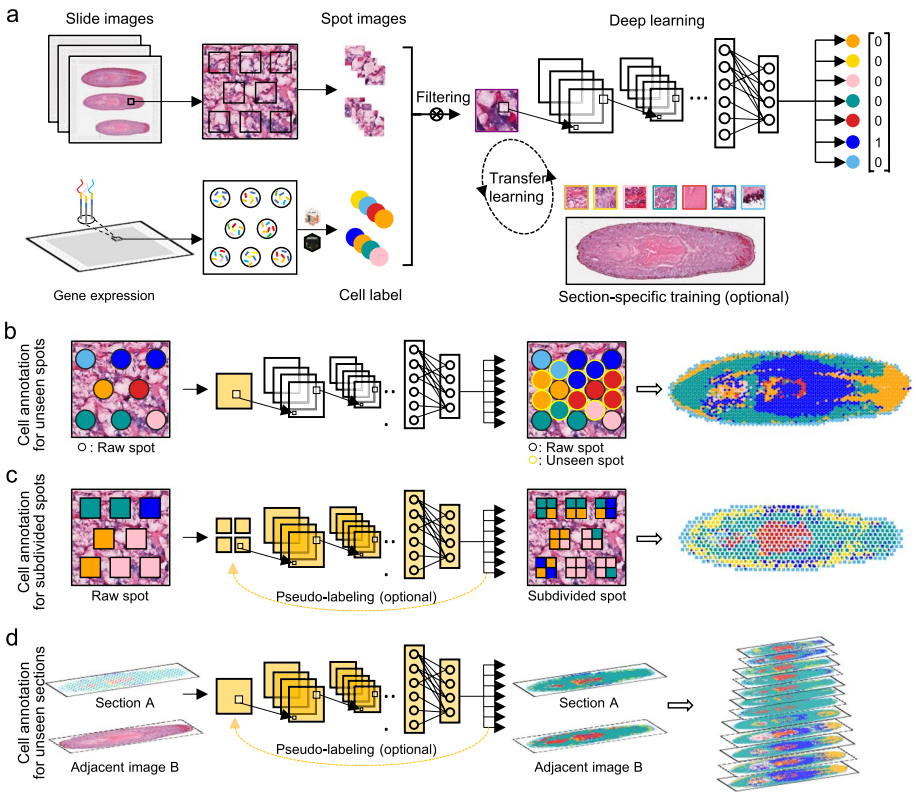

Fig. 1 Overview of STASCAN

STASCAN can accurately predict cell types solely based on histological images (Fig. 1a, Additional file 1: Fig. S1 and " Methods").

STASCAN is further designed into three application modules: (1) cell annotation for the embedded unseen spots in uncharted areas, which is based on learning image features from the measured raw spots, assigning predicted cell type for each unseen spot, and merging the unseen and raw spots to achieve a super-resolved cell distribution (Fig. 1b); (2) cell annotation for subdivided spots, utilizing features learned from subdivided-spot images with optional pseudo-labels to obtain a sub-resolution cell distribution (Fig. 1c); and (3) cell annotation for unseen sections, which learns the spot images from measured ST sections with optional pseudo-labels to predict the cell distribution on adjacent uncharted section images from the consecutive sections for constructing 3D cell models (Fig. 1d).

STASCAN enables more precise cell annotation and cell type prediction solely from images

To quantitatively evaluate the performance of STASCAN, we initially applied it to a comprehensive planarian (*Schmidtea mediterranea*) dataset generated by 10 × Visium technology, which includes ten sequenced ST sections (containing both spatial gene expression data and histological images) and nine unsequenced sections adjacent to ST sections (only containing histological images) [35]. Given the comprehensiveness of the planarian dataset, we first constructed a base model using 1829 spot images extracted from 10 collected sections to learn the features of 7 main cell types identified

by sequencing information, including epidermal, gut, muscle, neoblast, neuronal, parenchymal, and secretory cells (Additional file 1: Fig. S2a, b, Additional file 2: Table S1 and " Methods"). Although there was uncertainty in predicting neoblast and parenchymal cell types due to the scarcity of training samples of these two, most cell types were annotated with a recall rate of over 78% (Fig. 2a and Additional file 1: Fig. S2c). In addition, the learned model showed excellent accuracy in predicting cell types, with the area under the curve (AUC) calculated from the receiver operating characteristic (ROC) curves reaching as high as 0.936 to 0.996 (Fig. 2b). Besides, considering the potential batch effect among different ST sections, we performed section-specific training based on the base model (" Methods"). The results from the section-specific model showed a significant improvement in accuracy with higher AUC values compared with the base model, indicating that section-specific training is beneficial for the improved prediction performance of the whole frame (Fig. 2c).

We further compared the performance of STASCAN in predicting dominant cell types on raw spots with other methods, such as Cell2location [17], Seurat [18], and RCTD [19], using the planarian dataset. We initially annotated the cell types of each raw spot manually according to the morphologic features of corresponding spot images, which are considered the ground truth. We calculated the Kullback–Leibler divergences

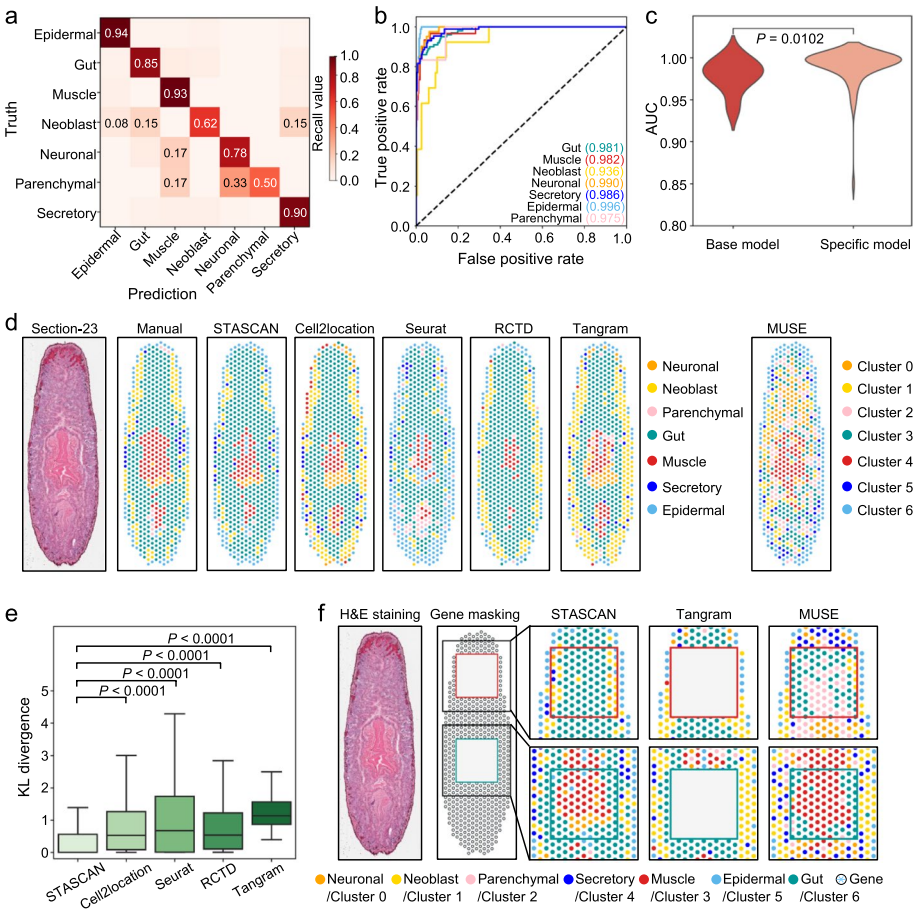

Fig. 2 Evaluation of STASCAN in the 10 × Visium planarian dataset

between the predicted cell distribution by different methods and the ground truth to evaluate the performance. STASCAN is highly consistent with manual annotations and significantly superior to other methods (Fig. 2d, e and " Methods"). We also observed that methods instead of STASCAN result in varied biases for cell annotation. For example, Cell2location could characterize most cell distribution but with a low sensitivity for epidermal cells; Seurat showed strong annotation bias for epidermal cells, leading to mislabeling for other cell types; and RCTD displayed some positive annotation but lost annotation information for most neuronal and secretory cells.

We also compared STASCAN with other methods utilizing both the morphological and transcriptional features for ST data analysis. Tangram [21] effectively illustrated the distribution of the majority of cells but exhibited a slight bias towards neuronal cells and a reduced sensitivity in detecting epidermal cells (Fig. 2d and e). On the other hand, MUSE [28] characterized tissue regions by identifying spot clusters, yet these clusters appeared relatively scattered and failed to represent corresponding structures in the planarian dataset (Fig. 2d). In contrast, our STASCAN displayed more precise performance in prediction, with the capability of accurately pinpointing the spatial distribution of seven main types of cells, in accord with their known biological functions [36] (Fig. 2d, and Additional file 1: Fig. S2d,e). For example, corresponding to clear tissue structures visible through hematoxylin and eosin (H&E) staining, the epidermal cells draw the contour of the planarian body, gut cells mark the location of the intestine, and muscle cells along with neuronal cells define the anatomy of the pharynx (Fig. 2d and Additional file 1: Fig. S2d).

Another significant advance of STASCAN compared to the existing methods is that STASCAN enables accurate cell-type prediction solely based on corresponding spot images. We compared the performance of STASCAN, Tangram, and MUSE in predicting cell types when morphological images are provided while gene expression information is masked (Fig. 2f). STASCAN achieved precise cell annotation predictions which were consistent with those made when both image and gene expression data were available. However, Tangram failed to predict cell types without gene expression data. Although MUSE achieved the characterization of cell clusters solely based on images, it was also disturbed by the absence of gene expression data, leading to incorrect predictions. For instance, MUSE identified two distinct cell clusters in the gut region that were disharmony with the manual annotations and also failed to identify the pattern of neuronal cells in the pharynx region (Fig. 2f). This comparison highlights the superiority of STASCAN and provides the basis for its utilization in three designed application modules in the subsequent steps.

STASCAN achieves super-resolution cellular patterns and improves 3D reconstruction in planarian

Next, we assessed the capabilities of STASCAN in different application modules using the planarian dataset [35]. When using Seurat and Cell2location to predict the spot cell types for pre-labeling, approximately half of the raw spots cannot be assigned with reliable cell labels (Additional file 1: Fig. S2a and " Methods"). This issue may be due to the noises generated by the complex signatures of gene expression in each spot, indicating the drawback of deconvolution in determining cell types (Fig. 3a). Choosing the other

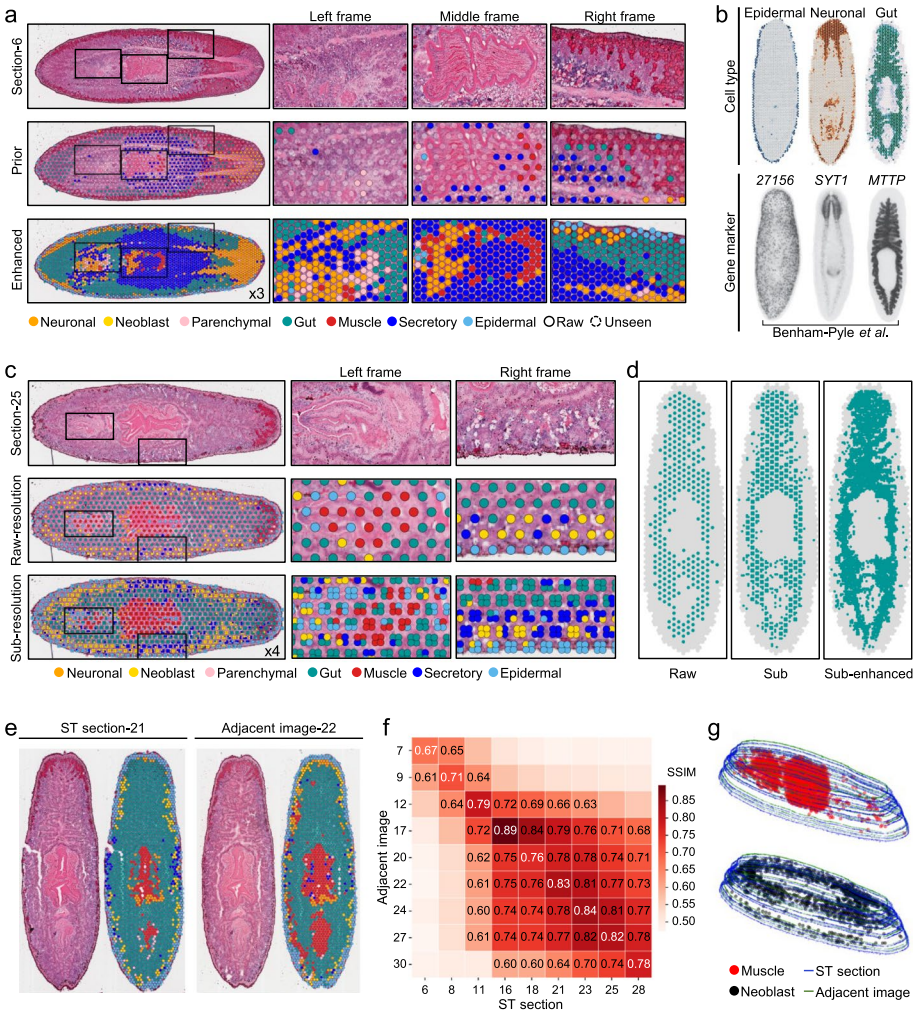

Fig. 3 STASCAN provides comprehensive and multidimensional cell annotation in the 10 × Visium planarian dataset

half of raw spots with credible labels as the prior spots to train the model, STASCAN achieved the reliable ability of cell type annotation based on images and depicted super-resolved cell distribution map (Fig. 3a). Firstly, STASCAN performed cell annotation for the unseen spots and demonstrated an enhanced resolution of cell distribution through combining both unseen spots and raw spots. The enhanced cell distribution map was highly consistent with H&E staining images, highlighting the relevant structures that were not shown at the raw resolution, such as the ventral nerve cord, genital chamber, pharynx, and contour (Fig. 3a and Additional file 1: Fig. S2d). Besides, it was highly consistent with the distribution of corresponding cell markers reported in previous literature [35, 36] (Fig. 3b and Additional file 1: Fig. S3a).

Furthermore, STASCAN pinpointed the composition of cell mixtures and their distinct locations at sub-resolution, effectively distinguishing cell types of each subdivided spot, and displaying a more detailed distribution of fine-grained cells (Fig. 3c). For example, STASCAN sensitively allocated secretory and neoblast cells around the contour into sub-divided positions according to the morphological differences. STASCAN also

identified muscle cells located at the junction of the pharynx and intestine at sub-resolution and were consistent with the biological priori information (Fig. 3c), which were not discovered at raw resolution from a group of gut cells. In addition, we utilized STASCAN to predict the enhanced sub-resolved distribution of gut cells, obtaining the fine-grained distribution of gut cells and reproducing the classical branching structure of the planarian intestinal tract (Fig. 3d). Moreover, we further compared the cell distributions generated by cell deconvolution methods on raw spots versus STASCAN on subdivided spots and found that STASCAN precisely assigns the fine-grained subdivided spots to their physically spatial locations with corresponding cell types, while the deconvolution method on the raw spots only resolved the composition of different cell types without determining the exact locations of the mixed cells (Additional file 1: Fig. S3b, c). Collectively, these results indicate that STASCAN significantly enhances cell granularity at sub-resolution, facilitating the depiction of instrumental sub-structure with fine-grained cells.

Last but not least, STASCAN achieved the prediction of cell distribution in the unseen sections only by H&E images using the learn features of the adjacent ST sections (Fig. 3e and Additional file 1: Fig. S4a-e). In line with the biological interpretation that the cell distribution between two consecutive sections should be similar, the structure similarity index measure [37] (SSIM) (ranging from 0.67 to 0.89) was linearly correlated with the spacing distance between adjacent images and ST sections, demonstrating the prediction accuracy of cell annotation for unseen sections (Fig. 3f, Additional file 1: Fig. S4a and " Methods"). Besides, we selected two adjacent ST sections that serve as the testing data (section-21) and ground truth (section-23). We trained two STASCAN models with section-21 and section-23, respectively, and employed these models to predict the cell distribution in section-23 solely based on the H&E staining image. Considering the manual cellular annotation in section-23 as ground truth, we observed that the model trained from section-21 enabled the prediction of cell distribution in section-23 and had a high correlation with both ground truth and the prediction generated by the model directly trained by section-23. These results firmly confirmed the reliability of STASCAN on cell annotation for unseen sections (Additional file 1: Fig. S4d, e). Finally, we generated raw and unseen spots from ST sections and adjacent images, applied STASCAN to predict cell types for those spots, and then reconstructed 3D models for different structures with cellular patterns (Fig. 3g and Additional file 1: Fig. S4b, c). The model displayed the cell distribution in three-dimensional, with improved cellular resolution in spatial and promoted utilization of staining images without ST sequencing.

STASCAN identifies well-defined boundaries of distinct cell layers in the human intestinal tissue

To further evaluate the performance of STASCAN for ST datasets on different tissue architectures, we applied STASCAN to human intestinal datasets generated by 10 × Visium technology. These datasets consisted of eight slides sampled at diverse sampling locations and time points [38]. We trained each slide using STASCAN with different sizes of prior spots, ranging from 297 to 1551, and observed stable performances across all sizes (Additional file 1: Fig. S5a-c, Additional file 2: Table S1 and " Methods").

Next, we applied STASCAN to predict cell types of unseen spots. Compared with the cell distribution of prior spots annotated by other methods, STASCAN stratified cell populations to finer regional layers. For example, in comparison with other methods only roughly distinguishing the distribution of different cells, STASCAN was able to delineate the borders of cell layers among the intestinal epithelium, fibroblasts, and muscularis, greatly enhancing the cellular spatial patterns (Fig. 4a and Additional file 1: Fig. S6).

Then, we used STASCAN to draw the spatial distribution map of fine cell subtypes in the human intestinal tissue (Fig. 4b). Actually, we labeled three anatomical layers of the intestinal tissue related to the morphological structures of H&E staining, including muscularis, fibroblasts, and epithelium layers, listed based on their distance to the intestinal edge (Fig. 4c). When evaluating four epithelium subtypes occupying an absolute

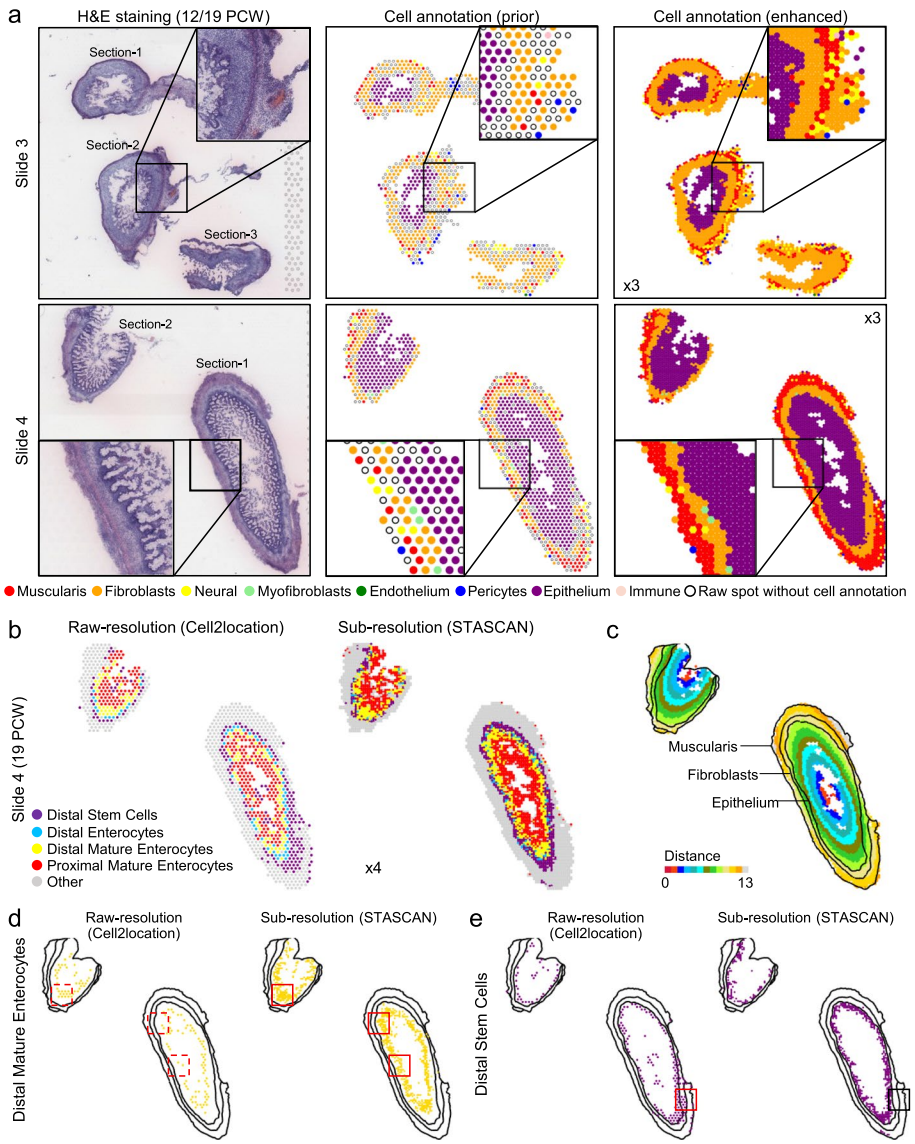

Fig. 4 STASCAN depicts spatial layers of distinct cell types in the 10 × Visium human intestinal dataset

proportion of the epithelium layer, we found that compared to the results of alternative method at raw resolution, STASCAN not only highlights the precise distribution of these subtypes but also accurately locates the positions of distinct subtype cells (Fig. 4b, d, e and Additional file 1: Fig. S7a). For instance, distal epithelium subtype cells tend to gather closer to the boundary of the epithelium layer and the fibroblast layer, and proximal epithelium subtype cells are prone to assemble to the surface of the epithelium layer at sub-resolution. Besides, distal stem cells were correctly predicted to be located in the epithelium layer at sub-resolution; however, at raw resolution, a part of distal stem cells was abnormally predicted to be located in the fibroblasts layer (Fig. 4d, e and Additional file 1: Fig. S7b).

In addition, we performed STASCAN on a pair of identical and adjacent sections of the intestinal tissue to valid the cell annotation for unseen sections (" Methods"), and the high correlation between them further confirmed the accuracy and reliability of the prediction (Additional file 1: Fig. S8a, b).

STASCAN uncovers a novel structure in the human lung tissue

Despite the limitations of raw spatial resolution of ST technologies, STASCAN assists in enhancing cellular patterns and rediscovering the micrometer-scale structure. Here, we applied STASCAN on the 10 × Visium human lung dataset [39], which sampled from the proximal airway. We previously redefined 13 reference cell types to better illustrate the organizational structure and annotated 822 ST spots with seven dominant cell types to train the STASCAN (Additional file 1: Fig. S9a-e, Additional file 2: Table S1 and " Methods"). With enhanced resolution, STASCAN showed more precise cellular and structural patterns of human lung tissue. Besides, we observed that STASCAN sensitively identified a micrometer-scale oval-shaped structure that was highly consistent with the H&E staining images, which was confirmed as the smooth muscle bundles adjacent to the tracheal wall. However, this structure was not evident at the raw resolution of prior spots, highlighting the capability of STASCAN to reveal refined structures of spatial regions (Fig. 5a and Additional file 1: Fig. S9c).

Furthermore, we compared the ability of STASCAN with other methods in revealing tissue structures with cellular patterns. The results showed that STASCAN could depict the silhouette of the airway with basal and neuroendocrine cells and the cricoid cartilage structure surrounded by goblet cells, mucous cells, smooth muscle cells, pericytes, etc. Moreover, the smooth muscle tissue traced by smooth muscle cells and pericytes in the left bottom of the slide was only identified by STASCAN (Fig. 5b and Additional file 1: Fig. S9c). Briefly, compared with the cell distribution pattern identified by other methods, STASCAN displayed superior advantages in identifying and characterizing spatial specific structures, which better reflects the anatomical structure after imputation.

STASCAN depicts the pathological spatial structural variations of human cardiac tissue after myocardial infarction

To explore whether further improved functional applications in ST data analysis can be achieved by STASCAN, we adopted this approach to reanalyze the 10 × Visium human cardiac datasets [40], which included 17 slides from normal hearts and the

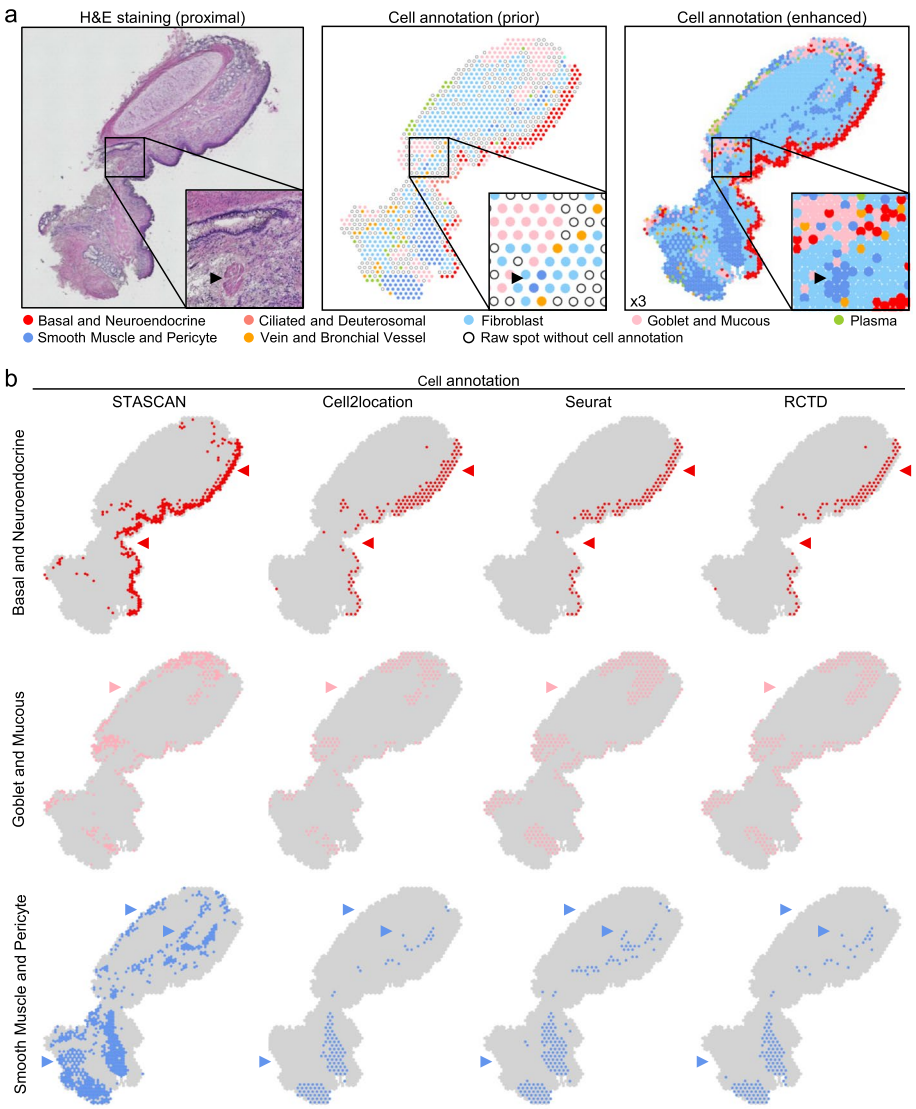

Fig. 5 STASCAN demonstrates the special structure from the 10 × Visium human lung data

pathological ones after myocardial infarction (Additional file 1: Fig. S10a-b, S11, Additional file 2: Table S1 and " Methods"). We first grouped these slides according to their sampling regions [40], including normal non-transplanted donor hearts as controls, necrotic (ischemic zone and border zone), and unaffected regions (remote zone), and regions at later stages after myocardial infarction (fibrotic zone).

Based on the ability of STASCAN in generating cell maps solely from histology images, we considered STASCAN not only a valuable algorithm to enhance the spatial cell distribution but also a constructive tool to imputing the cellular pattern of missing regions that failed to capture transcripts normally during ST sequencing. To evaluate the performance of STASCAN in imputing missing cellular distribution, we first selected half of the ST spots in slide_ACH003 as missing spots and then trained STASCAN with the other half spots. After that, we performed STASCAN to predict the whole cell distribution of the slide_ACH003. Considering cell annotations among

prior spots as ground truth, we observed that STASCAN could well reproduce the cell distribution among the missing spots, especially replicated the structure of vasculature surrounded by fibroblast and vSMC cells, which solidly indicated the reliability of STASCAN in imputing cellular patterns in the missing regions (Additional file 1: Fig. S12a, b).

After the evaluation, we focused on two slides with a lot of missing spots that were filtered out due to scanty genes and unique molecular identifiers (UMI) measured in the original literature [40] (Fig. 6a and Additional file 1: Fig. S13-S17a). In these two slides, STASCAN not only more accurately depicted the cell distribution pattern of tissue structures but also predicted the potential cell distributions in the missing areas only from images. Especially for the serious missing in the slide_ACH0010 sampled from the ischemic zone, STASCAN better imputed the reasonable diffusion of the cardiomyocyte,

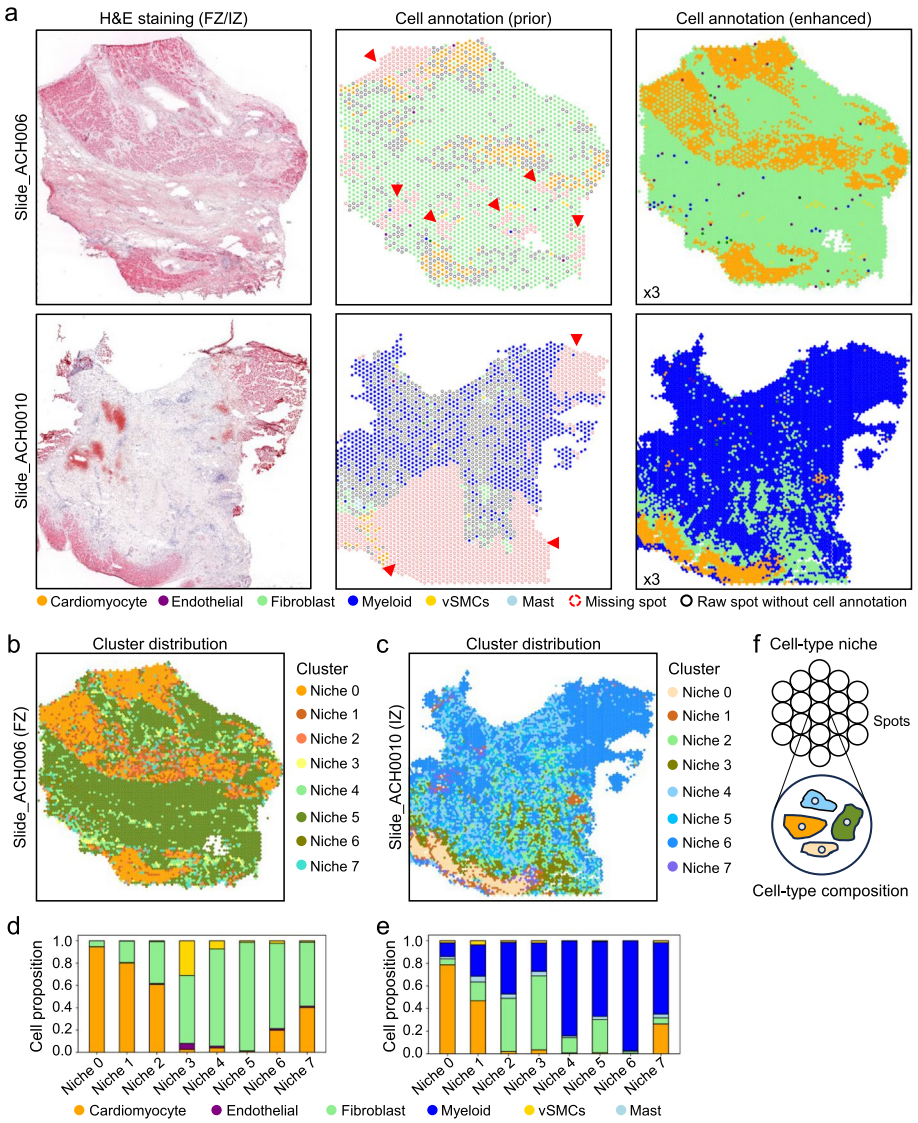

Fig. 6 STASCAN reveals cell-type niches in the 10 × Visium human cardiac data

fibroblast, and myeloid cells in line with the histological morphology (Fig. 6a). The proximity of the latter two cells indicated a strong dependence between them on the areas of immune cell infiltration and scar formation [40].

We then explore the spatial structural variations in these two slides by performing unsupervised clustering for spots based on the composition of cell annotations predicted by STASCAN and then mapped the clusters, defined as cell-type niches, to the spatial regions (Fig. 6b, c, f and Additional file 1: Fig. S13-S17b). Through redrawing the spatial distributions for these cell-type niches, the cardiac tissue manifested more delicate spatial patterns compared with the dominant cell annotation, which were consistent with histological morphology and detailed structural variations observed during physiological and pathological processes.

Besides, these cell-type niches based on diverse cell propositions revealed more elaborate cell-interacting microenvironments with potential biological insights (Fig. 6d, e and Additional file 1: Fig. S13-S17c). For example, we observed myogenic cell-type niches (0, 1, and 2) mainly displaying characteristics of cardiomyocyte cells and fibrotic cell-type niches (3, 4, 5, 6, and 7) mainly presenting characteristics of fibroblast cells, in the slide_ACH006 sampled from the fibrotic zones. On the aspect of spatial distributions, the myogenic cell-type niches could jointly characterize the myocardial structure and the fibrotic cell-type niches distinguished by the proportion of fibroblast cells indicated different fibrosis processes during the lesion. Especially, there was a measure of characteristics of vSMCs and endothelial cells in niches 3 and 4 for depicting the spatial structure of the cardiac vasculature. In addition, in slide_ACH0010, we observed inflammatory cell-type niches (4, 5, 6, and 7) which mainly exhibit characteristics of myeloid and mast cells, apart from myogenic cell-type niches (0 and 1) and fibrotic cell-type niches (2 and 3). These three types of niches took up distinct spatial regions, but there was niche 7 located at the intersection, in line with the proposition of cardiomyocyte, fibroblast, myeloid, and mast cells in niche 7. Especially, niches 2, 3, 4, and 5 showed co-enrichment between myeloid and fibroblast cells, in accordance with the role of macrophages in fibroblast activation [41] and fibroblast cells in macrophage attraction [42]. Overall, STASCAN expands the application of niche distribution and provides better insights into understanding cellular microenvironment interactions.

STASCAN deciphers the intricate tissue organization throughout the developmental stages of the mouse brain

We further test whether STASCAN is also applicable to ST data derived from various technologies. We first employed STASCAN on an embryonic mouse brain dataset from MISAR-seq [43], a microfluidic indexing-based spatial technology motivated by DBiT-seq [4] with both high-quality image and sequencing data (Additional file 1: Fig. S18a-c, Additional file 2: Table S1 and " Methods"). Notably, although the H&E images adopted in this dataset were obtained from the adjacent tissue slide which causes a partial disharmony between the actual gene expression pattern and morphological images, STASCAN still achieved excellent results. When compared with the cellular distribution annotated by RCTD [43], STASCAN significantly improved the cellular resolution with highlighted characteristics of tissue structures (Fig. 7a and Additional file 1: Fig. S18a-c). For example, the enhanced distribution pattern of forebrain GABAergic neurons was

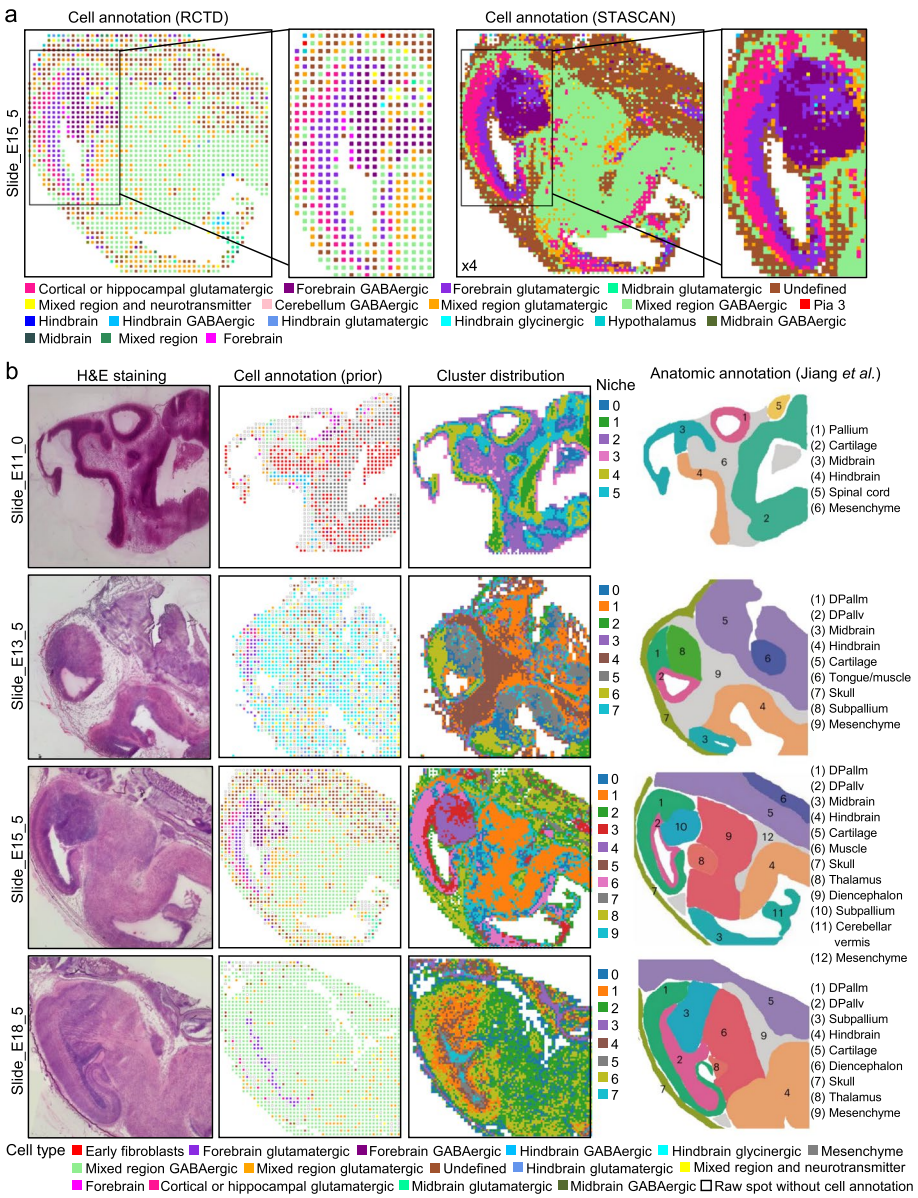

Fig. 7 STASCAN revealed major anatomical tissue regions in the mouse brain dataset generated from microfluidic technologies

associated with the subpallium, and a group of the forebrain glutamatergic and cortical or hippocampal glutamatergic neurons spotlighted the dorsal pallium of the forebrain at enhanced resolution (Fig. 7a).

Furthermore, we respectively generated cell-type niches for different development stages of mouse brain tissue and mapped them to spatial regions (Fig. 7b and " Methods"). Using the manual anatomical annotations of major tissue organizations from H&E images as the ground truth [43] (Fig. 7b), we compared the cluster distribution of cell-type niches to the raw resolution distributions of cell annotations on prior spots. The cluster distribution of cell-type niches generated by STASCAN more remarkably recapitulates the tissue organization in the developing mouse brain than the ones from

raw resolution. Especially for the E18.5 embryonic mouse brain tissue, in contrast to cell annotation at raw resolution which showed an unrecognized organization in nearly entire brain region, STASCAN clearly defined the major tissue domains by cell-type niches clustering (Fig. 7b). Collectively, these results illustrate the strength of STASCAN in highlighting tissue structures and redrawing finer organizations using ST data from various technologies.

Discussion

Although current spatial transcriptomics has achieved remarkable progress in deciphering the distribution and interaction of cells in tissues, spatial resolution limitations hinder its broader application. The development of computational algorithms is crucial for the analysis of spatial transcriptomic data. Here, we have developed STASCAN, a tool that integrates histological images and spatial gene expression to depict comprehensive cell atlases with enhanced spatial resolution.

Compared with the traditional image-based CNN models, STASCAN integrates gene expression and image information, in which gene expression helps to automatically label spots used for training without manual annotation. Meanwhile, images extend a new perspective and supplement the information that sequencing cannot provide. Additionally, STASCAN can accurately extract features and improve the accuracy of the model through transfer learning and pseudo-labeling. Thus, STASCAN effectively resolves images and uses the image information as the main reference for cell type determination. This approach helps to predict cell distribution and increases the number of spots in the point cloud representing cells, ultimately facilitating the construction of fine-resolved 3D spatial cell distribution maps.

Through fully utilizing both sequencing and image data, STASCAN not only accurately annotates the cell types on raw ST spots but also enables the prediction of cell distribution in the uncharted areas across tissues. Besides, STASCAN fully utilizes the image information of continuous sections to construct 3D tissue models without expensive ST experimental costs, laying the foundations for expanding the future frontier in 3D cellular atlas. We performed STASCAN in different species and tissues and observed a substantial advance over current methods in depicting spatial cellular patterns and tissue structures. In planarian, STASCAN successfully predicted cell distribution of unseen spots and sections in uncharted areas, leading to the construction of more detailed 3D cell distribution models. In addition, we found that STASCAN discovers precise tissue structures with enhanced cell annotation. In the human lung tissue, STASCAN enhanced spatial resolution and identified a micrometer-scale structure that was made up of a group of smooth muscle cells located near the tracheal wall and involved in the regulation of intra-pulmonary airway caliber [44], while the raw resolution is too coarse to resolve the same specific structure.

The low resolution of NGS-based spatial transcriptomics is also attributed to the spacing of captured spots, which are not small enough to achieve single-cell resolution [8]. Most current methods estimate the proportion or abundance of each cell type from the cell mixtures on the capture spots but are incapable of assigning each cell into exact locations within each spot. Although Tangram provides a computer vision module that segments cell nuclei based on histological images and predicts the

cell type for each cell in the spot, it performs random cell assignment for the segmented maskings within the spot and still cannot definitely resolve cell position [21]. In comparison, STASCAN introduces a predicting module at sub-resolution, allowing the assignment of cell types to sub-divided spots based on the tissue morphological features corresponding to the exact positions. In the human intestinal tissue, STASCAN precisely pinpointed the fine-grained cell subtypes between distinct cell layers.

Additionally, we are aware of some developing methodologies that aim to enhance gene expression resolution by integrating ST data and histology images. These methods usually interpret each gene's spatial expression as an image and improve gene expression resolution by regularly distributing the gene expression of each spot into corresponding pixels [10, 29, 30]. Logically, these methods can be assembled with the deconvolution approaches into the pipeline to finish the task of cell-type annotation. However, they are easily affected by the high dimensionality and sparsity of the raw data and usually lack high accuracy, especially when performing predictions in complex tissue types. Besides, the deconvolution methods are easily affected by the potential "dropout" of marker genes in ST data and the inaccuracies of single-cell reference data [20, 21]. Although some gene expression enhancing methods have achieved cell annotation by a marker-gene-based strategy [10, 29], they may still suffer from the inaccuracies resulting from the "dropout" and biases of marker genes like deconvolution methods (Additional file 1: Fig. S19). In contrast, STASCAN employs that cellular distributions can be reflected on the histological images, where a skillful pathologist can directly identify diverse cells from the H&E staining images. Given that the biological inference is more solid and direct, STASCAN may provide a better strategy to enhance the cellular resolution in ST data with higher interpretability and accuracy.

The advantages of STASCAN enable it to have a broader range of applications. For example, STASCAN can effectively predict cell distribution of some regions with missing sequencing data due to a low number of captured UMIs and genes. In human cardiac tissues, STASCAN could fill in the missing data and help to display an extensive spatial proximity distribution of fibroblast and myeloid cells in the vicinity of scar tissue. This provides a better understanding of the relationship between fibroblasts and immune cells in myocardial infarction, such as immune factors in stimulating fibroblast transformation [45]. Additionally, the cell proposition annotations generated by STASCAN are supported to cluster different cell-type niches, providing novel insights into the microenvironment at the cellular resolution. In human cardiac tissues, STASCAN generated distinct niches including inflammatory, myogenic, and fibrotic cell-type niches, with the best representation of the dynamic diversity of cells in the microenvironment during myocardial infarction. In the embryonic mouse brain tissue, STASCAN revealed major anatomical tissue organizations across the development stages, with gradually increased complexity of brain structures during brain development.

Logically, STASCAN can be employed in all ST technologies that utilize microarrays, microfluidics, or similar designs to annotate cell types for unseen spots. It can also be used in other NGS-based ST technologies to annotate cell types for subdivided spots and unseen sections, only if the technology provides both H&E staining

images and transcriptional data. Due to the limitation in collecting high-quality H&E staining images of published ST datasets, we only applied STASCAN in ST datasets generated by the 10 × Visium and MISAR-seq technologies, which respectively stand for microarray-based and microfluidic-based technologies. Nevertheless, the technical framework of STASCAN is well-designed and provides convenient interfaces for further implementation in other NGS-based ST technologies.

Although STASCAN shows notable superiority and advantages as a cell annotation tool with enhanced resolution for ST data, further refinements are certainly needed to improve its performance. One issue is the possible inaccurate prediction of rare cell types owing to fewer samples. Although we leveraged some methods to mitigate the effect of sample imbalance, such as under-sampling for the categories with larger sizes, there is still a need to incorporate more advanced strategies to solve this issue. Besides, tissue samples with severe necrotic regions may also lead to lower training accuracy of STASCAN due to the irregular changes of cellular morphology, which remain to be further improved.

Conclusions

In summary, STASCAN is a universal and precise cell-type prediction method, and it integrates image and expression information to determine complex spatial distributions in tissues. The diverse applications of STASCAN have demonstrated its superior performance in enhancing spatial resolution and discovering novel structures, offering the potential for resolving cell type distribution in various types of tissues across different stages of development, regeneration, and disease. The positional relationships of diverse cell types determined by STASCAN provide new insights into cell crosstalk critical in orchestrating organismal development and homeostasis. Furthermore, STASCAN fully utilizes more easily acquired image datasets under various biological conditions and can potentially be used to infer subtypes of pathological cells and further link spatial cellular distribution to disease diagnosis cost-efficiently through extensive data training, eliminating the need for sequencing.

Methods

Preparation of spot images

We implemented quality control for H&E staining images corresponding to ST data. We discarded the unevenly stained, unfocused, and low-resolution images. We then cropped the retained H&E staining images into a set of spot images based on the spot coordinates. Note that the side length of each spot image was equal to the diameter length of the spot. Finally, to maintain the purity of the training dataset, we only kept the spot images under the tissue. Specifically, each spot image used to train the model for subdivided-spot cell annotation was a subdivided quarter from the corresponding standard ST spot images. The spatial locations of the subdivided spots were assigned by their physical spatial coordinates, and the side length of the subdivided spot images was half of the raw spot images.

Pre-labeling of cell types and preparation of prior spots

Before training, the spot images need to be set with credible predefined labels. They were defined as prior spots. We provided two strategies to label the cell types for each spot as follows:

Annotated collectively by two different pre-labeling methods. Considering the differences in accuracy, generalizability, and efficiency of different methods [16], we used Cell2location [17] and Seurat [18] to jointly label the cell types for each ST spot by deconvolution. Strictly, we selected spots annotated as the same dominant cell types by both methods as prior ones and paired the images and the cell types of these spots.

Annotated by single pre-labeling method. Given that the stringent selection based on the shared annotation described above may result in fewer available spots, and the effectiveness of deep learning depends in part on the number of training samples, we also proposed a looser annotation strategy that used only a single method that suitable for the targeted dataset to annotate cell types for ST spots. The spots where the proportion of dominant cell types reached a user-defined threshold or the proportion of dominant cell types exceeded the given multiple of the proportion of secondary cell types were selected as prior ones. We considered a spot with a high proportion of dominant cell type, generally above 90%, as a reliable one that can be subdivided into four sub-spots with the same dominant cell type. That is considered a prior spot during cell annotation for subdivided spots.

Construction of the CNN model

The architecture of the base CNN model

In the field of visual recognition, CNN is a classic deep-learning method that can extract features from images with several convolution and max pooling layers and enable image classification after training. All CNN models used in this article were constructed employing TensorFlow [46] (version 2.9.0). We adopted the VGG16 architecture, which is a classical CNN model with sixteen weight layers [34]. The weights of convolution and max-pooling layers were trained with the ImageNet dataset by TensorFlow previously. The number of output features in VGG16 was adapted to the number of cell types. During the training process, the loss value was calculated with the categorical_crossentropy function, and the effect of learning was assessed by the accuracy metric. Moreover, the stochastic gradient descent (SGD) optimizer was set with the hyperparameters of momentum = 0.9 and learning rate = 1e − 3.

Training and evaluation

Firstly, we filtered out the prior spots with label-class sizes less than 10 (by default) and randomly selected 500 prior spots with label-class sizes more than 500 (by default) to mitigate the possible class imbalance. Next, we used 80% of the prior spots as a training dataset and 20% as a testing dataset, and the latter one was also considered as validation data during the training process. The CNN model automatically adjusted the training and testing images into a fixed size (default: 224×224 pixels) with an adjustable batch_size parameter (default: 32). Additionally, we applied image augmentation using the ImageDataGenerator function (rotation, shifting, rescaling, shear, zooming, and

horizontal flipping) to fully utilize the training dataset. Each training process was performed for 50 epochs by default to steady the loss value and accuracy. For evaluation, we calculated the recall (TPR) and AUC values for each cell type, which reflect the measurable effects of the training process. The recall and AUC values are calculated as follows:

$$\mathrm{TPR} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}}$$

$$\mathrm{FPR} = \frac{\mathrm{FP}}{\mathrm{FP} + \mathrm{TN}}$$

$$\mathrm{AUC} = \int_0^1 \mathrm{TPR}(\mathrm{FPR}^{-1}(x))dx$$

where TP, FN, FP, and TN are true positives, false negatives, false positives, and true negatives, respectively and FPR^{-1} is the inverse function of FPR.

Section-specific training (optional)

We provided optional section-specific training to learn the characteristics of multitudinous spot images from consecutive sections. We first trained a base CNN model with the settings and parameters mentioned above and evaluated its accuracy and whether there were biases between the partitioned training and testing datasets by a fivefold cross-validation. After finishing the model evaluation by fivefold cross-validation, we used all prior spots instead of 80% to train a new base model and no longer used the testing dataset during the model training process. After that, allowing for the potential batch effect between different sections, we adopted the transfer learning to leverage generalizable knowledge from the new base model and retrained the inherited model using prior spots extracted from corresponding sections for each section, respectively. This model can better characterize spot images from a particular section.

Pseudo-labeling (optional)

To improve the annotation accuracy for subdivided spots and unseen sections predicted by the base CNN model, we provided a function to get pseudo labels of these spot images. After constructing and training the basic model, we initially utilized the trained model to roughly infer cell types of subdivided spots or simulated spots on unseen slides. Next, we selected the spots with high scores for the probability of predicted cell types, which were over 0.90 by default, and relabeled these spots with the predicted ones. Lastly, we retrained the basic model with those new spots by transfer learning to improve its accuracy.

Transfer learning

Transfer learning was used in the optional section-specific training and pseudo-labeling involving cell annotation for subdivided spots and unseen sections. During the transfer learning phase, we retrained the inherited model without freezing the model.

Prediction under distinct modules

Cell annotation for unseen spots. At first, to remove negative impacts caused by engraving errors, we adjusted the coordinates of the spots and made them arranged in a line for each row. For 10 × Visium datasets, considering the area connected by every three adjacent ST spots between two rows as a standard equilateral triangle, we simulated an imputed spot that tangent to the three ST spots for each triangular area and defined it as an unseen spot. For other ST datasets based on microfluidic design, considering each single raw ST spot as the upper left quarter of a square, we simulated the other three quarters and defined them as unseen spots. We calculated the coordinates of all unseen spots on one slide and cropped the whole H&E image into spot images by the simulated coordinates. After that, we applied a trained model to predict cell types for these unseen spots and raw spots, excluding prior spots. Finally, we combined all raw spots with unseen spots to obtain the super-resolved cell distribution for slides.

Cell annotation for subdivided spots

We performed the trained model to predict the cell types for all sub-divided spots for slides. Pseudo-labeling is optional to use in this module to improve accuracy.

Cell annotation for unseen sections

Based on biological prior knowledge that cell distribution across successive sections should be continuous, we utilized the model trained by the ST section to predict the cell distribution of the adjacent unseen H&E staining images. First, we manually aligned the ST section image and the adjacent image using Photoshop (version CS6) and replicated the coordinates of spots on the ST section to generate simulated spots of the adjacent unseen image. After removing the simulated spots outside the tissue on the unseen image, we predicted cell types for the rest simulated spots and attained the predicted cell distribution of the unseen image.

Implementation of alternative cell annotation methods and image-utilized methods

We compared STASCAN with Cell2location [17], Seurat [18], RCTD [19], Tangram [21], and MUSE [28] to demonstrate the superiority of STASCAN. Besides, we used Cell2location [17] and Seurat [18] jointly, Cell2location [17] individually (the human cardiac dataset), or RCTD [19] individually (the embryonic mouse brain) to pre-label cell types before training. The parameters of each method were set as below:

- Cell2location [17]: the Cell2location workflow was followed with the official tutorial: https://cell2location.readthedocs.io/en/latest/notebooks/cell2location_tutorial.html. The hyperparameters were set by default according to the tutorial.
- Seurat [18]: SCTransform function was used to normalize the scRNA-seq and spatial transcriptomics data, and the FindTransferAnchors function was used to construct anchors between the two normalized data. Predicted assays of cell proportion were attained using the TransferData function.
- RCTD [19]: we first convert spatial transcriptomics data to a SpatialRNA object and scRNA-seq reference to a reference object following the RCTD tutorial on GitHub:

https://github.com/dmcable/spacexr. To compare with STASCAN, we performed RCTD with the hyperparameters of "doublet_mode = full" and other default settings. To pre-label the cell types of prior spots in the mouse brain dataset, we performed RCTD with the hyperparameters of "doublet_mode = multi" in agreement with the original literature [43].

- Tangram [21]: the Tangram workflow was followed with the GitHub repository: https://github.com/broadinstitute/Tangram. We performed the parameters of "mode = clusters" and other default settings.
- MUSE [28]: the MUSE workflow was followed with the official tutorial on the GitHub: https://github.com/AltschulerWu-Lab/MUSE. The hyperparameters were set by default according to the tutorial. Especially, we set the value of gene expression of those masked spots as 0 when performed gene-masking test.

Analysis of planarian dataset

Implementation of the STASCAN model

To construct STASCAN in the planarian dataset, we first used 1829 spot images from 10 ST sections to train the base model. We then trained section models for each section with different sets of spot images from each specific section ranging from 208 to 608 and employed them to predict cell types of raw and imputed spots. Specifically, the section for training each section-specific model was included in the ten sections for training the base model. We randomly selected some prior spot images of each specific section and some samples with the categories that have larger sizes than others randomly as the training data for the base model, making the validation for the section-specific model in the planarian dataset close to the out-of-sample validation where most of the samples in the testing dataset of the section-specific model were not involved in the training dataset of the base model. The sizes of spot images were both 60×60 pixels and all prior spot images were selected by the first strategy described above (Additional file 2: Table S1). For cell annotation for subdivided spots and unseen sections, we adopted pseudo-labeling based on the trained section models. After endowing pseudo labels, there were 1336 spot images selected to train the basic CNN model on cell annotation for subdivided spots, 1431 spot images selected to train a new base model, and different sets of spot images ranging from 182 to 596 selected to train section models on cell annotation for unseen sections. The sizes of sub-spot images were 30×30 pixels (Additional file 2: TableS1). The hyperparameters were set by default (Additional file 2: Table S2).

Validation and comparison of STASCAN with different methods for cell annotation

We validated the performance of STASCAN and compared it with other cell annotation methods at raw resolution in the planarian dataset. We manually labeled the spot image with the dominant cell type in section-23 as ground truth. During this process, we first determined the region of the organizational tissue to which the spots belong and then annotated the cell types of these spots according to the morphological features of spot images, such as the size and shape of the cell nucleus, cytoplasmic morphology, the colored degree of cells, and so on. After determining the cell types for 683 spots in section-23, we performed the one-hot encoding on these spots with annotated cell types

to construct a ground-truth cell-spot matrix, which is defined as the cell distribution. We performed STASCAN and other methods to annotate spots, respectively (Fig. 2d). The detailed procedure for implementation of these methods was described above (the " Implementation of alternative cell annotation methods and image-utilized methods" subsection in " Methods"). The Kullback–Leibler divergences were calculated between the ground-truth cell-spot matrix and the cell-spot matrix generated by STASCAN or other deconvolution methods (Fig. 2e) using custom Python code. The Kullback–Leibler divergence for each spot is calculated as follows:

$$D_{\mathrm{KL}}(P\|Q) = \sum_{i=1}^{n} P(x_i)\log\frac{P(x_i)}{Q(x_i)}$$

where n is the number of cell types, $P(x_i)$ and $Q(x_i)$ are the probability values of the cell type $i(i = 1, \cdots, n)$.

Performance evaluation for the adjacent section prediction

We evaluated the effect of cell annotation for unseen sections by the SSIM values, which could measure the similarity between two images from luminance, contrast, and structure. We calculated the SSIM values for each randomized pair using the Python package (scikit image) and visualized them with a heatmap. The formulation of SSIM is defined as:

$$\mathrm{SSIM}(x, y) = \frac{(2\mu_x\mu_y + c_1) \times (2\sigma_{xy} + c_2)}{({\mu_x}^2 + {\mu_y}^2 + c_1) \times ({\sigma_x}^2 + \mu{\sigma_y}^2 + c_1)}$$

where x and y are two images and (μ_x, μ_y) and (σ_x, σ_y) are the means and standard deviations of x and y, σ_{xy} is the covariance of x and y, $c_1 = (0.01 \times L)^2$ and $c_2 = (0.03 \times L)^2$ and L is the dynamic range of the pixel values.

Construction of the 3D model

We manually aligned the consecutive H&E staining images by Photoshop and selected two pairs of anchor points, so that we were able to calculate the variable values during the alignment, including rotation value and translation value. Meanwhile, we extracted the contour of tissues on H&E staining images by Photoshop and transcribed it into pixel coordinates by custom code. Applying rigid transformation according to those variable values, we obtained the coordinates of aligned spot arrays and contour lines and reconstructed the 3D model for the neoblast and muscle cells of the planarian (Fig. 3g).

Analysis of the human intestinal dataset

Implementation of STASCAN

We constructed STASCAN for unseen spots and subdivided spots. To perform cell annotation for unseen spots, we respectively trained base models for seven ST slides. The spot images were selected by the first strategy described above and the number of spot images for each ST slide ranged from 297 to 1551. Besides, we used 1104 spot images to train a base model for subdivided spots on slide 4. The prior spot images of slide 4 were selected by the second strategy described above due to the requirement for

a greater number of training samples. The size of spot images was 80×80 pixels, and the size of subdivided spot images was 40×40 pixels (Additional file 2: TableS1). The hyperparameter "learning rate = 1e-4" was used for cell annotation on unseen spots, and "learning rate = 1e-3" was used for cell annotation on subdivided spots. Other hyperparameters were set by default (Additional file 2: Table S2).

Manual annotation based on the H&E staining image

We manually labeled three anatomical layers of slide-4, including the muscularis, fibroblasts, and epithelium layers presenting magenta, purple, and dark purple stainings, respectively, and with distinct cell populations with different shapes and densities in the H&E staining image. These different stained colors and diverse cellular shapes and densities in the three layers are considered the discriminant conditions during manual annotation.

Distance-based analysis

We fitted a set of zoomed elliptic curves for the intestinal tissues on the slide 4 and calculated the Euclidean distance from each spot to the focuses of the ellipse which was the smallest one of a group of ellipses containing the measured spots. We visualized the distance of raw spots to the tissue edge and subdivided spots to the tissue edge and evaluated the significance of differences with the Mann–Whitney *U* test (Additional file 1: Fig. S7b). Besides, we manually annotated the layers of epidermal, fibroblast, and muscularis cells based on the H&E staining image of the tissue and calculated the average distance of different layer boundaries to the tissue edge (Fig. 4c and Additional file 1: Fig. S7b).

Validation of cell annotation for unseen sections

We chose two ST slides with similar H&E staining images, slides 6 and 7, to validate the reliability of cell annotation for unseen sections. We constructed two CNN models, one trained by spots from slide 6 and the other trained by spots from slide 7. We then respectively applied the two CNN models to predict cell distribution on slide 6 (Additional file 1: Fig. S8a). Considering the cell distribution annotated by the first model as ground truth, we calculated the Pearson correlations between cell types of spots predicted by these two models trained with different input data (Additional file 1: Fig. S8b).

Analysis of the human lung dataset

Redefining cell populations of single-cell reference data

To predict cell distribution more accurately and precisely, we re-annotated the single-cell data [39]. We performed Seurat [18] to cluster cell populations and used published marker genes to annotate cell populations (Additional file 1: Fig. S9a, b). We set the parameter dims = 1:30 for FindNeighbors function, the parameter resolution = 0.1 for FindClusters function, and the parameter dims = 1:30 for the RunUMAP function. Finally, we increased the number of populations from 4 to 13 and used the related annotation to predict the paired ST data.

Implementation of STASCAN

We integrated the redefined single-cell reference with ST data using Cell2location [17] and Seurat [18], respectively, and selected spot images to train the model with the first strategy described above. We used 822 spot images to train a basic model for imputing spot predicting. The size of the spot images was 30×30 pixels (Additional file 2: TableS1). The hyperparameters were set by default (Additional file 2: Table S2).

Analysis of the human cardiac dataset

Implementation of STASCAN

We respectively trained the base models for 17 ST slides on cell annotation for unseen spots. We chose the second strategy described above to select spot images because the first strategy was too strict to obtain sufficient training samples in the human cardiac dataset. In this step, we performed Cell2location [17] individually to pre-label the prior spots, which was consistent in the original literature [40]. The size of the spot images was about 200×200 pixels (Additional file 2: TableS1). The hyperparameter "learning rate = 1e-4" was used for cell annotation on unseen spots, and other hyperparameters were set by default (Additional file 2: Table S2).

Cell-type niche analysis

STASCAN can lead to a matrix containing the probabilities of different cell types for all raw spots and unseen spots. Performing an unsupervised *k*-means clustering algorithm ($k=8$) to cluster spots from this matrix, we generated eight populations indicating distinct cell-type niches and mapped them into spatial regions (Fig. 6b, c and Additional file 1: Fig. S13-S17b).

Analysis of the embryonic mouse brain dataset

Implementation of STASCAN

We respectively trained the base model for the 4 ST slides on cell annotation for unseen spots. For the same reason mentioned above, we chose the second strategy and performed RCTD individually, which was in agreement with the original literature [43]. The size of the spot images was about 16×16 pixels (Additional file 2: TableS1). The hyperparameters were set by default (Additional file 2: Table S2).

Cell-type niche analysis

We respectively performed the unsupervised *k*-means clustering ($k=6$, 8, 10, 8) for the cell-probability matrices generated by STASCAN to annotate different populations and mapped them into spatial regions (Fig. 7b).

Implementation of iStar and comparison with STASCAN

We compared the performance of STASCAN and iStar [29] in cell annotations with enhanced resolution in the planarian dataset. According to the tutorial, iStar achieves cell annotation by a built-in function, needing input from a list of marker genes. We prepared three different lists of marker genes to test the functions: (a) we used the FindMarkers function in Seurat [18] to identify markers for seven major cell types in the planarian ST data and sorted these identified genes based on the order of scores and chose the top 100 genes for each cell types as the list of marker genes; (b) we did the same as (a) but chose

top 300 genes; and (c) we collected the classical marker genes from previous reports (four genes for epidermal cells; two genes for gut cells; two genes for muscle cells; two genes for neoblast cells; two genes for parenchymal cells; three genes for secretory cells; four genes for neuronal cells). The other hyperparameters were set by default according to the tutorial (https://github.com/daviddaiweizhang/istar). After that, we compared STASCAN with iStar by training iStar with the different lists of marker genes, respectively, and performed the cell type annotation in section-23 (Additional file 1: Fig. S19).

Computational efficiency

We reported the running time of STASCAN in two ST datasets, which were also used as detailed guiding illustration on GitHub (Additional file 2: Table S3). All experiments were conducted on our GPU platform (NVIDIA Tesla V100s with 32 GB memory).

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s13059-024-03421-5.

Additional file 1: Supplementary Figures S1-S19. Fig. S1. Workflow of STASCAN. Fig. S2. Preparation and evaluation of the base model of STASCAN in the planarian dataset. Fig. S3. Verification of STASCAN for unseen spots and subdivided spots in the planarian dataset. Fig. S4. Verification of STASCAN for unseen sections in the planarian dataset. Fig. S5. Preparation and evaluation of STASCAN in the human intestinal dataset. Fig. S6. Super-resolution cell distribution by STASCAN in the human intestinal dataset. Fig. S7. Sub-resolution cell distribution of epithelium subtypes by STASCAN in the human intestinal dataset. Fig. S8. Verification of STASCAN for unseen sections in the human intestinal dataset. Fig. S9. Preparation and evaluation of STASCAN in the human lung dataset. Fig. S10. Preparation of the human cardiac dataset. Fig. S11. Evaluation of STASCAN in the human cardiac dataset. Fig. S12. Evaluation of STASCAN for imputing missing spots in the human heart dataset. Fig. S13. Super-resolution cell distribution of slides sampled from non-transplanted donor hearts in the human cardiac dataset by STASCAN. Fig. S14. Super-resolution cell distribution of slides sampled from the fibrotic zone in the human cardiac dataset by STASCAN. Fig. S15. Super-resolution cell distribution of slides sampled from the ischaemic zone in the human cardiac dataset by STASCAN. Fig. S16. Super-resolution cell distribution of slides sampled from the remote zone in the human cardiac dataset by STASCAN. Fig. S17. Super-resolution cell distribution of the slide sampled from the border zone in the human cardiac dataset by STASCAN. Fig. S18. Applications of STASCAN in the embryonic mouse brain dataset. Fig. S19. Comparisons between STASCAN and iStar.

Additional file 2: Supplementary Tables S1-3. Table S1: Dataset statistics. Table S2: The summary of user-defined hyperparameters used in corresponding datasets. Table S3: The running time of STASCAN.

Additional file 3: Review history.

Acknowledgements
We thank Prof. Zhihua Zhang and Prof. Jun Cai from the Beijing Institute of Genomics, Chinese Academy of Sciences and China National Center for Bioinformation for their thoughtful advice.

Authors' contributions
Y.G.Y. and S.Z. conceived and supervised the project. Y.W. developed and implemented the STASCAN algorithm. Y.W., J.Y.Z., B.Y., and G.C. discussed and integrated the data. J.Y.Z. and C.C.G. validated the methods. Y.G.Y., S.Z., Y.W., J.Y.Z., C.C.G., Y.Y., and Y.L.Z. wrote the manuscript. All authors read and approved the final manuscript.

Peer review information
Veronique van den Berghe was the primary editor of this article and managed its editorial process and peer review in collaboration with the rest of the editorial team.

Funding
This work has been supported by the National Key R&D Program of China (nos. 2023YFC3402200 to Y.G.Y, 2021YFA1302500 to S.Z.), the National Natural Science Foundation of China (nos. 32121001 to Y.G.Y, 32341013, 12326614, 12126605 to S.Z.), the Chinese Academy of Sciences Project for Young Scientists in Basic Research (nos. YSBR-073 to C.C.G, YSBR-034 to S.Z.], and the R&D project of Pazhou Lab (Huangpu) (no. 2023K0602 to S.Z.)

Data availability
The datasets analyzed in this study are all from publicly available datasets. Specifically, the 10 × Visium planarian dataset can be accessed in the OMIX database under accession numbers OMIX003867 [47] and OMIX003889 [48]. The 10 × Visium human intestinal dataset can be accessed from GEO under accession numbers GSE158328 [49] and GSE158702 [50], and from Mendeley Data at https://data.mendeley.com/datasets/gncg57p5x9/2 [51], respectively. The 10 × Visium human lung datasets can be accessed from GEO under accession numbers GSE178360 [52] and GSE178361 [53]. The 10 × Visium human cardiac datasets can be accessed at cellxgene https://cellxgene.cziscience.com/

collections/8191c283-0816-424b-9b61-c3e1d6258a77 [54] and at the Human Cell Atlas Data Portal https://data.humancellatlas.org/explore/projects/e9f36305-d857-44a3-93f0-df4e6007dc97 [55], respectively. The MISAR-seq embryonic mouse brain dataset can be accessed under the National Genomics Data Center accession number (OEP003285, www.biosino.org/node/project/detail/OEP003285)[56].

Code availability

An open-source Python implementation of the STASCAN package is available on GitHub (https://github.com/AbbyWY/STASCAN [57] https://github.com/zhanglabtools/STASCAN [58] and Zenodo (https://zenodo.org/records/13839138 [59]. The source code is released under the MIT license.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Review history

The review history is available as Additional File 3.

Competing interests

The authors declare no competing interests.

Author details

[1]China National Center for Bioinformation, Beijing 100101, China. [2]Beijing Institute of Genomics, Chinese Academy of Sciences, Beijing 100101, China. [3]University of Chinese Academy of Sciences, Beijing 100049, China. [4]College of Future Technology, University of Chinese Academy of Sciences, Beijing 100049, China. [5]NCMIS, CEMS, RCSDS, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China. [6]School of Mathematical Sciences, University of Chinese Academy of Sciences, Beijing 100049, China. [7]Key Laboratory of Systems Health Science of Zhejiang Province, School of Life Science, Hangzhou Institute for Advanced Study, University of Chinese Academy of Sciences, Hangzhou 310024, China. [8]Sino-Danish College, University of Chinese Academy of Sciences, Beijing 101408, China. [9]Institute of Stem Cell and Regeneration, Chinese Academy of Sciences, Beijing 100101, China.

Received: 12 January 2024 Accepted: 10 October 2024
Published online: 22 October 2024

References

1. Longo SK, Guo MG, Ji AL, Khavari PA. Integrating single-cell and spatial transcriptomics to elucidate intercellular tissue dynamics. Nat Rev Genet. 2021;22:627–44.
2. Rao A, Barkley D, Franca GS, Yanai I. Exploring tissue architecture using spatial transcriptomics. Nature. 2021;596:211–20.
3. Vickovic S, et al. High-definition spatial transcriptomics for in situ tissue profiling. Nat Methods. 2019;16:987–90.
4. Liu Y, et al. High-spatial-resolution multi-omics sequencing via deterministic barcoding in tissue. Cell. 2020;183:1665–1681.e18.
5. NanoString Technologies, Inc. CosMxTM spatial molecular imager grant support document [Brochure]. 2022. https://nanostring.com/wp-content/uploads/FL_MK3934_SMI-Grant-Package-V4-1.pdf.
6. Chen KH, Boettiger AN, Moffitt JR, Wang S, Zhuang X. Spatially resolved, highly multiplexed RNA profiling in single cells. Science. 2015;348:412–26.
7. Moses L, Pachter L. Museum of spatial transcriptomics. Nat Methods. 2022;19:534–46.
8. Stark R, Grzelak M, Hadfield J. RNA sequencing: the teenage years. Nat Rev Genet. 2019;20:631–56.
9. Ståhl P. Visualization and analysis of gene expression in tissue sections by spatial transcriptomics. Science. 2016;353:78–82.
10. Hu J, et al. Deciphering tumor ecosystems at super resolution from spatial transcriptomics with TESLA. Cell Syst. 2023;14:404–17.
11. Rodriques S, et al. Slide-seq: a scalable technology for measuring genome-wide expression at high spatial resolution. Science. 2019;363:1463–7.
12. Chen A, et al. Spatiotemporal transcriptomic atlas of mouse organogenesis using DNA nanoball-patterned arrays. Cell. 2022;185:1777–1792.e21.
13. Fang S, et al. Computational approaches and challenges in spatial transcriptomics. Genom Proteomics Bioinformatics. 2023;21:24–47.
14. Chen A, et al. Single-cell spatial transcriptome reveals cell-type organization in the macaque cortex. Cell. 2023;186:3726–3743.e24.
15. Asp M, et al. A spatiotemporal organ-wide gene expression and cell atlas of the developing human heart. Cell. 2019;179:1647–1660.e19.
16. Li B, et al. Benchmarking spatial and single-cell transcriptomics integration methods for transcript distribution prediction and cell type deconvolution. Nat Methods. 2022;19:662–70.
17. Kleshchevnikov V, et al. Cell 2location maps fine-grained cell types in spatial transcriptomics. Nat Biotechnol. 2022;40:661–71.
18. Stuart T, et al. Comprehensive integration of single-cell data. Cell. 2019;177:1888–1902.e21.
19. Cable DM, et al. Robust decomposition of cell type mixtures in spatial transcriptomics. Nat Biotechnol. 2022;40:517–26.

20. Maaskola J, et al. Charting tissue expression anatomy by spatial transcriptome deconvolution. BioRxiv. 2018. https://doi.org/10.1101/362624.
21. Biancalani T, et al. Deep learning and alignment of spatially resolved single-cell transcriptomes with Tangram. Nat Methods. 2021;18:1352–62.
22. Geras A, et al. Celloscope: a probabilistic model for marker-gene-driven cell type deconvolution in spatial transcriptomics data. Genome Biol. 2023;24:120.
23. Yang F, et al. scBERT as a large-scale pretrained deep language model for cell type annotation of single-cell RNA-seq data. Nat Mach Intell. 2022;4:852–66.
24. Huang Q, et al. Evaluation of cell type annotation R packages on single-cell RNA-seq data. Genomics Proteomics Bioinformatics. 2021;19:267–81.
25. Perlman ZE, et al. Multidimensional drug profiling by automated microscopy. Science. 2004;306:1194–8.
26. Feldman D, et al. Optical pooled screens in human cells. Cell. 2019;179:787–799.e717.
27. Loo LH, Wu LF, Altschuler SJ. Image-based multivariate profiling of drug responses from single cells. Nat Methods. 2007;4:445–53.
28. Bao F, et al. Integrative spatial analysis of cell morphologies and transcriptional states with MUSE. Nat Biotechnol. 2022;40:1200–2120.
29. Zhang D, et al. Inferring super-resolution tissue architecture by integrating spatial transcriptomics with histology. Nat Biotechnol. 2024. https://doi.org/10.1038/s41587-023-02019-9.
30. Bergenstråhle L, et al. Super-resolved spatial transcriptomics by deep data fusion. Nat Biotechnol. 2021;40:476–9.
31. He B, et al. Integrating spatial gene expression and breast tumour morphology via deep learning. Nat Biomed Eng. 2020;4:827–34.
32. Monjo T, Koido M, Nagasawa S, Suzuki Y, Kamatani Y. Efficient prediction of a spatial transcriptomics profile better characterizes breast cancer tissue sections without costly experimentation. Sci Rep. 2022;12:4133.
33. Li S, Gai K, Dong K, Zhang Y, Zhang S. High-density generation of spatial transcriptomics with STAGE. Nucleic Acids Res. 2024;52:4843–56.
34. Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. eprint arXiv:1409.1556 [cs.CV]. 2014.
35. Cui G, et al. Spatiotemporal transcriptomic atlas reveals the dynamic characteristics and key regulators of planarian regeneration. Nat Commun. 2023;14:3205.
36. Benham-Pyle BW, et al. Identification of rare, transient post-mitotic cell states that are induced by injury and required for whole-body regeneration in Schmidtea mediterranea. Nat Cell Biol. 2021;23:939–52.
37. Wang Z, Bovik AC, Sheikh HR, Simoncelli EP. Image quality assessment: from error visibility to structural similarity. IEEE Trans Image Process. 2004;13:600–12.
38. Fawkner-Corbett D, et al. Spatiotemporal analysis of human intestinal development at single-cell resolution. Cell. 2021;184:810–826.e823.
39. Kadur Lakshminarasimha Murthy, P. et al. Human distal lung maps and lineage hierarchies reveal a bipotent progenitor. Nature. 2022;604:111–119.
40. Kuppe C, et al. Spatial multi-omic map of human myocardial infarction. Nature. 2022;608:766–77.
41. Shook BA, et al. Myofibroblast proliferation and heterogeneity are supported by macrophages during skin repair. Science. 2018;362: eaar2971.
42. Pakshir P, et al. Dynamic fibroblast contractions attract remote macrophages in fibrillar collagen matrix. Nat Commun. 2019;10:1850.
43. Jiang F, et al. Simultaneous profiling of spatial gene expression and chromatin accessibility during mouse brain development. Nat Methods. 2023;20:1048–57.
44. An S, et al. Airway smooth muscle dynamics: a common pathway of airway obstruction in asthma. Eur Respir J. 2007;29:834–60.
45. Cho N, Razipour SE, McCain ML. Featured Article: TGF-β1 dominates extracellular matrix rigidity for inducing differentiation of human cardiac fibroblasts to myofibroblasts. Exp Biol Med. 2018;243:601–12.
46. Abadi M, et al. TensorFlow: large-scale machine learning on heterogeneous distributed systems. eprint arXiv:1603.04467 [cs.DC]. 2016.
47. Cui, G. et al. Spatiotemporal transcriptomic atlas reveals dynamic characteristics and key regulators of planarian regeneration. Datasets. OMIX. 2023. https://ngdc.cncb.ac.cn/omix/release/OMIX003867.
48. Cui, G. et al. Spatiotemporal transcriptomic atlas reveals dynamic characteristics and key regulators of planarian regeneration. Datasets. OMIX. 2023. https://ngdc.cncb.ac.cn/omix/release/OMIX003889.
49. Fawkner-Corbett, D. et al. Spatiotemporal analysis of human intestinal development at single cell resolution - spatial transcriptomics. Datasets. Gene Expression Omnibus. 2021. https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE158328.
50. Fawkner-Corbett, D. et al. Spatiotemporal analysis of human intestinal development at single cell resolution - scRNA-Seq. Datasets. Gene Expression Omnibus. 2021. https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE158702.
51. Fawkner-Corbett, D. et al. Spatiotemporal analysis of human intestinal development at single cell resolution: supplementary data. Datasets. Mendeley Data. 2021. https://data.mendeley.com/datasets/gncg57p5x9/2.
52. Kadur Lakshminarasimha Murthy, P. et al. Human distal lung maps and lineage hierarchies reveal a bipotent progenitor [single-cell RNA-seq]. Datasets. Gene Expression Omnibus. 2022. https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE178360.
53. Kadur Lakshminarasimha Murthy, P. et al. Human distal lung maps and lineage hierarchies reveal a bipotent progenitor [Spatial transcriptomics]. Datasets. Gene Expression Omnibus. 2022. https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE178361.
54. Kuppe, C. et al. Spatial multi-omic map of human myocardial infarction. Datasets. cellxgene. 2022. https://cellxgene.cziscience.com/collections/8191c283-0816-424b-9b61-c3e1d6258a77.
55. Kuppe, C. et al. Spatial multi-omic map of human myocardial infarction. Datasets. the Human Cell Atlas Data Portal. 2022. https://data.humancellatlas.org/explore/projects/e9f36305-d857-44a3-93f0-df4e6007dc97.

56. Jiang, F. et al. Simultaneously spatiotemporal gene expression and chromatin accessibility for mouse brain development. Datasets. National Genomics Data Center. 2023. www.biosino.org/node/project/detail/OEP003285.
57. Wu, Y. et al. STASCAN deciphers fine-resolution cell-distribution maps in spatial transcriptomics by deep learning. Github. 2024. https://github.com/AbbyWY/STASCAN.
58. Wu Y. et al. STASCAN deciphers fine-resolution cell-distribution maps in spatial transcriptomics by deep learning. Github. 2024. https://github.com/zhanglabtools/STASCAN.
59. Wu, Y. et al. STASCAN deciphers fine-resolution cell-distribution maps in spatial transcriptomics by deep learning. Zenodo. 2024.https://zenodo.org/records/13839138.

Publisher's Note