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

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

Affiliations

[1]CAS Key Laboratory of Genomic and Precision Medicine, Beijing Institute of Genomics, Chinese Academy of Sciences and China National Center for Bioinformation, Beijing 100101, China

[2]University of Chinese Academy of Sciences, Beijing 100049, China

[3]College of Future Technology, University of Chinese Academy of Sciences, Beijing 100049, China

[4]NCMIS, CEMS, RCSDS, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China

[5]School of Mathematical Sciences, University of Chinese Academy of Sciences, Beijing 100049, China

[6]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

[7]Sino-Danish College, University of Chinese Academy of Sciences, Beijing 101408, China

[8]Institute of Stem Cell and Regeneration, Chinese Academy of Sciences, Beijing 100101, China

[9]These authors jointly supervised this work: zsh@amss.ac.cn (S.Z.); ygyang@big.ac.cn (Y.G.Y.);

Abstract

Background: The spatial transcriptomics (ST) technologies have been widely applied to decode the spatial distribution of cells by resolving gene expression profiles in tissues. However, a fine-resolved spatial cell map is still limited by algorithmic tools and sequencing techniques.

Results: Here we develop a novel deep learning approach, STASCAN, which could define the spatial cellular distribution of both captured and uncharted areas by cell feature learning that combines gene expression profiles and histology images. STASCAN additionally adopts optional transfer learning and pseudo-labeling methods to improve the accuracy of the cell-type prediction from images. We have successfully applied STASCAN to enhance cell resolution, and revealed finer organizational structures across diverse datasets from various species and tissues generated from 10× Visium technology. STASCAN improves cell resolution of *Schmidtea mediterranea* datasets by six times and reconstructs more detailed 3D cell-type models. Furthermore, STASCAN could accurately pinpoint the boundaries of distinct cell layers in human intestinal tissue, specifically identify a micrometer-scale smooth muscle bundle structure in consistent with anatomic insights in human lung tissue, and redraw the spatial structural variation with enhanced cell patterns in human myocardial infarction tissue. Additionally, through STASCAN on embryonic mouse brain datasets generated by DBiT-derived MISAR-seq technology, the increased cellular resolution and distinct anatomical tissue domains with

cell-type niches are revealed. Collectively, STASCAN is compatible with different ST technologies and has notable advantages in generating cell maps solely from histology images, thereby enhancing the spatial cellular resolution.

Conclusions: In short, STASCAN displays significant advantages in deciphering higher-resolution cellular distribution, resolving enhanced organizational structures and demonstrating its potential applications in exploring cell-cell interactions within the tissue microenvironment.

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

Introduction

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].

Current ST technologies are typically divided into two categories: (1) imaging-based approaches that capture genes corresponding to the targeted probes through *in situ* sequencing or *in situ* hybridization, which can achieve single-cell resolution but are limited by low throughput, biased coverage for

71 pre-selected gene targets, and expensive specialized equipment; and (2)
72 next-generation sequencing (NGS)-based approaches that capture transcripts
73 from tissues combined with encoded spatial positional information before
74 sequencing[3]. By contrast, NGS-based approaches have been more widely
75 used due to their advantages of higher throughput, unbiased coverage for
76 transcriptome, and more accessible commercial products[4].

77 However, NGS-based approaches still suffer from several limitations. The
78 significant one is the spatial resolution which is restricted by the area and
79 sparsity of captured domains (defined as spots)[5]. For example, the spots of
80 10× Visium are 55 μm in diameter with a space of 100 μm, which are too large
81 to capture the cells at single-cell resolution and too sparse to measure cells in
82 the uncharted areas between the captured domains[6]. In addition, DBiT-seq[7]
83 was designed to produce smaller and more densely spots (10, 25, or 50 μm in
84 width). Even though it provides higher spatial resolution, the transcriptional
85 information in the uncharted areas is still inevitably lost. Other approaches,
86 such as Slide-seq[8] and Stereo-seq[9], also aim to achieve single-cell or
87 subcellular resolution by using small and densely packed spots. However, they
88 are also affected by the fact of multiple fractions of cells within a single spot.

89 Moreover, the uncaptured cells in uncharted areas lead to a limited spatial
90 cellular resolution in not only 2D but also 3D levels. Currently, ST technologies
91 enable us to depict cellular maps from a regional tissue section. By
92 continuously sectioning the tissue, 3D cellular maps of organs can be

constructed. However, the high-cost only allows a small portion of consecutive tissue sections be sequenced, incurring the problem of uncharted areas along the z-axis and ultimate low 3D resolution [10].

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[11], such as Cell2location[12], Seurat[13], and RCTD[14]. However, the deconvolution methods are easily affected by the potential 'dropout' of marker genes in ST data and the inaccuracies in single-cell reference data[15]. More importantly, these methods only aim to improve computationally cellular resolution within captured domains by inferring the proportion or abundance of cell type. But 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[16]. 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 increase the accuracy of cell annotation. Some emerging computational

methods integrating gene expression and morphological information have been developed for ST data. For instance, Tangram[17] synthesizes histological images to estimate the proportion of cells in each spot during single-cell deconvolution, and MUSE[18] combines transcriptional profiles and morphological features to characterize the cells and tissue regions. However, none of them involve the improvment 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[19]. Nevertheless, due to the limitation of cell-type deconvolution methods mentioned above and the need for refining gene expression-imputing models[20], there are still obstacles to directly use the enhanced gene expression profile to predict cell types.

To address this issue, 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 and 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, confirmed as a smooth muscle bundle near the tracheal wall, which has not been identified by other methods. Moreover, STASCAN provided a refined distribution of cell-type niches in human cardiac tissue and mouse embryonic tissue, facilitating the understanding 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, to depict a fine-resolution cell distribution map in tissues (**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 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, STASCAN can accurately predict cell types solely based on histological images (**Fig.1a**, **Supplementary Fig.1** 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 sub-spot images with optional pseudo-labels to obtain a sub-resolution cell distribution (**Fig. 1c**); (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 types prediction solely from images

To quantitatively evaluate the performance of STASCAN, we initially applied it in 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)[21]. Given the comprehensiveness of the planarian dataset, we first constructed a base model using 1,829 spot images extracted from 10 collected sections to learn the features of seven main cell types identified by seuqnecing

181 information, including epidermal, gut, muscle, neoblast, neuronal,
182 parenchymal, and secretory cells (**Supplementary Fig. 2a, b**, **Supplementary**
183 **Table 1** and **Methods**). Although there was uncertainty in predicting neoblast
184 and parenchymal cell types due to the scarcity of training samples of these two,
185 most cell types were annotated with a recall rate of over 78% (**Fig. 2a** and
186 **Supplementary Fig. 2c**). In addition, the learned model showed excellent
187 accuracy in predicting cell types, with the area under the curve (AUC)
188 calculated from the receiver operating characteristic (ROC) curves reaching as
189 high as 0.936 to 0.996 (**Fig. 2b**). Besides, considering the potential batch
190 effect among different ST sections, we performed section-specific training
191 based on the base model (**Methods**). The results from the section-specific
192 model showed a significant improvement in accuracy with higher AUC values
193 compared with the base model, indicating that section-specific training is
194 beneficial for the improved prediction performance of the whole frame (**Fig.**
195 **2c**).

196 To compare the performance of STASCAN in predicting dominant cell
197 types on raw spots with other methods, we similarly performed Cell2location[12],
198 Seurat[13], and RCTD[14] on the planarian dataset. We initially annotated the cell
199 types of each raw spot by manual according to the morphologic features of
200 corresponding spot images, which considered as the ground truth. The cell
201 distribution predicted by STASCAN is highly correlated with that of manual
202 annotations, significantly superior to other methods (**Fig. 2d, e** and **Methods**).

203 We also observed that alternative methods result in varied degrees of biases
204 for cell annotation. For example, Cell2location could characterize most cell
205 distribution but with a low sensitivity for epidermal cells; Seurat showed strong
206 annotation bias for epidermal cells, leading to mislabeling for other cell types;
207 and RCTD displayed some positive annotation but lost annotation information
208 for most neuronal and secretory cells.

209 We also compared STASCAN with other methods that utilize both
210 morphological and transcriptional features for ST data analysis. Tangram[17]
211 effectively illustrated the distribution of the majority of cells, but exhibited a
212 slight bias towards neuronal cells and a reduced sensitivity in detecting
213 epidermal cells (**Fig. 2d** and **e**). On the other hand, MUSE[18] characterized
214 tissue regions by identifying spot clusters, yet these clusters appeared
215 relatively scattered, and failed to represent corresponding structures in the
216 planarian dataset (**Fig. 2d**). In contrast, our STASCAN displayed more precise
217 performance in prediction, with the capability of accurately pinpointing the
218 spatial distribution of seven main types of cells, in accord with their known
219 biological functions[22] (**Fig. 2d**, and **Supplementary Fig. 2d,e**). For example,
220 corresponding to clear tissue structures visible through hematoxylin and eosin
221 (H&E) staining, the epidermal cells draw the contour of the planarian body, gut
222 cells mark the location of the intestine, and muscle cells along with neuronal
223 cells define the anatomy of the the pharynx (**Fig. 2d** and **Supplementary Fig.
224 2d**).

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 avaiable. 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[21]. 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 (**Supplementary Fig. 2a** 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 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 **Supplementary Fig. 2d**). Besides, it was highly consistent with the distribution of corresponding cell markers reported in previous literatures[23] (**Fig. 3b** and **Supplementary Fig. 3a**).

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**). These results indicate that STASCAN significantly enhances cell granularity at sub-resolution, facilitating 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 **Supplementary Fig. 3b**). In line with the biological interpretation that the cell distribution between two consecutive sections should be similar, the structure similarity index measure (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**, **Supplementary Fig. 3b** and **Methods**). Furthermore, 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**). The model displayed the cell distribution in three-dimensional, with improved cellular resolution in spatial and promoted utilization of staining images without ST sequencing.

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

293 To further evaluate the performance of STASCAN for ST datasets on different
294 tissue architectures, we applied STASCAN to human intestinal datasets
295 generated by 10× Visium technology. These datasets consisted of eight slides
296 sampled at diverse sampling locations and time points[24]. We trained each slide
297 using STASCAN with different sizes of prior spots, ranging from 297 to 1,551,
298 and observed stable performances across all sizes (**Supplementary Fig. 4a-c**,
299 **Supplementary Table 1** and **Methods**).

300 Next, we applied STASCAN to predict cell types of unseen spots.
301 Compared with the cell distribution of prior spots annotated by other methods,
302 STASCAN stratified cell populations to finer regional layers. For example, in
303 comparison with other methods only roughly distinguishing the distribution of
304 different cells, STASCAN was able to delineate the borders of cell layers
305 among the intestinal epithelium, fibroblasts, and muscularis, greatly enhancing
306 the cellular spatial patterns (**Fig. 4a** and **Supplementary Fig. 5**).

307 Then, we used STASCAN to draw the spatial distribution map of fine cell
308 subtypes in the human intestinal tissue (**Fig. 4b**). Actually, we labeled three
309 anatomical layers of the intestinal tissue related to the morphological
310 structures of H&E staining, including muscularis, fibroblasts and epithelium
311 layers, listed based on their distance to the intestinal edge (**Fig. 4c**). When
312 evaluating four epithelium subtypes occupying an absolute 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 **Supplementary Fig. 6a**). 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 were abnormally predicted to be located in the fibroblasts layer (**Fig. 4d, e** and **Supplementary Fig. 6b**).

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 (**Supplementary Fig. 7a, b**).

STASCAN uncovers a novel structure in the human lung tissue

Despite the limitations of raw spatial resolution of ST technologies, STASCAN provides assistance in enhancing cellular pattern and rediscovering the micrometer-scale structure. Here, we applied STASCAN on the 10× Visium human lung dataset[25], which sampled from the proximal airway. We previously redefined 13 reference cell types to bette illustrate the organizational structure

335 and annotated 822 ST spots with seven dominant cell types to train the
336 STASCAN (**Supplementary Fig. 8a-e**, **Supplementary Table 1** and
337 **Methods**). With enhanced resolution, STASCAN showed more precise cellular
338 and structural patterns of human lung tissue. Besides, we observed that
339 STASCAN sensitively identified a micrometer-scale oval-shaped structure that
340 was highly consistent with the H&E staining images, which was confirmed as
341 the smooth muscle bundles adjacent to the tracheal wall. However, this
342 structure was not evident at the raw resolution of prior spots, highlighting the
343 capability of STASCAN to reveal refined structures of spatial regions (**Fig. 5a**
344 and **Supplementary Fig. 8c**).

345 Furthermore, we compared the ability of STASCAN with other methods in
346 revealing tissue structures with cellular patterns. The results showed that
347 STASCAN could clearly depict the silhouette of the airway with basal and
348 neuroendocrine cells and the cricoid cartilage structure surrounded by goblet
349 cells, mucous cells, smooth muscle cells, pericytes, etc. Moreover, the smooth
350 muscle tissue traced by smooth muscle cells and pericytes in the left bottom of
351 the slide was only identified by STASCAN (**Fig. 5b** and **Supplementary Fig.**
352 **8c**). Briefly, compared with the cell distribution pattern identified by other
353 methods, STASCAN displayed superior advantages in identifying and
354 characterizing spatial specific structures, which better reflects the anatomical
355 structure after imputation.

356

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

359 To explore whether further improved functional applications in ST data analysis
360 can be achieved by STASCAN, we adopted this approach to reanalyze the 10×
361 Visium human cardiac datasets[26], which included 17 slides from normal hearts
362 and the pathological ones after myocardial infarction (**Supplementary Figs.**
363 **9a,b, 10**, **Supplementary Table 1** and **Methods**). We first grouped these
364 slides according to their sampling regions[26], including normal non-transplanted
365 donor hearts as controls, necrotic (ischaemic zone and border zone), and
366 unaffected regions (remote zone), and regions at later stages after myocardial
367 infarction (fibrotic zone). After performing STASCAN to enhance the spatial cell
368 distribution, we focused on two slides with a lot of missing spots that were
369 filtered out due to scanty genes and unique molecular identifiers (UMI)
370 measured in the original literature[26] (**Fig.6a** and **Supplementary Figs. 11-15a**).
371 In these two slides, STASCAN not only more accurately depicted the cell
372 distribution pattern of tissue structures, but also predicted the potential cell
373 distributions in the missing areas only from images. Especially for the serious
374 missing in the slide_ACH0010 sampled from the ischaemic zone, STASCAN
375 better imputed the reasonable diffusion of the cardiomyocyte, fibroblast, and
376 myeloid cells in line with the histological morphology (**Fig. 6a**). The proximity of
377 the latter two cells indicated a strong dependence between them on the areas
378 of immune cell infiltration and scar formation[26].

379 We then explore the spatial structural variations in these two slides by
380 performing unsupervised clustering for spots based on the composition of cell
381 annotations predicted by STASCAN, and then mapped the clusters, defined as
382 cell-type niches, to the spatial regions (**Fig. 6b, c, f** and **Supplementary Figs.**
383 **11-15b**). Through redrawing the spatial distributions for these cell-type niches,
384 the cardiac tissue manifested more delicate spatial patterns compared with the
385 dominant cell annotation, which were consistent with histological morphology
386 and detailed structural variations observed during physiological and
387 pathological processes.

388 Besides, these cell-type niches based on diverse cell propositions
389 revealed more elaborate cell interacting microenvironments with potential
390 biological insights (**Fig. 6d, e** and **Supplementary Figs. 11-15c**). For example,
391 we observed myogenic cell-type niches (0, 1, and 2) mainly displaying
392 characteristics of cardiomyocyte cells, and fibrotic cell-type niches (3, 4, 5, 6,
393 and 7) mainly presenting characteristics of fibroblast cells, in the
394 slide_ACH006 sampled from the fibrotic zones. On the aspect of spatial
395 distributions, the myogenic cell-type niches could jointly characterize the
396 myocardial structure, and the fibrotic cell-type niches distinguished by the
397 proprotion of fibroblast cells indicated different fibrosis processes during the
398 lesion. Especially, there was a measure of characteristics of vSMCs and
399 endothelial cells in niches 3 and 4 for depicting the spatial structure of the
400 cardiac vasculature. In addition, in slide_ACH0010, we observed inflammatory

401 cell-type niches (4, 5, 6, and 7) which mainly exhibit characteristics of myeloid
402 and mast cells, apart from myogenic cell-type niches (0 and 1) and fibrotic
403 cell-type niches (2 and 3). These three types of niches took up distinct spatial
404 regions, but there was niche 7 located at the intersection, in line with the
405 proposition of cardiomyocyte, fibroblast, myeloid, and mast cells in the niche 7.
406 Especially, niches 2, 3, 4, and 5 showed co-enrichment between myeloid and
407 fibroblast cells, in accordance with the role of macrophages in fibroblast
408 activation[27] and fibroblast cells in macrophage attraction[28]. Overall, STASCAN
409 expands the application of niche distribution and provides better insights into
410 understanding cellular microenvironment interactions.
411

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

414 We further test whether STASCAN is also applicable in ST data derived from
415 various technologies. We first employed STASCAN on an embryonic mouse
416 brain dataset from MISAR-seq[29], a microfluidic indexing-based spatial
417 technology motivated by DBiT-seq[7] with both high-quality image and
418 sequencing data (**Supplementary Fig. 16**, **Supplementary Table 1** and
419 **Methods**). Notably, although the H&E images adopted in this dataset were
420 obtained from the adjacent tissue slide which causes a partial disharmony
421 between the actual gene expression pattern and morphological images,
422 STASCAN still achieved excellent results. When compared with the cellular

423 distribution annotated by RCTD[29], STASCAN significantly improved the cellular
424 resolution with highlighted characteristics of tissue structures (**Fig. 7a** and
425 **Supplementary Fig. 16**). For example, the enhanced distribution pattern of
426 forebrain GABAergic neurons was associated with the subpallium, and a group
427 of the forebrain glutamatergic and cortical or hippocampal glutamatergic
428 neurons spotlighted the dorsal pallium of the forebrain at enhanced resolution
429 (**Fig. 7a**).

430 Furthermore, we respectively generated cell-type niches for different
431 development stages of mouse brain tissue and mapped them to spatial regions
432 (**Fig. 7b** and **Methods**). Using the manual anatomical annotations of major
433 tissue organizations from H&E images as the ground truth[29] (**Fig. 7b**), we
434 compared the cluster distribution of cell-type niches to the raw-resolution
435 distributions of cell annotations on prior spots. The cluster distribution of
436 cell-type niches generated by STASCAN more remarkably recapitulates the
437 tissue organization in the developing mouse brain than the ones from
438 raw-resolution. Especially for the E18.5 embryonic mouse brain tissue, in
439 contrast to cell annotation at raw resolution which showed an unrecognized
440 organization in nearly entire brain region, STASCAN clearly defined the major
441 tissue domains by cell-type niches clustering (**Fig. 7b**). Collectively, these
442 results illustrate the strength of STASCAN in highlighting tissue structures and
443 redrawing finer organizations using ST data from various technologies.

444

Discussion

Although current spatial transcriptomics has achieved remarkable progress in deciphering the distribution and interaction of cells in tissues, the 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 clould 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,

467 STASCAN fully utilizes the image information of continuous sections to
468 construct 3D tissue models without expensive ST experimental costs, laying
469 the foundations for expanding the future frontier in 3D cellular atlas. We
470 performed STASCAN in different species and tissues and observed a
471 substantial advance over current methods in depicting spatial cellular patterns
472 and tissue structures. In planarian, STASCAN successfully predicted cell
473 distribution of unseen spots and sections in the uncharted areas, leading to the
474 construction of more detailed 3D cell distribution models. In addition, we found
475 that STASCAN discovers precise tissue structures with enhanced cell
476 annotation. In human lung tissue, STASCAN enhanced spatial resolution and
477 identified a micrometer-scale structure that was made up of a group of smooth
478 muscle cells located near the tracheal wall and involved in the regulation of
479 intra-pulmonary airway caliber[30], while the raw resolution are too coarse to
480 resolve the same specific structure.

481 The low resolution of NGS-based spatial transcriptomics is also attributed
482 to the spacing of captured spots, which are not small enough to achieve
483 single-cell resolution[5]. Most current methods estimate the proportion or
484 abundance of each cell type from the cell mixtures on the capture spots but are
485 incapable of assigning each cell into exact locations within each spot. Although
486 Tangram provides a computer vision module that segments cell nuclei based
487 on histological images and predicts the cell type for each cell in the spot, it
488 performs random cell assignment for the segmented maskings within the spot

489 and still cannot definitely resolve cell position[17]. In comparison, STASCAN
490 introduces a predicting module at sub-resolution, allowing the assignment of
491 cell types to sub-divided spots based on the tissue morphological features
492 corresponding to the exact positions. In human intestinal tissue, STASCAN
493 precisely pinpointed the fine-grained cell subtypes between distinct cell layers.

494 The advantages of STASCAN enable it to have a broader range of
495 applications. For example, STASCAN can effectively predict cell distribution of
496 some regions with missing sequencing data due to a low number of captured
497 UMIs and genes. In human cardiac tissues, STASCAN could fill in the missing
498 data, and help to display an extensive spatial proximity distribution of fibroblast
499 and myeloid cells in the vicinity of scar tissue. This provides a better
500 understanding of the relationship between fibroblasts and immune cells in
501 myocardial infarction, such as immune factors in stimulating fibroblast
502 transformation[31]. Additionally, the cell proposition annotations generated by
503 STASCAN is supported to cluster different cell-type niches, providing novel
504 insights into the microenvironment at the cellular resolution. In human cardiac
505 tissues, STASCAN generated distinct niches including inflammatory, myogenic,
506 and fibrotic cell-type niches, with the best representation of the dynamic
507 diversity of cells in the microenvironment during myocardial infarction. In
508 embryonic mouse brain tissue, STASCAN clearly revealed major anatomical
509 tissue organizations across the development stages, with gradually increased
510 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 stands 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.

In summary, STASCAN is a universal and precise cell type prediction method, which 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 potential for resolving cell type distribution in various types of tissues across different stages of development, regeneration, and disease. The positional relationships of different 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

533 pathological cells and further link spatial cellular distribution to disease
534 diagnosis cost-efficiently, through extensive data training, eliminating the need
535 for sequencing.
536

537 **Reference**

538 1 Longo, S. K., Guo, M. G., Ji, A. L. & Khavari, P. A. Integrating single-cell
539 and spatial transcriptomics to elucidate intercellular tissue dynamics.
540 Nat Rev Genet. 2021;22:627-644.
541 2 Rao, A., Barkley, D., Franca, G. S. & Yanai, I. Exploring tissue
542 architecture using spatial transcriptomics. Nature. 2021;596:211-220.
543 3 Vickovic, S. et al. High-definition spatial transcriptomics for in situ tissue
544 profiling. Nat Methods. 2019;16:987-990.
545 4 Liu, Y. et al. High-spatial-resolution multi-omics sequencing via
546 deterministic barcoding in Tissue. Cell. 2020;183:1665-1681.e1618.
547 5 Moses, L. & Pachter, L. Museum of spatial transcriptomics. Nat
548 Methods. 2022;19:534-546.
549 6 Stark, R., Grzelak, M. & Hadfield, J. RNA sequencing: the teenage
550 years. Nat Rev Genet. 2019;20:631-656.
551 7 Ståhl, P. Visualization and analysis of gene expression in tissue
552 sections by spatial transcriptomics. Science. 2016;353:78-82.
553 8 Stickels, R. R. et al. Highly sensitive spatial transcriptomics at
554 near-cellular resolution with Slide-seqV2. Nat biotechnol.

555 2021;39:313-319.
556 9 Chen, A. et al. Spatiotemporal transcriptomic atlas of mouse
557 organogenesis using DNA nanoball-patterned arrays. Cell.
558 2022;185:1777-1792.e1721.
559 10 Asp, M. et al. A spatiotemporal organ-wide gene expression and cell
560 atlas of the developing human heart. Cell. 2019;179:1647-1660.e1619.
561 11 Chen, A. et al. Single-cell spatial transcriptome reveals cell-type
562 organization in the macaque cortex. Cell. 2023;186:3726-3743.e24.
563 12 Li, B. et al. Benchmarking spatial and single-cell transcriptomics
564 integration methods for transcript distribution prediction and cell type
565 deconvolution. Nat Methods. 2022;19:662-670.
566 13 Kleshchevnikov, V. et al. Cell2location maps fine-grained cell types in
567 spatial transcriptomics. Nat Biotechnol. 2022;40:661-671.
568 14 Stuart, T. et al. Comprehensive integration of single-cell data. Cell.
569 2019;177:1888-1902.e1821.
570 15 Cable, D. M. et al. Robust decomposition of cell type mixtures in spatial
571 transcriptomics. Nat Biotechnol. 2022;40:517-526.
572 16 Maaskola, J. et al. Charting tissue expression anatomy by spatial
573 transcriptome deconvolution. BioRxiv. 2018.
574 https://doi.org/10.1101/362624.
575 17 Biancalani, T. et al. Deep learning and alignment of spatially resolved
576 single-cell transcriptomes with Tangram. Nat Methods.

577 2021;18:1352-1362.
578 18 Perlman, Z. E. et al. Multidimensional drug profiling by automated
579 microscopy. Science. 2004;306:1194-1198.
580 19 Feldman, D. et al. Optical pooled screens in human cells. Cell.
581 2019;179:787-799.e717.
582 20 Loo, L. H., Wu, L. F. & Altschuler, S. J. Image-based multivariate
583 profiling of drug responses from single cells. Nat Methods.
584 2007;4:445-453.
585 21 Bao, F. et al. Integrative spatial analysis of cell morphologies and
586 transcriptional states with MUSE. Nat Biotechnol. 2022;40:1200-120.
587 22 He, B. et al. Integrating spatial gene expression and breast tumour
588 morphology via deep learning. Nat Biomed Eng. 2020;4:827-834.
589 23 Bergenstrahle, L. et al. Super-resolved spatial transcriptomics by deep
590 data fusion. Nat Biotechnol. 2021;40:476-479.
591 24 Monjo, T., Koido, M., Nagasawa, S., Suzuki, Y. & Kamatani, Y. Efficient
592 prediction of a spatial transcriptomics profile better characterizes breast
593 cancer tissue sections without costly experimentation. Sci Rep.
594 2022;12:4133.
595 25 Cui, G. et al. Spatiotemporal transcriptomic atlas reveals the dynamic
596 characteristics and key regulators of planarian regeneration. Nat
597 Commun. 2023;14:3205.
598 26 Benham-Pyle, B. W. et al. Identification of rare, transient post-mitotic

599 cell states that are induced by injury and required for whole-body
600 regeneration in Schmidtea mediterranea. Nat Cell Biol.
601 2021;23:939-952.
602 27 Fawkner-Corbett, D. et al. Spatiotemporal analysis of human intestinal
603 development at single-cell resolution. Cell. 2021;184:810-826.e823.
604 28 Kadur Lakshminarasimha Murthy, P. et al. Human distal lung maps and
605 lineage hierarchies reveal a bipotent progenitor. Nature.
606 2022;604:111-119.
607 29 Kuppe, C. et al. Spatial multi-omic map of human myocardial infarction.
608 Nature. 2020;608:766-777.
609 30 Shook, B. A. et al. Myofibroblast proliferation and heterogeneity are
610 supported by macrophages during skin repair. Science.
611 2018;362:eaar2971.
612 31 Pakshir, P. et al. Dynamic fibroblast contractions attract remote
613 macrophages in fibrillar collagen matrix. Nat Commun. 2019;10:1850.
614 32 Jiang, F. et al. Simultaneous profiling of spatial gene expression and
615 chromatin accessibility during mouse brain development. Nat Methods.
616 2023;20:1048-1057.
617 33 An, S. et al. Airway smooth muscle dynamics: a common pathway of
618 airway obstruction in asthma. Eur Respir J. 2007;29:834-860.
619 34 Cho, N., Razipour, S. E. & McCain, M. L. Featured Article: TGF-β1
620 dominates extracellular matrix rigidity for inducing differentiation of

621 human cardiac fibroblasts to myofibroblasts. Exp Biol Med.
622 2018;243:601-612.
623 35 Abadi, M. et al. TensorFlow: Large-scale machine learning on
624 heterogeneous distributed systems. ArXiv. 2016. abs/1603.04467.
625 36 Simonyan, K. Z., A. Very deep convolutional networks for large-scale
626 image recognition. CoRR. 2014. abs/1409.1556.
627
628

Figure legends

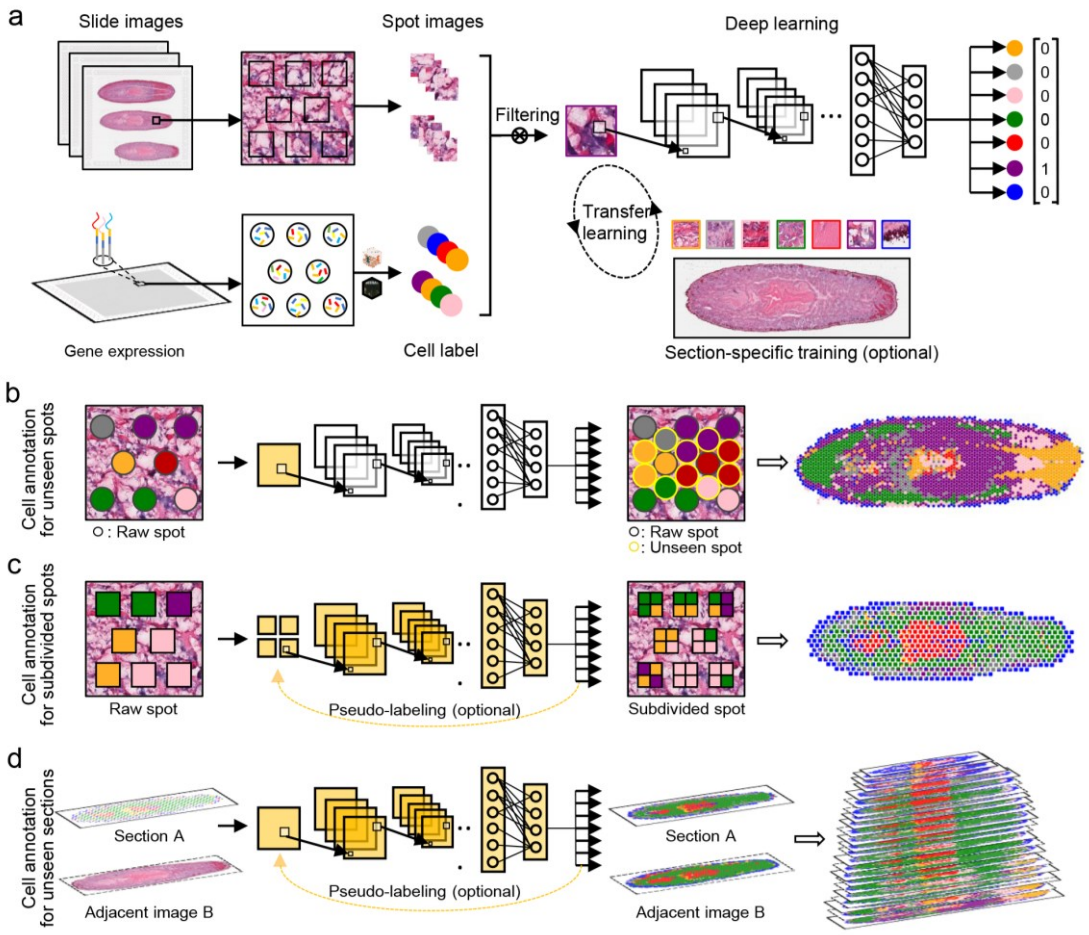

Fig. 1. Overview of STASCAN.

(a) Architecture of the STASCAN framework. STASCAN infers fine-resolution cell distribution with spatial transcriptomic (ST) data based on a deep learning model. Preprocessed spot images and dominant cell labels from slide images and gene expression data are taken as input to train a base CNN model. STASCAN provides an optional process to conduct section-specific training, which performs transfer learning to learn the characteristics of specific sections.

(b) Architecture of cell annotation for unseen spots. STASCAN infers cell types for unseen spot images to decipher the super-resolution cell distribution of a

ST slide. Unseen spots, the artificial spots among the unmeasured area.

(c) Architecture of cell annotation for subdivided spots. STASCAN infers cell types of subdivided spot images to decipher the sub-resolution cell distribution of a ST slide. Raw spots, spots at raw resolution. Subdivided spots, spots by dividing raw spots for high resolution.

(d) Architecture of cell annotation for unseen sections. STASCAN predicts cell types of simulated spots on adjacent images by learning characteristics of prior spot images from ST sections, depicting 3D cell distribution of tissue. Sections, ST sections with sequencing data. Adjacent images, adjacent H&E images without sequencing data.

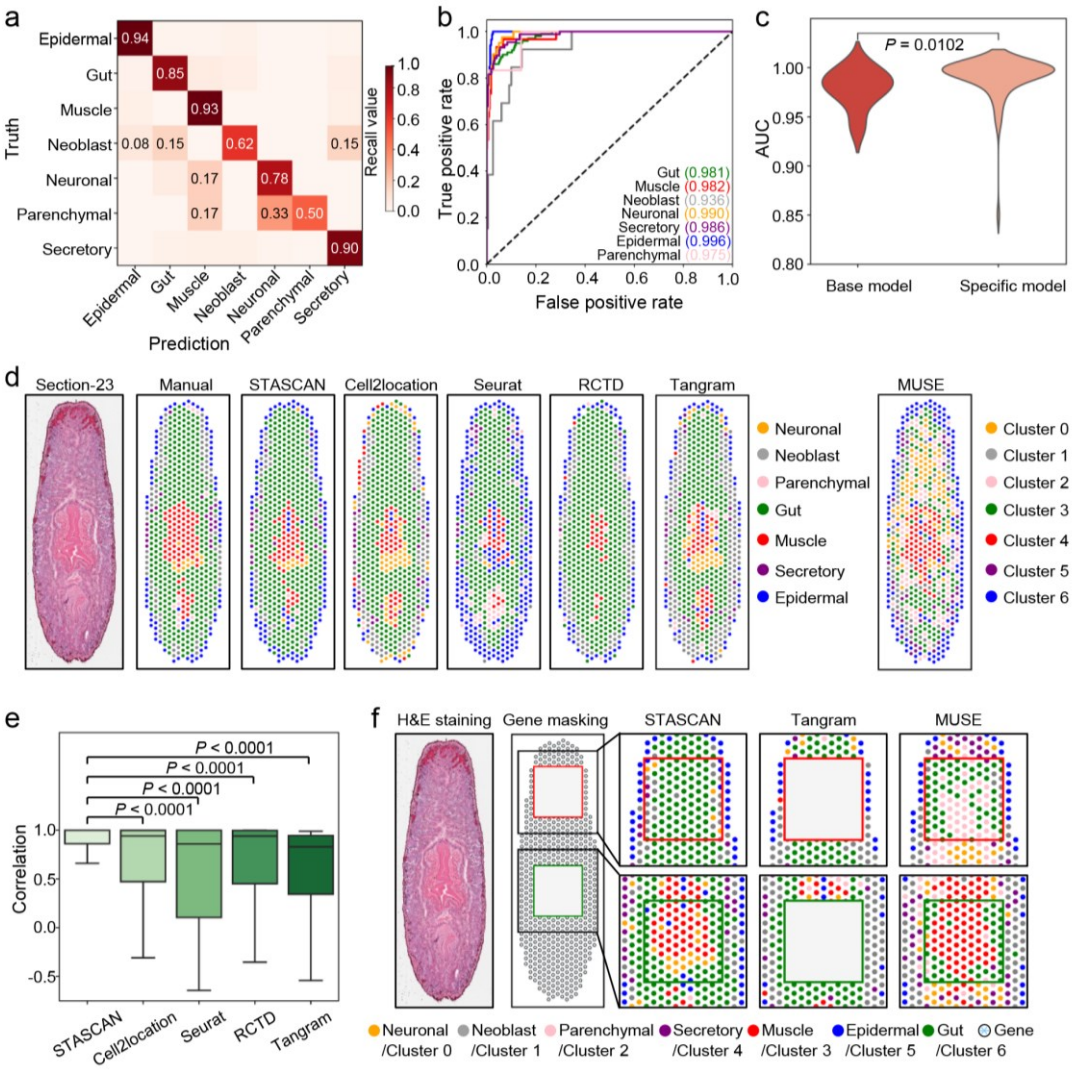

Fig. 2. Evaluation of STASCAN in 10× Visium planarian dataset.

(a) Heatmap showing the confusion matrix of the prediction and truth value of the base model for the 5-fold cross-validation in the planarian dataset. The number in each box represents the recall value.

(b) ROC curves of each major cell type of the base model in the planarian dataset. The number in each pair of parentheses represents the AUC value.

(c) Violin plot showing the AUC values of the base model (n = 7) and specific model (n = 70), respectively. P value is calculated using two-sided Wilcoxon signed-rank tests.

672 **(d)** Cellular distribution of section 23 in the planarian dataset. Left, H&E
673 staining image showing general morphology and spatial plots (x and y axis)
674 depicting the distribution of the major cell types annotated by manual
675 annotation, STASCAN, Cell2location, Seurat, RCTD, and Tangram. Right,
676 spatial plots (x and y axis) depicting the distribution of cell clusters assigned by
677 MUSE.
678 **(e)** Boxplot showing the Pearson correlation between cell distribution of section
679 23 annotated by manual and alternative methods, respectively (n = 683, 683,
680 683, 681, 683). *P* values are calculated using two-sided Wilcoxon signed-rank
681 tests.
682 **(f)** Cellular distribution of section 23 in the planarian dataset with partly gene
683 expression masking. Left, H&E staining image showing general morphology.
684 Middle, schematic of gene expression masking. Right, zoomed regions of
685 spatial plots (x and y axis) depicting the cell distribution predicted by
686 STASCAN, Tangram, and MUSE.
687
688
689
690
691
692
693

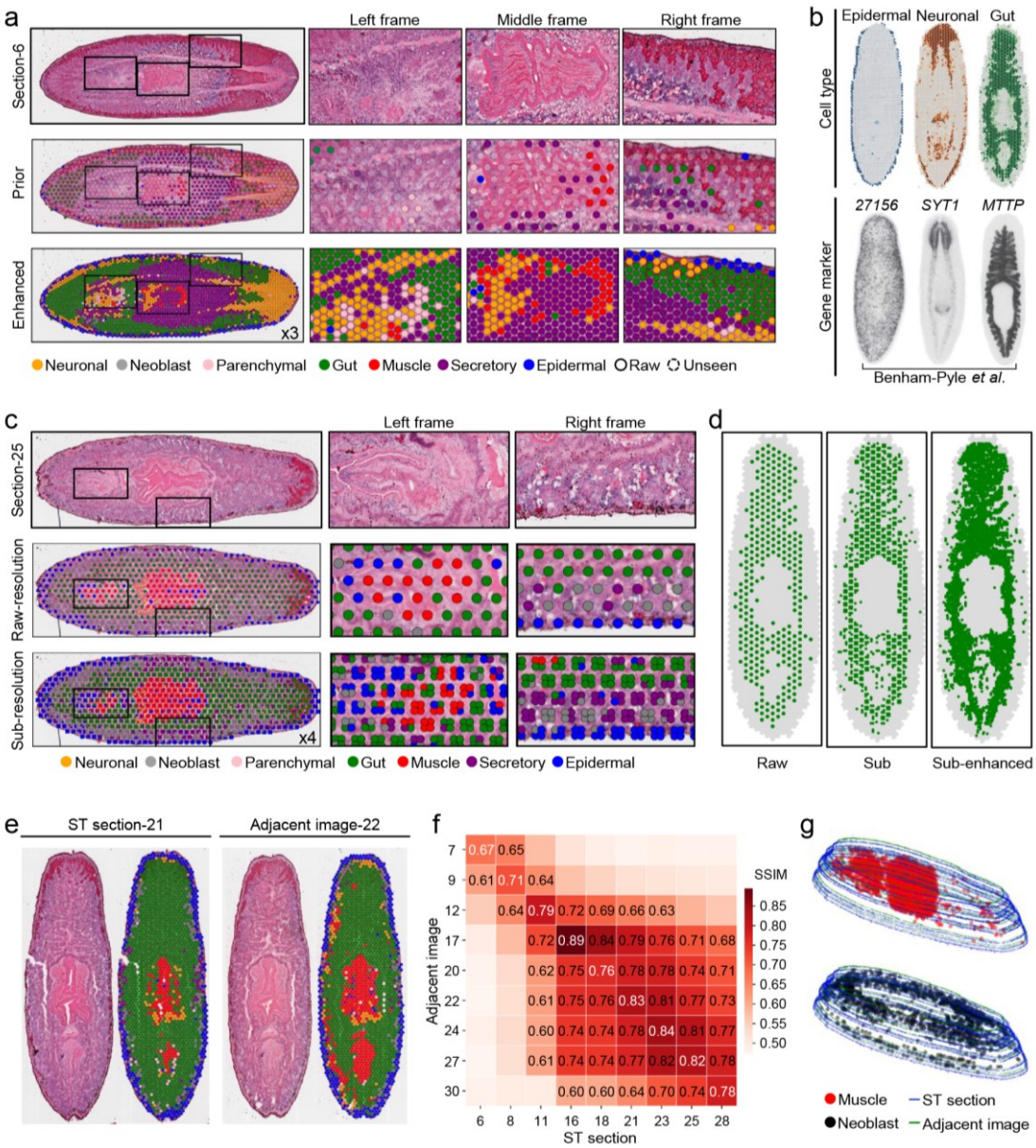

Fig. 3. STASCAN provides comprehensive and multidimensional cell annotation in 10x Visium planarian dataset.

(a) Tissue regions of the planarian in section 6. Left top, H&E staining image showing general morphology. Left center, spatial plot (x and y axis) covering the H&E staining image depicting distribution for the major cell types on the prior spots. Prior spots, the spots that are predicted to be the same cell types by both the Cell2location and Seurat at the raw resolution. Left bottom, spatial

plot (x and y axis) covering the H&E staining image depicting distribution for the major cell types on the fully-tiled spots. The number in the bottom right indicates a 3 times improvement in cellual resolution over the raw resolution. Fully-tiled spots, containing both the raw spots and unseen spots. Raw spots, containing prior spots and other spots at the raw resolution. The latter were assigned with cell types by STASCAN. Unseen spots, the artificial spots among unmeasured areas, inferred with cell types by STASCAN. Right, enlarged view of the black framed area, arranged by left, middle, and right frame.

(b) Visualizations of cell distribution for three major cell types and expression patterns of corresponding marker genes. Top, cell distribution annotated by STASCAN. Bottom, whole-mount in situ hybridization of classic gene markers (from published research[22]).

(c) Tissue regions of the planarian in section 25. Left top, H&E staining image showing general morphology. Left center, spatial plot (x and y axis) covering H&E staining image depicting cell distribution for the major types annotated by STASCAN at raw resolution. Left bottom, spatial plot (x and y axis) covering the H&E staining image to decipher cell distribution for the major types annotated by STASCAN at sub-resolution. The number in the bottom right indicates a 4 times improvement in cellual resolution over the raw resolution. Right, enlarged view of the black framed area, arranged by left and right frame.

723 **(d)** Spatial plots (*x* and *y* axis) depicting cell distribution for gut cells. Left, cell
724 distribution predicted by STASCAN at raw resolution. Middle, cell distribution
725 predicted by STASCAN at sub-resolution. Right, cell distribution predicted by
726 STASCAN at sub-enhanced resolution.
727 **(e)** Tissue region of planarian in section 21 and adjacent image 22. Left, H&E
728 staining image and spatial plots (*x* and *y* axis) depicting cell distribution
729 predicted by STASCAN of section 21. Right, H&E staining image and spatial
730 plots (*x* and *y* axis) to decipher cell distribution predicted by STASCAN of
731 adjacent image 22.
732 **(f)** Heatmap showing the values of SSIM between the cell distribution of
733 adjacent images and nether ST sections. Numbers on the coordinate axis
734 represent the order of ST sections and adjacent images from the same
735 consecutive sections set.
736 **(g)** 3D spatial plots (x, y, and z axis) depicting 3D distribution of two selected
737 cell types. Top, muscle cell. Bottom, neoblast cell. Both 3D models are
738 reconstructed with aligned outlines and spots from the ST section and adjacent
739 images.
740
741
742
743
744

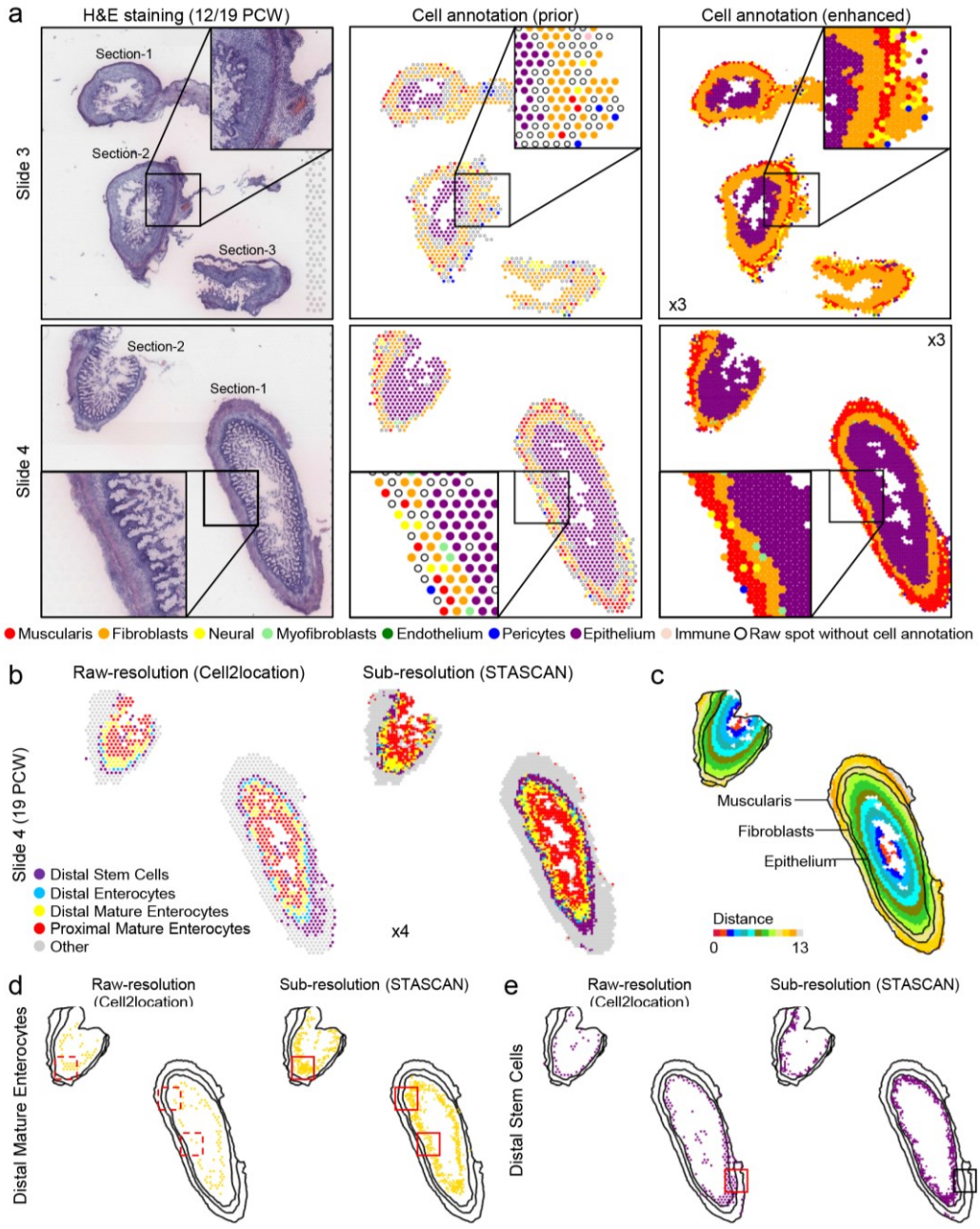

Fig. 4. STASCAN depicts spatial layers of distinct cell types in 10x Visium human intestinal dataset.

(a) Tissue regions of the human intestinal tissue for both the 12 PCW colon (slide 3) and 19 PCW colon (slide 4) respectively. Left, H&E staining image showing general morphology. Middle, spatial plot (x and y axis) depicting distribution for the major cell types of the prior spots. Right, spatial plot (x and y

752 axis) depicting super-resolution cell distribution for the major cell types of the
753 fully-tiled spots. Slide numbers are shown in columns. The numbers indicate a
754 3 times improvement in cellual resolution over the raw resolution.
755 **(b)** Tissue regions of the human intestinal tissue for 19 PCW colon. Left,
756 spatial plot (x and y axis) depicting cell distribution for epithelium subtypes
757 inferred by Cell2location at raw resolution. Right, spatial plot (x and y axis)
758 depicting cell distribution for epithelium subtypes inferred by STASCAN at
759 sub-resolution. The number in the bottom left indicates a 4 times improvement
760 in cellual resolution over the raw resolution.
761 **(c)** Schematic of muscularis, fibroblasts, and epithelium layers, colored in
762 gradients according to the distance from the spot to the tissue edge.
763 **(d)** Spatial plots (*x* and *y* axis) depicting cell distribution for distal mature
764 enterocytes. Left, cell distribution predicted by Cell2location at raw resolution.
765 Right, cell distribution inferred by STASCAN at sub-resolution. The black lines
766 represent the schematic of different morphological layers.
767 **(e)** Spatial plots (*x* and *y* axis) depicting cell distribution for distal stem cells.
768 Left, cell distribution inferred by Cell2location at raw resolution. Right, cell
769 distribution inferred by STASCAN at sub-resolution. The black lines represent
770 the schematic of different morphological layers.
771
772
773

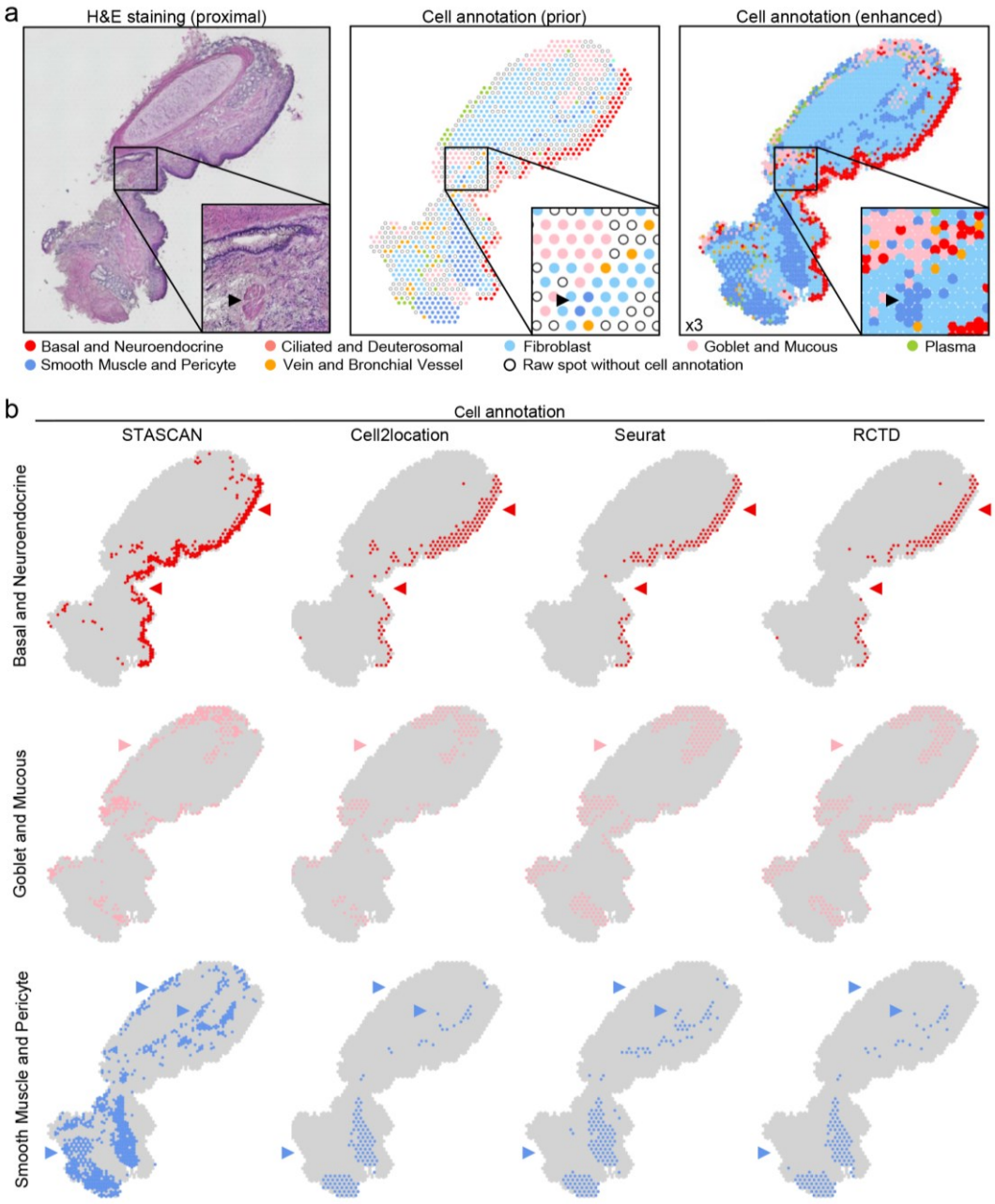

Fig. 5. STASCAN demonstrates the special structure from 10x Visium human lung data.

(a) Tissue regions of the human proximal lung. Left, H&E staining image showing general morphology. Middle, spatial plot (*x* and *y* axis) depicting cell distribution for the major types on prior spots. Right, spatial plot (x and y axis) depicting super-resolution cell distribution for the major types on the fully-tiled

spots. The number in the bottom left indicates a 3 times improvement in cellual resolution over the raw resolution.

(b) Spatial plots (*x* and *y* axis) depicting cell distribution for the three major types. Cell types are shown in columns, and alternative methods (STASCAN, Cell2locaiton, Seurat, and RCTD) used to annotate the corresponding cell type are shown in rows.

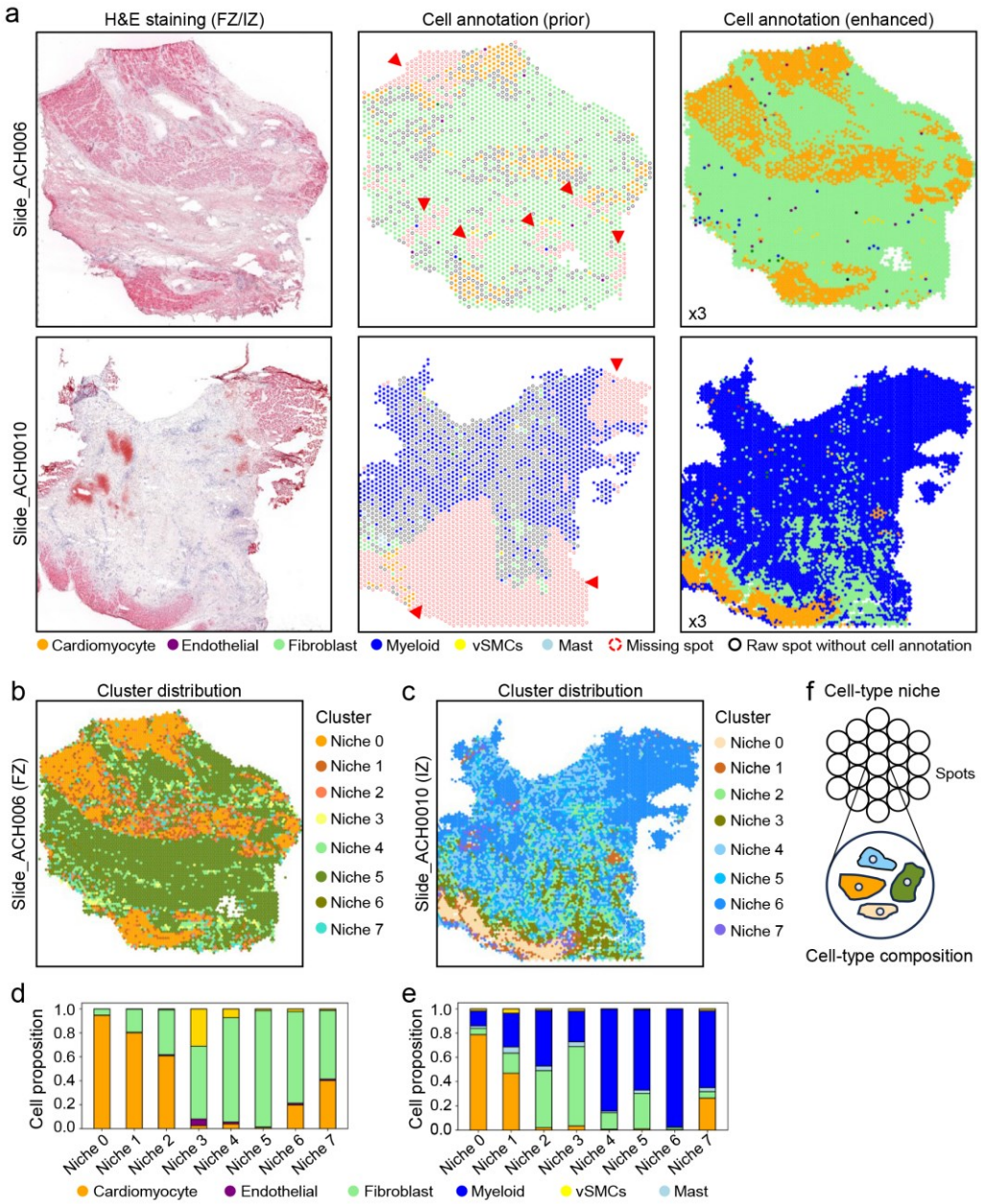

Fig. 6. STASCAN reveals the cell-type niches in 10x Visium human cardiac data.

(a) Tissue regions of the human cardiac tissue from the fibrotic zone (Slide_ACH006) and ischaemic zone (Slide_ACH0010) after myocardial infarction. Left, H&E staining image showing general morphology. Middle, spatial plot (*x* and *y* axis) depicting cell distribution for the major cell types on prior spots. Right, spatial plot (x and y axis) depicting super-resolution cell

distribution for the major cell types on the fully-tiled spots. The numbers in the bottom left indicate a 3 times improvement in cellual resolution over the raw resolution. Slide numbers are shown in columns. Missing spots, spots removed from raw spots in the original literature due to the small number of measured genes and UMIs. FZ, fibrotic zone. IZ, ischaemic zone.

(b) Spatial plots (*x* and *y* axis) depicting spatial distribution for cell-type niches in slide_ACH006, which are assigned by unsupervised clustering of spots with the cell-type compositions.

(c) Spatial plots (*x* and *y* axis) depicting spatial distribution for cell-type niches in slide_ACH0010, which are assigned by unsupervised clustering of spots with the cell-type compositions.

(d) Bar plot showing the cell proposition of cell-type niches in slide_ACH006.

(e) Bar plot showing the cell proposition of cell-type niches in slide_ACH0010.

(f) Schematic of cell-type niches.

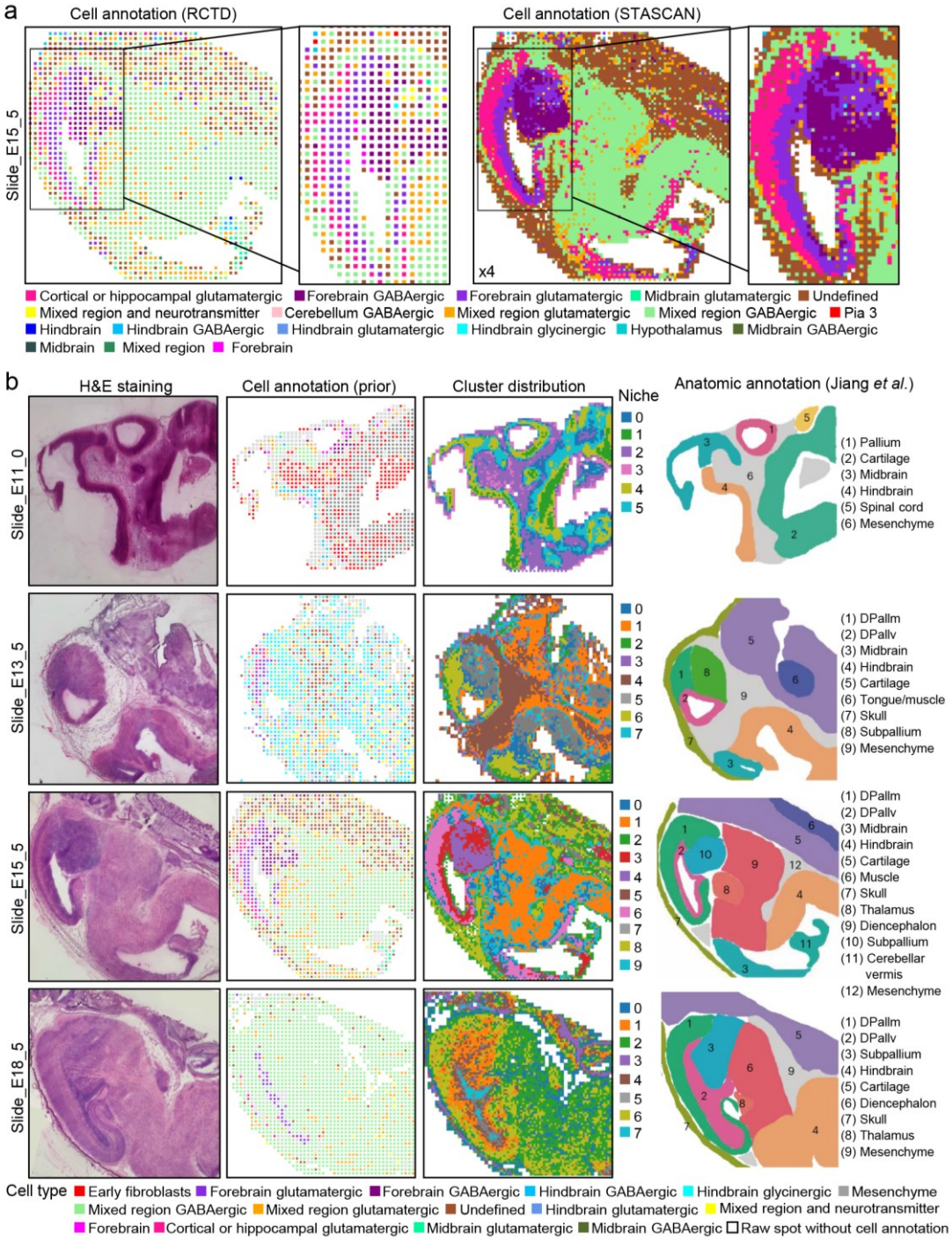

Fig. 7. STASCAN revealed major anatomical tissue regions in mouse brain dataset generated from microfluidic technologies.

(a) Tissue regions of the E15.5 embryonic mouse brain tissue. Left, spatial plot (*x* and *y* axis) depicting cell distribution annotated by RCTD. Right, spatial plot (x and y axis) depicting super-resolution cell distribution annotated by

STASCAN. The number in the bottom left indicates a 4 times improvement in cellual resolution over the raw resolution.

(b) Tissue regions of the embryonic mouse brain tissues across development stages. Left, H&E staining image showing general morphology. Middle left, spatial plot (x and y axis) depicting cell distribution for the major cell types on prior spots. Middle right, spatial plot (x and y axis) depicting spatial distribution for cell-type niches, which are assigned by unsupervised clustering of spots with the cell-type compositions. Slide numbers are shown in columns. Right, manual anatomic annotation of major tissue regions for different mouse brain stages from original literature[29]. DPallm, mantle zone of the dorsal pallium. DPallv, ventricular zone of the dorsal pallium.