Annotation of Spatially Resolved Single-cell Data with STELLAR

Maria Brbić[1,*], Kaidi Cao[1,*], John W. Hickey[2,*], Yuqi Tan[2], Michael P. Snyder[3], Garry P. Nolan[2,4,†], and Jure Leskovec[1,5,†]

[1] Department of Computer Science, Stanford University, Stanford, CA 94305, USA

[2] Baxter Laboratories Department of Microbiology and Immunology, Stanford University School of Medicine, Stanford University, Stanford, CA 94305, USA

[3] Department of Genetics, Stanford University School of Medicine, Stanford University, Stanford, CA 94305, USA

[4] Department of Pathology, Stanford University School of Medicine, Stanford University, Stanford, CA 94305, USA

[5] Chan Zuckerberg Biohub, San Francisco, CA 94158, USA

[*]These authors contributed equally.
[†]Corresponding author. Email: gnolan@stanford.edu, jure@cs.stanford.edu

Abstract

Spatial protein and RNA imaging technologies have been gaining rapid attention but current computational methods for annotating cells are based on techniques established for dissociated single-cell technologies and thus do not take spatial organization into account. Here we present STELLAR, a geometric deep learning method that utilizes spatial and molecular cell information to automatically assign cell types from an annotated reference set as well as discover new cell types and cell states. STELLAR transfers annotations across different dissection regions, tissues, and donors and detects higher-order tissue structures with dramatic time savings.

Main

Use of spatial protein and RNA imaging technologies has opened new opportunities for understanding location-dependent properties of cells and molecules [1–4]. Tissue-mapping consortia efforts such as the Human BioMolecular Atlas Program [5], Human Tumor Atlas Network [6] , and Human Cell Atlas [7] will generate comprehensive cell atlas datasets. These consortia efforts necessitate computational methods that can assist with robust cell type annotation of comprehensive spatial datasets and aid in understanding functional spatial biology. While tools for dissociated single-cell technologies have been used to annotate spatial datasets [4,8–10], these tools ignore the spatial information crucial for more accurate annotation of spatial datasets and typically require manual annotation. What is therefore needed are computational approaches that simultaneously leverage the additional spatial features as well as the molecular features of cells.

Here, we present STELLAR (SpaTial cELl LeARning), a geometric deep learning tool for cell-type discovery and identification in spatially resolved single-cell datasets that uses spatial and molecular features of single cells. STELLAR takes as input *(i)* an annotated spatially resolved single cell-dataset with cell types defined by molecular profiles (reference dataset), and *(ii)* a completely unannotated spatially resolved single-cell dataset with cell types that are unknown (Fig. 1). The reference and unannotated datasets do not need to be related and can belong to different dissection regions, different donors, or different tissue types.

STELLAR has two unique properties. First, using graph convolutional neural networks, STELLAR learns latent low-dimensional cell representations that jointly capture spatial and molecular similarities of cells (Methods). In this way, cells that are spatially close together and that have similar levels of gene or protein expression are embedded close to each other. Second, in a new completely unannotated dataset, STELLAR automatically assigns cells to cell types included in the reference set and also identifies cells with unique properties as belonging to a novel type that is not part of the reference set (Methods and Supplementary Fig. 1). Because of these two unique features STELLAR solves major limitations with current annotation tools. While supervised machine learning methods [11–13] can classify cells into preexisting cell types, they are not effective for transferring annotations across different biological contexts that do not necessarily share the cell types with the reference tissue. On the other hand, unsupervised machine learning methods [14,15] cannot automatically assign cell labels to cell types without expert reannotation. STELLAR has the unique ability to assign cells to one of the cell types seen in the reference set, or discover a novel cell type for previously uncharacterized cell type. From a practical standpoint this ability is

essential for leveraging reference sets generated in a disease-free state or in different tissues that may not have all the cell types represented in the target tissue but that may include most established cell types. STELLAR can be applied to both multiplexed protein and multiplexed RNA imaging datasets.

To demonstrate STELLAR's ability to assigns cells to one of the cell types seen in the reference set and discover a novel cell type, we applied STELLAR on CODEX multiplexed imaging data [9, 10, 16]. We used data from human tonsil as the reference set (Fig. 2a and Supplementary Fig. 2) and tissue from a patient with Barrett's esophagus (BE) as the unannotated target tissue (Fig. 2b). Tonsils are often used for testing antibody panels as they contain high numbers of immune cell markers. The BE tissue contains differentiated epithelial cells, but tonsils do not. Indeed, three subtypes of epithelial cells appear only in the BE dataset, while B cells appear only in tonsil data. Moreover, the two datasets have different distributions of cell types. For example, smooth muscle cells are major cluster in BE, but form the smallest group in tonsil (Supplementary Fig. 3). Despite these major differences, we found that STELLAR accurately assigned cell types to 92% of cells in the BE data that are present in the reference dataset from tonsil. Moreover, STELLAR identified three BE-specific subtypes of epithelial cells as a newly discovered cell type (Fig. 2c,d).

We also compared STELLAR to different machine learning methods trained on the tonsil dataset, and evaluated on the BE dataset. STELLAR substantially outperformed the best alternative method, achieving 60% higher accuracy (Fig. 2e and Supplementary Fig. 4). STELLAR was also the best performer based on other evaluation metrics, such as F1-score, precision, and recall (Supplementary Fig. 5), and STELLAR outperformed other alternative methods by 6% even when evaluated only on cell types shared between tonsil and BE datasets (Supplementary Fig. 6). We further directly measure gains obtained by including spatial information and find that spatial information in STELLAR achieves 11% improvement over only molecular information (Fig. 2f and Methods). Additionally, we evaluate STELLAR on a spatial transcriptomics MERFISH dataset from mouse primary motor cortex and show that STELLAR achieves 95% accuracy and it is robust to the number of novel cell types (Supplementary Fig. 7).

A critical current bottleneck in analysis of the spatially resolved single cells for consortia efforts is an accurate assignment of granular cell-type labels across donors of the same tissue. For example, a typical CODEX dataset with a 48-marker panel used to analyze four tissues of the intestine requires at least 25 hours of work to cluster, merge, recluster, subcluster, and assign cell types based on average marker expression and location within CODEX images. As a part of the

HuBMAP consortium, we have generated CODEX imaging data for tissues from three donors for a total of 24 tissues from a healthy intestine. We first evaluated the performance of STELLAR on this more granular data and used expert-annotated cell-type labels of images of different regions of healthy colon from one donor. We evaluated STELLAR in the leave-one-region-out setting where we trained on three regions and predicted annotations on the fourth region. STELLAR had high accuracy in cell-type label transfers (Fig. 3a) and had higher performance (F-score of 0.8) than tested unsupervised clustering methods without manual intervention (F-score of 0.3 across different methods) [17].

Encouraged by these results, we next used our expert-annotated samples from one donor as training data and applied STELLAR to unannotated samples from two other donors. We find that the fluorescent CODEX data of new donors (Fig. 3b and Supplementary Fig. 8) agree well with the cell-type map predicted by STELLAR (Fig. 3c and Supplementary Fig. 9). We additionally confirmed the quality of STELLAR's predictions by looking at average marker expression profiles of predicted cell types for the samples from a different donor (Fig. 3d). Protein marker distributions match expert hand-annotated profiles in fine-grained cell types predicted by STELLAR. For example, CD4 is expressed only in cell types annotated as $CD4^+$ T cells, whereas CD8 is expressed exclusively in $CD8^+$ T cells. This additionally confirms that STELLAR predictions are reliable. Expert annotation of these images would require approximately 150 hours of manual labor, while with STELLAR it took the expert only 2 hours to annotate the images.

The classification of cell-cell interactions enabled by multiplexed imaging of many cell types simultaneously unlocks the ability to identify and discover multicellular structures. Characterization of multicellular modules like tertiary lymphoid structures is critical to understanding of tissue function, its relationship with disease, and how to design of effective therapeutics [18]. We hypothesized that the latent cell embedding space defined by STELLAR could reveal biologically meaningful information about the tissue organization. To explore this, we clustered cells in the STELLAR embedding space using Louvain clustering [19] and analyzed the local microenvironment for each resultant cluster. We found that the clusters represent major multicellular structural features of the intestine such as the *immune follicle* structure, which was enriched for $CD4^+$ T cells, B cells, and dendritic cells, as well as *secretory epithelial* which was enriched for goblet cells, neuroendocrine cells, and transit amplifying cells (Fig. 3e-f and Supplementary Fig. 10). Similarly structural information was obtained upon STELLAR analysis of the MERFISH data (Supplementary Fig. 11). Thus, the embedding space from STELLAR also provides a method for

identifying multicellular structures in tissues.

Development of STELLAR was motivated by a growing need to leverage spatial and molecular information across different biological contexts. Multiplexed imaging technologies will drive future efforts to understand both healthy and diseased tissue processes, enabled by large consortia efforts that generate comprehensive datasets and standardize computational methods. Novel insights will come from the rich, but still unexplored and underused information about spatial cell organization. We anticipate that STELLAR will alleviate future annotation efforts of large-scale spatial cell atlases, identify major multicellular structures in tissues, and reveal how cells cooperatively coordinate to enable tissues to function.

Methods

Overview of STELLAR. STELLAR learns spatial and molecular cell similarities that are transferable across different biological contexts, such as different dissections regions, donors, or tissues. Across different contexts, STELLAR learns to automatically assign cells to cell types seen in the annotated reference set, or forms novel cell types if cells have unique properties that are not present in the reference dataset.

Specifically, STELLAR starts with an annotated reference cell graph $\mathcal{G}_a = \{\mathcal{V}_a, \mathcal{E}_a\}$ with molecular features for all nodes $\mathbf{x}_{v_i} \in \mathbb{R}^D, \forall v_i \in \mathcal{V}_a$, and an unannotated cell graph $\mathcal{G}_u = \{\mathcal{V}_u, \mathcal{E}_u\}$ with molecular features $\mathbf{x}_{v_i} \in \mathbb{R}^D, \forall v_i \in \mathcal{V}_u$. The nodes in each graph correspond to cells and cells are connected if they are spatially close. Node features correspond to gene or protein expressions of cells where D denotes the total number of measured genes and proteins. For the reference graph we assume we are given a vector of cell annotations $\mathbf{y} = \{y_i \in 1, ..., K_a\}_{i=1}^{|\mathcal{V}_a|}$ that assigns each cell to one of the K_a cell types (or some other annotations).

Given reference graph $\mathcal{G}_a$ and unannotated graph $\mathcal{G}_u$, STELLAR first applies the encoder function $f_\theta : \mathbb{R}^D \rightarrow \mathbb{R}^d$ that maps cells from both graphs into a joint embedding space that captures spatial and molecular similarities between the cells. The cell embedding encoder function f_θ is parameterized by learnable parameters θ of a graph convolutional neural network (GCN) [20]. The encoder function f_θ generates d-dimensional cell embeddings $\mathbf{z}_i \in \mathbb{R}^d, \forall v_i \in \mathcal{V}_a \cup \mathcal{V}_u$. On top of the encoder function, we add a single linear layer parameterized by a weight matrix $\mathbf{W} : \mathbb{R}^D \rightarrow \mathbb{R}^{K_a+K_n}$, where K_n corresponds to expected number of novel cell types that need to be discovered. Next, a softmax layer is added that assigns each cell to one of the $K_a + K_n$ cell types.

Graph construction. Given spatial cell coordinates of each cell, STELLAR first calculates Euclidean distances $d_{i,j}$ for each pair of cells (v_i, v_j) from the same region, and edge (v_i, v_j) is added to the edge set $\mathcal{E}$ if $d_{i,j} < \tau$, where τ is a tunable threshold. In all experiments, we select the value of τ so that each cell has an average of 5 edges connected. Graph construction step is independent of the subsequent method and can be changed as long as the constructed graph meaningfully reflects spatial similarities between cells.

STELLAR encoder. The STELLAR encoder contains one fully connected layer followed by nonlinear activation function:

$$\mathbf{h}_i^{(1)} = \phi(\mathbf{W}^{(0)}\mathbf{h}_i^{(0)} + \mathbf{b}^{(0)}), \tag{1}$$

where $\mathbf{h}_i^{(k)}$ is the hidden state of node v_i in k-th layer of the neural network and $k = 0, 1$. $\mathbf{W}$ is parameter matrix, $\mathbf{b}$ is bias vector and ϕ denotes nonlinear activation function. The hidden state $\mathbf{h}_i^{(0)}$ in layer 0 is set to node features $\mathbf{x}_{v_i}$ *i.e.*, gene or protein expression vector. The Rectified Linear Unit (ReLU) is used as the as activation function ϕ: $\text{ReLU}(\cdot) = \max(0, \cdot)$.

We then use a graph convolutional layer [20] to enable message passing among nearby cells:

$$\mathbf{h}_i^{(2)} = \phi\big(\sum_{j \in \mathcal{N}^i} \mathbf{W}^{(1)} \mathbf{h}_j^{(1)} + \mathbf{h}_i^{(1)}\big), \tag{2}$$

where $\mathcal{N}^i$ denotes neighborhood on node v_i using ReLU as the activation function. The final embedding of node v_i is $\mathbf{z}_i = \mathbf{h}_i^{(2)}$. In non-spatial STELLAR (Fig. 2f), we replace graph convolutional layer with another fully connected layer.

STELLAR objective function. The objective function in STELLAR assigns cells to cell types from the reference set or discovers novel cell types. Inspired by the recent work [21], the objective function consists of two main components: *(i)* a component for discovering novel cell types, and *(ii)* a component for learning to recognize cell types from the reference set (Supplementary Fig. 1).

In the component for novel cell type discovery, we use an objective term that predicts pairwise similarities given cell embeddings obtained using STELLAR encoder, *i.e.*, we predict whether two cells are similar or not. For the reference graph, we use ground-truth annotations to learn to predict similarity between two cells, that is two cells are similar if they belong to the same cell type. For the unannotated graph, pseudo-labels are generated based on the distances between cells in the embedding space. In particular, for each cell within the mini-batch, we identify the most similar nearest neighbor cell and generate pseudo-labels for the given pair. In that way, pseudo-labels are generated only for the pairs in which there is the most confidence. To find nearest neighbors, cell embeddings before the graph convolutional layer $h^{(1)}$ are used to allow cells from the reference graph to be selected as the neighbors of cells from the unannotated graph. Formally, the component for discovering cell types (DCT) minimizes the following term:

$$\mathcal{L}_{\text{DCT}} = \frac{1}{N_a + N_u} \sum_{\substack{\mathbf{z}_i, \mathbf{z}'_i \in \\ (\mathcal{Z}_a \cup \mathcal{Z}_u, \mathcal{Z}'_a \cup \mathcal{Z}'_u)}} - \log\langle \sigma(\mathbf{W}^T \cdot \mathbf{z}_i), \sigma(\mathbf{W}^T \cdot \mathbf{z}'_i) \rangle, \tag{3}$$

where $\mathbf{W}$ denotes linear layer weight matrix, σ denotes softmax function, N_a and N_u denote numbers of cells for annotated and unannotated graphs, respectively. $\mathcal{Z}_a$ and $\mathcal{Z}_u$ denote cell embeddings

for annotated and unannotated graphs, and $\mathcal{Z}'_a$ and $\mathcal{Z}'_u$ denote the set of closest neighbors of $\mathcal{Z}_a$ and $\mathcal{Z}_u$ used for generating pseudo-labels.

The second component for recognizing cell types learns to distinguish cell types from the reference graph using ground-truth cell annotations $\mathbf{y}$. Standard cross-entropy classification loss was enhanced with an adaptive margin mechanism that controls the learning speed of cell types from the reference set compared to novel cell types. Formally, STELLAR minimizes the following objective term to learn to recognize cell types (RCT) in the reference graph:

$$\mathcal{L}_{\text{RCT}} = \frac{1}{N_a} \sum_{z_i \in \mathcal{Z}_a} -\log \frac{e^{s(\mathbf{W}_{y_i}^T \cdot \mathbf{z}_i + \bar{u})}}{e^{s(\mathbf{W}_{y_i}^T \cdot \mathbf{z}_i + \bar{u})} + \sum_{j \neq i} e^{s\mathbf{W}_{y_j}^T \cdot \mathbf{z}_i}}, \tag{4}$$

where s is temperature scaling parameter and $\bar{u}$ is uncertainty. The uncertainty is estimated as the average confidence of unlabeled examples computed from the output of the softmax function:

$$\bar{u} = \frac{1}{N_u} \sum_{v_i \in \mathcal{V}_u} 1 - \max_k \Pr(y_i = k | v_i), \tag{5}$$

where k goes over all reference and novel cell types. At the start of the training uncertainty is large, leading to a larger margin and forcing larger intra-class variance [22]. As training proceeds, the margin becomes smaller and the objective boils down to standard cross-entropy.

Additionally, we use maximum entropy regularization term to avoid a trivial solution of assigning all cells to the same cell type. In particular, the regularization term is the following:

$$\mathcal{R} = \frac{1}{N_a + N_u} \sum_{\mathbf{z}_i \in \mathcal{Z}_l \cup \mathcal{Z}_u} \sum_k \sigma_k(\mathbf{W}^T \cdot \mathbf{z}_i) \log \sigma_k(\mathbf{W}^T \cdot \mathbf{z}_i), \tag{6}$$

where k goes over all reference and novel cell types and σ_k denotes kth cell type value of the softmax output.

Finally, the objective function in STELLAR combines reference cell type recognition, novel cell type discovery, and regularization components:

$$\mathcal{L}_* = \mathcal{L}_{RCT} + \eta_1 \mathcal{L}_{DCT} + \eta_2 \mathcal{R}, \tag{7}$$

where η_1 and η_2 are regularization parameters.

Architecture and hyperparameters. The encoder network in STELLAR consists of one fully-connected layer with ReLU activation and a graph convolutional layer with a hidden dimension of 128 in all layers. It uses the Adam optimizer with an initial learning rate of 10^{-3} and weight decay

0. The model is trained with a batch size of 512 for 20 epochs. A cluster sampler [23] first clusters input graphs into sub-graphs and then assigns the sub-graphs to mini-batches. The temperature scaling parameter s in eq. (4) is set to 10. Regularization parameters in (7), η_1 add η_2, are set to 1 and 0.3, respectively. These hyperparameters were used across all experiments.

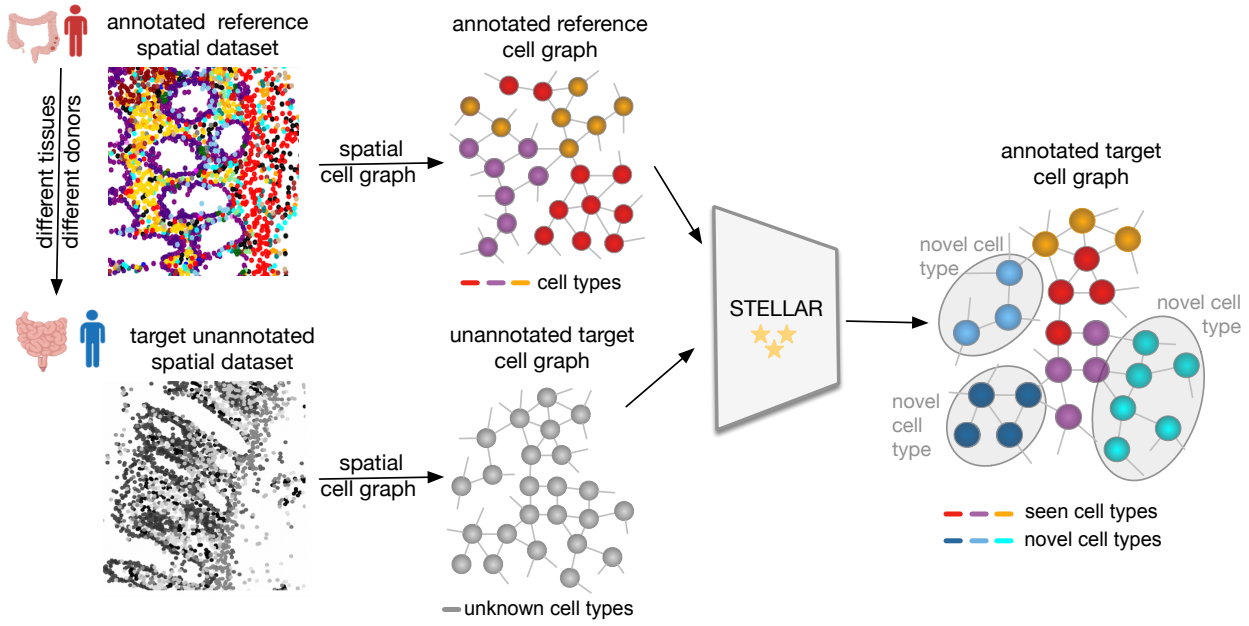

Figure 1: STELLAR is a geometric deep learning framework for annotating spatially resolved single-cell datasets. Given a reference spatially resolved single-cell dataset in which cells are annotated according to their cell types, STELLAR assigns cells in an unannotated spatial single-cell dataset to cell types included in the reference dataset or discovers novel cell types as a group of cells with unique properties not present in the reference dataset. STELLAR uses a graph convolutional encoder to learn low-dimensional cell embeddings that capture cell topology as well as their molecular profiles. Reference and unannotated datasets can originate from different sources as STELLAR transfers information across different tissues and different donors.

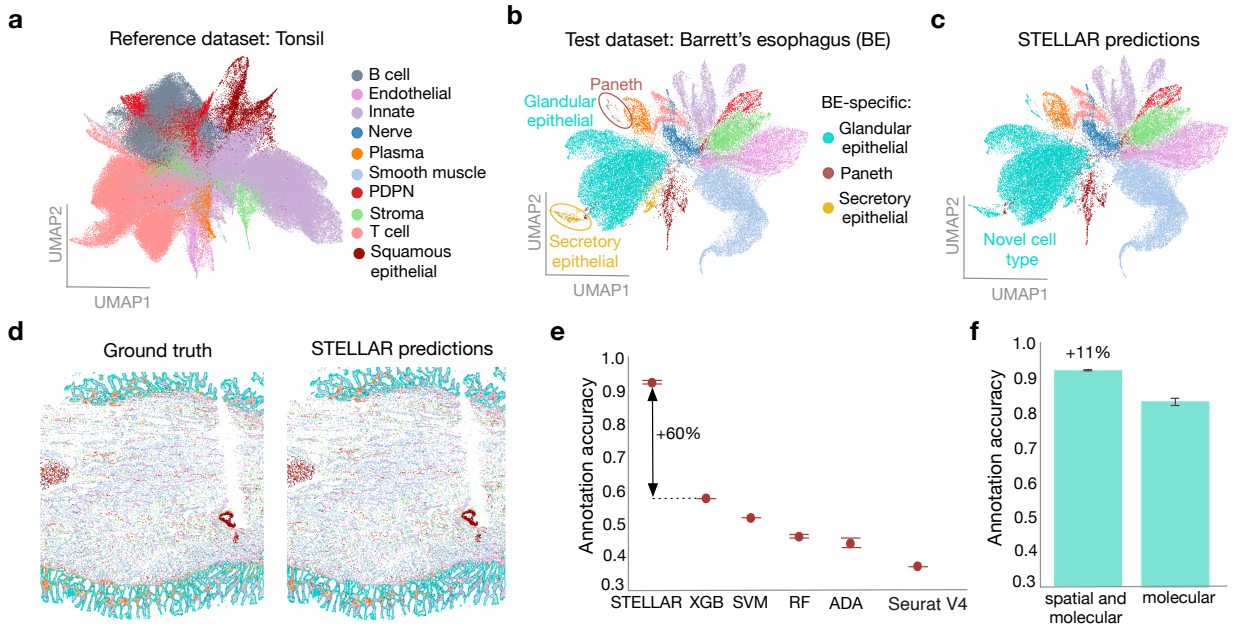

Figure 2: STELLAR accurately identifies cell types from the reference set and discovers novel cell types that have never been characterized in the reference set. **(a)** UMAP embedding of the healthy tonsil data used as the reference dataset. Colors denote ground truth cell types annotations used to train STELLAR. **(b, c)** UMAP embedding of the Barrett's esophagus (BE) data used as the test set. Three subtypes of epithelial cells are not found in the tonsil reference data. Colors denote **(b)** ground-truth cell-type annotations and **(c)** STELLAR predictions. **(d)** CODEX image of BE in spatial coordinates colored according to ground-truth annotations (left) and STELLAR predictions (right). **(e)** Accuracy of cell-type assignments in BE dataset by STELLAR compared to alternative approaches (XGBoost [24], SVM [25], Random Forest [26], AdaBoost [27], Seurat V4 [28]). Performance was evaluated as a mean accuracy score across five runs of each method. Error bars are from standard deviation. **(f)** Accuracy of cell-type assignments in BE dataset by STELLAR and when modifying STELLAR to rely only on the molecular information in cell type assignment (Methods). Performance was evaluated as a mean accuracy score across five runs of each method. Error bars are from standard deviation.

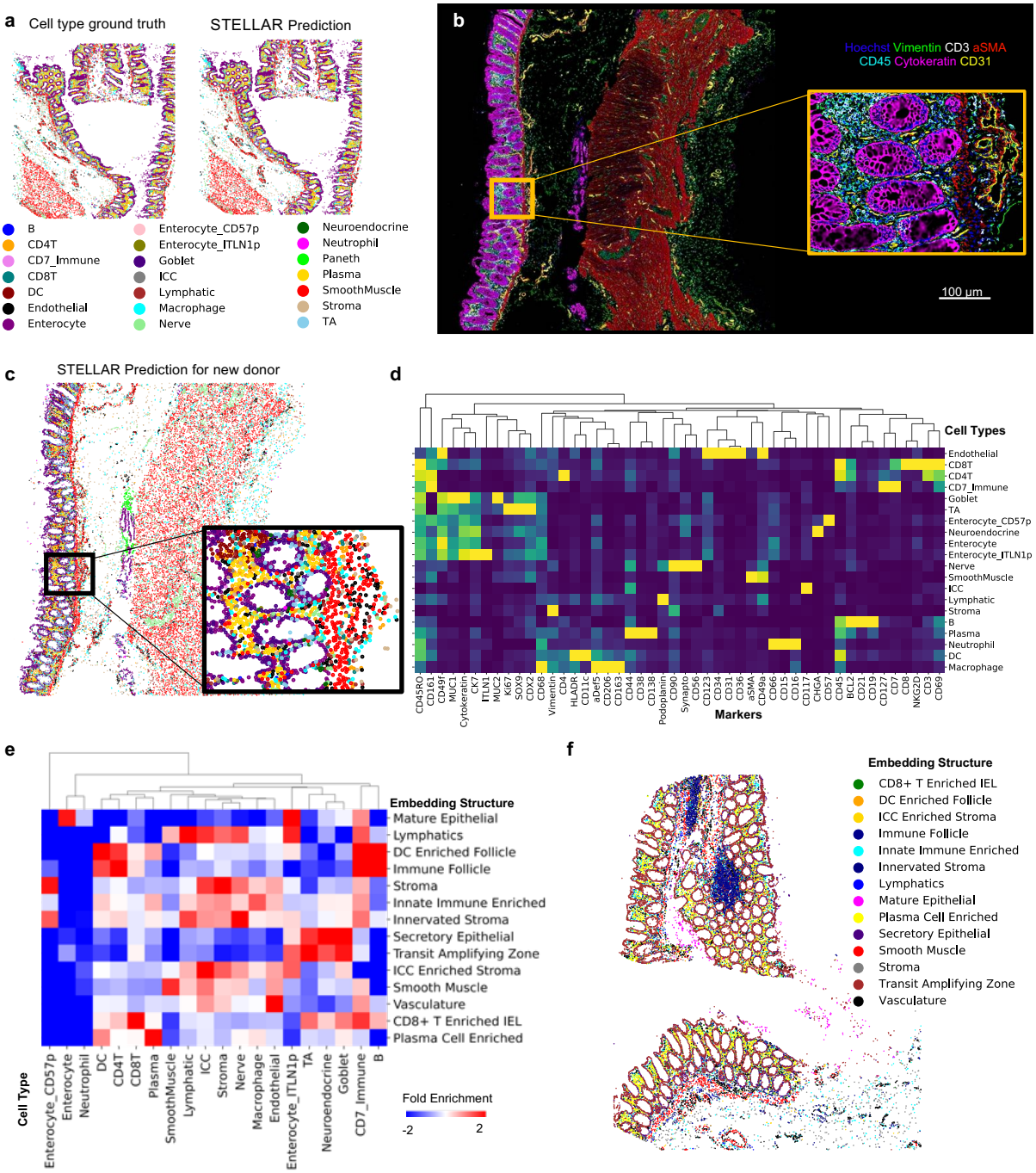
a
Cell type ground truth
STELLAR Prediction
B
CD4T
CD7_Immune
CD8T
DC
Endothelial
Enterocyte
Enterocyte_CD57p
Enterocyte_ITLN1p
Goblet
ICC
Lymphatic
Macrophage
Nerve
Neuroendocrine
Neutrophil
Paneth
Plasma
SmoothMuscle
Stroma
TA
b
Hoechst Vimentin CD3 aSMA
CD45 Cytokeratin CD31
100 µm
c
STELLAR Prediction for new donor
d
Cell Types
Endothelial
CD8T
CD4T
CD7_Immune
Goblet
TA
Enterocyte_CD57p
Neuroendocrine
Enterocyte
Enterocyte_ITLN1p
Nerve
SmoothMuscle
ICC
Lymphatic
Stroma
B
Plasma
Neutrophil
DC
Macrophage
CD45RO CD161 CD49f MUC1 Cytokeratin CK7 ITLN1 MUC2 Ki67 SOX9 CDX2 CD68 Vimentin CD4 HLA-DR CD15 aDef5 CD206 CD163 CD44 CD38 CD138 Podoplanin CD90 Synapto CD56 CD123 CD34 CD31 CD36 aSMA CD49a CD66 CD15 CD16 CD117 CHGA CD57 CD45 BCL2 CD21 CD19 CD127 CD7 CD8 NKG2D CD3 CD69
Markers
e
Embedding Structure
Mature Epithelial
Lymphatics
DC Enriched Follicle
Immune Follicle
Stroma
Innate Immune Enriched
Innervated Stroma
Secretory Epithelial
Transit Amplifying Zone
ICC Enriched Stroma
Smooth Muscle
Vasculature
CD8+ T Enriched IEL
Plasma Cell Enriched
Cell Type
Enterocyte_CD57p Enterocyte Neutrophil DC CD4T CD8T Plasma SmoothMuscle Lymphatic ICC Stroma Nerve Macrophage Endothelial Enterocyte_ITLN1p TA Neuroendocrine Goblet CD7_Immune B
Fold Enrichment
-2
2
f
Embedding Structure
CD8+ T Enriched IEL
DC Enriched Follicle
ICC Enriched Stroma
Immune Follicle
Innate Immune Enriched
Innervated Stroma
Lymphatics
Mature Epithelial
Plasma Cell Enriched
Secretory Epithelial
Smooth Muscle
Stroma
Transit Amplifying Zone
Vasculature

Figure 3 *(preceding page)*: STELLAR transfers granular cell-type labels across tissue regions and donors from HuBMAP data and identifies major structures of healthy human intestine tissue. (a) Cell-type maps with ground-truth and STELLAR predictions made after training on CODEX multiplexed imaging of three different intestine regions across the colon. **(b-f)** We used data from one donor as the reference set, and applied STELLAR to annotate new, completely unannotated donor datasets. **(b)** Fluorescent images of 7 of 54 markers used in the CODEX multiplexed imaging of healthy intestine from a different donor. Hoechst, blue; vimentin, green; CD3, white; aSMA, red; CD45, cyan; cytokeratin, magenta; and CD31, yellow. Scale bar of the zoomed out image is 100 micrometers. **(c)** STELLAR predictions of cell types after training on data from a different donor mapped to CODEX spatial coordinates. **(d)** Average marker expression for cell types predicted by STELLAR for intestinal samples. **(e)** Clusters in STELLAR's embedding space reveal that the embeddings capture major structures within the intestinal tissue. Heatmap shows enrichment of cell types assigned by STELLAR within the structural clusters as compared to tissue average percentages. **(f)** Graphical display of the embedding structures found for one of the regions of the colon. Major structures such as immune follicles were identified by STELLAR.

Data availability

The datasets presented in this study can be found in the online repository Dryad. Specifically, the quantified single-cell data are provided (with cells in rows and protein expression, x/y position, and cell type labels in columns). Additionally, the datasets used to transfer from the tonsil to BE tissue is there (BE_Tonsil_dryad.csv) and expert-annotated healthy human intestine (B004_training_dryad.csv), which was used to test the accuracy of STELLAR across the four regions of the colon regions of this dataset and also for training for transferring cell-type labels to unlabeled donors (B0056_unnannotated_dryad.csv) are provided.

Code availability

STELLAR was written in Python using the PyTorch library. The source code is available on Github at https://github.com/snap-stanford/stellar.

Acknowledgements

This work was supported by the U.S. National Institutes of Health (2U19AI057229-16, 5P01HL10879707, 5R01GM10983604, 5R33CA18365403, 5U01AI101984-07, 5UH2AR06767604, 5R01CA19665703, 5U54CA20997103, 5F99CA212231-02, 1F32CA233203-01, 5U01AI140498-02, 1U54HG010426-01, 5U19AI100627-07, 1R01HL120724-01A1, R33CA183692, R01HL128173-04, 5P01AI131374-02, 5UG3DK114937-02, 1U19AI135976-01, IDIQ17X149, 1U2CCA233238-01, 1U2CCA233195-01); Cancer Research UK (C27165/A29073); the Parker Institute for Cancer Immunotherapy. JWH was supported by an NIH T32 Fellowship (T32CA196585) and an American Cancer Society - Roaring Fork Valley Postdoctoral Fellowship (PF-20-032-01-CSM). We also gratefully acknowledge the support of DARPA under Nos. HR00112190039 (TAMI), N660011924033 (MCS); ARO

under Nos. W911NF-16-1-0342 (MURI), W911NF-16-1-0171 (DURIP); NSF under Nos. OAC-1835598 (CINES), OAC-1934578 (HDR), CCF-1918940 (Expeditions), IIS-2030477 (RAPID), NIH under No. R56LM013365; Stanford Data Science Initiative, Wu Tsai Neurosciences Institute, Chan Zuckerberg Biohub, Amazon, JPMorgan Chase, Docomo, Hitachi, Intel, JD.com, KDDI, NVIDIA, Dell, Toshiba, Visa, and UnitedHealth Group. JL is a Chan Zuckerberg Biohub investigator.

Author information

Contributions. MB, KC, JWH and JL conceived the research. MB, KC, JWH and YT performed research and analyzed results. KC, MB, and JL contributed new analytical tools and created the algorithmic framework. JWH, MPS, and GPN generated and analyzed the data. JL, GPN and MPS supervised the research. All authors participated in interpretation and wrote the manuscript.

References

1. Lewis, S. M. *et al.* Spatial omics and multiplexed imaging to explore cancer biology. *Nature Methods* 1–16 (2021).

2. Bodenmiller, B. Multiplexed epitope-based tissue imaging for discovery and healthcare applications. *Cell systems* **2**, 225–238 (2016).

3. Chen, K. H., Boettiger, A. N., Moffitt, J. R., Wang, S. & Zhuang, X. Spatially resolved, highly multiplexed RNA profiling in single cells. *Science* **348** (2015).

4. Hickey, J. W. *et al.* Spatial mapping of protein composition and tissue organization: a primer for multiplexed antibody-based imaging. *Nature Methods* (2021).

5. Consortium, H. *et al.* The human body at cellular resolution: the NIH Human Biomolecular Atlas Program. *Nature* **574**, 187 (2019).

6. Rozenblatt-Rosen, O. *et al.* The Human Tumor Atlas Network: charting tumor transitions across space and time at single-cell resolution. *Cell* **181**, 236–249 (2020).

7. Regev, A. *et al.* Science forum: The Human Cell Atlas. *eLife* **6**, e27041 (2017).

8. Zhang, M. *et al.* Spatially resolved cell atlas of the mouse primary motor cortex by MERFISH. *Nature* **598**, 137–143 (2021).

9. Black, S. *et al.* CODEX multiplexed tissue imaging with DNA-conjugated antibodies. *Nature Protocols* 1–36 (2021).

10. Goltsev, Y. *et al.* Deep profiling of mouse splenic architecture with CODEX multiplexed imaging. *Cell* **174**, 968–981 (2018).

11. Pliner, H. A., Shendure, J. & Trapnell, C. Supervised classification enables rapid annotation of cell atlases. *Nature Methods* **16**, 983–986 (2019).

12. Kimmel, J. C. & Kelley, D. R. Semi-supervised adversarial neural networks for single-cell classification. *Genome Research* gr–268581 (2021).

13. Tan, Y. & Cahan, P. SingleCellNet: A computational tool to classify single cell RNA-seq data across platforms and across species. *Cell Systems* **9**, 207–213 (2019).

14. Brbić, M. *et al.* MARS: Discovering novel cell types across heterogeneous single-cell experiments. *Nature Methods* **17**, 1200–1206 (2020).

15. Teng, H., Yuan, Y. & Bar-Joseph, Z. Clustering spatial transcriptomics data. *Bioinformatics* (2021).

16. Kennedy-Darling, J. *et al.* Highly multiplexed tissue imaging using repeated oligonucleotide exchange reaction. *European Journal of Immunology* **51**, 1262–1277 (2021).

17. Hickey, J. W., Tan, Y., Nolan, G. P. & Goltsev, Y. Strategies for accurate cell type identification in CODEX multiplexed imaging data. *Frontiers in Immunology* 3317 (2021).

18. Sautès-Fridman, C., Petitprez, F., Calderaro, J. & Fridman, W. H. Tertiary lymphoid structures in the era of cancer immunotherapy. *Nature Reviews Cancer* **19**, 307–325 (2019).

19. Blondel, V. D., Guillaume, J.-L., Lambiotte, R. & Lefebvre, E. Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment* **2008**, P10008 (2008).

20. Kipf, T. N. & Welling, M. Semi-supervised classification with graph convolutional networks. In *International Conference on Learning Representations* (2017).

21. Cao, K., Brbic, M. & Leskovec, J. Open-world semi-supervised learning. *arXiv preprint arXiv:2102.03526* (2021).

22. Liu, B. *et al.* Negative margin matters: Understanding margin in few-shot classification. In *European Conference on Computer Vision* (2020).

23. Chiang, W.-L. *et al.* Cluster-GCN: An efficient algorithm for training deep and large graph convolutional networks. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 257–266 (2019).

24. Chen, T. & Guestrin, C. XGBoost: A scalable tree boosting system. In *ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 785–794 (2016).

25. Cortes, C. & Vapnik, V. Support-vector networks. *Machine learning* **20**, 273–297 (1995).

26. Breiman, L. Random forests. *Machine learning* **45**, 5–32 (2001).

27. Freund, Y. & Schapire, R. E. A decision-theoretic generalization of on-line learning and an application to boosting. *Journal of Computer and System Sciences* **55**, 119–139 (1997).

28. Hao, Y. *et al.* Integrated analysis of multimodal single-cell data. *Cell* (2021).