Tissue reassembly with generative AI

Tingyang Yu[1], Chanakya Ekbote[1], Nikita Morozov[1], Jiashuo Fan[1], Pascal Frossard[2], Stéphane d'Ascoli[1,3] and Maria Brbić[1,4,5†]

[1] School of Computer and Communication Sciences, EPFL, Lausanne, Switzerland

[2] School of Engineering, EPFL, Lausanne, Switzerland

[3] Meta AI, Paris, France

[4] School of Life Sciences, EPFL, Lausanne, Switzerland

[5] Swiss Institute of Bioinformatics, Switzerland

[†]Corresponding author. Email: mbrbic@epfl.ch

The spatial arrangement of cells plays a critical role in determining their functions and interactions within tissues. However, single-cell RNA sequencing (scRNA-seq) dissociates cells from their native tissue context, resulting in a loss of spatial information. Here, we show that complex tissue structures can be reassembled from the gene expression profiles of dissociated cells. To achieve this, we developed LUNA, a generative AI model that reconstructs tissues conditioned solely on gene expressions of cells by learning spatial priors over existing spatially resolved datasets. We show that LUNA effectively reconstructs slices from the MERFISH whole mouse brain atlas with over 1.2 million cells, including cells from cell types never seen during model training. Applying LUNA to the mouse central nervous system scRNA-seq atlas, we show that LUNA is applicable for de novo generation of tissue structures. Additionally, LUNA can infer locations of nuclei lost during cell profiling with Slide-tags technology and correctly places cells belonging to spatially distinct compartments in a human metastatic melanoma sample. We envision that AI-driven tissue reassembly can help to overcome current technological limitations and advance our understanding of tissue organization and function, paving the way towards virtual tissue models.

Introduction

Single-cell RNA sequencing (scRNA-seq) technologies have enabled the high-throughput profiling of cellular transcriptomes, making it possible to identify and explore unique transcriptional profiles within individual cells [1, 2]. These technologies have revolutionized our understanding of cellular diversity providing unprecedented insights into the complexity of cells and tissues [3–8]. However, a key limitation of scRNA-seq technologies is the loss of spatial context, as cells are dissociated from their native microenvironments during the sequencing process. Spatial context is crucial for understanding how cells interact with their neighbors and the surrounding microenvironment, influencing their roles in healthy tissue function and disease development. Spatially resolved sequencing technologies aim to address this limitation [9–11] but they are limited in the number of genes that can be measured.

To address the limitations of individual technologies, computational methods have been developed to infer spatial context from gene expressions of single-cells. Early approaches relied on in situ hybridization (ISH)-based reference atlases of landmark gene expressions and leveraged them to infer a spatial location of cells in scRNA-seq datasets [12, 13]. More recently, reference mapping methods such as Tangram [14], CytoSPACE [15], and CeLEry [16], proposed mapping dissociated tissue data onto reference spatial datasets. While these methods have shown promising results, they rely on an exact spatial reference for mapping scRNA-seq datasets, thus primarily functioning as data integration tools. On the other hand, novoSpaRc [17, 18] enabled a *de novo* spatial reconstruction of single-cell gene expression. However, novoSpaRc assumes that cells in close physical proximity exhibit similar gene expression profiles — a condition that is often violated in complex tissue architectures.

Here, we present LUNA (Location reconstrUction using geNerative Ai), a generative AI model that reassembles complex tissue structures from gene expressions of cells by learning spatial priors over spatial transcriptomics datasets. LUNA learns cell representations that capture cellular interactions globally and locally across the entire tissue slice, enabled via an attention mechanism [19, 20] that takes into consideration interactions across all cells. LUNA operates as a diffusion model [21–23] – during training it learns to denoise corrupted cell coordinates, while during inference it starts from random noise and reconstructs physical locations of cells *de novo* solely from their gene expressions.

We apply LUNA to reassemble the MERFISH whole mouse brain atlas, which consists of 1.23 million cells spanning all brain regions [24], as well as the scRNA-seq mouse central nervous

system atlas of 1.08 million cells [25]. LUNA effectively reconstructs complex mouse brain tissue across a wide range of brain regions and captures spatial gene expression patterns. LUNA exhibits a strong zero-shot ability by generalizing to cells from previously unseen cell types, accurately predicting their spatial positions within the tissue structure without explicit training on those specific cell types. Furthermore, LUNA enables the enrichment of Slide-tags datasets [26] by inferring the tissue locations of nuclei lost during cell profiling. LUNA correctly placed different subtypes of neuronal cells in mouse embryonic tissue as well as tumor cells belonging to spatially distinct compartments in a human metastatic melanoma sample, enriching the sparse Slide-tags data by up to 100%. LUNA is highly scalable, with linear time and memory complexity relative to the number of input cells and it can infer the spatial coordinates of tens of thousands of cells in less than a few minutes on a single GPU.

Results

Overview of LUNA. Given gene expressions of dissociated cells, we aim to address the challenge of reassembling cells into compact tissue structures (Fig. 1a). To achieve that, we designed a diffusion model [21–23], named LUNA, that generates complex tissue structures by predicting the physical locations of individual cells based on their gene expressions. The key challenge in generating tissue structures *de novo* arises from the inherent complexity of tissues, which are composed of a vast array of heterogeneous cells engaged in intricate cell-cell communication. LUNA overcomes these challenges by enabling each cell to "communicate" with all other cells guided by the learned weights, *i.e.*, attention scores (Methods). In this way, LUNA learns cell embeddings that capture cellular interactions on both the global and local levels, allowing LUNA to directly map cell embeddings to spatial locations. During training, LUNA learns spatial priors on existing spatial transcriptomics data. In the inference stage, LUNA generates complex tissue structures solely from gene expressions of cells.

The neural network in LUNA consists of a multi-head self-attention mechanism [19,20] that allows each cell to focus on gene expressions of specific cells when making predictions (Fig. 1b). The cell embeddings are then mapped from the latent space to the physical space via a fully connected layer. To ensure robustness to rotations and translations of input slices, the loss function in LUNA is specially designed to be invariant to these transformations and preserve the relative spatial relationships of the cells (Methods). The training stage of LUNA comprises two processes:

corruption and denoising, both of which are repeated over a range of diffusion times (Methods). During the corruption process, a controlled amount of Gaussian noise is added to the ground truth cell coordinates. Using the noised cell coordinates generated from this process and conditioned on the gene expression data, LUNA then learns to recover the original cell coordinates through the denoising process. During the inference stage, LUNA only takes gene expression profiles as model input and reconstructs tissue structures by denoising cell coordinates in an autoregressive manner. Initially, cell coordinates are generated from random Gaussian noise, and LUNA iteratively removes noise from these coordinates, predicting spatial locations of cells *de novo* (Fig. 1c, Supplementary Fig. 1). LUNA is highly scalable, with linear time and memory complexity relative to the number of cells. It infers physical locations of tens of thousands of cells in just a few minutes on a single GPU. For example, LUNA inferred locations of $40,000$ cells from the Allen Brain Cell (ABC) MERFISH mouse brain atlas [24] in approximately four minutes using around 6 GB of GPU memory (Fig. 1d).

LUNA accurately reconstructs the whole mouse brain atlas. LUNA effectively scales to atlas-scale large datasets and accurately reconstructed a whole mouse brain of the Allen Brain Cell (ABC) MERFISH mouse brain atlas [24]. We trained LUNA on 2.85 million cells across 147 slices from one mouse (Supplementary Note 1). We then applied it to reassemble cells from the whole brain of another mouse, never seen during model training, consisting of 1.23 million cells and 338 identified subclasses in 66 slices (Fig. 2a).

We evaluated LUNA's predictions using the ground-truth cell locations. We found that LUNA's predictions agree well with the ground-truth cell locations across 11 major regions identified in the ABC atlas, despite their distinct structural characteristics (Supplementary Figs. 2-4). For example, LUNA accurately captured the circular structure of the olfactory bulb region, the layered organization of the isocortex region, and the anatomical separation between the brain stem and cerebrum (Fig. 2b). In the slice from the olfactory bulb region affected by sequencing artifacts, LUNA inferred the spatial distribution of cells in areas lacking ground truth spatial information. To further evaluate the generation results, we compared LUNA's predictions across various levels of categorization, ranging from neurotransmitter level (4 types) to cell supertype level ($1,201$ types) and cell cluster level ($5,322$ types). LUNA accurately positioned cells at all levels of categorization, from coarse to the finest level (Fig. 2c and Supplementary Fig. 5).

We next assessed LUNA's ability to infer the spatial distribution of gene expressions and

correctly position cells within a cell type based on their gene expression profiles. We identified spatially variable genes by computing Moran's I statistic [27] and visualized the expression values of genes with the highest Moran's I values (Supplementary Note 3). LUNA successfully captured the spatial expression patterns of genes within individual cell types, demonstrating its ability to preserve biological information. For example, LUNA accurately positioned cells expressing *Rxfp1* and *Rorb* in extratelencephalic and intratelencephalic neurons containing the neurotransmitter Glut (Fig. 2d). We also observed high reconstruction quality across non-neuronal cells, including the correct physical locations of a small group of 85 cells out of $3,725$ astrocyte-ependymal cells expressing the *Cfap206* gene (Supplementary Fig. 6).

To quantitatively evaluate the conservation of spatial gene expression patterns, we computed spatial autocorrelation using the Moran's I statistic for all $1,122$ genes across the 66 slices. We compared Moran's I values derived from LUNA's predicted locations with those from the ground-truth cell locations. Our analysis revealed a strong correlation between the Moran's I values based on LUNA's predictions and those based on the true locations, achieving an average Pearson correlation coefficient of 0.95 across all genes (Supplementary Fig. 7). For instance, the spatial autocorrelation for the spatially variable *Rorb* gene was consistently well-preserved across all slices, with a near perfect Pearson's correlation coefficient (Fig. 2e). These results demonstrate that LUNA effectively preserves spatial gene expression patterns throughout the mouse brain.

LUNA generalizes to unseen cell classes. We next evaluated zero-shot capability of the model to predict the spatial locations of cells from classes unseen during model training without any additional training or fine-tuning of the model. We trained LUNA on all slices from the Animal 1 in the ABC atlas, excluding NP-CT-L6b Glut cells ($n = 69,641$). The NP-CT-L6b Glut class was randomly chosen from 34 cell classes identified in [24]. We then applied LUNA to predict tissue structures for all cell classes in Animal 2, including unseen NP-CT-L6b Glut cells ($n = 29,262$) that have never been seen during model training (Fig. 2f). The predictions closely matched the ground truth locations in all cell classes (Fig. 2g). Across different slices, LUNA correctly positioned NP-CT-L6b Glut cells within the spatial architecture of the tissue (Fig. 2h,i). We additionally examined the spatial expression patterns of spatially variable genes in NP-CT-L6b Glut cells. We identified genes *Tshz2* and *Syt6* as genes with the highest Moran's I values among $1,122$ genes based on the ground-truth cell locations. We find that within the unseen NP-CT-L6b Glut class, LUNA successfully positioned cells according to the *Tshz2* and *Syt6* expression pat-

terns (Fig. 2h,i). Similarly, we observed comparable performance for other randomly excluded cell classes, such as the HY GABA cell class (Supplementary Fig. 8) and the Vascular cell class (Supplementary Fig. 9). These results demonstrate that LUNA preserves key spatial gene expression patterns of novel cell types, showcasing strong zero-shot capabilities.

LUNA outperforms other methods by a large margin. LUNA is a unique method in its capability to reconstruct complex tissue structures. Existing methods are reference mapping methods and require the pre-selection of a single reference slice to which then new cells are mapped based on the similarity of the gene expressions between the cells in the reference slice and the input gene expression profiles (Fig. 3a). This mapping is performed either by explicitly learning a mapping matrix, as in novoSpaRc [17], Tangram [14], and CytoSPACE [15], or by implicitly regressing from gene expression space to the physical space, as in CeLEry [16] (Supplementary Note 2).

We compared the performance of LUNA to alternative methods on the spatially resolved cell atlas of the mouse primary motor cortex sequenced by MERFISH [28] (Supplementary Note 1). We used 31 slices ($118,036$ cells) from one animal as the test set and trained LUNA on 33 slices ($158,379$ cells) from another mouse to learn spatial priors across multiple slices. During inference, all methods were conditioned only on the gene expressions. Since existing methods can not accept multiple slices as input, we randomly selected a single slice from the training mouse to serve as the reference, repeating the procedure for each slice of the testing mouse. To ensure a fair comparison, all methods used the same seed for random slice selection.

We evaluated the results using the Spearman's rank correlation between the predicted and ground truth locations, averaging the performance across all slices from the unseen mouse (Supplementary Note 3). LUNA achieved an average correlation coefficient of 44.8%, outperforming the best alternative baseline CeLEry by 100% (Fig. 3b). Performance gains of LUNA were retained using other evaluation metrics, such as precision and RSSD (Supplementary Note 3, Supplementary Fig. 10). CeLEry is the only other method besides LUNA capable of potentially leveraging information from multiple slices. We reran CeLEry using all slices from the training mouse as the training dataset. However, because CeLEry is a regression model that does not account for geometric constraints, it is not robust to variations in slice orientation or cutting shapes within the training set. Consequently, we observed a significant performance decline when when CeLEry was trained with multiple slices compared to using a single slice (Supplementary Fig. 10).

We visualized the spatial locations predicted by different methods on a sample slice from the

unseen test mouse (Fig. 3c). Compared to other methods, LUNA not only accurately predicted the layer-wise structure of mouse primary motor cortex tissues, but is also the only approach capable of correctly inferring the spatial priors of the tissue, such as the contour of the tissue. The limitations of other methods stem from their reliance on pre-selected reference slices, which constrains their generalization capabilities. To systematically evaluate the sensitivity of existing methods on the reference slice, we selected different slices as a reference training slice and evaluated performance on a fixed test slice. We computed the average performance of these methods across different reference slices. LUNA achieved a 125.2% higher Spearman's rank correlation coefficient compared to the average performance of the best alternative method, CeLEry, where the average is calculated over different reference slices (Supplementary Fig. 11).

We further compared the spatial gene expression patterns of cell type specific marker genes. LUNA distinguished cells from different cell classes, accurately reflecting the spatial gene expression patterns (Fig. 3d-f). In contrast, other methods struggled due to the dissimilarity between the reference and test slices (Supplementary Fig. 12). This consistently demonstrates that LUNA effectively preserves the spatial gene expression flow.

De novo reconstruction of scRNA-seq data. We next applied LUNA for *de novo* generation of tissue structures of 1.08 million dissociated single cells across 13 coronal slices from the mouse central nervous system (CNS) scRNA-seq atlas [25]. We used the model trained on the ABC MERFISH mouse brain atlas to predict spatial locations for cells in the scRNA-seq mouse CNS atlas. Since the scRNA-seq atlas lacks spatial information, we validated LUNA's performance by using estimated cell locations derived from the integration of the scRNA-seq atlas with the STARmap PLUS CNS spatial atlas [29] (Supplementary Note 1).

LUNA 's predictions aligned closely with cell locations estimated through the integration data with the STARmap PLUS atlas at both the coarse cell class level (23 types; Fig. 4a; Supplementary Fig. 13) and the finer submolecular class level, which includes 216 distinct types (Fig. 4a; Supplementary Fig. 14). The agreement is particularly pronounced in the isocortex, hippocampal formation, thalamus, and hypothalamus regions. We further evaluated LUNA's performance in reconstructing mouse CNS tissue architecture across various cell classes (Supplementary Fig. 15). Additionally, we examined LUNA's predictions for specific neuronal subtypes including di- and mesencephalon excitatory neurons (Fig. 4b), di- and mesencephalon inhibitory neurons (Fig. 4c), and telencephalon projecting excitatory neurons (Fig. 4d), along with their respective

sub-molecular classes. The results indicate that LUNA not only accurately placed cells across major cell classes but also precisely predicted the spatial relationships of complex sub-molecular classes within these neuronal cells. This underscores LUNA's capability to capture both broad and intricate cellular architectures within the CNS.

To evaluate the conservation of spatial gene expression patterns, we analyzed the spatial distribution of three highly variable genes—*LPL*, *IQGAP2*, and *MEF2C*, within telencephalon-projecting excitatory neurons (Fig. 4e). These genes were selected based on their high Moran's I values, calculated from a dataset of $10,844$ genes using LUNA's predictions. Our analysis reveals that the spatial gene expression patterns predicted by LUNA closely align with the estimated ground truth, demonstrating LUNA's effectiveness in capturing gene expression gradients at the cellular level. Specifically, *IQGAP2* gene shows a gradual downregulation across the Ammon's horn region (CA2) of the hippocampus, while *LPL* displays an opposite trend with gradual upregulation in this region. For *MEF2C* gene, expression is minimal in telencephalon-projecting excitatory neurons within the CA2 region but substiantially upregulated in the isocortex and olfactory areas. Overall, these results confirm that LUNA effectively reconstructs the complex spatial architecture of the mouse CNS scRNA-seq atlas and captures gene expression spatial changes. This underscores LUNA's applicability in reconstructing cell locations for scRNA-seq atlases.

LUNA infers locations of spatially unmapped nuclei in the Slide-tags data. LUNA can be used to predict the tissue locations of nuclei lost during cell profiling with Slide-tags technology [26]. Slide-tags enables profiling single-cell and spatially resolved transcriptome by tagging nuclei with spatial barcode oligonucleotides derived from DNA-barcoded beads with known positions and then using tagged nuclei as an input to single-nucleus profiling assays [26]. However, many nuclei are lost during the barcoding process due to a combination of dissociation and microfluidic losses. To compensate for the sparsity in Slide-tags data, we applied LUNA to assign spatial locations to the spatially unmapped nuclei in the Slide-tags data.

We first applied LUNA to the sagittal section of the embryonic mouse brain at embryonic day 14 (E14) dataset consisting of $4,614$ spatially mapped with Slide-tags (Fig. 5a) (Supplementary Note 1). To evaluate LUNA's performance using ground truth nuclei locations, we used $4,152$ of spatially mapped nuclei as the train set and evaluated LUNA's performance on remaining 462 cells. LUNA's predictions agree well with the ground truth locations of cells obtained with Slide-tags (Fig. 5b). Without using any cell class information, LUNA accurately predicted cell locations,

even for different neuronal subclasses. We further validated these predictions by computing spatial autocorrelation using Moran's I index across all genes used for training and observe high agreement between LUNA's predictions and true nuclei locations (Fig. 5c).

Encouraged by these results, we applied LUNA to infer locations of spatially unmapped nuclei due to nuclei losses with Slide-tags. We trained LUNA on spatially mapped cells of mouse E14 tissue and applied the model to infer locations of $4,414$ spatially unmapped nuclei. To validate predictions of LUNA on the cell class level, we ran linear support vector classifier to infer cell classes of spatially unmapped nuclei since annotations for spatially unmapped nuclei were missing in the original dataset (Supplementary Note 1). We found that LUNA placed spatially unmapped nuclei in proximity of cells with the same class, successfully enriching sparse Slide-tags dataset. We further examined the spatial distribution of specific cell types, additionally confirming that LUNA placed nuclei from both neuronal and non-neuronal cells to the correct spatial regions placed (Fig. 5e, Supplementary Fig. 16).

We next applied LUNA to a human metastatic melanoma sample obtained with Slide-tags technology [26]. This sample includes a diverse array of cells, including immune cells and tumor cells. We trained LUNA on $4,804$ spatially mapped nuclei and applied it to generate locations for the spatially unmapped nuclei. We again transferred cell type annotations of spatially unmapped nuclei with the linear support vector classifier to validate LUNA's predictions (Supplementary Note 1). We found that LUNA successfully enriched the dataset to the total of $6,466$ spatially mapped cells (Fig. 5f). LUNA correctly placed two tumor subpopulations into spatially distinct compartments according to predicted cell class annotations (Fig. 5g), as well as non-tumor cells (Supplementary Fig. 17). To further validate LUNA's predictions, we investigated the the heterogeneity of these two tumor populations by identifying highly differential genes between the two cell classes. Cells expressing *CLU* gene are placed by LUNA in different spatial compartments than cells *DISC2FP1* gene, which is also the case in the original dataset of spatially mapped nuclei of the melanoma sample (Fig. 5i). We applied LUNA to other Slide-tags datasets, including samples from human cortex (Supplementary Fig. 18) and human tonsil (Supplementary Fig. 19). These results confirm LUNA's ability to correctly place spatially unmapped nuclei in Slide-tags datasets.

Discussion

LUNA is a generative AI model that enables *de novo* generation of tissue architectures from their

gene expressions. By learning spatial priors from existing spatial transcriptomics datasets, LUNA captures the underlying patterns of tissue organization, enabling it to reconstruct complex tissue architectures. The model learns cell representations that allow for accurate positioning of cells within the tissue context, even in the absence of direct spatial information. Different from previous methods, LUNA's generative nature allows it to operate in a continuous physical space, unrestricted by a finite number of physical spots, thus enabling the generation of tissue structures with arbitrary spatial density.

LUNA has broad implications in the field of spatial biology and addresses existing technological limitations. LUNA can infer 2D spatial locations for dissociated single cells from scRNA-seq datasets, enabling reconstruction of complex tissue architectures *de novo*. In addition, LUNA can generate locations for cells or spots missing spatial information in the existing spatial technologies such as Slide-tags [26]. Furthermore, LUNA offers an effective solution for spatial transcriptomics by reducing the need for comprehensive imaging of all tissue slices. A smaller number of slices can be imaged and LUNA can be used to predict spatial information for the remaining slices, thus lowering experimental costs while maintaining high-resolution spatial insights. Finally, LUNA is a complementary tool for analyzing tissue spatial structures, gene expression gradients and cell-cell communication along with existing tools for spatial analysis such as Squidpy [30], Seurat V5 [31], STELLAR [32] and STalign [33].

Currently, LUNA operates within a two-dimensional (2D) spatial framework, which does not fully capture the inherently three-dimensional (3D) nature of tissue architectures. This limitation stems primarily from the constraints of existing sequencing technologies, which are not yet capable of measuring tissue structures in full 3D detail. However, with ongoing advancements in technology [34], datasets that provide rich molecular organization of tissues in 3D are likely to become more widely available in the near future. Our framework can be flexibly extended beyond 2D spatial reconstruction.

Our current analysis has mainly focused on mouse brain cell atlases due to the abundance of available data. However, as our results on the Slide-tags human tissues show, LUNA is a versatile framework that can be readily applied to other organs and species. Future research could extend the use of LUNA to various tissues and model organisms, enabling to uncover organizational principles underlying tissue architectures across various biological systems.

LUNA's scalability, accuracy, and versatility position it as a powerful resource for studying tissue spatial organization across a wide range of biological contexts, paving the way towards the

construction of comprehensive, spatially resolved single-cell transcriptome atlases and generation of virtual tissue models. We envision that AI-driven tissue reassembly holds great potential for advancing our understanding of complex biological processes and can drive future discoveries in cellular and tissue architecture, ultimately enhancing our ability to map and interpret the spatial structure across multiple biological systems in healthy and disease states.

Methods

Overview of LUNA. LUNA is a diffusion-based generative model designed to generate spatial locations of cells from their gene expressions. Using attention-based mechanism, LUNA learns how each individual cell should attend to other cells based on their relevance to a cell's own location and molecular features. During training, LUNA leverages spatial transcriptomics data with ground truth cell locations to learn to generate cell locations conditioned on the gene expressions. During inference, the input to LUNA are the gene expressions of dissociated cells and the model then generates cell locations from pure noise conditioned on the gene expressions, enabling *de novo* inference of cell locations.

In the training phase, the input to the LUNA is spatial transcriptomics data from single or multiple slices. LUNA learns cell embeddings that capture both the local and global tissue structure using a multi-head self-attention mechanism [19, 20] that enables cells to learn to attend to other cells. We designed the loss in LUNA as an SE(2)-invariant function [35–38], *i.e.*, it is robust to arbitrary rotations, translations and reflections of the predicted locations, only enforcing the preservation of relative distances between cells within each slice.

Training LUNA involves two main steps: *(i)* a corruption process that iteratively adds noise to the ground truth cell locations, and *(ii)* a denoising process that generates the ground truth locations from the noise-corrupted locations. During the denoising process, LUNA learns to reverse the corruption process, gradually transitioning from noisier to less noisy cell locations, ultimately reconstructing true cell locations from pure noise.

Learning cell representations. LUNA learns cell representations that capture the relationship between the molecular features of individual cells and their spatial context within tissue. These embeddings form the foundation for predicting the locations of cells.

The model is built on the premise that a cell's location is influenced not only by its own molecular features but also by those of other cells in the tissue. To effectively model these interactions, LUNA employs a multi-head self-attention mechanism [19, 20], which dynamically integrates information from all cells in the slice. This mechanism allows the model to assign attention scores to different cells based on their relevance to the prediction of each cell's location, capturing both local and global cellular interactions. Once the cell embeddings are learned, LUNA projects each cell's embedding into a 2-dimensional space to predict its locations.

Formally, LUNA takes as input *(i)* molecular features of cells $\mathbf{X}_s \in \mathbb{R}^{m \times d}$ in a given slice

s where m denotes the number of cells and d denotes the number of genes, *(ii)* time step $t \in \mathbb{N}$ uniformly sampled from 0 to the maximum diffusion time *T*, and *(iii)* corrupted cell locations $\tilde{\mathbf{R}}_{s,t} \in \mathbb{R}^{m\times 2}$ at time step t. The output of the model are the denoised cell locations $\hat{\mathbf{R}}_s \in \mathbb{R}^{m\times 2}$ for a given slice s. We omit slice index s for the ease of notation.

LUNA is composed of multiple transformer layers, each integrating fully connected layers followed by the self-attention block. At each layer l given a fixed diffusion time t, LUNA learns *(i)* cell embeddings $\mathbf{F}^{(l)} \in \mathbb{R}^{m\times d_x^{(l)}}$, *(ii)* a diffusion time embedding $\gamma_t^{(l)} \in \mathbb{R}$, and *(iii)* cell location embeddings $\mathbf{R}^{(l)} \in \mathbb{R}^{m\times d_r^{(l)}}$. The cell embeddings $\mathbf{F}^{(0)}$ in layer 0 are initialized via fully connected neural network layer that maps the molecular features of cells into the latent embedding space with dimensionality $d_x^{(0)}$:

$$\mathbf{F}^{(0)} = \mathit{FCN}(\mathbf{X}_s), \tag{1}$$

where *FCN* denotes a fully connected neural network layer. The diffusion time embedding $\gamma_t^{(0)}$ is initialized to the value of the diffusion time itself:

$$\gamma_t^{(0)} = t. \tag{2}$$

Finally, the cell location embeddings are initialized as the corrupted cell locations obtained from the corruption process at time step t:

$$\mathbf{R}^{(0)} = \tilde{\mathbf{R}}_t. \tag{3}$$

Then, at each layer $l \in \{1, \ldots, L\}$, LUNA first transforms cell embeddings, a diffusion time embedding and cell location embeddings using fully connected layers:

$$\mathbf{F}_f^{(l)} = \mathit{FCN}\left(\mathbf{F}^{(l-1)}\right), \tag{4}$$

$$\gamma_t^{(l)} = \mathit{FCN}\left(\gamma_t^{(l-1)}\right), \tag{5}$$

$$\mathbf{R}_r^{(l)} = \mathit{FCN}\left(\mathbf{R}^{(l-1)}\right). \tag{6}$$

The transformed cell embeddings, diffusion time embedding and cell location embeddings are concatenated to form a unified representation for each cell:

$$\mathbf{H}^{(l)} = \left[\mathbf{F}_f^{(l)}, \gamma_t^{(l)}\mathbf{1}_{m\times 1}, \mathbf{R}_r^{(l)}\right], \tag{7}$$

where $\mathbf{1}_{m\times 1}$ is a column vector of ones used to replicate the diffusion time embedding $\gamma_t^{(l)}$ across all m cells. To obtain final cell embeddings $\mathbf{F}^{(l)}$ in layer l, the unified representation $\mathbf{H}^{(l)}$ is passed through a self-attention block which computes attention weights between all pairs of cells [20]:

$$\mathbf{F}^{(l)} = \text{SelfAttention}^{(l)}\left(\mathbf{H}^{(l)}\right). \tag{8}$$

This self-attention operation enables the model to aggregate information from all cells in a slice, capturing how each cell's location is influenced by others. To ensure scalability, we implemented the efficient attention [20] – an approximation of attention computation with a linear complexity. This approach reduces memory usage and computational complexity by circumventing the need for explicit pairwise interactions between every query-key pair (Supplementary Note 5).

Finally, the cell embeddings $\mathbf{F}^{(l)}$ are used to predict the locations of each cell. The cell locations are generated from the embeddings using a fully connected layer that projects the location of each cell from the latent embedding space to the two-dimensional physical space:

$$\mathbf{R}^{(l)} = FCN\left(\mathbf{F}^{(l)}\right). \tag{9}$$

At each layer, the updated locations are centered by subtracting the mean, ensuring that the cell locations remain stable and consistent relative to the tissue structure. At the final layer, the output cell locations are the predicted clean cell locations $\hat{\mathbf{R}} = \mathbf{R}^{(L)}$.

Loss objective. To optimize LUNA, we introduce a specific loss objective that is SE(2)-invariant. This geometric invariance is crucial because the spatial arrangement of cells in the tissue may undergo transformations due to experimental artifacts, while the underlying gene expression profiles remain unchanged.

To ensure SE(2)-invariance, LUNA introduces a pairwise loss function which focuses on preserving the relative distances between all cells in the slice defined as follows:

$$\mathcal{L}_{\text{diff}} = \frac{1}{m^2}\|\Delta(\hat{\mathbf{R}}) - \Delta(\mathbf{R}_{gt})\|_F^2, \tag{10}$$

where $\Delta(\cdot) : \mathbb{R}^{m\times 2} \to \mathbb{R}^{m\times m}$ computes the pairwise squared Euclidean distances between the locations of all cells and $\mathbf{R}_{gt} \in \mathbb{R}^{m\times 2}$ represents the ground true cell locations.

The function $\Delta(\hat{\mathbf{R}})_{ij} = \|\hat{r}_i - \hat{r}_j\|_2^2$ computes the squared distance between the predicted locations of cells i and j, and similarly, $\Delta(\mathbf{R}_{gt})_{ij} = \|r_i - r_j\|_2^2$ computes the squared distance between the true locations. The pairwise loss then becomes:

$$\mathcal{L}_{\text{diff}} = \frac{1}{m^2}\sum_{i=1}^{m}\sum_{j=1}^{m}(\|\hat{r}_i - \hat{r}_j\|_2^2 - \|r_i - r_j\|_2^2)^2, \tag{11}$$

which penalizes deviations between the predicted and true pairwise distances, ensuring that the spatial relationships between cells are preserved.

The final loss is computed over all slices $s \in \{1, ...S\}$ where S denotes the total number of slices:

$$\mathcal{L} = \frac{1}{S} \sum_{s=1}^{S} \mathcal{L}_{\text{diff}}^{(s)}. \tag{12}$$

Corruption process. The diffusion model training begins with a corruption process, where LUNA progressively introduces noise into the true cell locations in an autoregressive manner. This step gradually transitions the true cell locations into pure noise, with the resulting corrupted locations serving as inputs for the subsequent denoising process.

We denote the initial diffusion time step, where no noise has been introduced to the cell locations, as time step 0, where ground truth spatial information is fully retained. Conversely, the maximum diffusion time step, denoted as T, represents the state where no ground-truth information is preserved, and the locations are sampled from pure Gaussian noise. The time step t is randomly selected from $\{0, 1, 2, \ldots, T\}$. The corrupted cell locations at time step t for slice s are denoted as $\tilde{\mathbf{R}}_t \in \mathbb{R}^{m \times 2}$ where again we omit slice index s for the ease of notation. From time step $t-1$ to t, Gaussian noise is added to the corrupted cell locations $\tilde{\mathbf{R}}_{t-1}$, yielding noisier locations $\tilde{\mathbf{R}}_t$ according to a predefined noise scheduler [21,22,39]. At time step T, the true cell locations $\mathbf{R}_{gt} = \tilde{\mathbf{R}}_0$ are transformed into corrupted cell locations $\tilde{\mathbf{R}}_T$ that have standard Gaussian distribution. The conditional distribution is defined as:

$$q_\phi(\tilde{\mathbf{R}}_t | \tilde{\mathbf{R}}_{t-1}) = \mathcal{N}(\tilde{\mathbf{R}}_t | \sqrt{\alpha_t} \tilde{\mathbf{R}}_{t-1}, (1 - \alpha_t)\mathbf{I}), \tag{13}$$

where $q_\phi(\cdot|\cdot)$ represents the corruption process with a noise scheduler ϕ that conditionally transitions the cell location distribution from less corrupted to more corrupted states. $\mathcal{N}(\cdot|\cdot, \cdot)$ denotes a Gaussian distribution, while $\mathbf{I}$ is the identity matrix. The noise scheduler ϕ, regulated by hyperparameters $\{\alpha_t\}$, controls the balance between retaining the original cell locations and adding noise.

To mitigate computational complexity, we compute the conditional distribution of corrupted cell locations given the true cell locations $\mathbf{R}_{gt}$, rather than from the previous diffusion step [22,40]:

$$q_\phi(\tilde{\mathbf{R}}_t | \tilde{\mathbf{R}}_0) = \mathcal{N}(\tilde{\mathbf{R}}_t | \sqrt{\overline{\alpha}_t} \tilde{\mathbf{R}}_0, (1 - \overline{\alpha}_t)\mathbf{I}). \tag{14}$$

Sampling from this conditional distribution is equivalent to adding Gaussian noise to the true cell locations. As shown in [40], the corrupted cell locations are computed as:

$$\tilde{\mathbf{R}}_t = \sqrt{\overline{\alpha}_t}\tilde{\mathbf{R}}_0 + \sqrt{(1-\overline{\alpha}_t)}\epsilon, \tag{15}$$

where $\overline{\alpha}_t = \prod_{i=1}^{t} \alpha_i$, and ϵ is drawn from a Gaussian distribution with mean 0 and variance $\mathbf{I}$, *i.e.*, $\epsilon \sim \mathcal{N}(0, \mathbf{I})$. As t progresses from 0 to T, the cell locations become increasingly corrupted until they are indistinguishable from white noise. At $t = T$, the corrupted locations $\tilde{\mathbf{R}}_T$ are sampled purely from Gaussian noise, with $\overline{\alpha}_T = 0$. To ensure robustness to translations, we additionally subtract the center of mass from the noise ϵ and the data slices $\tilde{\mathbf{R}}_0$, thus working with distributions defined on the subspace where the center of mass is fixed to 0 [41, 42].

Denoising process. Given the corrupted locations, we train LUNA model to recover the true cell locations by reversing the corruption process using the gene expression profiles. We denote the model as $\mu_\theta(\cdot, \cdot, \cdot)$ where θ represents all model parameters. LUNA progressively denoises the corrupted cell locations $\tilde{\mathbf{R}}_t$ producing less corrupted locations $\tilde{\mathbf{R}}_{t-1}$ given time step t, more corrupted locations $\tilde{\mathbf{R}}_t$ at time step t, and the molecular features $\mathbf{X}$.

Eventually, LUNA learns to generate the clean cell locations from pure white noise, transitioning from $t = T$ to $t = 0$. To achieve this, the model first predicts the clean cell locations $\hat{\mathbf{R}}$ using the learned network $\mu_\theta(\cdot, \cdot, \cdot)$. It then interpolates between the corrupted locations $\tilde{\mathbf{R}}_t$ and the predicted clean locations $\hat{\mathbf{R}}$, producing less corrupted locations for the next diffusion step:

$$\tilde{\mathbf{R}}_{t-1} = \frac{(1-\overline{\alpha}_{t-1})\sqrt{\alpha_t}}{1-\overline{\alpha}_t}\tilde{\mathbf{R}}_t + \frac{(1-\alpha_t)\sqrt{\overline{\alpha}_{t-1}}}{1-\overline{\alpha}_t}\hat{\mathbf{R}}. \tag{16}$$

Inference phase. Once the model $\mu_\theta(\cdot, \cdot, \cdot)$ is trained, LUNA generates cell locations based on the molecular features of cells in the test set $\mathbf{X}_{ts} \in \mathbb{R}^{m_{ts} \times d}$. The inference process starts by sampling locations from a normal distribution $\epsilon \sim \mathcal{N}(0, \mathbf{I})$, where $\epsilon \in \mathbb{R}^{m_{ts} \times 2}$. The noisy locations, $\mathbf{R}_{ts,T} = \epsilon$ are then refined through a sequence of diffusion steps.

At each step t, the model takes the noisy locations $\mathbf{R}_{ts,t}$, the molecular features $\mathbf{X}_{ts}$, and the time step t as inputs to predict the denoised locations for the next time step $t-1$:

$$\mathbf{R}_{ts,t-1} = \frac{(1-\overline{\alpha}_{t-1})\sqrt{\alpha_t}}{1-\overline{\alpha}_t}\mathbf{R}_{ts,t} + \frac{(1-\alpha_t)\sqrt{\overline{\alpha}_{t-1}}}{1-\overline{\alpha}_t}\mu_\theta(\mathbf{R}_{ts,t}, \mathbf{X}_{ts}, t). \tag{17}$$

This process is repeated for each time step, iterating from $t = \{T, ..., 0\}$ until the model generates the final clean cell locations at $t = 0$.

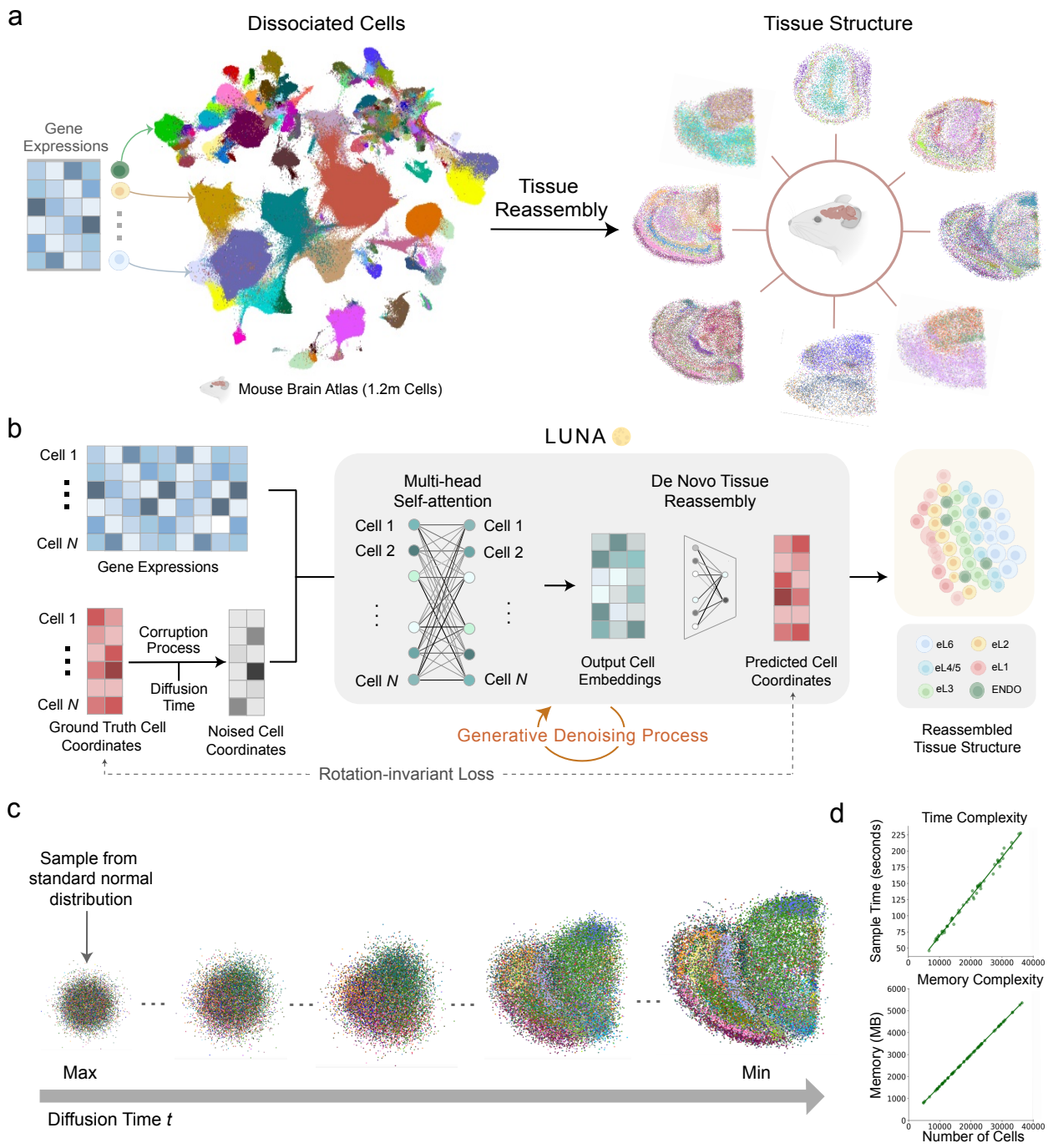
a
Dissociated Cells
Tissue Structure
Gene Expressions
Tissue Reassembly
Mouse Brain Atlas (1.2m Cells)
b
LUNA
Cell 1
Cell N
Gene Expressions
Cell 1
Cell N
Corruption Process
Diffusion Time
Ground Truth Cell Coordinates
Noised Cell Coordinates
Multi-head Self-attention
Cell 1
Cell 2
Cell N
Cell 1
Cell 2
Cell N
De Novo Tissue Reassembly
Output Cell Embeddings
Predicted Cell Coordinates
eL6
eL2
eL4/5
eL1
eL3
ENDO
Reassembled Tissue Structure
Generative Denoising Process
Rotation-invariant Loss
c
Sample from standard normal distribution
Max
Min
Diffusion Time t
d
Time Complexity
Sample Time (seconds)
225
200
175
150
125
100
75
50
0
10000
20000
30000
40000
Memory Complexity
Memory (MB)
6000
5000
4000
3000
2000
1000
0
10000
20000
30000
40000
Number of Cells

Figure 1 *(preceding page)*: LUNA is a generative AI model that reassembles cells into complex tissue structures from their gene expressions by learning spatial priors over existing spatially resolved datasets. (a) Given gene expressions of cells, the tissue reassembly process involves predicting the complex tissue architecture. We visualize the UMAP projection of cells based on their gene expression profiles ($1,122$ genes) from the MERFISH whole mouse Allen Brain Cell (ABC) Atlas [24] (left). Each cell is colored according to one of 338 distinct subclasses. LUNA's predictions of the spatial tissue structure is displayed on the right with eight slices randomly chosen from a total of 66 slices. LUNA was trained on spatial transcriptomics dataset of another mouse. **(b)** Overview of the training stage of LUNA. LUNA takes as input spatial transcriptomics data and corrupts the spatial locations of cells by adding noise. Using the attention mechanism, LUNA learns how to position each individual cell with respect to other cells based on their gene expressions. The loss function in LUNA ensures invariance to the rotations and reflections of the predicted slice. **(c)** Visualization of the denoising during the inference stage in LUNA across different diffusion steps on an example slice ($29,079$ cells) from the ABC atlas. The noise state is initialized from a standard normal distribution. Conditioned on the gene expressions of the cells, LUNA progressively removes noise at each diffusion step. **(d)** LUNA achieves linear time and memory complexity during inference relative to the number of input cells. Correlations between inference time and the number of input cells (upper) and between reserved memory of the GPU and the number of input cells (lower) across all 66 slices from the ABC atlas. Each point corresponds to a slice, correlating the cell count with either the time or memory used during the inference stage by LUNA.

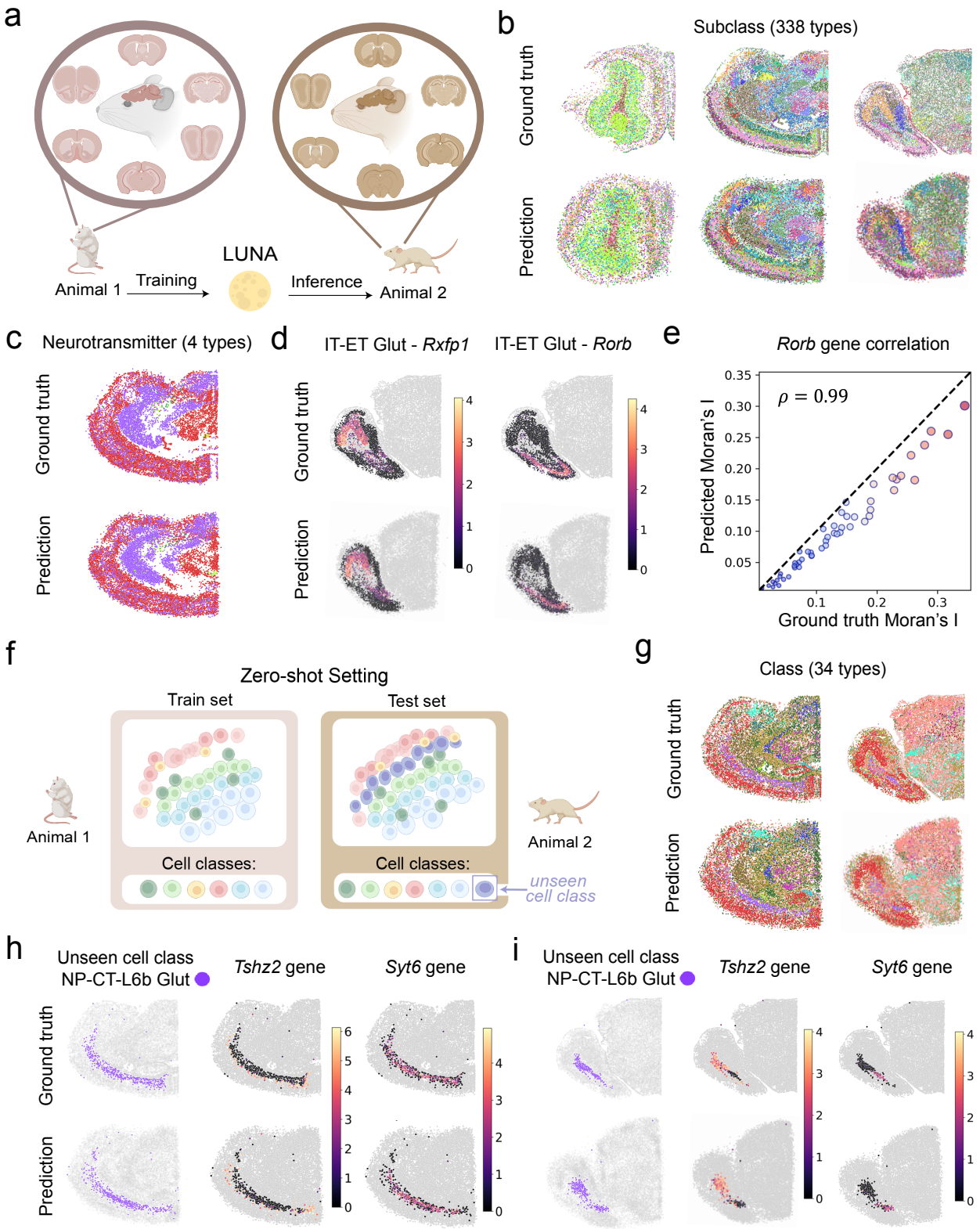
a
LUNA
Animal 1
Training
Inference
Animal 2
b
Subclass (338 types)
Ground truth
Prediction
c
Neurotransmitter (4 types)
Ground truth
Prediction
d
IT-ET Glut - Rxfp1
IT-ET Glut - Rorb
Ground truth
Prediction
e
Rorb gene correlation
ρ = 0.99
Predicted Moran's I
Ground truth Moran's I
f
Zero-shot Setting
Train set
Test set
Animal 1
Animal 2
Cell classes:
Cell classes:
unseen cell class
g
Class (34 types)
Ground truth
Prediction
h
Unseen cell class NP-CT-L6b Glut
Tshz2 gene
Syt6 gene
Ground truth
Prediction
i
Unseen cell class NP-CT-L6b Glut
Tshz2 gene
Syt6 gene

Figure 2 *(preceding page)*: LUNA accurately reconstructs the entire mouse brain atlas of 1.23 million cells. (a) LUNA was trained on all slices from the Animal 1 in the ABC atlas (2.85 million cells) and then applied to predict the tissue structure of cells from the Animal 2 unseen during training (1.23 million cells). **(b)** Examples of tissue reassembly results using LUNA on three representative slices from distinct major brain regions of the MERFISH whole mouse Allen Brain Cell (ABC) atlas [24], comprising $12,610$ cells, $30,931$ cells and $22,010$ cells, respectively. For visualization purposes, three slices were randomly selected from the ABC atlas. Cells are coloured according to their cell type annotations with 338 distinct cell subclasses, showing ground truth (top) and predicted (bottom) locations. **(c)** LUNA's predictions across 4 neurotransmitter types. Ground truth locations (top) and LUNA's predictions (bottom) for a slice with $35,738$ cells . **(d)** Spatial expression patterns of two spatially variable genes in the IT-ET Glut cell type. To select spatially variable genes, we calculated Moran's I values for all genes ($1,122$ in total) based on ground truth locations and identified those with the highest Moran's I value. Representative genes include *Rxfp1* and *Rorb* for IT-ET Glut. Cells from other cell types are shown in gray color. **(e)** Spatial autocorrelation for the *Rorb* gene computed using the Moran's I on the ground truth locations and locations predicted by LUNA. Each point represents a single slice across 66 slices in total. Slices are color-coded from blue (low) to red (high) according to the true Moran's I values for *Rorb*. Points near the diagonal show better spatial pattern preservation by LUNA. Pearson's correlation coefficient (ρ) is displayed in the top left. **(f)** Zero-shot setting evaluation of LUNA. We exclude all cells from a randomly chosen cell class during model training and evaluate LUNA's predictions on the cells from a held-out cell class without any further training or refinement of the model. **(g)** Tissue reassembly results using LUNA on two representative slices when NP-CT-L6b Glut cell class ($69,641$ cells in Animal 1) was excluded during training. The model was evaluated on Animal 2 including the NP-CT-L6b Glut cell class ($29,262$ cells). Cells are color-coded by 34 cell classes. **(h, i)** LUNA effectively generalizes to unseen cell types in a zero-shot setting, accurately capturing spatial gene expression patterns. Spatial distribution for the novel class (left) and the spatial expression patterns for genes *Tshz2* and *Syt6* in the NP-CT-L6b Glut cell type across two representative slices (middle and left). Notably, *Tshz2* and *Syt6* were identified as the most spatially variable among $1,122$ genes, selected based on the highest Moran's I values calculated from ground truth locations.

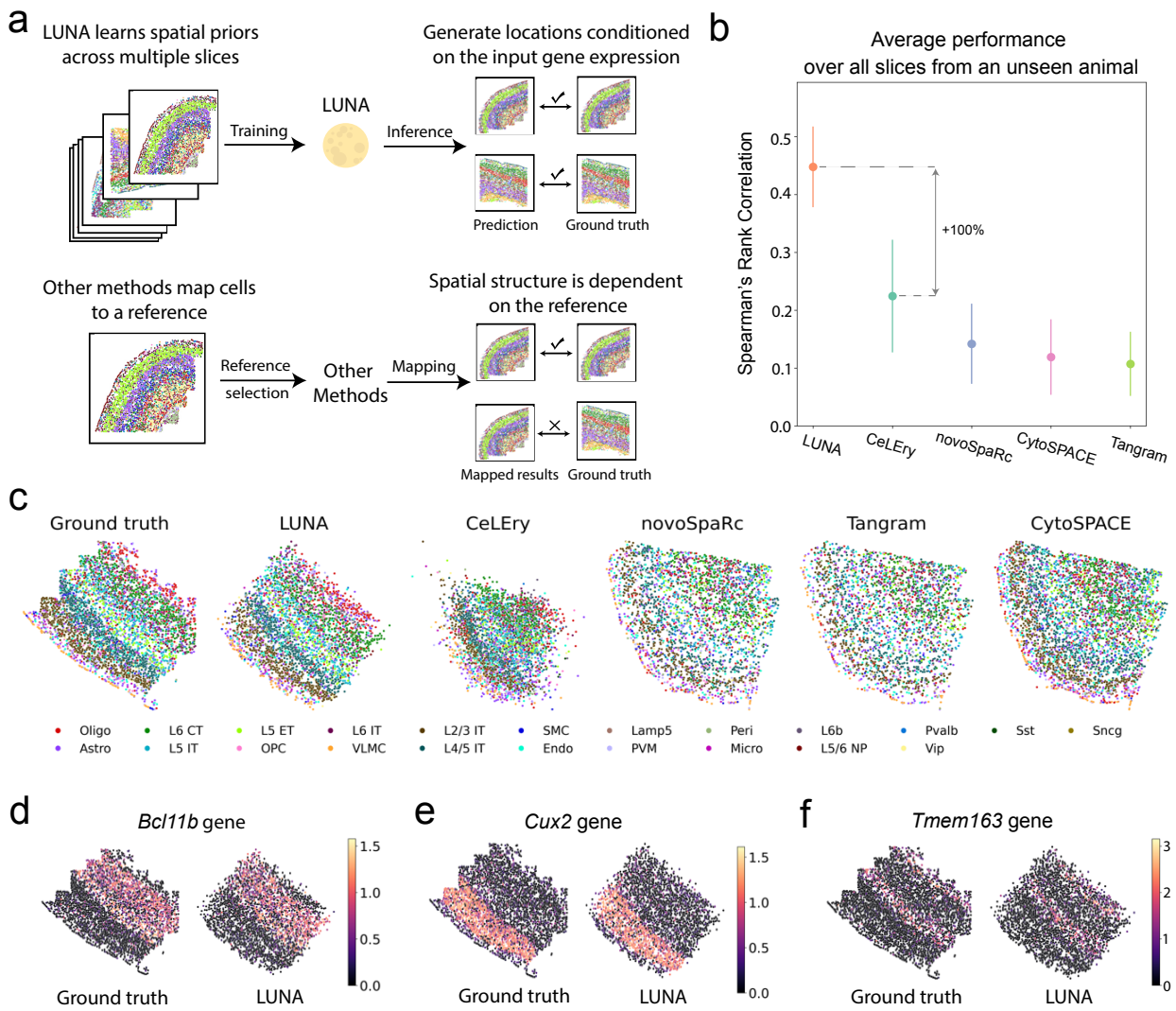
a
LUNA learns spatial priors across multiple slices
Training
LUNA
Inference
Generate locations conditioned on the input gene expression
Prediction
Ground truth
Other methods map cells to a reference
Reference selection
Other Methods
Mapping
Spatial structure is dependent on the reference
Mapped results
Ground truth
b
Average performance over all slices from an unseen animal
Spearman's Rank Correlation
+100%
LUNA
CeLEry
novoSpaRc
CytoSPACE
Tangram
c
Ground truth
LUNA
CeLEry
novoSpaRc
Tangram
CytoSPACE
Oligo
Astro
L6 CT
L5 IT
L5 ET
OPC
L6 IT
VLMC
L2/3 IT
L4/5 IT
SMC
Endo
Lamp5
PVM
Peri
Micro
L6b
L5/6 NP
Pvalb
Vip
Sst
Sncg
d
Bcl11b gene
Ground truth
LUNA
e
Cux2 gene
Ground truth
LUNA
f
Tmem163 gene
Ground truth
LUNA

Figure 3 *(preceding page)*: LUNA outperforms alternative baselines on the MERFISH mouse primary motor cortex atlas by a large margin. (a) LUNA fundamentally differs from the existing methods. LUNA takes as input different slices and learns spatial priors across all slices during training. In the inference stage, LUNA generates cell locations conditioned on their gene expressions by leveraging learnt spatial priors. In contrast, other methods require the pre-selection of a single reference slice and then map cells to the reference slice. Consequently, the spatial structure of their results heavily depends on the chosen reference slice. **(b)** Performance comparison on the cross-animal generalization task between LUNA and alternative baselines on the MERFISH mouse primary motor cortex atlas [28]. The performance is measured as the median Spearman's rank correlation between the ground truth spatial coordinates and the predicted spatial coordinates over all slices from an unseen mice (31 slices, $118,036$ cells). Higher correlation indicate better prediction performance. Error bars represent standard deviation across 31 slices. **(c)** Prediction of cell locations for a single slice with $5,024$ cells using LUNA and alternative baselines. Colors denote cell class labels. Ground truth locations are shown on the far left. LUNA can accurately predict the layered structure of the mouse cortex. **(d, e, f)** Spatial expression patterns of **(d)** the *Cux2* gene, a marker for layer 2/3 (L2/3 IT), **(e)** the *Bcl11b* gene, a marker for layer 6 corticothalamic (L6 CT), and **(f)** the *Tmem163* gene, a marker for layer 6b (L6b). Ground truth cell locations are shown in the left plot, while the predictions made by LUNA are displayed in the right plot. LUNA correctly captures gene expression gradients.

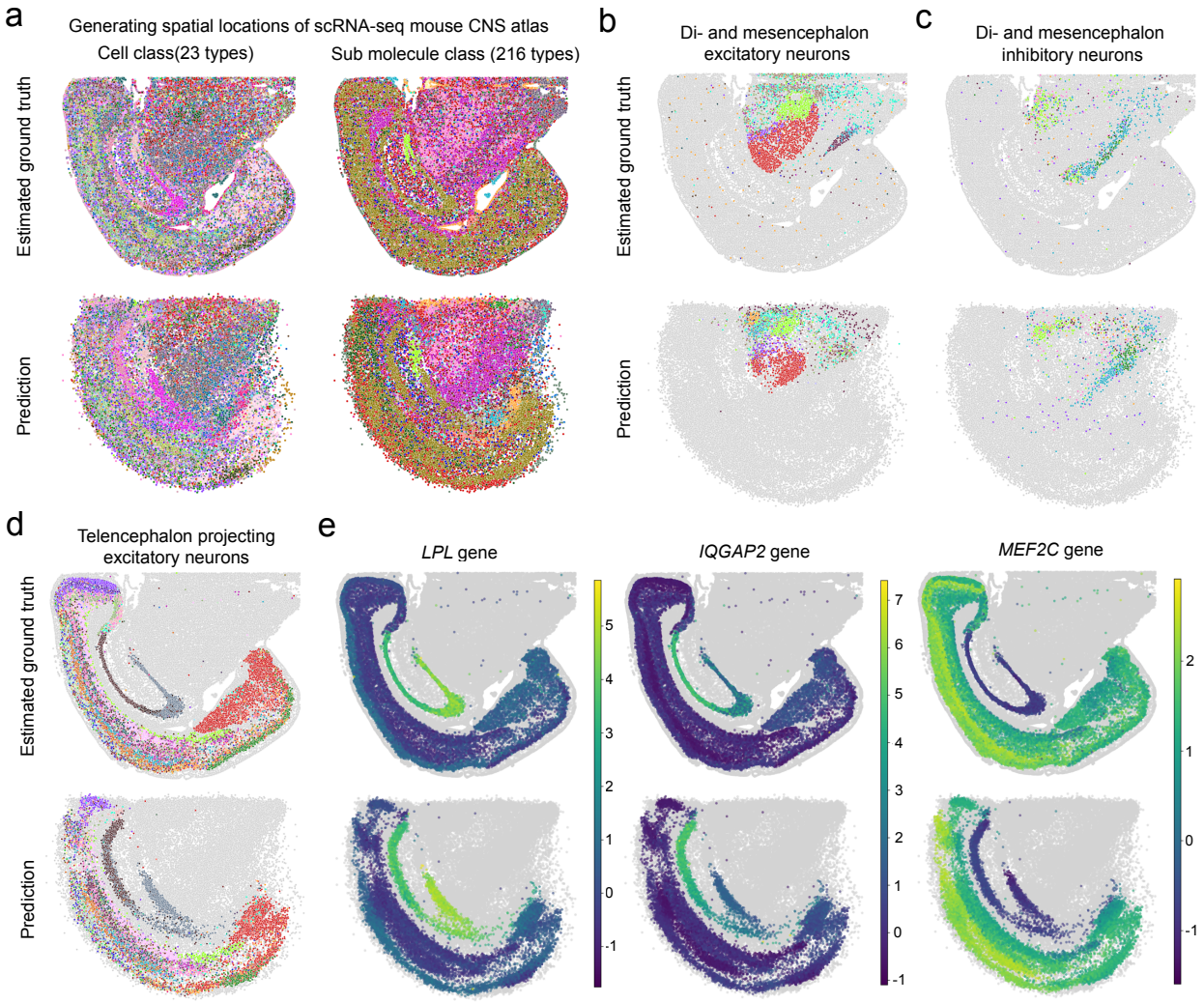
a
Generating spatial locations of scRNA-seq mouse CNS atlas
Cell class(23 types)
Sub molecule class (216 types)
Estimated ground truth
Prediction
b
Di- and mesencephalon excitatory neurons
c
Di- and mesencephalon inhibitory neurons
d
Telencephalon projecting excitatory neurons
e
LPL gene
IQGAP2 gene
MEF2C gene

Figure 4 *(preceding page)*: LUNA reassemble the tissue structure of 1.08 million cells from the scRNA-seq mouse central nervous atlas *de novo*. (a) Tissue reassembly of the scRNA-seq atlas of the mouse central nervous system atlas using LUNA [25]. LUNA was trained on all the cells from Animal 1 of the ABC atlas (2.85 million cells) and applied to generate cell locations for a scRNA-seq atlas from the mouse central nervous system (1.08 million cells, 13 coronal slices). One slice containing $35,738$ cells is visualized and cells are colored based on the cell classes (23 types: left) and sub-molecule classes (216 types; right). Ground truth spatial locations are estimated by integrating the CNS scRNA-seq dataset with the STARmap PLUS dataset [29]. **(b-d)** Tissue reassembly of the scRNA-seq mouse central nervous system atlas using LUNA for **(b)** di- and mesencephalon excitatory neurons, **(c)** di- and mesencephalon inhibitory neurons, and **(d)** telencephalon projecting excitatory neurons. The top plot shows estimated ground truth locations obtained by aligning scRNA-seq atlas with the STARmap atlas and the bottom plot shows LUNA's predictions. Cells from each cell type are colored by their sub-molecule class. LUNA accurately places cells and preserves the spatial neighborhood of single cells at a very fine-grained level. Cells from other cell types are shown in gray color. **(e)** Spatial expression patterns of three spatially variable genes across telencephalon-projecting excitatory neurons. To identify these genes, we calculated Moran's I values for all $10,844$ genes based on LUNA's predictions, selecting three genes with distinct spatial patterns from the top 10 genes with the highest Moran's I values. The representative genes include *LPL*, *IQGAP2*, and *MEF2C*. Cells from other cell types are depicted in gray.

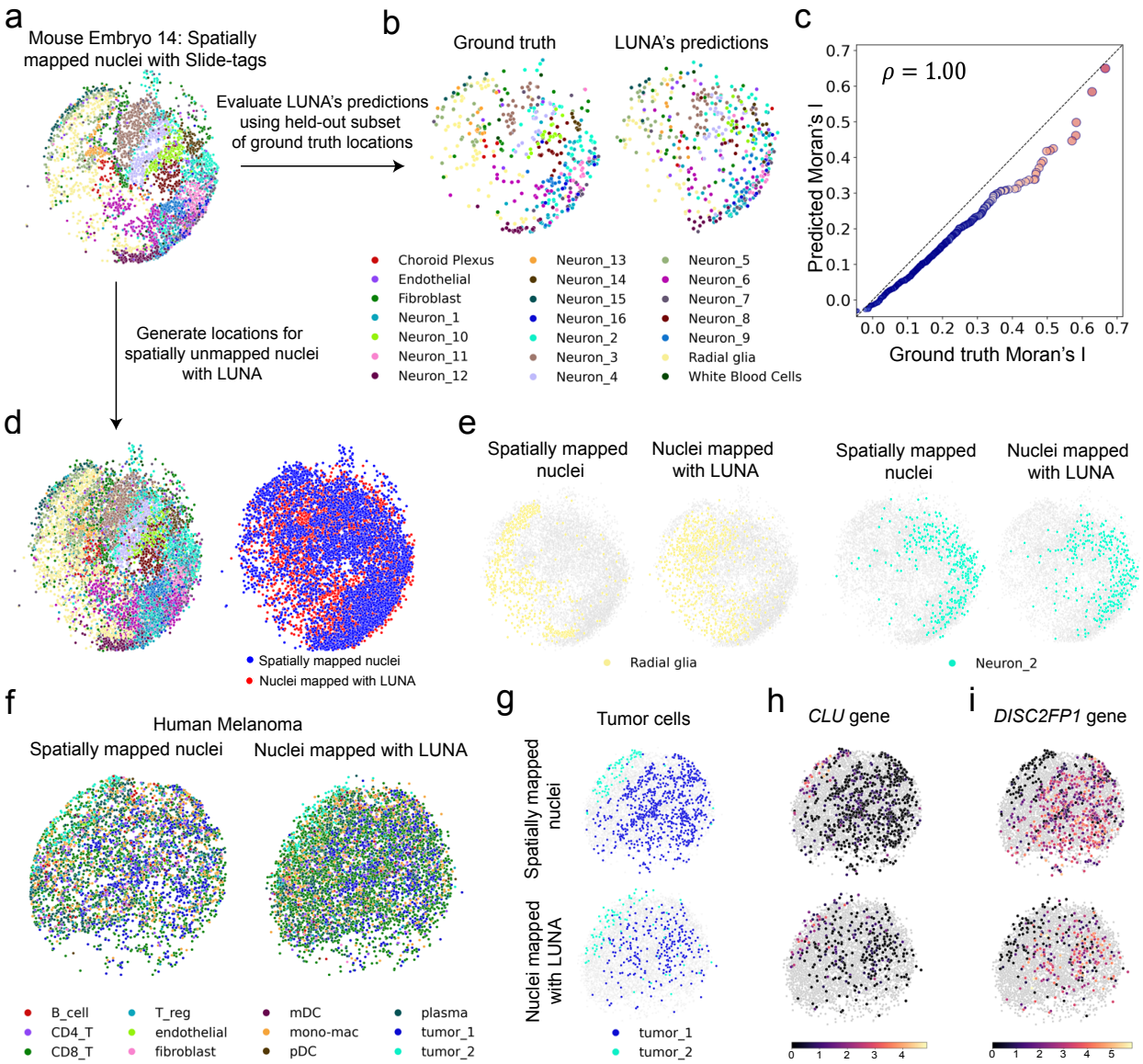
a
Mouse Embryo 14: Spatially mapped nuclei with Slide-tags
Evaluate LUNA's predictions using held-out subset of ground truth locations
Generate locations for spatially unmapped nuclei with LUNA
b
Ground truth
LUNA's predictions
Choroid Plexus
Endothelial
Fibroblast
Neuron_1
Neuron_10
Neuron_11
Neuron_12
Neuron_13
Neuron_14
Neuron_15
Neuron_16
Neuron_2
Neuron_3
Neuron_4
Neuron_5
Neuron_6
Neuron_7
Neuron_8
Neuron_9
Radial glia
White Blood Cells
c
$\rho = 1.00$
Predicted Moran's I
Ground truth Moran's I
0.0 0.1 0.2 0.3 0.4 0.5 0.6 0.7
d
Spatially mapped nuclei
Nuclei mapped with LUNA
e
Spatially mapped nuclei
Nuclei mapped with LUNA
Spatially mapped nuclei
Nuclei mapped with LUNA
Radial glia
Neuron_2
f
Human Melanoma
Spatially mapped nuclei
Nuclei mapped with LUNA
B_cell
CD4_T
CD8_T
T_reg
endothelial
fibroblast
mDC
mono-mac
pDC
plasma
tumor_1
tumor_2
g
Tumor cells
Spatially mapped nuclei
Nuclei mapped with LUNA
tumor_1
tumor_2
h
CLU gene
0 1 2 3 4
i
DISC2FP1 gene
0 1 2 3 4 5

Figure 5 *(preceding page)*: LUNA infers locations of spatially unmapped nuclei in the Slide-tags data and correctly predicts the spatial compartments of tumor cells. **(a)** Spatially mapped nuclei using Slide-tags from the embryonic mouse brain at embryonic day 14 (E14) containing $4,614$ cells [26]. Cells are coloured according to their cell type annotation. **(b)** LUNA's predictions of nuclei locations agree well with the ground truth locations from Slide-tags. Ground truth locations (left) and LUNA's predictions (right) on the subset of the held-out dataset of the mouse E14 embryonic brain. LUNA was trained on $4,152$ cells with ground truth locations from spatially mapped nuclei and we evaluated its predictions on the held-test set of 462 cells. Cells are coloured according to their cell type annotation. **(c)** Spatial autocorrelation computed using the Moran's I on the ground truth locations and locations predicted by LUNA on the held-out test set of mouse E14 embryonic brain. Each point represents a gene and 700 highly differentially expressed genes are visualized. Genes are colored on a gradient from blue (low) to red (high) based on their ground truth Moran's I value. Points closer to the diagonal indicate better preservation of spatial patterns by LUNA. The Pearson's correlation coefficient (ρ) is reported on the top left. **(d)** LUNA generated locations of $4,414$ initially unmapped nuclei of the mouse E14 embryonic brain. LUNA was trained on spatially mapped nuclei and then applied to infer locations for the unmapped nuclei. The left panel shows the combined locations of nuclei mapped by LUNA and those spatially mapped by Slide-tags with cells coloured according to their cell type annotation. The right panel differentiates the source of cell locations: nuclei mapped by Slide-tags (blue) and those mapped by LUNA (red). **(e)** Nuclei mapped by Slide-tags (left) and nuclei mapped by LUNA for Neuron_2 (yellow color) and Radial glia (cyan color) cell types. All other cells are shown in gray color. **(f)** LUNA enriched the Slide-tags human metastatic melanoma sample from $4,804$ cells to $6,466$ cells by generating locations for spatially unmapped nuclei. The left panel shows the ground truth locations of spatially mapped nuclei used as the training dataset, while the right panel combines these with the nuclei locations predicted by LUNA. **(g)** LUNA correctly inferred the spatial compartment of two tumor clusters segregated into spatially distinct compartments. Spatially mapped nuclei by Slide-tags (top) and locations of unmapped nuclei generated by LUNA (bottom). Tumor cells are colored according to their type, while other cells are shown in gray color. **(h, i)** Spatial expression patterns of **(h)** the *CLU* gene and **(i)** the *DISC2FP1* gene across spatial compartments of tumor cells. We selected these genes as spatially variable genes for each tumor type: the *CLU* gene for tumor_1 and the *DISC2FP1* gene for tumor_2 (i). Spatially mapped nuclei by Slide-tags (top) and locations of unmapped nuclei generated by LUNA (bottom). Tumor cells are colored according to expression levels of these genes. Other cells are shown in gray color.

Acknowledgements

We are grateful to Ramon Viñas Torné, Shuyang Fan, Clément Vignac, Yiming Qin, Gioele la Manno, Philippe Schwaller, Evan Macosko and Mor Nitzan for valuable discussions. We gratefully acknowledge the support of the Swiss National Science Foundation (grant IC00I0-231922 and starting grant TMSGI2_226252) and Zeiss. Figure elements, including icons of species, were created with BioRender.com.

Author Contributions Statement

T.Y., C.E. and M.B. designed the study, performed research, contributed new analytical tools and analyzed data. T. Y. and C.E. performed experiments and developed the software. N.M., J.F. and S.A. contributed to the codebase. P.F. contributed to the algorithmic framework. T.Y. and M.B. wrote the manuscript with the input from other authors. M.B. supervised the research.

Competing Interests Statement

We declare no competing interests.

Data availability

All analyzed datasets are publicly available. MERFISH Whole Mouse Brain Atlas (ABC Atlas) for Animal 1 is available at https://alleninstitute.github.io/abc_atlas_access/descriptions/Zhuang-ABCA-1.html and for Animal 2 is available at https://alleninstitute.github.io/abc_atlas_access/descriptions/Zhuang-ABCA-2.html. MERFISH Mouse Primary Motor Cortex Atlas is available at the Brain Image Library: https://doi.brainimagelibrary.org/doi/10.35077/g.21. scRNA-seq Mouse

Central Nervous System Atlas is available at the Single Cell Portal: https://singlecell.broadinstitute.org/single_cell/study/SCP1830. Slide-tags Datasets for all the tissues are available at the Broad Institute Single Cell Portal under the following accession numbers: SCP2170 (mouse E14), SCP2171 (human melanoma), SCP2169 (human tonsil) and SCP2167 (human brain).

Code availability

LUNA was written in Python 3.9 using the PyTorch library. The source code is available on Github at https://github.com/mlbio-epfl/LUNA. The project website with links to data and code can be accessed at https://brbiclab.epfl.ch/projects/LUNA/.

References

1. Tang, F. *et al.* mRNA-Seq whole-transcriptome analysis of a single cell. *Nature Methods* **6**, 377–382 (2009).

2. Macosko, E. Z. *et al.* Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets. *Cell* **161**, 1202–1214 (2015).

3. Patel, A. P. *et al.* Single-cell RNA-seq highlights intratumoral heterogeneity in primary glioblastoma. *Science* **344**, 1396–1401 (2014).

4. Consortium*, T. T. S. *et al.* The Tabula Sapiens: A multiple-organ, single-cell transcriptomic atlas of humans. *Science* **376**, eabl4896 (2022).

5. Lu, T.-C. *et al.* Aging Fly Cell Atlas identifies exhaustive aging features at cellular resolution. *Science* **380**, eadg0934 (2023).

6. Li, H. *et al.* Fly Cell Atlas: A single-nucleus transcriptomic atlas of the adult fruit fly. *Science* **375**, eabk2432 (2022).

7. Sikkema, L. *et al.* An integrated cell atlas of the lung in health and disease. *Nature Medicine* **29**, 1563–1577 (2023).

8. Hickey, J. W. *et al.* Organization of the human intestine at single-cell resolution. *Nature* **619**, 572–584 (2023).

9. Chen, K. H., Boettiger, A. N., Moffitt, J. R., Wang, S. & Zhuang, X. Spatially resolved, highly multiplexed RNA profiling in single cells. *Science* **348**, aaa6090 (2015).

10. Codeluppi, S. *et al.* Spatial organization of the somatosensory cortex revealed by osmFISH. *Nature Methods* **15**, 932–935 (2018).

11. Wang, X. *et al.* Three-dimensional intact-tissue sequencing of single-cell transcriptional states. *Science* **361**, eaat5691 (2018).

12. Achim, K. *et al.* High-throughput spatial mapping of single-cell RNA-seq data to tissue of origin. *Nature Biotechnology* **33**, 503–509 (2015).

13. Satija, R., Farrell, J. A., Gennert, D., Schier, A. F. & Regev, A. Spatial reconstruction of single-cell gene expression data. *Nature Biotechnology* **33**, 495–502 (2015).

14. Biancalani, T. *et al.* Deep learning and alignment of spatially resolved single-cell transcriptomes with Tangram. *Nature Methods* **18**, 1352–1362 (2021).

15. Vahid, A. *et al.* High-resolution mapping of the spatial organization of single cells in tissues using CytoSPACE. *Nature Methods* **20**, 362–373 (2023).

16. Zhang, W. *et al.* Leveraging single-cell spatial transcriptomics with CeLEry for spatial reconstruction. *Nature Biotechnology* **41**, 245–256 (2023).

17. Nitzan, M., Karaiskos, N., Friedman, N. & Rajewsky, N. Gene expression cartography. *Nature* **576**, 132–137 (2019).

18. Moriel, N. *et al.* NovoSpaRc: Flexible spatial reconstruction of single-cell gene expression with optimal transport. *Nature Communications* **12**, 1–12 (2021).

19. Vaswani, A. *et al.* Attention is all you need. *Advances in Neural Information Processing Systems* (2017).

20. Shen, Z., Zhang, M., Zhao, H., Yi, S. & Li, H. Efficient attention: Attention with linear complexities. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision* (2021).

21. Sohl-Dickstein, J., Weiss, E., Maheswaranathan, N. & Ganguli, S. Deep unsupervised learning using nonequilibrium thermodynamics. In *International Conference on Machine Learning* (2015).

22. Ho, J., Jain, A. & Abbeel, P. Denoising diffusion probabilistic models. *Advances in Neural Information Processing Systems* (2020).

23. Song, Y. & Ermon, S. Generative modeling by estimating gradients of the data distribution. *Advances in Neural Information Processing Systems* (2019).

24. Zhang, M. *et al.* Molecularly defined and spatially resolved cell atlas of the whole mouse brain. *Nature* **624**, 343–354 (2023).

25. Zeisel, A. *et al.* Molecular architecture of the mouse nervous system. *Cell* **174**, 999–1014 (2018).

26. Russell, A. J. *et al.* Slide-tags enables single-nucleus barcoding for multimodal spatial genomics. *Nature* **625**, 101–109 (2024).

27. Moran, P. A. Notes on continuous stochastic phenomena. *Biometrika* **37**, 17–23 (1950).

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

29. Shi, H. *et al.* Spatial atlas of the mouse central nervous system at molecular resolution. *Nature* **622**, 552–561 (2023).

30. Palla, G. *et al.* Squidpy: A scalable framework for spatial omics analysis. *Nature Methods* **19**, 171–178 (2022).

31. Hao, Y. *et al.* Dictionary learning for integrative, multimodal and scalable single-cell analysis. *Nature Biotechnology* **42**, 293–304 (2024).

32. Brbić, M. *et al.* Annotation of spatially resolved single-cell data with STELLAR. *Nature Methods* **19**, 1411–1418 (2022).

33. Clifton, K. *et al.* STalign: Alignment of spatial transcriptomics data using diffeomorphic metric mapping. *Nature Communications* **14**, 8123 (2023).

34. Schott, M. *et al.* Open-ST: High-resolution spatial transcriptomics in 3D. *Cell* **187**, 3953–3972 (2024).

35. Segol, N. & Lipman, Y. On universal equivariant set networks. In *International Conference on Learning Representations* (2020).

36. Torres-Mendez, L. A., Ruiz-Suarez, J. C., Sucar, L. E. & Gómez, G. Translation, rotation, and scale-invariant object recognition. *IEEE Transactions on Systems, Man, and Cybernetics* **30**, 125–130 (2000).

37. Romero, D., Bekkers, E., Tomczak, J. & Hoogendoorn, M. Attentive group equivariant convolutional networks. In *International Conference on Machine Learning* (2020).

38. Wu, Y., Löwe, H., Carricato, M. & Li, Z. Inversion symmetry of the Euclidean group: theory and application to robot kinematics. *IEEE Transactions on Robotics* **32**, 312–326 (2016).

39. Song, J., Meng, C. & Ermon, S. Denoising diffusion implicit models. In *International Conference on Learning Representations* (2021).

40. Chan, S. H. Tutorial on diffusion models for imaging and vision. *arXiv preprint arXiv:2403.18103* (2024).

41. Xu, M. *et al.* GeoDiff: A geometric diffusion model for molecular conformation generation. In *International Conference on Learning Representations* (2022).

42. Hoogeboom, E., Satorras, V. G., Vignac, C. & Welling, M. Equivariant diffusion for molecule generation in 3D. In *International Conference on Machine Learning* (2022).