NOLAN: SELF-SUPERVISED FRAMEWORK FOR MAPPING CONTINUOUS TISSUE ORGANIZATION

Artemii Bakulin
Weizmann Institute of Science
Rehovot, Israel
artemii.bakulin@weizmann.ac.il

Nathan Levy & Jonas Maaskola
Weizmann Institute of Science
Rehovot, Israel
{nathan.levy, jonas.maaskola}@weizmann.ac.il

Can Ergen
University of California, Berkeley
Berkeley, CA, USA
cergen@berkeley.edu

Nir Yosef
Weizmann Institute of Science &
University of California, Berkeley
nir.yosef@weizmann.ac.il

ABSTRACT

Spatial transcriptomics offers unprecedented insights into tissue organization, yet current methods often overlook transitional zones between cellular niches. We introduce NOLAN, a self-supervised framework that goes beyond detecting discrete niches to capture the continuous spectrum of tissue organization patterns. NOLAN learns cell representations informed by their neighborhoods, capturing variation within niches and across their boundaries. Using these representations, NOLAN constructs a graph-based abstraction of the tissue, modeling it as a network of interconnected regions bridged by transitional zones. Applying NOLAN to a multi-cancer spatial transcriptomics atlas, we uncover a landscape of both tissue-specific and shared cellular niches. Crucially, NOLAN reveals the continuous gradients of gene expression and cell type composition across these transitional zones, showcasing the ability of NOLAN to build a common coordinate system of tissues in an integrative analysis.

1 INTRODUCTION

Understanding tissue structure and organization is critical for deciphering cellular function and interactions in health and disease. Spatial transcriptomics provides a powerful means to profile gene expression at near-cellular resolution, allowing us to study cells within their native tissue context. A common analysis strategy of the spatial data involves clustering cellular neighborhoods to identify spatial domains, also referred to as cellular niches. Several computational methods have been to autmoatically detect these niches (Hu et al., 2021; Yuan, 2024; Varrone et al., 2024; Singhal et al., 2024). However, many existing approaches assume a simplistic tissue model partitioned into discrete regions. Recent work (Bhate et al., 2022) demonstrates that tissue organization is considerably more complex, with gradual transitions between the niches. This highlights a critical need for computational methods that move beyond discrete partitioning and explicitly model the continuous nature of tissue organization.

To address this, we introduce NOLAN (NO Label Analysis of Niches), a self-supervised model for spatial transcriptomics analysis. NOLAN constructs a rich representation of each cell, incorporating information from its spatially adjacent neighbors. The self-supervised strategy used by NOLAN was adapted from DINOv2 (Oquab et al., 2024) by modifying the architecture, tokenization method and augemntations. This strategy allows NOLAN to capture spatial prototypes with conserved gene expression and cell type distribution, as well as transitionary zones between them.

NOLAN's representations of the spatial transcriptomic data, sensitive to both intra-niche variation and inter-niche transitions, enable a powerful new way to describe tissue architecture. We conceptualize this model as a tissue graph, where nodes represent individual cells, and edges connect cells based on the similarity of their NOLAN representations. Crucially, the strength of these connections reflects the abruptness of the transition between niches and the size of their interface. To assess the

accuracy of node assignments within the tissue graph, we developed novel biologically motivated metrics. These metrics quantify the structural and cell type composition consistency of the identified niches. We used these metrics across diverse spatial transcriptomic datasets, showcasing NOLAN's superior performance.

To demonstrate the broad applicability and power of NOLAN, we performed an integrated analysis of a large and diverse collection of cancer spatial transcriptomics data, spanning multiple tissue origins. In complex tissues like cancer, spatial organization is often obscure, yet it plays a crucial role in driving tumor progression and heterogeneity, as many pathological traits exhibit localized patterns. Leveraging NOLAN's unique ability to define both distinct niches and their transitional zones within a unified tissue graph framework, we directly addressed the challenge of comparative analysis across these heterogeneous samples. Applying NOLAN, we identified niches with diverse cell type distribution, revealing both shared and unique niche compositions across different tumor types. Critically, by mapping the transitional regions across samples, we were able to use them as anchor points to describe spatially-associated changes in gene expression and cell-type distribution.

2 RELATED WORK

2.1 IDENTIFICATION OF TISSUE STRUCTURE

Several computational methods have been developed for tissue segmentation from spatial transcriptomics data, reflecting the rapid advancements in this field. We discuss a few recent examples that are relevant to our benchmarking analysis.

MENDER (Yuan, 2024) employs a two-stage clustering approach. First, it clusters cells into subtypes based on their gene expression profiles. Then, it calculates the frequency of each subtype within small neighborhoods around each cell. These neighborhood compositions are subsequently used to cluster cellular environments, defining tissue domains. A consideration with MENDER is its reliance on discrete cell-type assignments, which may not fully capture the continuous spectrum of cell states often observed in vivo.

BANKSY (Singhal et al., 2024) adopts a different strategy, directly analyzing gene expression gradients using filter theory. Spatial kernels, including azimuthal Gabor filters, are used to encode the transcriptomic environment around each cell. This approach allows BANKSY to integrate cell typing and tissue domain segmentation into a single framework and offers scalability for large datasets. One aspect to consider is that BANKSY does not inherently include batch effect correction, which can be important for analyzing data from multiple samples or experiments.

CellCharter (Varrone et al., 2024) leverages a graph neural network (GNN) architecture for spatial domain identification. It uses scVI embeddings (Lopez et al., 2018) of gene expression as input to the GNN, allowing it to learn complex spatial relationships between cells.

Existing methods, including those described above, have primarily focused on identifying discrete tissue niches or domains. A key distinction of our approach, NOLAN, is its explicit modeling and characterization of the transitional regions between niches, represented within a tissue graph framework. This capability allows for a more nuanced understanding of tissue organization, going beyond a simple partitioning into discrete regions.

3 METHODS

3.1 BACKGROUND

3.1.1 SELF-SUPERVISED LEARNING

Self-supervised learning (SSL) has emerged as a powerful paradigm in machine learning, encompassing two primary approaches: contrastive learning and self-distillation. These methods aim to learn meaningful representations from unlabeled data, reducing the need for extensive labeled datasets.

Self-distillation strategies leverage the model's own predictions to generate supervisory signals, often within a teacher-student framework. In this setting, the student network is trained to match

the predictions of a teacher network. Let $p_s(x)$ and $p_t(x)$ denote the output probability distributions from the student and teacher networks, respectively. A common loss function used to enforce consistency is the cross-entropy between these two distributions:

$$\mathcal{L}_{\text{distill}} = -\sum_{c=1}^{C} p_{t,c}(x) \log p_{s,c}(x), \tag{1}$$

where C is the number of classes (or feature dimensions) and $p_{t,c}(x)$ and $p_{s,c}(x)$ are the probabilities corresponding to class c.

DINO (DIstillation with NO labels) (Caron et al., 2021) and its successor, DINOv2 (Oquab et al., 2024), both outperform most self-supervised models, such as BYOL (Grill et al., 2020), as well as contrastive methods like MoCov2 (Chen et al., 2003). In these models, the teacher's parameters θ_t are updated using an exponential moving average of the student's parameters θ_s, ensuring stable training:

$$\theta_t \leftarrow \lambda\theta_t + (1-\lambda)\theta_s, \tag{2}$$

where $\lambda \in [0, 1]$ is a smoothing coefficient. The student network learns to produce consistent predictions across different augmented views of the same input, guided by the teacher's outputs.

DINOv2 additionally applies a masked image modelling loss, also referred to as iBOT (Zhou et al., 2021). To compute this loss DINOv2 evaluates the cross-entropy between the softmax-normalized teacher and student representations of the masked patches of the image.

3.2 Overview of NOLAN

NOLAN is a self-supervised framework that uses knowledge distillation, a concept rooted in image analysis to learn informative representations of cell neighborhoods (**Figure 1**). At both training and inference time, the model samples cells as well as their cell neighborhoods, encodes them as a sequence of tokens, appends a learnable classification token to this sequence, and passes it through a transformer encoder network. Following this encoder network, the classification token is passed through a projection head and is converted into a niche-aware representation of the center cell, used for niche assignment.

Unlike DINOv2, which was designed for image analysis, NOLAN is specifically tailored for spatial transcriptomics data, which can be represented as 2D point clouds of cells. Our tokenization strategy departs significantly from DINOv2's image-based approach. Instead of image patches, we sample cells and their surrounding neighborhoods, treating each cell as a token. As an augmentation strategy, we perform random rotation, flipping of the neighborhoods and stochastic masking of cells.

NOLAN's training, like DINOv2, employs a teacher-student framework with two networks that are architectural clones, but have different weight initializations. The teacher network processes larger, "global" views of the tissue, using cell neighborhoods typically with a 120 µm radius. The student network uses both these global views and smaller, "local" views (30 µm radius), providing a multi-scale perspective. To accelerate training, we leverage pre-computed scVI embeddings of gene expression (Lopez et al., 2018). These embeddings are projected to NOLAN's dimensionality using a linear layer with spectral initialization. Positional information is encoded using frequency-based positional encodings (Dufter et al., 2022), relative to the center of each neighborhood. The resulting positional and gene expression encodings are summed, and a learnable classification token, analogous to DINOv2, is appended to the sequence of cell tokens and passed through a shallow transformer encoder and a DINO head. Finally, mirroring DINOv2, we compute both classification token and masked feature cross-entropy losses during training.

See the details about how we construct the tissue graph in the Supplementary.

3.3 Evaluation metrics

Because NOLAN is, to our knowledge, the first method to explicitly model transitions between niches, our primary benchmarking focus was on evaluating the quality of the identified niches themselves. A commonly used benchmark for spatial transcriptomics methods involves comparing predicted niches to manual annotations. Following established practice, we compute Normalized Mutual Information (NMI) between the predicted niches and the annotations from a MERFISH brain

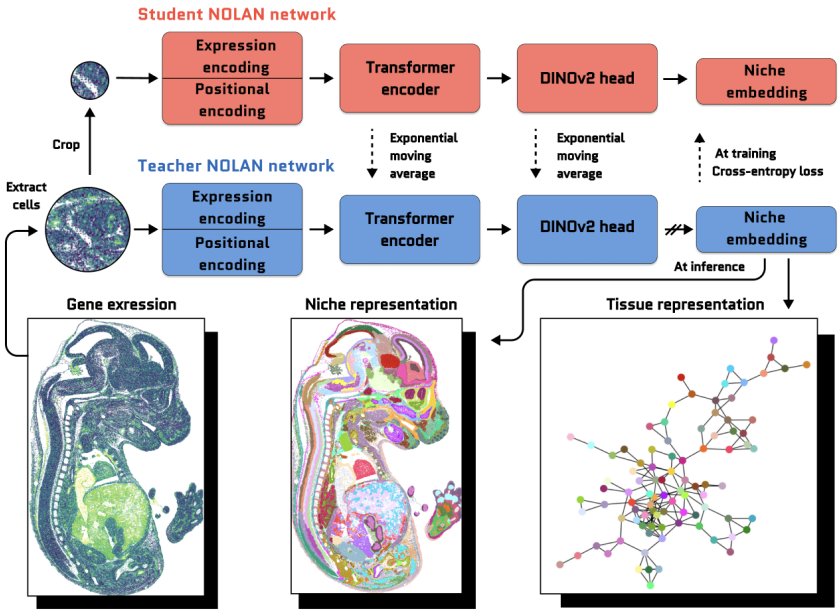

Figure 1: Overview of NOLAN. NOLAN is designed to learn niche representation of the cell based on the gene expression distribution in its neighborhood. To learn robust representations, NOLAN uses a self-supervised learning. Student and teacher networks are trained using correspondingly big and small neighborhoods. The objective of the training is for the student to mimic the teacher's output. At inference the niche embeddings associated with each cell are collected and used to build tissue graph and assign niche labels.

atlas. However, due to the low resolution of the atlas annotation and the possible inaccuracies in how these annotations were constructed, we extend our evaluation suite with additional biologically motivated metrics.

First, we evaluated the degree of complexity of the predicted annotations. We measure **niche complexity** as the Shannon entropy of the distribution of niche label (n) assignment ($a(c)$) within each tissue slice (S) (Formula 3). Higher entropy values indicate a more refined and detailed segmentation, capturing a greater diversity of spatial patterns.

$$NC = -\sum_{n \in N} p(n) \log_2 p(n), \quad p(n) = \frac{|\{c \in S \mid a(c) = n\}|}{|S|} \tag{3}$$

However, high complexity alone is insufficient; it must be balanced with spatial coherence. To prevent degenerate solutions—where each cell is trivially assigned to its own unique niche, creating a fragmented, biologically unrealistic pattern – we introduce **tissue continuity** (Fomula 4). For each cell, we calculate the fraction of its ten nearest spatial neighbors that share the same niche label and then report its mean value over a dataset.

$$TC = \frac{1}{|S|} \sum_{c \in S} f(c), \quad f(c) = \frac{|\{c' \in KNN(c, k) \mid a(c') = a(c)\}|}{k} \tag{4}$$

Beyond spatial coherence, we also assess **sample integration** using the Adjusted Rand Index (ARI) (Yang, 2016). ARI measures the agreement between the predicted niche labels and the known sample origins, providing insight into whether the identified niches are consistent across different biological samples. A low ARI suggests that the identified niches are not tissue specific and the dataset is well integrated.

Finally, we measured **cell type consistency** within niches. For each cell, we consider its local neighborhood (a 100-micrometer radius) and compare its cell type distribution to the overall cell

type distribution in the niche. Specifically, we use a two-sided t-test for proportions and Benjamini-Hochberg correction (Benjamini & Hochberg, 1995) to identify cell types that are significantly enriched or depleted within the neighborhood. A lower number of significantly different cell types indicates higher consistency. Note that in the benchmarking section we report the success rate over a series of trials with randomly selected neighborhoods where the success means that the method identifies fewer differentially abundant cell types than the rest. And so for this metric the higher values are better, suggesting that the identified niches maintain a stable cell-type composition across their spatial extent.

4 RESULTS

4.1 NOLAN PROVIDES GRAPH ABSTRACTION OF CONTINUOUS TISSUE VARIATION

A defining characteristic of NOLAN is its ability to capture the continuous variation of cell states, reflecting the nuanced nature of tissue organization. We can visualize this by examining the manifold structure learned by NOLAN. Using a UMAP projection of the cell embeddings, we observe a topology composed of distinct clusters connected by thin "bridges" (see **Figure 2A**). These clusters, we find, correspond to the identified niches, and the connecting bridges represent the transitional zones between them, reflecting the smooth transitions in cellular states. This manifold structure is directly mirrored in our tissue graph abstraction: the UMAP clusters are mapped to nodes, each representing a distinct niche, while the connecting bridges between clusters are represented by the edges connecting these nodes in the graph (see **Figure 2B**). Furthermore, we can project the niche assignments directly onto the spatial coordinates of the tissue section, revealing a coherent and spatially organized structure (see **Figure 2C**).

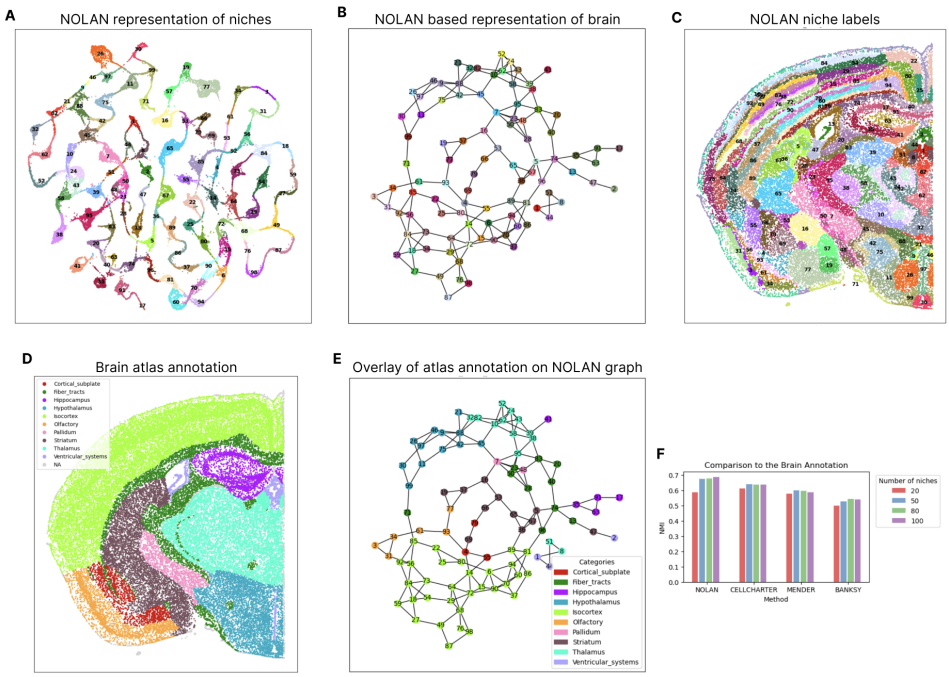

Figure 2: NOLAN captures continuous niche variation. A. UMAP representation of the NOLAN embedding space. B. Tissue graph colored by the NOLAN derived niche labels. C. Brain slice colored by the the niche labels. D. Brain slice colored according to the atlas annotation. E. Tissue graph colored according to the correspondence of nodes to regions from brain annotation. F. NMI scores between NOLAN annotation and atlas annotation, the results are reported for different clustering levels.

Notably, the connectivity structure of the tissue graph, reflecting the relationships between niches, exhibits a strong coherence with the spatial colocalization of niches. To demonstrate this, we color

the graph nodes by the corresponding labels from the brain atlas annotation and show that these atlas-labeled niches form connected communities on the graph (see **Figure 2E**). This indicates that niches from the same anatomical region tend to be interconnected, which demonstrates the accuracy of our representation of the spatial distribution of biological zones. Finally, we quantify the concordance between the identified niche labels and the brain atlas annotation using the Normalized Mutual Information (NMI), demonstrating a high degree of correspondence (see **Figure 2F**). By visualizing tissue architecture as a graph, NOLAN provides a powerful and interpretable framework for understanding the spatial relationships between niches.

4.2 BENCHMARKING

We conducted a comprehensive evaluation of our method to demonstrate two key aspects: (1) confirming that the identified niches correspond to biologically meaningful tissue domains, and (2) demonstrating NOLAN's superior ability to recover these structures.

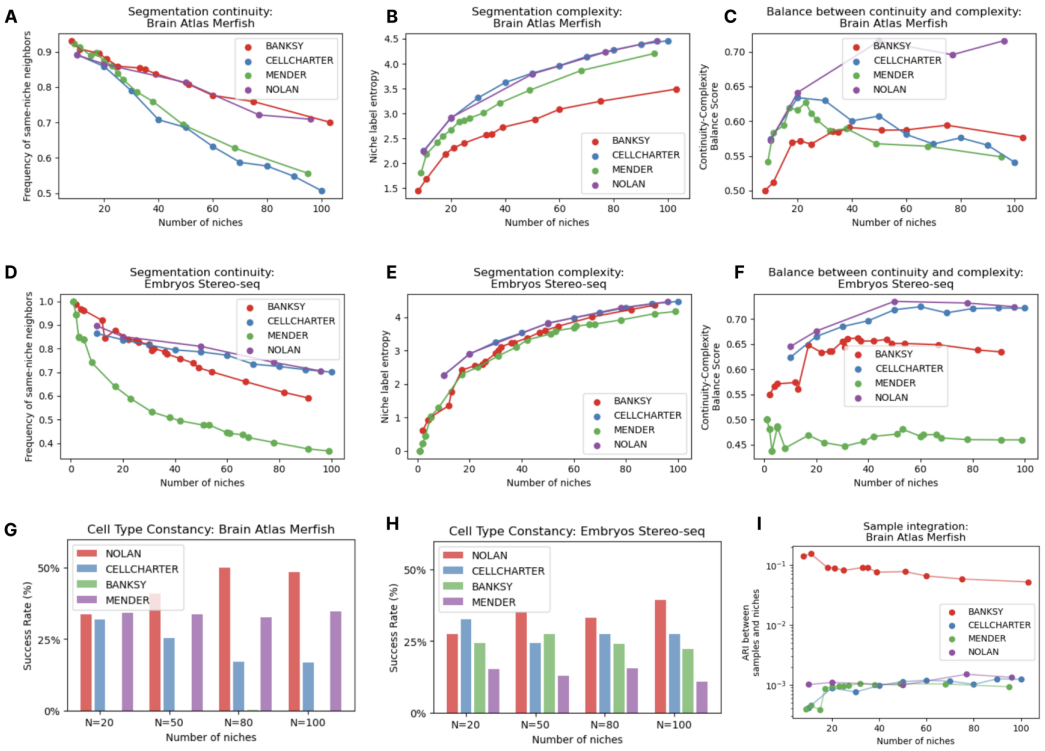

Figure 3: NOLAN outperforms other methods across diverse metrics. A-F. Metrics that measure structural properties of the niches. G-H. metrics that measure constancy of cell-type content. I. Metric that measures quality of batch integration.

We first examined the balance between niche complexity and tissue continuity (**Figure 3A-B**), two contrasting metrics that capture different aspects of tissue organization. The results revealed that NOLAN consistently achieved an optimal balance, ranking as the top or joint-top method across all datasets and evaluation criteria. For example, in cases where NOLAN shared the top rank with BANKSY in terms of tissue continuity (**Figure 3A**), BANKSY exhibited significantly lower performance in terms of tissue complexity (**Figure 3B**), highlighting its limited ability to capture the fine-grained structure of the tissue. Conversely, while CellCharter excelled in tissue complexity (**Figure 3B**), it underperformed in tissue continuity (**Figure 3A**), indicating a trade-off between capturing detailed structure and maintaining spatial coherence. These findings collectively demonstrate NOLAN's ability to achieve a robust representation of tissue organization, simultaneously preserving both spatial coherence and detailed tissue differentiation.

Next, we evaluated the consistency of cell-type composition within the identified niches, providing a direct link between the identified niches and their biological interpretation. To do this, we assessed the consistency of cell type content across the identified niches (**Figure 3 G-H**). The results demonstrated that NOLAN consistently outperformed other methods across both datasets, with the performance gap widening as the number of identified niches increased. The only exception occurred in the embryo dataset, where CellCharter surpassed NOLAN when the number of identified niches was set low (**Figure 3H**). Overall, these results provide strong evidence that NOLAN's identified niches are directly interpretable from a perspective of cell type conservation and highlight NOLAN's superior performance compared to state-of-the-art methods in capturing the spatial organization of cell types.

Finally, we assessed the quality of batch integration across different methods (**Figure 3 I**). Notably, most methods demonstrated high levels of batch integration; only BANKSY, which does not provide batch correction, achieved a worse score. This underscores NOLAN's utility in integrating multiple slides and identifying shared niches effectively.

4.3 Analysis of pan-cancer data

Having demonstrated the robust performance of NOLAN in the benchmarking, we applied our framework to a cohort of cancer biopsies representing diverse tissue origins. Our aim was to move beyond simple tissue segmentation and perform a comparative analysis of tumor organization, uncovering both conserved architectural principles and tissue-specific variations across different cancers.

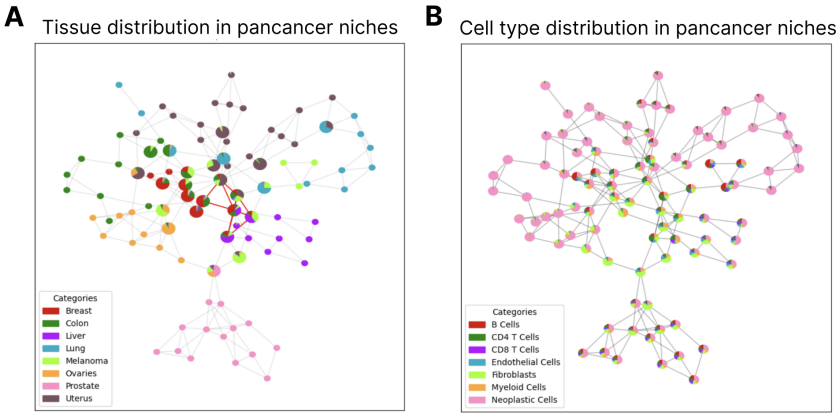

Figure 4: Integrated view of the tissue structure in pan-cancer dataset. A. Tissue graph colored by the abundance of different cancer types. B. Tissue graph colored according to the distribution of cell types.

Our analysis begins with an integrated pan-cancer niche graph, constructed from NOLAN's learned embeddings (see **Figure 4A**). This global view of tumor structure immediately reveals a striking pattern: while, as expected, the majority of identified niches are tissue-specific—reflecting the well-known heterogeneity of cancers—a distinct core network of interconnected niches emerges, shared across multiple tumor types. This observation strongly suggests the existence of conserved organizational features within the tumor microenvironment, despite the diverse origins of the cancers. To understand the biological basis of this shared structure, we examined the cell-type composition of the niches within the pan-cancer graph (**Figure 4B**). This revealed that the tissue-specific niches were predominantly composed of neoplastic cells, which, as expected, exhibit highly divergent transcriptional profiles. And in contrast, the shared niches were enriched in stromal components – fibroblasts, endothelial cells, and immune cells – cell types known to play conserved roles in the tumor microenvironment.

Intrigued by these shared niches, we zoomed in, on a few of them (nodes connected by the red edges in **Figure 4A**). These niches, present in at least four different cancer types (**Figure S4A**), showed subtle yet significant variations in the proportions of CD4 T cells, B cells, and fibroblasts (**Figure**

S4B), suggesting a dynamic interplay between these cell types. The spatial localization of these niches confirmed their location in the peritumoral stroma, the interface between the tumor and the surrounding tissue (**Figure S4C**).

To investigate the dynamics of cellular interactions within the shared tumor microenvironment, we focused on a specific transition between two adjacent niches: niche 54, characterized by a low abundance of CD4 T cells ("CD4-poor"), and niche 64, enriched in these cells ("CD4-rich") (**Figure S5A**). Analyzing the compositional changes along this transition across multiple samples (**Figure S5B**) revealed a marked shift: a decrease in fibroblasts, lymphatic, and arterial endothelial cells, coupled with a concurrent increase in multiple immune cell subtypes, including naive, Th1, and regulatory T cells, as well as memory and plasma B cells. This pattern suggests a functional specialization of these niches: niche 54 likely represents a vascularized region, potentially serving as an entry point for immune cells, while niche 64 constitutes a more immunologically active zone where these cells accumulate and exert their functions.

To further refine this model, we examined molecular changes within key cell types across this same transition. In CD4 T cells, moving from niche 54 to niche 64, we observed a significant upregulation of genes critical for T cell activation and function (TAPBP, CD2, CD27, and TNFRSF13C; (Papadas et al., 2023; Skartsis et al., 2022) (**Figure S5C**), supporting the notion of increased immune activity in niche 64. Concurrently, in endothelial cells, we detected changes suggesting an active role in modulating the immune response (**Figure S5D**). Increased expression of NFKBIA and STAT1 indicated activation of inflammatory pathways, potentially promoting immune cell recruitment, while decreased expression of FLT4 and CCND1 suggested a shift in endothelial function from proliferation towards facilitating immune cell trafficking (Zhang et al., 2011; de Prati et al., 2005; Lee et al., 1996; Pi et al., 2018). This shows that in these regions immune modulation in endothelial cells is associated with the activation of CD4 T cells and is negatively correlated with the remodeling of the vasculature.

Overall, this work highlights the unique capabilities of NOLAN. We have shown that NOLAN can not only identify spatially distinct niches across diverse cancer samples, revealing both shared and cancer-specific organizational features, but also and critically characterize the transitions between these niches.

5 Conclusion

This paper introduces NOLAN, a self-supervised learning framework specifically designed for the analysis of tissue architecture from spatial transcriptomics data. NOLAN's key innovation is its ability to model not only distinct tissue niches but also the continuous transitions between them, providing a more nuanced and biologically relevant representation of tissue organization than methods that rely on discrete partitioning. To effectively capture and visualize this continuous variation, we introduce a tissue graph representation, where nodes represent distinct niches and edges represent the transitional zones connecting them.

We demonstrate the utility of NOLAN through a comprehensive analysis of a pan-cancer dataset, providing an integrated, cross-tumor perspective on tissue composition and organization. This analysis revealed distinct tissue-specific niches, identified a core network of conserved niches shared across multiple tumor types, and allowed for a detailed characterization of changes in their cellular and molecular profiles in transitional zones. This integrated study highlights the potential of NOLAN to serve as a foundational model by identifying generalizable principles of tissue organization.

6 Meaningfulness statement

In this paper, we introduce NOLAN, a model designed to learn representations of tissue by capturing its basic building blocks, the niches, and the transitional regions between them. Using these representations, we construct a graph-based abstraction of the tissue that facilitates detailed analysis of its topology and supports comparisons across multiple samples. This framework allows us to study molecular properties of the niches in detail while also providing a comprehensive view of the overall tissue organization.

REFERENCES

Yoav Benjamini and Yosef Hochberg. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal statistical society: series B (Methodological)*, 57(1):289–300, 1995.

Salil S Bhate, Graham L Barlow, Christian M Schürch, and Garry P Nolan. Tissue schematics map the specialization of immune tissue motifs and their appropriation by tumors. *Cell Systems*, 13 (2):109–130, 2022.

Mathilde Caron, Hugo Touvron, Ishan Misra, Hervé Jégou, Julien Mairal, Piotr Bojanowski, and Armand Joulin. Emerging properties in self-supervised vision transformers. In *Proceedings of the IEEE/CVF international conference on computer vision*, pp. 9650–9660, 2021.

Ao Chen, Sha Liao, Mengnan Cheng, Kailong Ma, Liang Wu, Yiwei Lai, Xiaojie Qiu, Jin Yang, Jiangshan Xu, Shijie Hao, et al. Spatiotemporal transcriptomic atlas of mouse organogenesis using dna nanoball-patterned arrays. *Cell*, 185(10):1777–1792, 2022.

Xinlei Chen, Haoqi Fan, Ross Girshick, and Kaiming He. Improved baselines with momentum contrastive learning. arxiv 2020. *arXiv preprint arXiv:2003.04297*, 2003.

Alessandra C de Prati, Anna R Ciampa, Elisabetta Cavalieri, Raffaela Zaffini, Elena Darra, Marta Menegazzi, Hisanori Suzuki, and Sofia Mariotto. Stat1 as a new molecular target of anti-inflammatory treatment. *Current medicinal chemistry*, 12(16):1819–1828, 2005.

Philipp Dufter, Martin Schmitt, and Hinrich Schütze. Position information in transformers: An overview. *Computational Linguistics*, 48(3):733–763, 2022.

Jean-Bastien Grill, Florian Strub, Florent Altché, Corentin Tallec, Pierre Richemond, Elena Buchatskaya, Carl Doersch, Bernardo Avila Pires, Zhaohan Guo, Mohammad Gheshlaghi Azar, et al. Bootstrap your own latent-a new approach to self-supervised learning. *Advances in neural information processing systems*, 33:21271–21284, 2020.

Jian Hu, Xiangjie Li, Kyle Coleman, Amelia Schroeder, Nan Ma, David J Irwin, Edward B Lee, Russell T Shinohara, and Mingyao Li. Spagcn: Integrating gene expression, spatial location and histology to identify spatial domains and spatially variable genes by graph convolutional network. *Nature methods*, 18(11):1342–1351, 2021.

Daniel C Jones, Anna E Elz, Azadeh Hadadianpour, Heeju Ryu, David R Glass, and Evan W Newell. Cell simulation as cell segmentation. *bioRxiv*, 2024.

James Lee, Alane Gray, Jean Yuan, Shiuh-Ming Luoh, Hava Avraham, and William I Wood. Vascular endothelial growth factor-related protein: a ligand and specific activator of the tyrosine kinase receptor flt4. *Proceedings of the National Academy of Sciences*, 93(5):1988–1992, 1996.

Romain Lopez, Jeffrey Regier, Michael B Cole, Michael I Jordan, and Nir Yosef. Deep generative modeling for single-cell transcriptomics. *Nature methods*, 15(12):1053–1058, 2018.

Maxime Oquab, Timothée Darcet, Théo Moutakanni, Huy Vo, Marc Szafraniec, Vasil Khalidov, Pierre Fernandez, Daniel Haziza, Francisco Massa, Alaaeldin El-Nouby, Mahmoud Assran, Nicolas Ballas, Wojciech Galuba, Russell Howes, Po-Yao Huang, Shang-Wen Li, Ishan Misra, Michael Rabbat, Vasu Sharma, Gabriel Synnaeve, Hu Xu, Hervé Jegou, Julien Mairal, Patrick Labatut, Armand Joulin, and Piotr Bojanowski. Dinov2: Learning robust visual features without supervision, 2024. URL `https://arxiv.org/abs/2304.07193`.

Athanasios Papadas, Yun Huang, Alexander Cicala, Yaling Dou, Matteo Fields, Alicia Gibbons, Duncan Hong, Daniel J Lagal, Victoria Quintana, Alejandro Rizo, et al. Emerging roles for tumor stroma in antigen presentation and anti-cancer immunity. *Biochemical Society Transactions*, 51 (6):2017–2028, 2023.

Jingjiang Pi, Jie Liu, Tao Zhuang, Lin Zhang, Huimin Sun, Xiaoli Chen, Qian Zhao, Yashu Kuang, Sheng Peng, Xiaohui Zhou, et al. Elevated expression of mir302-367 in endothelial cells inhibits developmental angiogenesis via cdc42/ccnd1 mediated signaling pathways. *Theranostics*, 8(6): 1511, 2018.

Vipul Singhal, Nigel Chou, Joseph Lee, Yifei Yue, Jinyue Liu, Wan Kee Chock, Li Lin, Yun-Ching Chang, Erica Mei Ling Teo, Jonathan Aow, et al. Banksy unifies cell typing and tissue domain segmentation for scalable spatial omics data analysis. *Nature Genetics*, 56(3):431–441, 2024.

Nikolaos Skartsis, Leonardo MR Ferreira, and Qizhi Tang. The dichotomous outcomes of tnfα signaling in cd4+ t cells. *Frontiers in Immunology*, 13:1042622, 2022.

Marco Varrone, Daniele Tavernari, Albert Santamaria-Martínez, Logan A Walsh, and Giovanni Ciriello. Cellcharter reveals spatial cell niches associated with tissue remodeling and cell plasticity. *Nature Genetics*, 56(1):74–84, 2024.

Vizgen. MERSCOPE PanCancer Pathways Human Panel, 2025. URL `https://info.vizgen.com/pancancerpathwayspanel`. Accessed: 2025-02-12.

Yun Yang. *Temporal data mining via unsupervised ensemble learning*. Elsevier, 2016.

Zhiyuan Yuan. Mender: fast and scalable tissue structure identification in spatial omics data. *Nature Communications*, 15(1):207, 2024.

Guo-Long Zhang, Yan-Feng Zou, Xiao-Liang Feng, He-Jian Shi, Xu-Feng Du, Min-Hua Shao, Yong Gu, and Qing Zhou. Association of the nfkbia gene polymorphisms with susceptibility to autoimmune and inflammatory diseases: a meta-analysis. *Inflammation research*, 60:11–18, 2011.

Meng Zhang, Xingjie Pan, Won Jung, Aaron R Halpern, Stephen W Eichhorn, Zhiyun Lei, Limor Cohen, Kimberly A Smith, Bosiljka Tasic, Zizhen Yao, et al. Molecularly defined and spatially resolved cell atlas of the whole mouse brain. *Nature*, 624(7991):343–354, 2023.

Jinghao Zhou, Chen Wei, Huiyu Wang, Wei Shen, Cihang Xie, Alan Yuille, and Tao Kong. ibot: Image bert pre-training with online tokenizer. *arXiv preprint arXiv:2111.07832*, 2021.

A SUPPLEMENTARY

A.1 SUPPLEMENTARY METHODS

A.1.1 CONSTRUCTING TISSUE GRAPH

During inference, we process every cell *c* in the dataset *C*. For each cell, we extract a representation of its neighborhood from the teacher network. In contrast to common practice in computer vision, where backbone features are often used, we leverage the representations from the network's projection head. We found that the "pseudolabels" (cluster assignments) generated by this head closely reflect known biological zonation patterns in tissues. Furthermore, our implementation uses a relatively low-dimensional projection head, where the output dimensionality directly corresponds to the number of identified niches *N*. This makes the head representations more interpretable and directly relevant for downstream analysis compared to the higher-dimensional backbone features. The embedding matrix *E*, contains these representations, where $\mathrm{E}_{c,n}$ represents the activation of cell c for niche n. To assign a primary niche label $a(c)$ to each cell, we perform an argmax operation across the niche dimensions of the classification token:

$$a(c) \;=\; \underset{n \in N}{\arg\max}\, E_{c,n}. \tag{5}$$

Notably, cells within transitional regions tend to exhibit high activation values across multiple embedding dimensions, reflecting their intermediate state between niches. To identify these transitional cells and infer connectivity between niches i and j, we first calculate the difference, d_c, between the activation of cell c for niche j and its activation for niche i:

$$d_c = E_{c,j} - E_{c,i} \tag{6}$$

We consider all (i, j) pairs s.t. $i < j$.

We then determine the minimum (d_{min}) and maximum (d_{max}) values of this difference across all cells:

$$d_{\min} = \min_{c \in C} d_c, \quad d_{\max} = \max_{c \in C} d_c. \tag{7}$$

To define a threshold for identifying transitional cells, we use a parameter α, such that $0 < \alpha < 0.5$ (smaller values would results in the assignment of more cells to the transitionary regions), to calculate lower and upper bounds (d_{left} and d_{right}) on this difference:

$$d_{\text{left}} = d_{\min} + (d_{\max} - d_{\min}) \times \alpha, \tag{8}$$

$$d_{\text{right}} = d_{\min} + (d_{\max} - d_{\min}) \times (1 - \alpha). \tag{9}$$

Finally, we define the set of cells cells2edges(i, j), that constitute the transitional region (edge) between niches i and j as those cells whose primary assigned niche is either i or j, and whose activation difference d_c falls between d_{left} and d_{right}:

$$\text{cells2edges}(i, j) = \{c \in C \mid a(c) \in \{i, j\},\ d_{\text{left}} \leq d_c \leq d_{\text{right}}\}. \tag{10}$$

This process yields a tissue graph representation: where $a(c)$ nodes represent distinct niches, and edges, defined by the cells2edges sets, represent the transitional regions connecting them.

A.1.2 Datasets used in the benchmarking

Since various spatial technologies have different limitations, we made sure to include in our benchmarking datasets from two principally different experimental methods: MERFISH and Stereo-Seq. The MERFISH dataset that we selected was murine brain atlas (Zhang et al., 2023), which we subsampled to five mid-coronal sections. The sequencing-based dataset that we included in out analysis consisted of one sagittal slice of a mouse embryo binned with cell-level resolution Chen et al. (2022).

A.1.3 Details on running the algorithms for benchmarking

To ensure a robust comparison, we selected three state-of-the-art programs recently published in Q1 journals as benchmarks: MENDER, BANKSY, and CellCharter. Among these, MENDER relies on prior cell-type labels, while CellCharter requires scVI embeddings as input. Similarly, NOLAN requires embeddings from scVI or a comparable method. Therefore, we first ran scVI on the input data to generate embeddings, which were subsequently used as input for both NOLAN and CellCharter. For MENDER, we further performed clustering of the cell types based on these scVI embeddings and provided these clusters as input to ensure a consistent and fair comparison.

Recognizing that all methods in our analysis are based on clustering, we ensured a fair comparison by evaluating each method across multiple resolution levels.

A.1.4 Pancancer dataset

As a test case of the method utility for the analysis of large-scale data, we collected a variety of openly available Vizgen MERSCOPE cancer datasets provided by (Vizgen, 2025), performed resegmentation using ProSeg (Jones et al., 2024) of this data yielding a high-quality cross-tissue spatial transcriptomics dataset with 12 million cells. The dataset consists of various tumors: melanoma, uterine cancer, prostate cancer, breast cancer, colon cancer, lung cancer, and liver cancer. Except for breast cancer, all other cancers had two replicates. For this data we performed 2 level hierarchical cell typing based on manually curated marker gene signatures.

A.2 Suppelementary figures

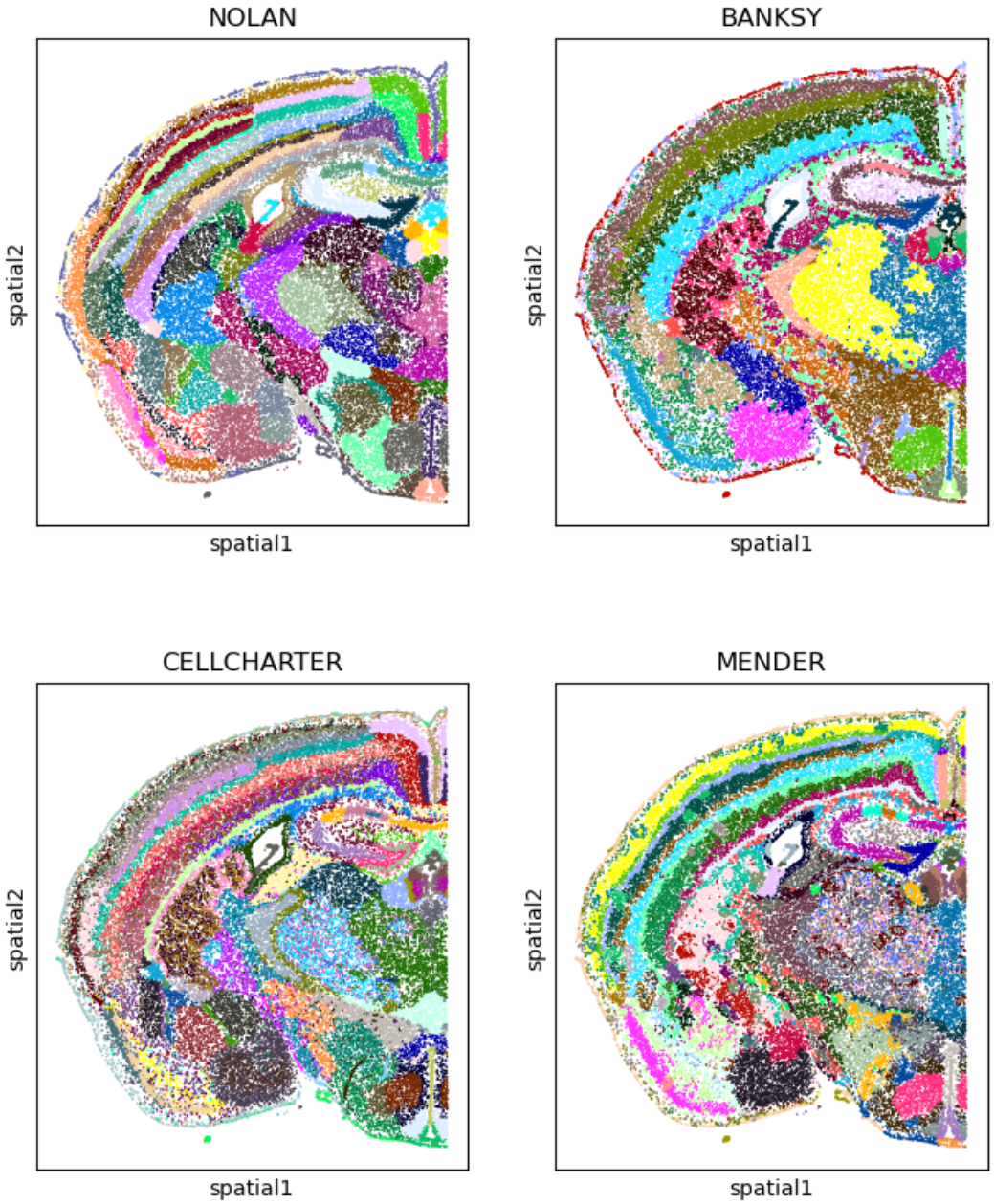

Figure S1: Brain atlas niche annotation with various methods

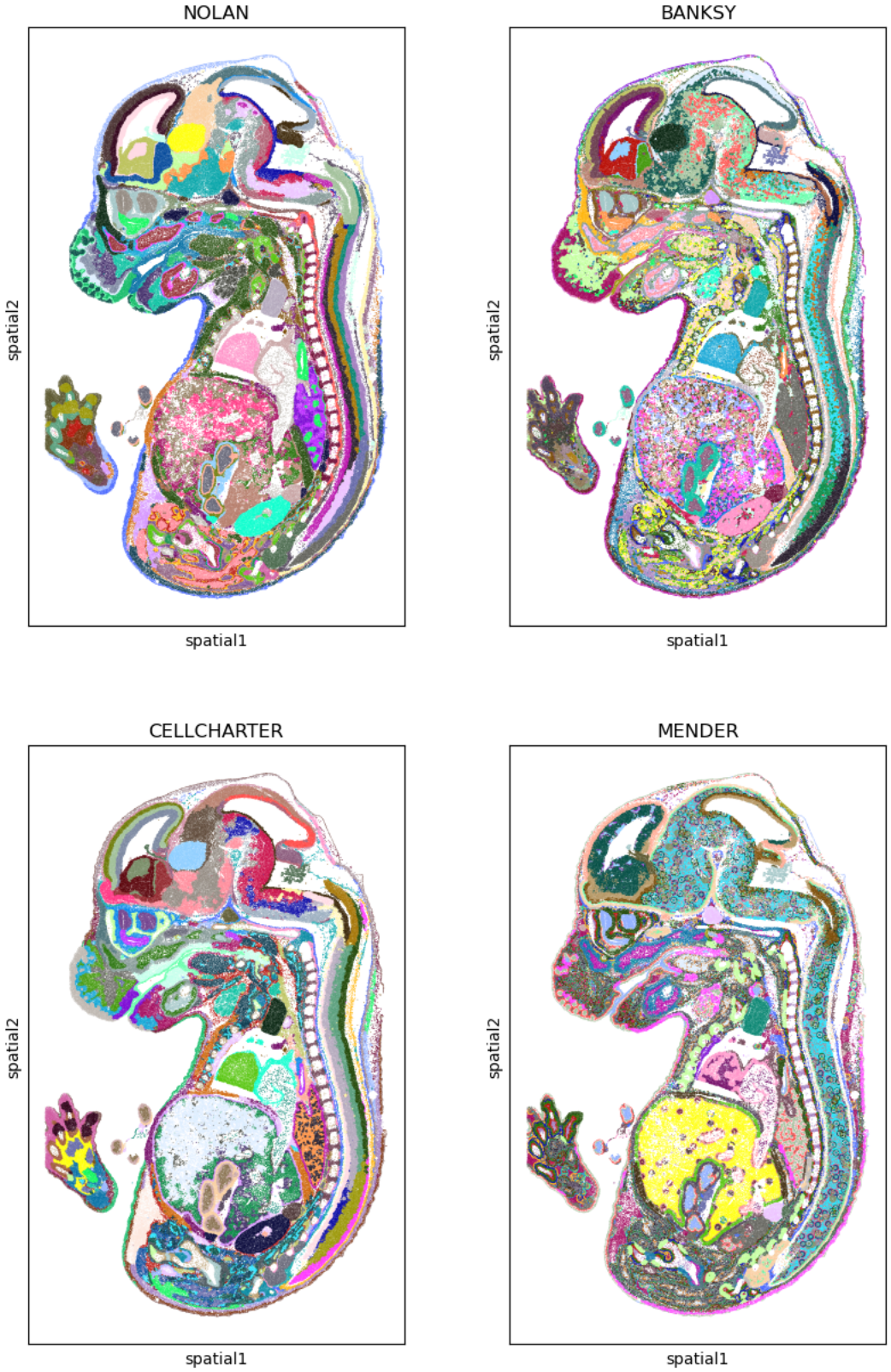

Figure S2: Embryo niche annotation with various methods

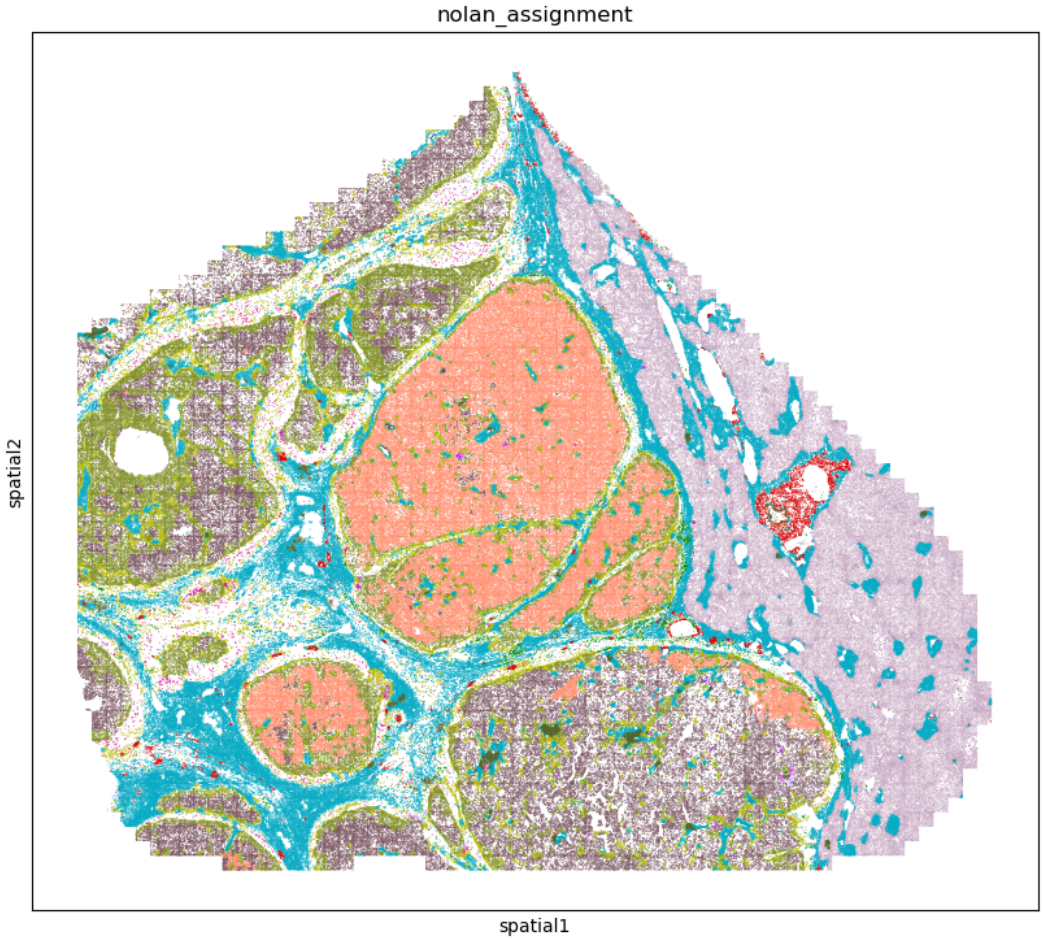

Figure S3: Liver cancer niche annotation

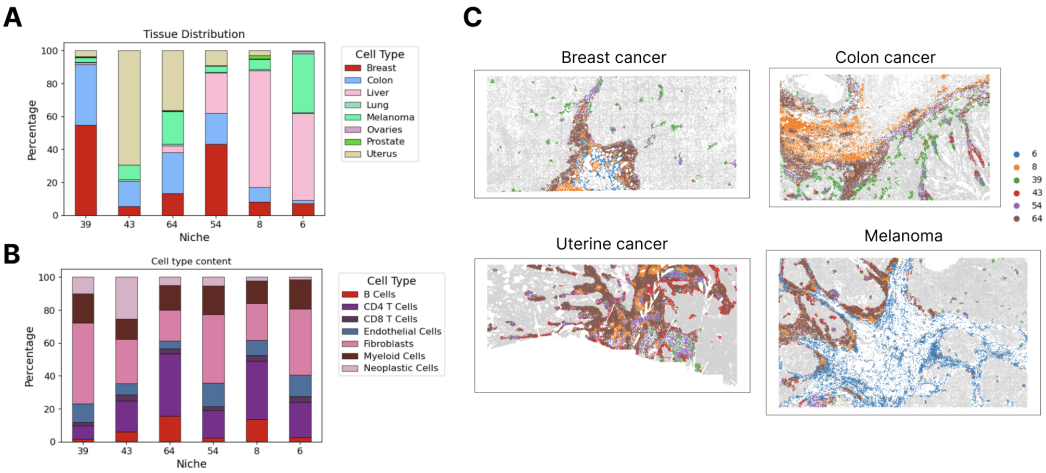

Figure S4: Properties of the subnetwork shared by tumors. A. Tissue distribution in detected niches. B. Cell-type distribution. C. Spatial visualization of the tumor network.

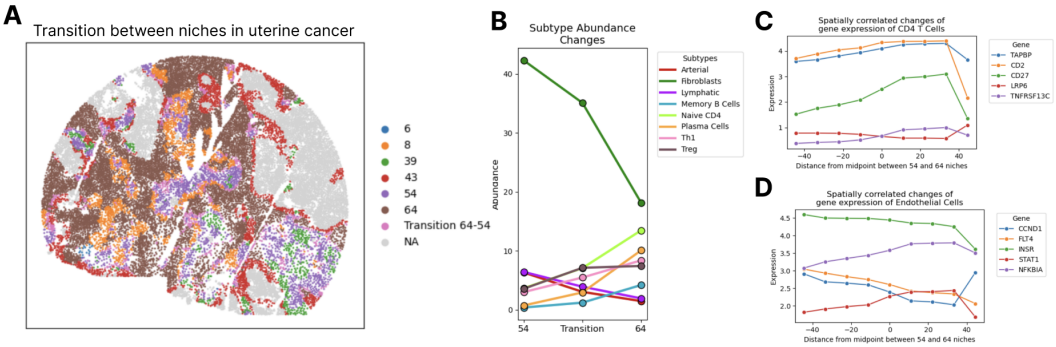

Figure S5: Transitional regions identified by NOLAN. A. Visualization of transition between niches 54 and 64 in uterine cancer. B. Changes in subtype content associated with the transition between niches. C. Changes in gene expression associated with transition between niches in CD4 T cells. D. Changes in gene expression associated with transition between niches in endothelial cells.