Reasoning Across Space: Tiny Recursive Models for Spatial Omics

Anonymous Authors[1]

Abstract

Spatial omics has enabled the study of gene expression within intact tissue, but existing models are limited by local processing, scarce annotations, and poor generalization across tissue sections. Current approaches perform spatial reasoning in a single pass over local regions, restricting the integration of long-range dependencies that underlie many biological signals. In this work, we investigate whether recursive computation provides a natural inductive bias for spatial omics, enabling global reasoning over entire tissue samples without test-time training. We apply Tiny Recursive Models (TRM) for spatial omics tasks and show that it consistently outperforms parameter-matched supervised baselines across robust validation splits. TRM also surpasses several slice-specific unsupervised methods in domain identification and achieves state-of-the-art performance among supervised methods in gene expression prediction. Further analysis reveals that recursion enables test-time scaling and adaptive computation, and that the model is partially generalizable across tissue types for domain identification. Overall, our results demonstrate that recursive computation provides a scalable, data-efficient framework for spatial omics, overcoming the locality constraints of existing approaches without sacrificing generalization.

1. Introduction

Spatial omics has transformed biology, enabling the study of gene expression directly within intact tissue and allowing the elucidation of molecular programs with spatial resolution (Ståhl et al., 2016; Chen et al., 2020). Spatial omics presents a challenging learning regime. First, annotations are scarce and expensive, making fully supervised learning difficult at scale. Second, models must explicitly capture spatial structure and dependencies across cells and tissue regions, rather than treating observations as independent samples.

This has motivated a range of models that operate on spatial graphs or localized regions of tissue. For example, spatial domain identification methods typically train individual graph-based models per tissue section, propagating information through local neighborhoods (Dai et al., 2026; Long et al., 2023; Dong & Zhang, 2022; Hu et al., 2021). Similarly, recent spatial gene expression prediction methods often partition data into independent patches and integrate spatial neighborhoods and histological features (Jia et al., 2024). While effective, most of these approaches remain fundamentally local: spatial reasoning is effectively performed in a single pass over local neighborhoods, without global iterative refinement. This restricts the ability to integrate long-range tissue interactions, that are essential for many biological mechanisms. As a result, most methods often rely on fitting directly to the test sample itself, struggling to generalize across tissue slices or tissue types. Ultimately, current models face a trade-off between scalability and the inductive bias required for global spatial reasoning.

Recent work has shown that recurrent forward passes provide an alternative axis of scaling deep neural networks by increasing the effective depth of the model without increasing parameter count (Geiping et al., 2025; Goyal et al., 2026): rather than learning deeper networks, the same block is repeatedly applied to iteratively refine intermediate representations. Tiny Recursive Models (TRM) (Jolicoeur-Martineau, 2025) additionally demonstrate that such depth recurrence across multiple levels offers additional benefits in terms of reasoning ability in data-scarce regimes, where a small shared network repeatedly refines its own predictions through iterative self-correction.

In this work, we investigate whether recursive computation provides a natural inductive bias for spatial omics. We focus on two representative tasks: spatial-domain identification from gene expression and histology-to-gene-expression prediction. We examine whether TRMs can reason jointly over entire spatial samples using a single shared recurrent model, rather than operating on local neighborhoods or independent patches. We evaluate both tasks under a strict leave-one-slice-out (LOSO) setting, where models must generalize to unseen tissue sections and patient biology without test-time training. Recursive reasoning substantially improves robustness in this regime, outperforming parameter-matched supervised baselines across all LOSO folds and exceeding most established per-slice unsupervised methods in domain identification and all supervised methods in gene expression prediction. We further analyze how recursion affects

performance through test-time scaling and adaptive computation (ACT), show that the reasoning latent space learned by TRM encourages generalization across tissue types in domain identification, and propose practical ACT strategies for continuous prediction tasks.

2. Method

Tasks: We consider two representative spatial transcriptomics (ST) tasks that are both widely used benchmarks, making them suitable for direct comparison with prior work:

1. **Spatial Domain Identification (SDI) task:** Given spot-level gene expression measurements from a 2D Visium section, cluster spots into biologically meaningful spatial domains.
2. **Histology-to-Gene expression prediction (H2G) task:** Given an H&E-stained histology image, predict the spatially resolved gene expression profile at each spot conditioned on paired H&E and ST measurements.

Both tasks provide a natural testbed for evaluating recursive models, as they require learning global spatial dependencies from high-dimensional, sample-limited data.

Architecture: We use the Tiny Recursive Model (TRM) as our core architecture, as it is designed to improve reasoning in data-scarce settings through iterative refinement rather than increased parameter count. Instead of producing the final answer in one forward pass, the TRM recursively refines its predictions using an auxiliary latent reasoning state. Intuitively, the latent state acts as an internal scratchpad that accumulates task-relevant information, whereas the prediction state represents the model's current best answer– for our tasks, spatial domain assignments (SDI task) or gene expression matrices (H2G task). Starting from randomly initialized noise, the same network updates both states, gradually transforming an initially uncertain prediction into a more accurate output. Across recursion cycles, the same shared network is reused, allowing the model to increase effective depth without increasing the number of parameters. This provides a strong inductive bias for our setting, where learning robust global structure is more important than scaling model size alone. The TRM's backbone is a 2-layer transformer with 896K trainable parameters, N_{Z_L}=6 latent steps, N_{Z_H}=3 deep steps, and N_{outer}=8 outer steps. In a single forward step, our TRM thus makes $(N_{Z_L}+1)\cdot N_{Z_H}\cdot 2 = 42$ forward passes. The maximum TRM rollout is $42 \cdot N_{outer} = 336$. Exact pseudocode is provided in Appendix A (Figure 5).

Adaptive computation time (ACT) strategy:

Classical ACT (Graves, 2016) was designed for tasks with a clear "solved/not-solved" signal (e.g., digit addition). Spatial gene-expression prediction is a continuous task: the output is a high-dimensional regression target, and quality improves *continuously* with depth. We study four halt recipes on TRM using (N_{z_H}=1, N_{z_L}=3, $K_{\max}$=8). Three are slide-level halt-target recipes; **PCC-halt**, **RVD-halt**, and **MSE-halt** that supervise the halt head against the panel-wide PCC / RVD / MSE on a 20% per-step held-out subset of the slide's supervised spot features. The fourth, **per-class PCC-halt**, supervises a separate halt logit per gene with the halt-prob threshold calibrated to sit *inside* the per-input distribution of converged halt probabilities.

3. Experiments and Results

3.1. Baselines

We compare TRM with two supervised baselines. Our first baseline, **No Recursion 2-Layer**, is a 2-layer non-recursive transformer, equivalent to setting N_{Z_L}=1, N_{Z_H}=1, and N_{outer}=1. This controls for per-step model capacity while removing iterative refinement entirely. Our second baseline is a **Universal Transformer** (Dehghani et al., 2019) without ACT: a 2-layer shared-weight transformer repeated 21 times, thus bringing its effective depth to 42 (equivalent to the effective depth of a single TRM forward pass). This provides a depth- and weight-sharing-matched control for repeated computation, allowing us to test whether TRM's gains stem from its structured separation of latent-state refinement and solution refinement, rather than from recurrence depth alone. A fully depth-matched baseline, that is, training a model with depth equivalent to the maximum TRM rollout, consistently overfitted and failed to generalize under the same training data, and hence we omit it from the comparison.

For both tasks, we additionally compare against several established state-of-the-art per-slice methods: for the SDI task, SpaGCN (Hu et al., 2021), STAGATE (Dong & Zhang, 2022), GraphST (Long et al., 2023), and 3d-OT (Dai et al., 2026), whereas for the H2G task, STEM (Zhu et al., 2025), BLEEP (Xie et al., 2023), TRIPLEX (Chung et al., 2024), and HisToGene (Pang et al., 2021).

3.2. Spatial Domain Identification

We use the dorsolateral prefrontal cortex (DLPFC) dataset (Maynard et al., 2021), which includes 12 human DLPFC Visium sections with expert annotations for seven cortical layers. We map each sample to a 2D grid and featurize gene expression with PCA over 3,000 highly variable genes (HVGs). We adopt a strict **leave-one-slice-out (LOSO)** protocol: each fold trains on 11 slices and evaluates zero-shot on the held-out 12th slice. Importantly, the 12th sample is excluded from PCA fitting and is projected into the training PCA space only at test time. Full implementation details are provided in Appendix B.

Comparison to baselines. We evaluate all methods using

Table 1. Per-slice ARI for all 12 LOSO SDI folds. TRM outperforms both baselines on all 12 slices. †Published per-slice results from Dai et al. (2026) source data (unsupervised methods that train on the test slice).

	151507	151508	151509	151510	151669	151670	151671	151672	151673	151674	151675	151676	Mean
No Rec. 2-Layer (LOSO)	.412	.366	.466	.337	.566	.580	.578	.546	.447	.406	.419	.413	.461
Univ. Transformer (LOSO)	.462	.407	.458	.408	.636	.612	.541	.547	.437	.442	.410	.439	.483
TRM (LOSO)	.519	.492	.555	.525	**.695**	**.736**	.665	.591	.520	.520	.466	.426	.559
SpaGCN†	.414	.369	.354	.336	.423	.406	.380	.363	.340	.352	.292	.293	.360
STAGATE†	.485	.471	.518	.483	.570	.535	.573	.510	.478	.483	.404	.376	.491
GraphST†	.504	.527	.589	.547	.621	.588	.571	.547	.547	.480	.417	.440	.532
3d-OT†	**.643**	**.639**	**.685**	**.622**	.684	.633	**.849**	**.798**	**.703**	**.722**	**.646**	**.631**	**.688**

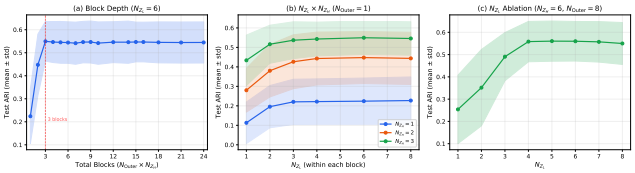

Figure 1. DLPFC test-time recursion ablation (mean ± std, 12 SDI LOSO folds). ARI increases with total processing blocks, saturating around 3. Latent recursion steps further improve performance.

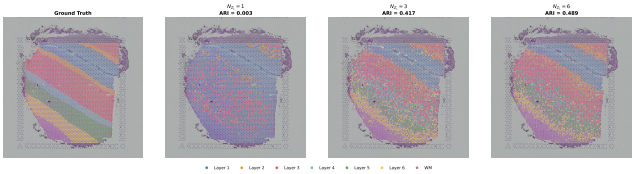

Figure 2. H&E overlay of ground truth and TRM predictions with 1, 3, and 6 latent recursion steps (slice 151508, LOSO). Increasing the number of latent recursion steps sharpens cortical layer boundaries.

the Adjusted Rand Index (ARI) between predicted cluster assignments and the ground-truth cortical layer annotations. Table 1 shows that TRM substantially outperforms both supervised baselines on all test slices, demonstrating that recursive refinement provides benefits beyond increased effective depth alone. Compared to established, per-slice methods, TRM achieves a mean LOSO ARI of 0.559, outperforming SpaGCN (0.360) and STAGATE (0.491), and GraphST (0.532), despite *never observing the test slice during training*. We note that this comparison is not strictly apples-to-apples: these methods are unsupervised and are fit explicitly on the test slice itself, while the TRM is trained on other slices. Nevertheless, the fact that a cross-slice model matches or surpasses per-slice methods underscores its generalization. Interestingly, the only method that substantially outperforms TRM-LOSO is 3d-OT (0.688), which combines a PointNet++ backbone with an optimal transport alignment step. 3d-OT encodes spatial structure through a fixed hierarchical decomposition, recursively over nested spatial partitions to build multi-scale representations. This provides a strong built-in inductive bias for spatial data, further strengthened by explicit access to the test slice through optimal transport alignment. Conversely, the TRM operates under a stricter setting where is does not impose an explicit spatial hierarchy, but rather learns to construct it dynamically through repeatedly updating its predictions via recursive computation.

Do recursions help performance? To isolate the role of the three abstractions of recursion, we perform inference-time ablations while keeping the model frozen. We observe that the model needs ∼3 total processing blocks to reach peak ARI, where each block comprises multiple latent recursion steps followed by one deep recursion update (Figure 1(a)). These blocks can arise from any combination of deep supervision steps and deep recursion cycles–the two axes are interchangeable at test time because the outer step boundary performs only read-only operations (Appendix C). Next, we compute the ARI vs N_{Z_L} for different values of N_{Z_H}. At $N_{Z_H} = 1$, improvement is capped at 3 N_{Z_L} steps, indicating that further deep recursion is necessary to improve latent recursion performance, which is validated by improvement for more N_{Z_L} steps at $N_{Z_H} \in \{2, 3\}$. Furthermore, at full N_{outer} and N_{Z_H} steps, we see significant improvement when increasing N_{Z_L} (Figure 1(c)), reducing the number of latent recursions from 6 to 1 drops mean ARI by up to 0.25. Figure 2 qualitatively visualizes the combined effect of recursive refinement on one example slice: with only 1 latent recursion step, the model produces near-random predictions (ARI $= 0.003$); at 3 steps it recovers the gross laminar structure (ARI $= 0.417$); and at 6 steps it achieves the sharpest layer boundaries (ARI $= 0.489$). Further, to test the generalization ability of TRM, we test zero-shot transfer to a human breast cancer Visium dataset with 20 annotated classes. TRM achieves 0.38 ARI, exceeding a few unsupervised baselines, and ARI increases to 0.52 after adding spatial context to its predictions (Appendix D).

3.3. Spatial gene expression prediction

We evaluate on the HER2ST dataset (Andersson et al., 2021), which contains 36 samples with spot-level gene expression

Table 2. HER2ST B1 holdout under the Zhu et al. (2025) protocol (300 HMHVG, log-transformed counts). †Published source data from Zhu et al. (2025).

	PCC-10	PCC-300	RVD
No Rec. 2-Layer	0.865	0.762	0.257
Univ. Transformer	0.830	0.745	1.088
TRM (N_{outer}=8)	**0.883**	**0.785**	0.354
HisToGene†	0.6812	0.5250	10.34
BLEEP†	0.7727	0.5652	0.6025
TRIPLEX†	0.7907	0.5766	0.6428
STEM†	0.8298	0.5984	**0.0693**

data and paired H&E images. We map each sample to a 2D spot grid and featurize each spot with a frozen Gigapath (Xu et al., 2024) embedding (1536-d) extracted from a 224×224 H&E patch centered on the spot; per-sample grids are padded to a square 35×35 union shape and accompanied by a binary tissue mask. Following (Zhu et al., 2025), targets are log-transformed values of the top 300 high mean and high variance (HMHVG) genes selected on the union of train-slide highly-variable genes per the protocol of Zhu et al. (2025). All models are trained under a single-slide-holdout protocol, matching Zhu et al. (2025): B1 is the test slide, B6 is the validation slide for checkpoint selection, and the remaining 34 slides form the training set. We also repeat the recursion-performance analysis for this task in Appendix E.

Comparison to baselines. We evaluate with top-k mean Pearson correlation in log space (PCC-k) and relative variation distance (RVD) (Table 2). We see that TRM (fixed-K ACT) outperforms all methods on PCC across the whole panel. STEM has a diffusion objective and leads in RVD, but for regression (our task) we find that this RVD gap can be mitigated by using ACT-halting (Table 6). More detailed metrics are provided in Appendix B.

Alignment with expert annotations. Following the downstream evaluation protocol used by STEM (Zhu et al., 2025), we also assess whether the predicted expression profiles preserve biologically meaningful spatial structure. We cluster TRM predictions and compare the resulting clusters with the expert annotations of the tissue regions. Figure 3 shows that the predicted expression from TRM yields clusters that align well with the major annotated regions.

4. ACT in the Non-Discrete Setting

We showcase the best-performing ACT recipe on the H2G task: **per-class PCC-halt**: A halt head emits one logit per gene, supervised against the per-gene PCC computed on a held-out subset of the slide's spot FM features (the remainder of the spots supervise the regression loss). The held-out

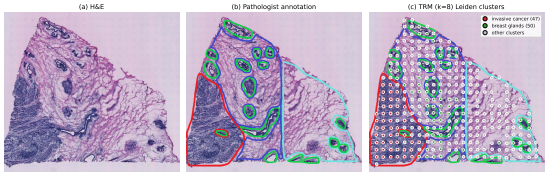

Figure 3. (a) The H&E held out slide, (b) Manual pathologist annotations, (c) TRM predicted gene panel clusters align well with expert annotations

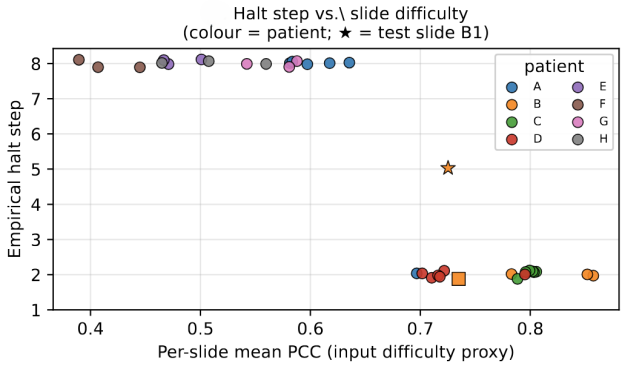

Figure 4. HER2ST: halt-step distribution across slides under per-class PCC-halt.

subset is re-sampled independently every training iteration, so over an epoch each spot contributes to both objectives but never within the same forward pass. Figure 4 shows the halt-step distribution across HER2ST slides (B1 is the test slide). The distribution is bimodal at $\{2, 8\}$, and halt step is inversely correlated with slide-level PCC, indicating that the model recurses more on harder inputs. This makes per-gene PCC an informative supervision signal for the halt head. On the test slide, the model halts at step 5, between the two training-set modes, providing evidence that the per-gene PCC strategy learns an input-adaptive halt policy. By contrast, the three slide-level supervision strategies (MSE-, RVD-, and PCC-halt) collapse to the minimum or maximum allowed ACT depth. Even so, these slide-level halt losses provide variance stabilization, with MSE-halt reaching an RVD of 0.109 ($\sim 3.2\times$ lower than fixed-depth TRM). A detailed analysis of all methods including studying the ACT behaviour for per-gene PCC across other test folds is provided in Appendix F.2.

5. Conclusion

In this work, we show that spatial reasoning improves performance and generalization on spatial omics tasks, and delineate the role of recursion and test-time computation. Our results confirm that recursive reasoning holds great potential for spatial omics tasks, pointing to recursive architectures as a promising direction for scalable, generalizable spatial modeling across tissues and modalities.

References

Albergaria, A., Paredes, J., Sousa, B., Milanezi, F., Carneiro, V., Bastos, J., Costa, S., Vieira, D., Lopes, N., Lam, E. W., Lunet, N., and Schmitt, F. Expression of FOXA1 and GATA-3 in breast cancer: the prognostic significance in hormone receptor-negative tumours. *Breast Cancer Research*, 11(3):R40, 2009. doi: 10.1186/bcr2327.

Andersson, A., Larsson, L., Stenbeck, L., Salmén, F., Ehinger, A., Wu, S. Z., Al-Eryani, G., Roden, D., Swarbrick, A., Borg, Å., et al. Spatial deconvolution of her2-positive breast cancer delineates tumor-associated cell type interactions. *Nature communications*, 12(1):6012, 2021.

Bassi, C., Fortin, J., Snow, B. E., Wakeham, A., Ho, J., Haight, J., You-Ten, A., Cianci, E., Buckler, L., Gorrini, C., Stambolic, V., and Mak, T. W. The PTEN and ATM axis controls the G1/S cell cycle checkpoint and tumorigenesis in HER2-positive breast cancer. *Cell Death & Differentiation*, 28(11):3036–3051, 2021. doi: 10.1038/s41418-021-00799-8.

Chen, W.-T., Lu, A., Craessaerts, K., Pavie, B., Frigerio, C. S., Corthout, N., Qian, X., Laláková, J., Kühnemund, M., Voytyuk, I., et al. Spatial transcriptomics and in situ sequencing to study alzheimer's disease. *Cell*, 182(4): 976–991, 2020.

Chung, Y., Ha, J. H., Im, K. C., and Lee, J. S. Accurate spatial gene expression prediction by integrating multi-resolution features. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 11591–11600, 2024.

Dai, B., Yi, L., Wang, P., Li, H., Hu, P., Song, Y., Xing, J., Feng, Z., Yuan, Z., and Zuo, Y. 3d-ot: a deep geometry-aware framework for heterogeneous slices alignment of spatial multi-omics. *Nature Methods*, pp. 1–12, 2026.

Dehghani, M., Gouws, S., Vinyals, O., Uszkoreit, J., and Kaiser, Ł. Universal transformers. *International Conference on Learning Representations*, 2019.

Dong, K. and Zhang, S. Deciphering spatial domains from spatially resolved transcriptomics with an adaptive graph attention auto-encoder. *Nature communications*, 13(1): 1739, 2022.

Geiping, J., McLeish, S., Jain, N., Kirchenbauer, J., Singh, S., Bartoldson, B. R., Kailkhura, B., Bhatele, A., and Goldstein, T. Scaling up test-time compute with latent reasoning: A recurrent depth approach. *arXiv preprint arXiv:2502.05171*, 2025.

Goyal, S., Agrawal, S., Anil, G. G., Jain, P., Paul, S., and Kusupati, A. Elt: Elastic looped transformers for visual generation. *arXiv preprint arXiv:2604.09168*, 2026.

Graves, A. Adaptive computation time for recurrent neural networks. *arXiv preprint arXiv:1603.08983*, 2016.

Hu, J., Li, X., Coleman, K., Schroeder, A., Ma, N., Irwin, D. J., Lee, E. B., Shinohara, R. T., and Li, M. 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.

Jia, Y., Liu, J., Chen, L., Zhao, T., and Wang, Y. Thitogene: a deep learning method for predicting spatial transcriptomics from histological images. *Briefings in Bioinformatics*, 25(1):bbad464, 2024.

Jolicoeur-Martineau, A. Less is more: Recursive reasoning with tiny networks, 2025. URL https://arxiv.org/abs/2510.04871.

Long, Y., Ang, K. S., Li, M., Chong, K. L. K., Sethi, R., Zhong, C., Xu, H., Ong, Z., Sachaphibulkij, K., Chen, A., et al. Spatially informed clustering, integration, and deconvolution of spatial transcriptomics with graphst. *Nature communications*, 14(1):1155, 2023.

Maynard, K. R., Collado-Torres, L., Weber, L. M., Uytingco, C., Barry, B. K., Williams, S. R., Catallini, J. L., Tran, M. N., Besich, Z., Tippani, M., et al. Transcriptome-scale spatial gene expression in the human dorsolateral prefrontal cortex. *Nature neuroscience*, 24(3):425–436, 2021.

Nielsen, T. O., Leung, S. C. Y., Rimm, D. L., Dodson, A., Acs, B., Badve, S., Denkert, C., Ellis, M. J., Fineberg, S., Flowers, M., Kreipe, H. H., Laenkholm, A.-V., Pan, H., Penault-Llorca, F. M., Polley, M.-Y., Salgado, R., Smith, I. E., Sugie, T., Bartlett, J. M. S., McShane, L. M., Dowsett, M., and Hayes, D. F. Assessment of Ki67 in breast cancer: Updated recommendations from the international Ki67 in breast cancer working group. *Journal of the National Cancer Institute*, 113(7):808–819, 2021. doi: 10.1093/jnci/djaa201.

Nieto, M. A., Huang, R. Y.-J., Jackson, R. A., and Thiery, J. P. EMT: 2016. *Cell*, 166(1):21–45, 2016. doi: 10.1016/j.cell.2016.06.028.

Pang, M., Su, K., and Li, M. Leveraging information in spatial transcriptomics to predict super-resolution gene expression from histology images in tumors. *BioRxiv*, pp. 2021–11, 2021.

Perou, C. M., Sørlie, T., Eisen, M. B., van de Rijn, M., Jeffrey, S. S., Rees, C. A., Pollack, J. R., Ross, D. T., Johnsen, H., Akslen, L. A., Fluge, Ø., Pergamenschikov, A., Williams, C., Zhu, S. X., Lønning, P. E., Børresen-Dale, A.-L., Brown, P. O., and Botstein, D. Molecular

portraits of human breast tumours. *Nature*, 406(6797): 747–752, 2000. doi: 10.1038/35021093.

Sahlberg, K. K., Hongisto, V., Edgren, H., Mäkelä, R., Hellström, K., Due, E. U., Vollan, H. K. M., Sahlberg, N., Wolf, M., Børresen-Dale, A.-L., Perälä, M., and Kallioniemi, O. The HER2 amplicon includes several genes required for the growth and survival of HER2 positive breast cancer cells. *Molecular Oncology*, 7(3):392–401, 2013. doi: 10.1016/j.molonc.2012.10.012.

Shiloh, Y. ATM and related protein kinases: safeguarding genome integrity. *Nature Reviews Cancer*, 3(3):155–168, 2003. doi: 10.1038/nrc1011.

Sørlie, T., Perou, C. M., Tibshirani, R., Aas, T., Geisler, S., Johnsen, H., Hastie, T., Eisen, M. B., van de Rijn, M., Jeffrey, S. S., Thorsen, T., Quist, H., Matese, J. C., Brown, P. O., Botstein, D., Lønning, P. E., and Børresen-Dale, A.-L. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proceedings of the National Academy of Sciences*, 98(19): 10869–10874, 2001. doi: 10.1073/pnas.191367098.

Ståhl, P. L., Salmén, F., Vickovic, S., Lundmark, A., Navarro, J. F., Magnusson, J., Giacomello, S., Asp, M., Westholm, J. O., Huss, M., et al. Visualization and analysis of gene expression in tissue sections by spatial transcriptomics. *Science*, 353(6294):78–82, 2016.

Stein, D., Wu, J., Fuqua, S. A., Roonprapunt, C., Yajnik, V., D'Eustachio, P., Moskow, J. J., Buchberg, A. M., Osborne, C. K., and Margolis, B. The SH2 domain protein GRB-7 is co-amplified, overexpressed and in a tight complex with HER2 in breast cancer. *The EMBO Journal*, 13(6):1331–1340, 1994. doi: 10.1002/j.1460-2075.1994.tb06386.x.

Theodorou, V., Stark, R., Menon, S., and Carroll, J. S. GATA3 acts upstream of FOXA1 in mediating ESR1 binding by shaping enhancer accessibility. *Genome Research*, 23(1):12–22, 2013. doi: 10.1101/gr.139469.112.

Tobin, N. P. and Bergh, J. Analysis of cyclin D1 in breast cancer: A call to arms. *Current Breast Cancer Reports*, 4 (3):171–173, 2012. doi: 10.1007/s12609-012-0083-7.

Xie, R., Pang, K., Chung, S., Perciani, C., MacParland, S., Wang, B., and Bader, G. Spatially resolved gene expression prediction from histology images via bi-modal contrastive learning. *Advances in Neural Information Processing Systems*, 36:70626–70637, 2023.

Xu, H., Usuyama, N., Bagga, J., Zhang, S., Rao, R., Naumann, T., Wong, C., Gero, Z., González, J., Gu, Y., et al. A whole-slide foundation model for digital pathology from real-world data. *Nature*, 630(8015):181–188, 2024.

Zhu, S., Zhu, Y., Tao, M., and Qiu, P. Diffusion generative modeling for spatially resolved gene expression inference from histology images. In *The Thirteenth International Conference on Learning Representations*, 2025. URL https://openreview.net/forum?id=FtjLUHyZAO.

A. The TRM Architecture

The pseudocode follows Figure 3 of Jolicoeur-Martineau (2025) exactly; : we only replace the discrete ACT target ($\hat{y}$=y_{true}) with a soft loss-proxy target $q^\star$.

```
def latent_recursion(x, y, z, n = 6):
  for i in range(n):                                  latent reasoning
    z = net(x, y, z)
  y = net(y, z)                                       refine output answer
  return y, z

def deep_recursion(x, y, z, n = 6, T = 3):
  recursing T − 1 times to improve y and z
  no gradients needed
  with torch.no_grad():
    for j in range(T − 1):
      y, z = latent_recursion(x, y, z, n)
  recursing once to improve y and z
  y, z = latent_recursion(x, y, z, n)
  return (y.detach(), z.detach()),
    output_head(y), Q_head(y)

Deep Supervision
for x_input, y_true in train_dataloader:
  y, z = y_init, z_init
  for step in range(N_supervision):
    x = input_embedding(x_input)
    (y, z), ŷ, q̂ = deep_recursion(x, y, z)
    loss = softmax_cross_entropy(ŷ, y_true)
    q* = σ((τ − CE_holdout) · s)
    loss+ = binary_cross_entropy(q̂, q*)
    loss.backward()
    opt.step()
    opt.zero_grad()
    if q̂ > 0:                                         early-stopping
      break
```

Figure 5. TRM training pseudocode. We follow the original TRM recursion structure and replace the discrete ACT target with a soft target $q^\star$.

B. Task Formulation and Data Overview

Datasets

DLPFC (Maynard et al., 2021). This dataset contains 12 dorsolateral prefrontal cortex sections from 3 neurotypical donors, profiled on the 10× Visium platform. Each slide carries ∼3,600 valid under-tissue spots, and each spot is manually annotated by neuropathologists into one of seven cortical regions (white matter plus layers 1–6); spots without a confident layer call are masked out. The dataset also distributes the H&E image accompanying each section.

HER2$^+$ ST (Andersson et al., 2021). This dataset contains thirty-six legacy Spatial Transcriptomics sections from eight HER2$^+$ breast cancer patients (lettered A–H, with up to six serial sections per patient). Each section is captured on the original ST 1k-spot array at a 100 μm spot size, yielding ∼176–712 valid spots per slide. Each spot is paired with a corresponding H&E patch from the section's whole-slide image, which is released alongside the count data. Slide B1 (and every patient's first section) carries free-form pathologist annotations of breast-cancer histology classes (invasive cancer, breast glands, connective tissue, adipose, undetermined).

Task Definitions

We study two complementary tasks on spatial transcriptomics (ST) data. In both settings, a slide is represented on a discrete spot grid with binary tissue mask $M_s \in \{0,1\}^{H_s \times W_s}$, and we write $\Omega_s = \{(i,j) : M_s(i,j) = 1\}$ for the valid tissue positions.

Task 1: Cross-slice Spatial Domain I (SDI)dentification

In the DLPFC data, each slide provides gene-expression features $X_s \in \mathbb{R}^{H_s \times W_s \times d_{\text{in}}}$ together with layer annotations $Y_s \in \{1,\dots,K\}^{H_s \times W_s}$ on Ω_s, where in our setup $d_{\text{in}} = 50$ PCA components computed from the top 3,000 batch-aware HVGs after $\log_2(\text{counts}+1)$ and $K = 7$ cortical layers. The goal is to learn a mapping

$$f_\theta : (X, M) \mapsto \hat{Y} \in \{1,\dots,K\}^{H \times W}$$

that predicts a domain label at every valid spot. Training minimizes masked per-spot cross-entropy,

$$\mathcal{L}_{\text{cls}}(\theta) = \frac{1}{|\Omega|} \sum_{(i,j)\in\Omega} \text{CE}\big(f_\theta(X, M)_{i,j}, Y_{i,j}\big),$$

The input grid is placed on a shared 78×128 canvas, linearly projected to hidden dimension d=192, RMS-normalized, and augmented with 2D sinusoidal positional encodings plus a learnable context token. During training, each slice is presented as 20 augmented copies per epoch, with joint transformations applied to the feature grid, label grid, and mask: random 180° rotation, horizontal and vertical flips, Gaussian noise $\mathcal{N}(0, 0.1^2)$ on valid spots, and feature dropout with probability 0.1 per PCA dimension. Validation and test evaluation use the unaugmented grid, and evaluation follows leave-one-slice-out (LOSO): for each held-out slide $s^\star$, the model is trained on all other slides and evaluated zero-shot on $s^\star$.

Task 2: Histology-to-gene-expression prediction (H2G)

For the histology-to-expression task, the input at inference time is only the H&E image. For each valid spot $(i,j) \in \Omega_s$, we extract a 224×224 image patch centered on that spot and embed it with a frozen pathology foundation model ϕ (Gigapath, 1536 dimensions), producing a feature tensor $V_s \in \mathbb{R}^{H_s \times W_s \times d_{\text{fm}}}$. The prediction target is a per-spot gene-expression vector $Y_s(i,j) \in \mathbb{R}^G$ of log-normalized count

Table 3. Full HER2ST B1 holdout metrics under the Zhu et al. (2025) evaluation protocol (300 HMHVG, $\log_2(\text{counts}+1)$, single test slide). †Published source data from Zhu et al. (2025).

	PCC-10	PCC-50	PCC-300	MAE	MSE	RVD
No Rec. 2-Layer	0.865	0.820	0.762	0.690	0.872	0.257
Univ. Transformer	0.830	0.792	0.745	0.708	0.822	1.088
TRM (N_{outer}=4)	0.869	0.824	0.771	0.689	0.836	0.383
TRM (N_{outer}=8)	**0.883**	**0.841**	**0.785**	**0.678**	**0.810**	0.354
HisToGene†	0.6812	0.6345	0.5250	0.9367	1.3468	10.34
BLEEP†	0.7727	0.7141	0.5652	0.8328	1.2428	0.6025
TRIPLEX†	0.7907	0.7394	0.5766	0.9311	1.3456	0.6428
STEM†	0.8298	0.7726	0.5984	0.7547	1.0742	**0.0693**

values over a fixed gene panel; for HER2ST, we use the top-302 HMHVG panel selected from the training slides, following Zhu et al. (2025). The model learns

$$g_\theta : (V, M) \mapsto \hat{Y} \in \mathbb{R}^{H\times W\times G},$$

trained with masked mean squared error,

$$\mathcal{L}_{\text{reg}}(\theta) = \frac{1}{|\Omega|G} \sum_{(i,j)\in\Omega} \sum_{c=1}^{G} \big(g_\theta(V, M)_{i,j,c} - Y_{i,j,c}\big)^2.$$

The spot embeddings are padded to a shared 35×35 grid together with the tissue mask. Following Zhu et al. (2025), we precompute Gigapath embeddings for all eight elements of the dihedral group D_4 and sample one coherent view per slide during training; we additionally use 20 augmented copies per sample per epoch, with Gaussian noise (σ=0.05) and feature dropout (p=0.05) applied to the embedding tensor during training only. Validation and test evaluation use the unaugmented upright view. Evaluation uses a single-slide holdout: one slide is reserved for testing (HER2ST: B1), the model is trained on the remaining slides, and predictions are scored on the held-out valid spots. For each gene c, we compute the across-spot Pearson correlation ρ_c between prediction and target, then report top-k mean Pearson correlation $\text{PCC}{-}k$ for $k \in \{10, 50, 200\}$, along with MSE, MAE, and the relative variation distance (RVD),

$$\text{RVD} = \frac{1}{G} \sum_{c=1}^{G} \frac{\left(\sigma_c^{2,\text{pred}} - \sigma_c^{2,\text{gt}}\right)^2}{\left(\sigma_c^{2,\text{gt}}\right)^2},$$

which measures how well predicted across-spot variance matches the true biological heterogeneity.

Observations about performance on HER2ST. Table 3 shows a detailed comparison between TRM and baselines for the HER2ST dataset. We see that the No-recursion variant already beats the baselines in terms of convergence error and PCC. However, on RVD, Stem's per-spot diffusion sampling reaches 0.069 on HER2ST B1, an order of magnitude below our best regression result. Two things follow.

(i) Joint-prediction over the whole slide significantly increases performance over baseline work. Even without recursion, predicting the entire spot grid jointly under a strong regulariser already exceeds prior regression baselines, which recursion further improves on.

(ii) The RVD gap is structural, not a tuning issue. By the law of total variance, $\text{Var}_x(y) = \mathbb{E}_x[\text{Var}(y \mid x)] + \text{Var}_x(\mathbb{E}[y \mid x])$. Any MSE-trained deterministic regressor $\hat{y}(x) = \mathbb{E}[y \mid x]$ emits only the second term and discards the spot-conditional residual $\mathbb{E}_x[\text{Var}(y \mid x)]$, which dominates on the 302-HMHVG panels and is exactly what RVD scores. PCC-k is invariant to global rescaling, so the same conditional-mean collapse can preserve top-k ranking while flattening variance—hence we beat Stem on PCC-k and lose by an order of magnitude on RVD. Zhu et al. (2025) introduce RVD precisely because "PCC in log-transformed space would be surprisingly high if the prediction is simply the mean expression across all genes," and motivate diffusion on the prior grounds that the image→expression map is intrinsically one-to-many. Drawing $S\approx20$ samples per spot from $p_\theta(y \mid x)$ and averaging restores both variance terms—spot-to-spot diversity preserves $\text{Var}_x(\mathbb{E}[y \mid x])$, sample-to-sample stochasticity reinjects $\mathbb{E}_x[\text{Var}(y \mid x)]$— collapsing the RVD ratio while keeping PCC-k competitive. However, using different ACT halting strategies along with scaling down N_{Z_L} and N_{Z_H}, we observed that different ACT objectives added variance-stabilizing properties to the final predictions.

C. Ablation Methodology

Procedure. All recursion ablations use models trained with default hyperparameters (deep recursion cycles = 3, latent recursion steps = 6, `halt_max_steps`= 8) in the LOSO (SDI) or held-out-B1 slice (H2G) setting. At test time, we override one or more parameters and evaluate on the held-out slice. For each of the 12 LOSO folds, we evaluate all parameter combinations and report mean ± std across folds.

Block equivalence. A TRM forward pass is parameterized by three counters, all referring to repeated applications of the shared transformer $\mathcal{F}_\theta$:

- N_{z_L} *(latent recursion, also denoted ℓ)*: how many times the latent state z_L is refreshed in a row before the next deep update.
- N_{z_H} *(deep recursion, also denoted h)*: how many such latent groups are stacked, each followed by one deep update of z_H.
- N_{outer} *(outer / deep-supervision steps, also denoted n)*: how many times the model traverses the entire deep-recursion stack, with a (read-only) outer-step boundary

in between.

A single *processing block* is the basic unit of Figure 1a: N_{z_L} latent updates of z_L (each a call to $\mathcal{F}_\theta$ with context z_H+x), followed by one deep update of z_H from z_L. One inner pass executes N_{z_H} such blocks back-to-back; one full forward pass executes N_{outer} inner passes. At test time, the outer-step boundary performs only read-only operations (logit decoding, halt-head evaluation) and does *not* reset or modify the carried (z_H, z_L) state. The sequence of $\mathcal{F}_\theta$ calls—and therefore the hidden-state trajectory—is controlled by N_{outer} and N_{z_H} only through their *product*: any two configurations with the same $N_{\text{outer}} \cdot N_{z_H}$ produce a bit-identical test-time forward pass. We verified this empirically: across all 12 folds, configurations with equal $N_{\text{outer}} \cdot N_{z_H}$ produce identical ARI to six decimal places (Table 4).

Table 4. Block equivalence verification (slice 151670). Configurations with the same total blocks ($n \times h$) yield identical ARI.

Config	Total blocks	ARI
n=1, h=2	2	0.6769
n=2, h=1	2	0.6769
n=1, h=3	3	0.7117
n=3, h=1	3	0.7117
n=2, h=3	6	0.7169
n=3, h=2	6	0.7169
n=6, h=1	6	0.7169

Grid ablation. For Figure 1a, we evaluate all (n, h) combinations with $n \in \{1, \ldots, 8\}$ and $h \in \{1, 2, 3\}$ at default latent recursion steps ℓ=6, then collapse results by total blocks $n \times h$. For Figure 1b, we fix n=1 and vary deep recursion cycles $h \in \{1, 2, 3\}$ with latent recursion steps $\ell \in \{1, 2, 3, 4, 6, 8\}$. For Figure 1c, we use the standard ablation at default n=3, h=3 with $\ell \in \{1, \ldots, 8\}$.

D. DLPFC: Zero-shot transfer to human breast cancer

We evaluate zero-shot generalization of DLPFC-trained TRM on the 10x Visium human breast cancer dataset of Andersson et al. (2021). This is a structurally distinct tissue type (breast carcinoma rather than cortex), but it shares the same human gene space: 2,872 of the 3,000 DLPFC HVGs are present in the breast cancer dataset, and the remaining 128 genes are zero-filled. No breast cancer-specific training is performed. Instead, we project breast cancer expression onto the DLPFC shared PCA space and pass the resulting 50-dimensional features directly to all 12 DLPFC LOSO checkpoints.

Setup: We use the fine annotation labels (20 histological classes spanning IDC, DCIS/LCIS, tumor edge, and healthy tissue) and report ARI and NMI. Cluster assignments are obtained by k-means with k=20 on the z_L latent, which consistently outperforms z_H in our DLPFC experiments. We report the best ARI over the 12 DLPFC LOSO checkpoints for the zero-shot neural baselines. We also test how much spatial information influences clustering as follows: For the TRM latent Z_L, we concatenate the spatial coordinates with Z_L and cluster it. This results in a +0.15 increase in ARI, which suggests that Z_L can benefit from spatial information for OOD prediction. Interestingly, the ARI also increases significantly when we repeat the same with the non-recursive baseline. Additionally, increasing N_{outer} degrades zero-shot performance, indicating that the model becomes less confident in its predictions as refinement proceeds. Overall, we find that TRM moderately generalizes to the breast cancer dataset, and adding spatial context further improves generalization.

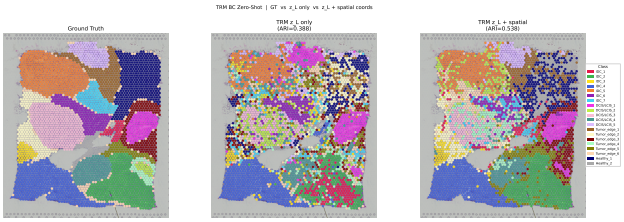

Figure 6. Zero-shot transfer to human breast cancer. TRM latent is significantly enhanced by spatial information.

Table 5. Human breast cancer spatial clustering (max ARI, 20 fine classes). Zero-shot methods are trained on LOSO DLPFC data. † Values taken from Dai et al. (2026). § L2-normalised z_L.

Method	Max ARI
PCA k-means (DLPFC HVG set (BC) projection)	0.289
No Rec. 2-Layer	0.343
No Rec. 2-Layer + spatial	0.458
Univ. Transformer	0.206
TRM (z_L + spatial§)	**0.538**
TRM (z_L)	0.388
SpaGCN	0.390
STAGATE	0.572
GraphST	0.577
3d-OT†	0.683

E. HER2ST test-time recursion ablation

We repeat the test-time recursion ablation of Fig. 1's gene-expression analog on the HER2ST B1 holdout, using the converged TRM (k=8, 302-HMHVG) checkpoint, following the same protocol as the DLPFC block-equivalence procedure. Figure 7 reports PCC-mean, PCC-50, and RVD; following Zhu et al. (2025), we omit error bars because this is a single-slide holdout.

The qualitative shape mirrors DLPFC: roughly three process-

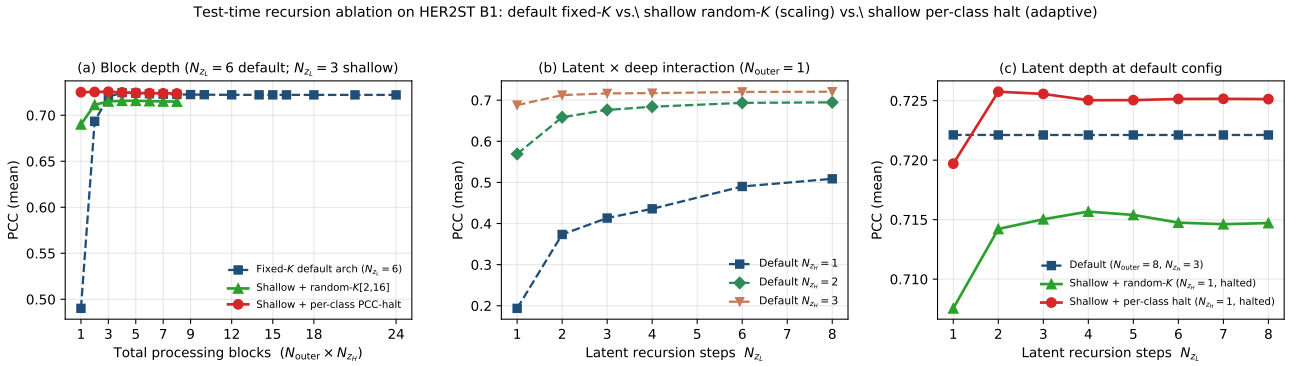

Figure 7. Test-time recursion ablation on the HER2ST B1 holdout . **(a)** Block-depth ablation at fixed ℓ=6: PCC-mean rises from 0.49 (1 block) to 0.72 by 3 blocks and saturates; RVD drops from 0.85 to a minimum of 0.26 at 3 blocks before stabilizing near 0.35. As in DLPFC, n and h are interchangeable at fixed $n{\cdot}h$. **(b)** Latent $\times$ deep interaction (n=1): with a single deep recursion (h=1), latent depth alone caps at PCC-mean ≈ 0.51 even at ℓ=8; h=2 reaches 0.69 and h=3 reaches 0.72, showing that the deep-recursion update increases the PCC significantly. **(c)** Latent depth at the default config (n=8, h=3): all $\ell \in \{1, \ldots, 8\}$ give identical PCC-mean $= 0.7221$, PCC-50 $= 0.8414$, and RVD $= 0.354$ to four decimal places, indicating that the trained model has fully saturated and additional latent applications make no measurable difference.

ing blocks suffice to reach peak PCC, and the deep-recursion update is the load-bearing component of the latent→deep →latent message-passing loop. However, in Figure 7(c), we see that at normal inference, increasing N_{Z_L} does not improve the PCC; it does, however, stabilize the variance. This might be due to the fact that the gene-expression target is a smoother, lower-frequency function of the foundation-model features than the cortical-layer label, so once the deep-recursion loop has refined the spatial features like in Figure 7(c), additional latent passes have nothing left to improve. This finding also indicates that with the default TRM config ($N_{Z_H} = 3, N_{ZL} = 6$), ACT would not work as outer recursions would not add any useful improvements to performance, and hence we try increasing the importance of outer recursions, leading to effective ACT halting in Appendix F.2

F. ACT details

Here, we provide details on the ACT recipes, calibration choices, formulae, and results referenced in Section 4, separated by dataset. Section F.1 covers the DLPFC LOSO setting (the original ACT-on-domain-identification analysis); Section F.2 covers the HER2ST gene-expression setting (the regression analog and the input-adaptive halt result). Across all ACT recipes, the halt head is supervised via a per-step, randomized within-sample spot holdout, disjoint from the regression-loss spots. Note that these spots (FM features) are seen by the model during training but are cut off from backpropagation via the regression loss, instead being supervised by the halt head.

Table 6. ACT strategy comparison on LOSO fold 151673 (DLPFC).

Strategy	Test ARI	Mean steps	Best epoch
Loss-Proxy ACT	**0.520**	8.0	6
Weighted BCE	0.465	2.0	13
Random-K [2,8]	0.457	4.0	100
Random-K [2,16]	0.462	8.0	100

F.1. DLPFC LOSO ACT

Classical ACT (Graves, 2016) was designed for tasks with a clear "solved/not-solved" signal per input, e.g., digit addition, where the model emits one scalar answer it can self-verify. Spatial-domain identification is technically still discrete — each spot is assigned to one of K cortical layers, but the output is a *joint multi-class distribution over all* $\sim$ 4000 *spots in the sample*, so partial correctness is the rule and a binary correct/not-correct signal per input is essentially never achievable. We compare three practical recipes on slice 151673 LOSO — training on all slices except 151673 and evaluating on it — with results in Table 6.

Loss-Proxy ACT. The halt head is supervised by a soft target derived from held-out cross-entropy. We want the ACT controller to encourage generalization:

$$q^{\star}_{\text{halt}} = \sigma\big((\tau - \text{CE}_{\text{holdout}}) \cdot s\big), \tag{1}$$

where τ=0.7 is a quality threshold and s=4 controls sharpness. This is also the formulation summarized in the pseudo code in Figure 5.

Weighted BCE. This is a calibrated variant of the same halt-head supervision, with the halt threshold lowered to

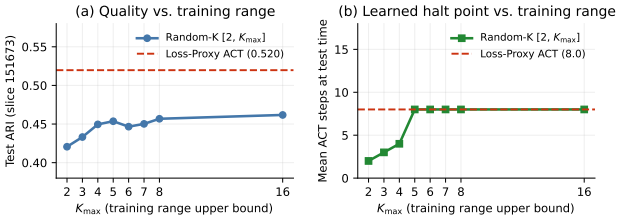

Figure 8. Random-K sweep over training range $[2, K_{\max}]$ on fold 151673. **(a)** Test ARI increases with $K_{\max}$ up to ≈4, then plateaus below the Loss-Proxy ACT baseline (dashed). **(b)** The learned halt point is pinned to $K_{\max}$ for $K_{\max} \leq 4$; for $K_{\max} \geq 5$ the model abandons early halting and runs all 8 steps.

0.75 and early-stopping patience relaxed to 8 epochs. We also bias the target to account for class imbalance, thereby promoting good performance across all classes. These changes bias training toward earlier stopping and produce the 2.0 mean-step solution reported in Table 6 and Figure 9.

Random-K. Each training sample is assigned a random target depth $K \sim \text{Uniform}[2, K_{\max}]$, where the model must always take at least two exploratory steps, and the halt head is trained to stop exactly at step K. We sweep $K_{\max} \in \{2, \ldots, 8\}$ and also with $K_{\max}$=16 (to make it a fair comparison to the other methods with mean steps 8) to study how the width of the training distribution affects both test ARI and the learned halting point (learned using the random-K values as supervision) (Figure 8).

Loss-Proxy ACT achieves the best ARI (0.520) on a held-out set by optimizing for halt head emph answer quality, but it almost always uses the full 8-step budget. Weighted BCE instead biases the controller toward *good-enough* solutions that account for class imbalance and is the only setting that yields real anytime behavior: mean steps drop to 2.0 (a 75% compute reduction) while retaining moderate quality (ARI 0.465). Random-K, where K corresponds to the choice of N_{outer}, trains over a random depth budget between $[2, K_{\max}]$ and acts as a controller-free regularized alternative; in practice, however, it does not match Loss-Proxy ACT. The key takeaway (Table 6) is that these recipes occupy different points on the quality–compute frontier: Loss-Proxy ACT is best for accuracy and encourages generalization, Weighted BCE offers a practical, compute-saving tradeoff, and Random-K is useful for budget diversification but tricky to train.

F.2. HER2ST ACT

The H2G task predicts gene expression vectors and is a continuous task. Here, we provide details for different ACT methods applied to this task. All experiments use the shallow-inner architecture (N_{z_H}=1, N_{z_L}=3, $K_{\max}$=8) trained on the Stem 2025 HER2ST protocol (Zhu et al., 2025) with 302 HMHVG genes. The models use an ACT halting head to choose N_{outer} steps out of a range of $[2, K_{max} = 8]$

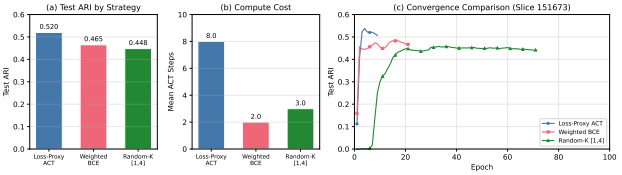

Figure 9. ACT strategy comparison on fold 151673. **(a)** Test ARI: Loss-Proxy achieves the highest quality. **(b)** Compute cost: Weighted BCE and Random-K learn to halt early. **(c)** Convergence: Loss-Proxy is fastest; Random-K is considerably slower.

Table 7. ACT comparison on HER2ST B1 holdout. $\bar{n}_{\text{halt}}$ is the empirical mean halt step over the 36 slides; values $< K_{\max}$ indicate halting fires.

Strategy	PCC-300	RVD	$\bar{n}_{\text{halt}}$
TRM fixed N_{outer}=8	**0.785**	0.354	8.0
PCC-halt (slide-level, never fires)	0.723	0.163	8.0
RVD-halt (slide-level, never fires)	0.723	0.164	8.0
MSE-halt (slide-level, always fires)	0.779	**0.109**	2.0
Per-class PCC-halt	0.780	0.158	**4.9**
STEM (Zhu et al., 2025)	0.598	**0.069**	—

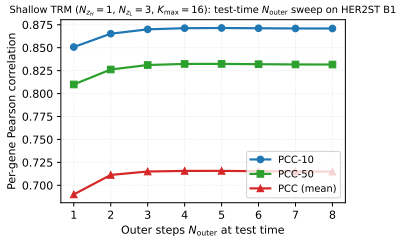

Figure 10. Test-time scaling on the small-recursion architecture trained with random-$K \sim \text{Uniform}[2, 16]$. PCC increases from 0.69 at N_{outer}=1 to 0.72 at N_{outer}=4 before saturating.

The intuition for scaling down N_{z_H} and N_{z_L} The default TRM (N_{z_H}=3, N_{z_L}=6) saturates after a single inner pass: sweeping N_{outer} at test time is flat from step 1 (Appendix E). This indicates that the expressivity of the inner recursions saturates performance meaning more N_{outer} steps cannot improve. We hence scale down the expressivity of the latent ($N_{Z_L} = 3$) and deep ($N_{Z_H} = 1$) recursions and, by doing so, observe input-adaptive halting in the continuous setting. To test whether this scaled-down architecture provides a learnable gradient for increasing N_{outer}, we trained a model with a random $K \sim \text{Uniform}[2, 16]$ ACT strategy. This model exhibits a real test-time scaling increase (PCC 0.69 at N_{outer}=1 → 0.72 at N_{outer}=4, Figure 10), and hence gives ACT a non-trivial outer-recursion gradient to operate on.

Halt-head supervision. We test three halt-target families and one per-class extension for the halt head. At each step k, the held-out per-batch quality estimate is mapped to a halt target via a sigmoid:

$$\textbf{PCC-halt} \quad q^\star_{\text{halt}} = \sigma\big((\overline{\text{PCC}}_k - \tau)\, s\big), \quad \tau{=}0.65,\ s{=}20,$$
$$\textbf{RVD-halt} \quad q^\star_{\text{halt}} = \sigma\big((\tau - \text{RVD}_k)\, s\big), \quad \tau{=}0.4,\ s{=}4,$$
$$\textbf{MSE-halt} \quad q^\star_{\text{halt}} = \sigma\big((\tau - \text{MSE}_k)\, s\big), \quad \tau{=}0.9,\ s{=}10,$$

trained against the halt logit via BCE. Per-class PCC-halt extends supervision to a halt logit per gene: $q^\star_{\text{halt},g} = \sigma((\text{PCC}_{g,k} - \tau)\, s)$, with the slide-level halt decision $q_{\text{halt}} = \frac{1}{G}\sum_g q_{\text{halt},g}$ (unchanged forward path) but a per-gene gradient signal driving the halt head during training. We note that the hyperparameters τ, s were empirically selected after measuring the performance of a fixed-K=8 TRM variant, and this selection is dataset specific.

PCC-halt and RVD-halt do not fire. Table 6 shows that the model runs for K_{max} steps for these ACT strategies. Logically, the model halts at step k if $\sigma(q_{\text{halt}}) > \theta$ for some halt-prob threshold θ (default θ=0.7). PCC-halt's converged mean halt probability is 0.27 across all 36 HER2ST slides; RVD-halt's is 0.07. In both cases, the halt head's output collapses to a value well below θ *uniformly across inputs*, so the controller never crosses threshold, and the model defaults to $K_{\max}$=8 steps on every slide. The reason is a calibration mismatch: PCC-halt's halt target $\sigma((\overline{\text{PCC}} - 0.65)\cdot 20)$ saturates near 0.83 at the converged-PCC ceiling ($\overline{\text{PCC}} \approx 0.72$), but the BCE loss is dominated by the early steps where the target is near 0, pulling the halt logit's converged value down. RVD-halt has the same dynamic with even more skew, since RVD has a wider range and the target sits closer to 1 at convergence than to 0. The remedy is either (i) recalibrate the threshold below the per-input converged distribution (Figure 11c), or (ii) modify the supervision target so the halt head converges higher — both of which we exploit below. Despite never firing, both recipes still cut RVD by $\sim 2\times$ over fixed-depth TRM (Table 7): the halt-loss BCE term acts as a strong variance-aware regularizer on the prediction head whether or not the controller eventually halts.

MSE-halt halts immediately. MSE-halt's sigmoid $\sigma((\tau - \text{MSE}_k), s)$ at $\tau = 0.9, s = 10$ places the midpoint at the converged held-out MSE ($\sim$ 0.7–1.0), so the supervision target rises sharply across training-time steps from ~ 0 at $k = 1$ to ~ 0.95 at saturation. The halt head therefore converges to mean halt probability 0.997, comfortably above the threshold $\theta = 0.7$ on every input, and halts at the earliest step the configuration allows.

Despite halting at the configured minimum, MSE-halt achieves the lowest slide-level RVD across all four recipes (0.109, 3.2$\times$ better than fixed-depth TRM at matched PCC). Two factors account for this. First, MSE-halt's halt-target trajectory has the largest dynamic range across steps of the three slide-level recipes: PCC saturates at $\approx$ 0.8; RVD-halt's target saturates at only ≈ 0.5, since τ sits right at the converged RVD; and MSE-halt's target saturates near 0.95. Its BCE halt loss is therefore the strongest signal pushing the gradient back through z_H during training.

Second, and more importantly, MSE is the exact same loss used to supervise the prediction head, so the halt loss adds a sample-level view of the same quality signal the prediction head receives at the spot level. This double supervision forces z_H to encode a per-sample coherence signal in addition to per-spot fit, which empirically biases the prediction head toward preserving across-spot variance — exactly what RVD penalizes. PCC-halt and RVD-halt provide weaker BCE gradients, with saturated targets of ≈ 0.8 and ≈ 0.5, respectively, and use quality estimators that are not the prediction loss. As a result, the same regularization channel is correspondingly weaker.

Per-class PCC-halt with calibrated thresholding reveals input-adaptive halting. The reason slide-level PCC-halt's halt head never fires is that $\sigma(q_{\text{halt}})$ converges to a value well below θ=0.7 uniformly across slides. Lowering the halt-prob threshold to θ=0.4 moves the threshold *inside* the natural per-input distribution of converged halt probabilities, so the upper half crosses while the lower half does not. We also enable per-class supervision: each of the 302 gene halt logits is supervised against its own held-out PCC trajectory, providing the halt head with a richer per-gene signal whose mean determines the slide-level halt decision. Evaluating the trained checkpoint on every available slide independently (Figure 11) reveals that the empirical halt step is bimodal at $\{2, 8\}$ across the 36 HER2ST slides — 17 slides halt at the floor with halt probability ≥ 0.69, 17 slides have halt probability collapsed below threshold and run to $K_{\max}$, and the test slide B1 sits between the two clusters at step 5. The halt step is inversely correlated with per-slide PCC (Figure 11): *high-quality slides halt early, low-quality slides keep computing*. This indicates that the empirical halt step varies meaningfully across inputs, suggesting that ACT-as-adaptive-computation is achievable for continuous, structured outputs, given a halt-prob threshold calibrated within the per-input halt-prob distribution.

Cross-fold validation. To rule out a single-fold artifact, we retrain the same recipe on three additional held-out test slides drawn from different patients (B6, C1, D1). The bimodality replicates on every fold (Table 8 and Figure 12) with cluster sizes varying mildly with fold (early-cluster size 13–19 of 36; late-cluster size 17–23 of 36), but the qualitative structure is unchanged. Across all four folds, the held-out test slide always lands in the cluster consistent with

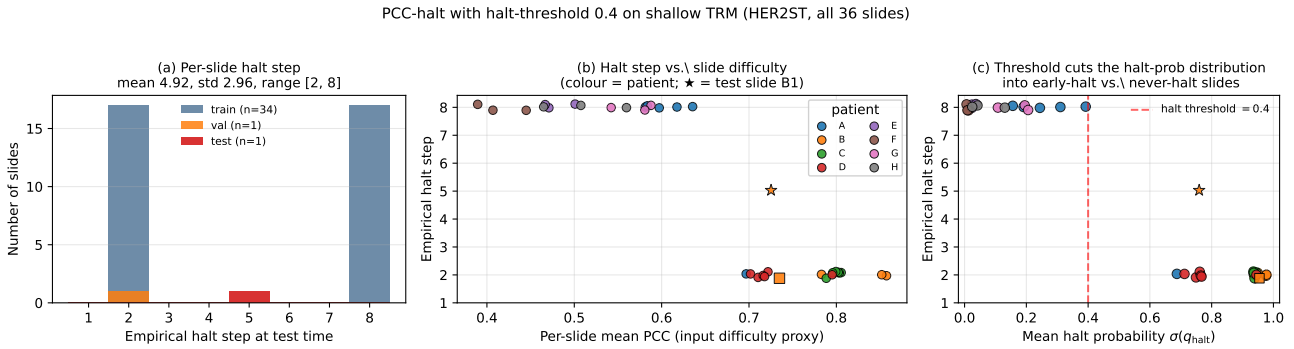

Figure 11. Input-adaptive halting on the shallow + per-class PCC-halt checkpoint with halt threshold 0.4, evaluated per-slide on all 36 HER2ST slides. **(a)** Histogram of empirical halt step — bimodal at $\{2, 8\}$ with the test slide B1 at step 5, mean 4.92, std 2.96. **(b)** Halt step vs. per-slide PCC, coloured by patient: easy slides (PCC $\gtrsim 0.70$) halt at the `min_halt_steps` floor; hard slides (PCC $\lesssim 0.55$) never cross threshold and run to $K_{\max}$. The test slide B1 (red star) sits at the boundary and halts at step 5. **(c)** Halt step vs. converged halt probability: the 0.4 threshold (red dashed) cleanly partitions the halt-prob distribution into the two halt-step clusters.

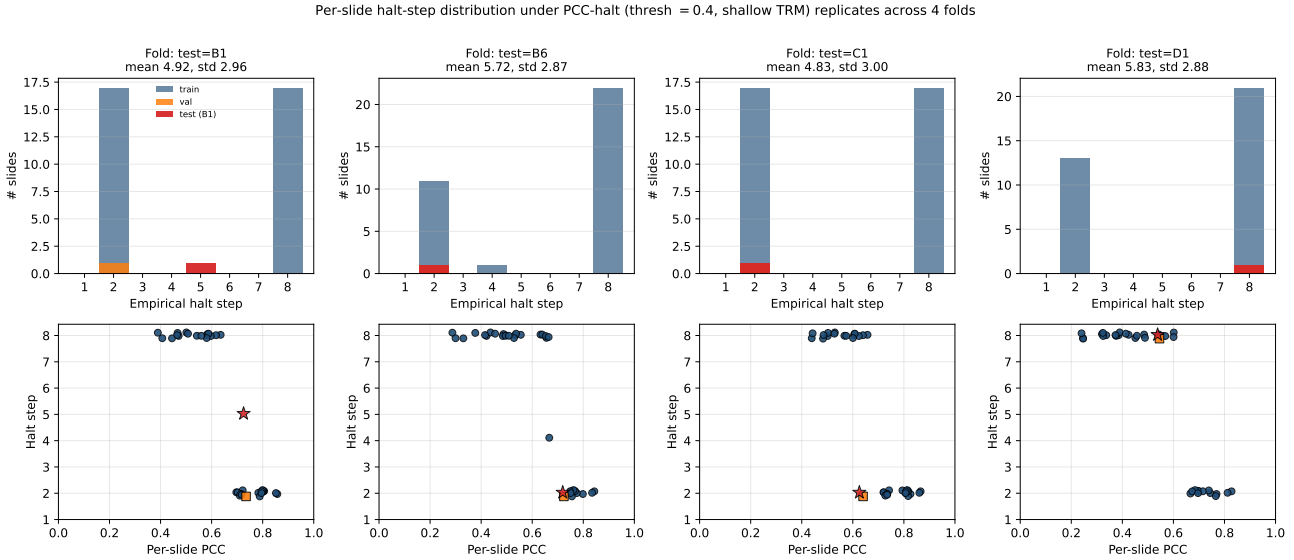

Figure 12. Per-slide halt-step distributions under PCC-halt (threshold 0.4, shallow TRM) across four independently retrained checkpoints with different held-out test slides. **Top row**: histogram of empirical halt step over all 36 slides, split by train/val/test. **Bottom row**: halt step versus per-slide PCC; the test slide is shown as a red star. Bimodality at $\{$`min_halt_steps`$=2, K_{\max}=8\}$ replicates on every fold, and the test slide always lands in the cluster predicted by its difficulty.

Table 8. Cluster sizes and test-slide outcomes across four HER2ST holdout folds, all trained with shallow TRM + per-class PCC-halt at threshold 0.4. "Early" = empirical halt step ≤ 4; "Late" = halt step $= K_{\max}$=8.

Fold	Test slide	Test PCC	RVD	Test halt step	Early (≤ 4)	Late ($= 8$)
0 (original)	B1	0.7253	0.158	5 (boundary)	18 / 36	17 / 36
1	B6	0.7196	0.144	2 (early)	14 / 36	22 / 36
2	C1	0.6257	0.186	2 (early)	19 / 36	17 / 36
3	D1	0.5381	0.286	8 (late)	13 / 36	23 / 36

its own converged PCC: high-PCC test slides (B1, B6, C1) halt early; the low-PCC holdout (D1, PCC 0.54) sits in the late cluster and runs to $K_{\max}$; only B1 happens to land at the boundary at step 5.

Compute–quality trade-off. Aggregating $\bar{n}_{\text{halt}}$ across the 36 slides on each per-class checkpoint gives a $\sim$ 42% mean compute reduction relative to fixed N_{outer}=8, at matched PCC and 2.2× lower RVD (0.158 vs. 0.354). At the more aggressive operating point with `min_halt_steps`=1 the saving rises to 50.7% at a 0.4% PCC drop; lowering the halt-prob threshold further (to θ=0.3) regresses on PCC because the under-trained halt head fires before the model has converged.

G. Marker-Gene Panel on HER2ST B1

G.1. HER2ST: Qualitative Marker-Gene Validation

The marker-gene grid in Figure 13 provides a qualitative sanity check on whether the model's predictions are biologically faithful, rather than merely statistically well-correlated. The HER2-amplicon row supplies the strongest such check: ERBB2 and GRB7 should localize to the same spots within the pathologist-annotated invasive cancer regions, because GRB7 is frequently co-amplified *and co-expressed* with ERBB2/HER2 in breast cancer (Stein et al., 1994; Sahlberg et al., 2013). Both ground-truth and predicted patterns concentrate inside those annotations. The cell-cycle (CCND1, MKI67) row extends this to graded, non-amplicon-driven proliferation signals, with CCND1 and Ki67/MKI67 has been established as a proliferation-associated marker in breast cancer (Tobin & Bergh, 2012; Nielsen et al., 2021). The luminal row (ESR1, GATA3, FOXA1, MUC1) probes more diffuse expression patterns, since ESR1, GATA3, FOXA1, and the associated epithelial/luminal programs define luminal breast cancer identity (Perou et al., 2000; Sørlie et al., 2001; Theodorou et al., 2013; Albergaria et al., 2009). The tumor-suppressor row (ATM, PTEN), grouped because both have established roles in the DNA-damage response in breast cancer (Shiloh, 2003; Bassi et al., 2021), and the EMT row (VIM) (Nieto et al., 2016) probes mesenchymal/EMT-associated signal, which is biologically distinct from the sharply localized HER2-amplicon markers and may be influenced by stromal or mixed-cell composition. We emphasize that the figure is not cherry-picked. The twelve genes were specified a priori from the breast cancer biology before any predictions were inspected, and the checkpoint ($k = 8$) was likewise fixed before this analysis.

H. Limitations

We consider only two representative tasks in spatial omics, so the impact of recursive reasoning at the scale of foundation models remains unexplored. Additionally, we propose heuristic solutions to repurpose ACT for continuous domains; inventing smarter approaches for ACT is the next step as our proposed solutions require delicate dataset-specific hyperparameter tuning. Additionally, the lack of absolute uncertainty or prediction confidence limits the application of the recursive reasoning paradigm to real-world clinical tasks.

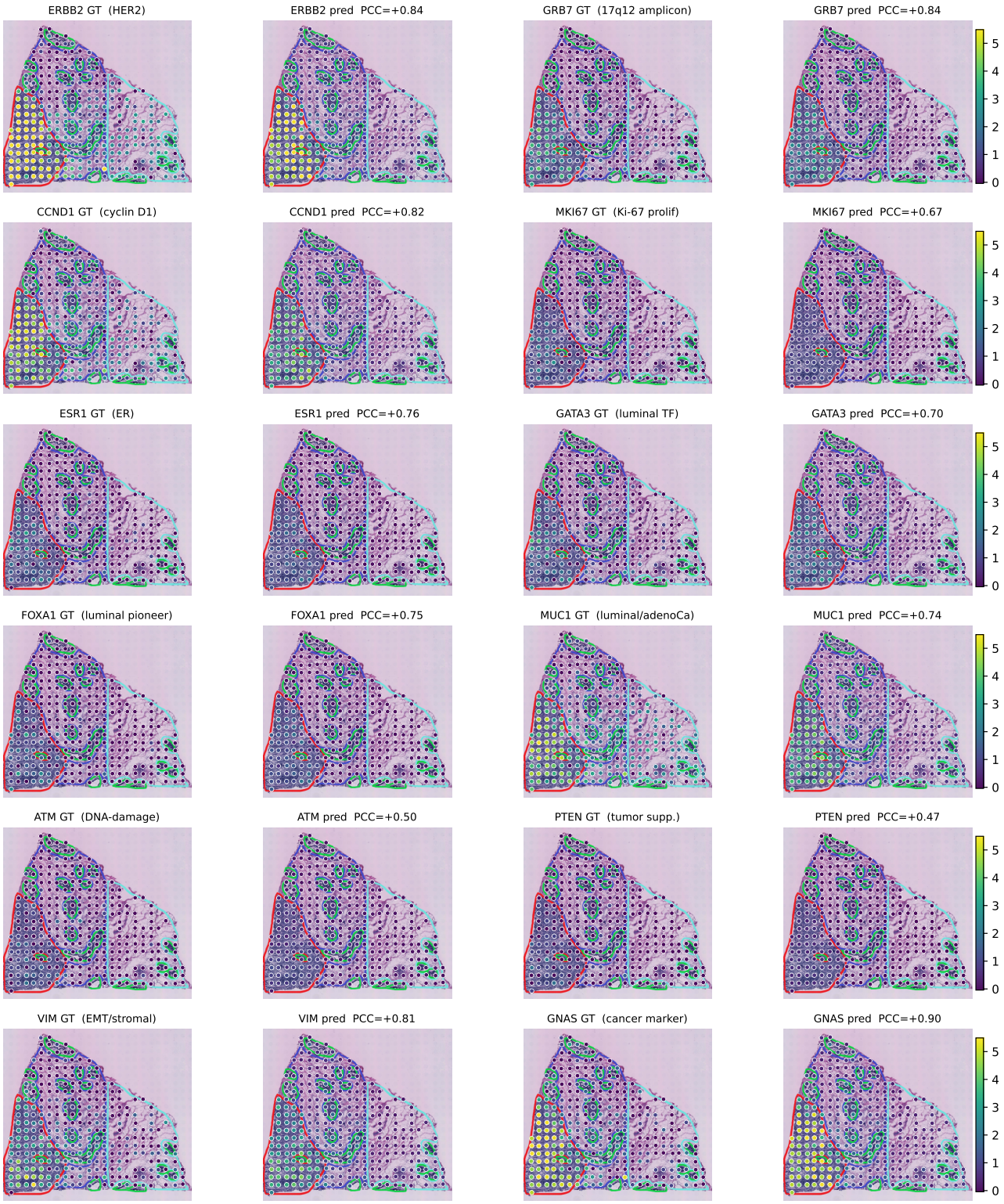

Figure 13. HER2ST B1 marker-gene grid. All expression values are normalized together to allow for direct comparison. Rows are grouped by biological axis: HER2 amplicon (*ERBB2*, *GRB7*); cell cycle (*CCND1*, *MKI67*); hormone receptor / luminal (*ESR1*, *GATA3*); luminal / adenocarcinoma (*FOXA1*, *MUC1*); DNA-damage response/tumor suppressor (*ATM*, *PTEN*); EMT / cancer marker (*VIM*, *GNAS*).