IS H&E IMAGE-TO-SPATIAL TRANSCRIPTOMICS SIMPLER THAN IT LOOKS?

Anonymous authors
Paper under double-blind review

ABSTRACT

Predicting spatial gene expression from routine H&E histology offers a scalable route toward spatial molecular profiling. Recent work has pursued increasingly sophisticated architectures to capture spatial context and richer expression structure. At the same time, simple estimators have shown strong performance in several studies, but what they already solve and where additional complexity is needed remain unclear. We study this behavior through the structure of prediction error under the mean-squared error (MSE) objective. Differences in average expression across genes can account for a substantial part of aggregate prediction performance, while a key unresolved error lies in recovering variation within each slide. Decomposing MSE into slide-level and within-slide components, we find that the within-slide component has lower residual-normalized parameter sensitivity in controlled neural experiments. This motivates Component-Guided Loss (CGL), which increases supervision of the within-slide component. CGL-Linear is a closed-form affine instantiation that achieves overall state-of-the-art performance across HEST-1k cohorts and gene-panel sizes. The same within-slide supervision improves existing neural models. These results suggest that substantial gains can come from aligning the training objective with prediction-error structure rather than increasing model complexity.

1 INTRODUCTION

Spatial transcriptomics measures gene expression at known locations in a tissue section, linking molecular variation to tissue morphology (Ståhl et al., 2016). Its cost and experimental requirements motivate predicting these measurements from routine hematoxylin-and-eosin (H&E) images (He et al., 2020). Accurate prediction would make spatial molecular profiling more accessible and enable the study of tissue samples for which expression measurements are unavailable.

Recent approaches use spatial attention, retrieval, and generative modeling to capture richer relationships between histology and expression (Xie et al., 2023; Huang et al., 2025; Wu et al., 2026; Hu et al., 2026; Byeon & Kwak, 2026). Yet lightweight predictors remain competitive in representation studies and broader benchmarks (Jaume et al., 2024; Pan et al., 2025; Wang et al., 2025). Their performance raises a basic question. What do simple predictors already capture, and how much of their remaining error can be reduced without increasing their capacity?

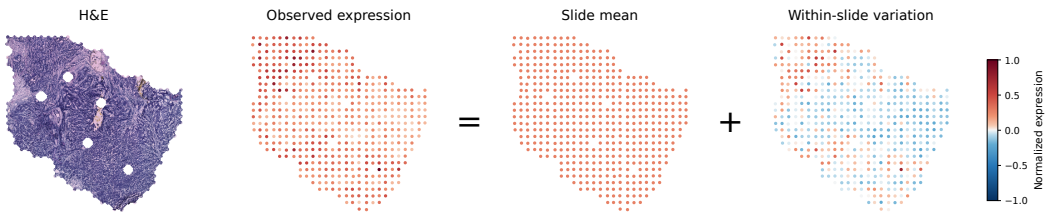

Figure 1: **Slide-wise expression structure.** For one gene on one slide, observed expression can be viewed as a slide-wise mean together with variation around that mean. Each dot denotes a measured spatial location, and the H&E image is shown for spatial context.

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

We approach this question by distinguishing expression shared at the slide level from variation within a slide. Figure 1 illustrates this structure for a single gene. A predictor that assigns the training mean of each gene to every location can obtain a strong aggregate score when all spot–gene measurements are evaluated together. Such a prediction preserves differences in average expression across genes but contains no spatial variation. This mean-only reference shows why evaluating a predictor's contribution requires examining what it recovers beyond gene-specific baselines.

The same distinction carries into learning under mean-squared error (MSE), a widely used objective for H&E-to-spatial-transcriptomics regression (He et al., 2020; Wu et al., 2026; Hu et al., 2026). Locations are grouped within slides, so prediction error can arise from an incorrect slide-wise mean or from inaccurate variation around that mean. We separate MSE exactly into these two components. This gives a common basis for examining both what a predictor recovers and which errors remain difficult to reduce.

Although the two components enter MSE with equal coefficients, they are optimized through shared parameters. Equal weighting therefore need not give equal progress. We characterize their response through residual-normalized parameter sensitivity, which accounts for differences in error magnitude. In controlled neural experiments, the within-slide component has lower sensitivity under standard MSE training. The same ordering appears on held-out slides. This finding motivates Component-Guided Loss (CGL), which increases the weight of within-slide error. Reweighting changes the fitting trade-off and improves held-out spatial recovery.

For an affine readout on frozen image representations, CGL remains a regularized least-squares problem. Its closed-form solution, CGL-Linear, uses the same predictor class as ordinary ridge regression. CGL-Linear improves on this baseline and achieves overall state-of-the-art performance across HEST-1k cohorts and gene-panel sizes. Since both affine estimators are solved directly, their comparison shows that additional within-slide signal can be recovered by changing the fitted objective at fixed readout capacity. The same supervision term also improves spatial recovery in existing neural methods. Together, these results identify a source of improvement available within established predictors and provide a practical reference for assessing the value of additional model complexity.

2 Related Work

Histology-to-expression prediction. Patch-level regression (He et al., 2020) has developed alongside models that use broader tissue context. HiST builds a sparse spatial hierarchy (Wu et al., 2026), HistoPrism combines cancer-type conditioning with Transformer-based feature aggregation (Hu et al., 2026), and STFlow models whole-slide expression through conditional flow matching (Huang et al., 2025). Representation quality provides another route to improved prediction. HEST-1k evaluates pretrained encoders with PCA and ridge regression (Jaume et al., 2024), while PEKA uses a lightweight readout after transcriptomic adaptation (Pan et al., 2025). Benchmark studies examine the limits of both approaches. Wang et al. (2025) report that more complex architectures are not consistently superior, and HESCAPE finds that stronger cross-modal alignment does not necessarily improve direct expression prediction (Gindra et al., 2025).

Spatial variation. Prior work has explicitly considered the gap between predicting expression levels and preserving variation. BLEEP learns a joint image–expression space for retrieval and reports that regression baselines can reproduce gene means while underestimating their variance (Xie et al., 2023). SEPAL predicts deviations from training-set gene means and adds graph-based spatial correction (Mejia et al., 2023). HEXST uses geometry-aware attention and an auxiliary head that matches standardized expression deviations (Byeon & Kwak, 2026). Our analysis focuses on the slide-wise components of the final prediction error under MSE. We study how these components interact with shared parameters and evaluate a penalty applied directly to within-slide prediction error. The resulting objective supports both a closed-form affine readout and augmentation of existing neural objectives.

3 METHOD

3.1 SLIDE-WISE ERROR DECOMPOSITION

Consider M locations from S tissue slides. Location i has an image representation $\mathbf{q}_i \in \mathbb{R}^d$, an expression vector $\mathbf{y}_i \in \mathbb{R}^G$, and slide membership s_i. We stack the representations and targets as $\mathbf{Q} \in \mathbb{R}^{M\times d}$ and $\mathbf{Y} \in \mathbb{R}^{M\times G}$. For predictions $\widehat{\mathbf{Y}}_\theta$, let $\mathbf{E} = \widehat{\mathbf{Y}}_\theta - \mathbf{Y}$. Standard MSE is $\mathcal{L}_{\mathrm{MSE}} = \|\mathbf{E}\|_F^2/(MG)$.

The grouping of locations distinguishes errors in a slide's mean expression from errors in its internal variation. Ordering locations by slide, define

$$\mathbf{P}_s = \frac{1}{n_s}\mathbf{1}_{n_s}\mathbf{1}_{n_s}^\top, \qquad \mathbf{P} = \mathrm{blockdiag}(\mathbf{P}_1, \ldots, \mathbf{P}_S), \qquad \mathbf{C} = \mathbf{I} - \mathbf{P}, \tag{1}$$

where n_s is the number of locations in slide s. For any matrix $\mathbf{Z}$, $\mathbf{PZ}$ repeats the mean of each slide at its locations and $\mathbf{CZ}$ retains deviations from that mean. These complementary orthogonal projections give $\mathbf{E} = \mathbf{PE} + \mathbf{CE}$ and

$$\mathcal{L}_{\mathrm{MSE}} = \underbrace{\frac{\|\mathbf{PE}\|_F^2}{MG}}_{\mathcal{L}_P} + \underbrace{\frac{\|\mathbf{CE}\|_F^2}{MG}}_{\mathcal{L}_C}. \tag{2}$$

Here, $\mathcal{L}_P$ measures error in the mean expression of each gene within each slide. The term $\mathcal{L}_C$ measures error in deviations around those means. A prediction can therefore have small slide-mean error while still missing within-slide variation. The within-slide component is invariant to additive slide-constant shifts, since $\mathbf{C}(\mathbf{Y} + \mathbf{D}) = \mathbf{CY}$ whenever $\mathbf{PD} = \mathbf{D}$. Appendix A.1 gives the derivation.

3.2 COMPONENT-GUIDED LOSS

Orthogonality in prediction space does not make the two errors equally responsive to shared parameters. Let $\mathbf{e}_P = \mathrm{vec}(\mathbf{PE})$ and $\mathbf{e}_C = \mathrm{vec}(\mathbf{CE})$. Denote the prediction Jacobian by $\mathbf{J}_\theta = \partial\,\mathrm{vec}(\widehat{\mathbf{Y}}_\theta)/\partial\theta$. For $j \in \{P, C\}$, the component gradient is $\mathbf{g}_j = \nabla_\theta\mathcal{L}_j = 2\mathbf{J}_\theta^\top\mathbf{e}_j/(MG)$. Its magnitude depends on both residual size and the model's parameter response. For $\mathcal{L}_j > 0$, we separate these effects through the residual-normalized parameter sensitivity

$$\kappa_j = \frac{\|\mathbf{J}_\theta^\top\mathbf{e}_j\|_2^2}{\|\mathbf{e}_j\|_2^2} = \frac{MG}{4}\frac{\|\mathbf{g}_j\|_2^2}{\mathcal{L}_j}. \tag{3}$$

The identity $\|\mathbf{g}_j\|_2^2 = 4\mathcal{L}_j\kappa_j/(MG)$ shows that a component with lower sensitivity produces a smaller gradient for equal residual energy. The relative sensitivities depend on the current residuals and the model Jacobian.

Under MSE gradient flow, $\dot\theta = -(\mathbf{g}_P + \mathbf{g}_C)$, the component losses satisfy

$$\dot{\mathcal{L}}_P = -\|\mathbf{g}_P\|_2^2 - \langle\mathbf{g}_P, \mathbf{g}_C\rangle, \qquad \dot{\mathcal{L}}_C = -\|\mathbf{g}_C\|_2^2 - \langle\mathbf{g}_P, \mathbf{g}_C\rangle. \tag{4}$$

The interaction term contributes equally to the absolute decrease rates. Their difference is therefore $\|\mathbf{g}_P\|_2^2 - \|\mathbf{g}_C\|_2^2$. Equal coefficients in MSE need not yield equal component-wise progress, even though the residuals are orthogonal.

Slide-wise centering gives a useful interpretation of this asymmetry. A parameter change that shifts predictions together within a slide can act on its mean error. The same shared response cancels from the within-slide gradient because centered residuals sum to zero. Reducing $\mathcal{L}_C$ therefore depends on how parameter responses differ among locations. When these differences are weak, within-slide error may produce little gradient despite substantial residual energy. Appendix A.2 derives this cancellation and the component dynamics.

This potential disadvantage motivates placing additional weight on within-slide recovery. We define *Component-Guided Loss* (CGL) as

$$\boxed{\mathcal{L}_{\mathrm{CGL}} = \mathcal{L}_P + (1+\lambda)\mathcal{L}_C = \mathcal{L}_{\mathrm{MSE}} + \lambda\mathcal{L}_C, \qquad \lambda \geq 0.} \tag{5}$$

Its gradient is $\mathbf{g}_P + (1+\lambda)\mathbf{g}_C$. At the same parameter state, CGL adds $\lambda\|\mathbf{g}_C\|_2^2$ to the gradient-flow decrease rate of $\mathcal{L}_C$. This provides a direct within-slide descent contribution while retaining supervision of the slide mean. Appendix A.3 gives the proof. The weight is fixed during fitting and selected on validation data.

3.3 CGL-LINEAR

CGL admits a closed-form solution for a shared affine readout on frozen features. We call this estimator *CGL-Linear*. For training data, let $\boldsymbol{\mu}_Q$ and $\boldsymbol{\mu}_Y$ be the pooled feature and expression means. Define $\mathbf{X} = \mathbf{Q} - \mathbf{1}\boldsymbol{\mu}_Q^\top$ and $\mathbf{Y}_c = \mathbf{Y} - \mathbf{1}\boldsymbol{\mu}_Y^\top$. A single coefficient matrix $\mathbf{A} \in \mathbb{R}^{d\times G}$ gives predictions $\widehat{\mathbf{Y}} = \mathbf{1}\boldsymbol{\mu}_Y^\top + \mathbf{X}\mathbf{A}$.

For residual $\mathbf{R} = \mathbf{X}\mathbf{A} - \mathbf{Y}_c$, the unnormalized data-fitting objective is

$$\mathcal{E}_\lambda(\mathbf{A}) = \|\mathbf{P}\mathbf{R}\|_F^2 + (1+\lambda)\|\mathbf{C}\mathbf{R}\|_F^2 = \langle \mathbf{R}, (\mathbf{I}+\lambda\mathbf{C})\mathbf{R}\rangle_F. \tag{6}$$

Both components are fitted through the same $\mathbf{A}$. Adding L_2 regularization with coefficient $\alpha_{\text{raw}} > 0$ yields

$$\mathbf{A}_{\lambda,\alpha} = \left[\mathbf{X}^\top(\mathbf{I}+\lambda\mathbf{C})\mathbf{X} + \alpha_{\text{raw}}\mathbf{I}\right]^{-1}\mathbf{X}^\top(\mathbf{I}+\lambda\mathbf{C})\mathbf{Y}_c. \tag{7}$$

The coefficients are obtained by a single linear solve. Appendix A.4 derives the solution with an unregularized intercept and specifies the regularization scale.

For new features $\mathbf{Q}_*$, prediction is

$$\widehat{\mathbf{Y}}_* = \mathbf{1}\boldsymbol{\mu}_Y^\top + (\mathbf{Q}_* - \mathbf{1}\boldsymbol{\mu}_Q^\top)\mathbf{A}_{\lambda,\alpha}. \tag{8}$$

Slide membership is needed only for fitting. Inference is pointwise and uses the training means. Setting $\lambda = 0$ recovers ordinary affine ridge regression, so the two estimators have the same predictor class and parameter count. The component weight changes which affine function is fitted within that class. The selection protocol is given in Appendix B.

4 EXPERIMENTS

4.1 EXPERIMENTAL SETUP

Data. We evaluate ten organ-level cohorts from HEST-1k (Jaume et al., 2024), covering brain, breast, colon, heart, kidney, liver, lung, prostate, skin, and uterus. HEST-1k aggregates spatial transcriptomics profiles from public and internal cohorts collected across multiple sources, and the cohorts used here retain this diversity in data origin. Models are fitted separately for each cohort using approximately $70/15/15$ training, validation, and test splits, constructed to avoid patient-level leakage across splits. Training data determine nested panels of $G \in \{500, 1000, 1500, 2000\}$ highly variable genes. Appendix B describes the cohort construction, splits, and preprocessing.

Compared methods. We compare CGL-Linear and ordinary affine ridge regression, denoted Linear, with HiST (Wu et al., 2026), HistoPrism (Hu et al., 2026), HEXST (Byeon & Kwak, 2026), STFlow (Huang et al., 2025), BLEEP (Xie et al., 2023), and SEPAL (Mejia et al., 2023). All methods use the same data splits and targets. We repeat the benchmark with frozen UNI, a pathology foundation model (Chen et al., 2024); DINOv2-Large, a general-purpose self-supervised vision model (Oquab et al., 2024); and ResNet18, a standard convolutional vision model (He et al., 2016). Neural methods are averaged over three seeds, while Linear and CGL-Linear are deterministic. Hyperparameters, including component weights, are selected using validation data following Appendix B.

Evaluation. Following prior H&E-to-spatial-transcriptomics work, we report the Pearson correlation coefficient (PCC) to measure association between predicted and measured expression (Wu et al., 2026; Hu et al., 2026). Flat PCC correlates the complete flattened spot–gene matrices. Gene PCC computes correlation separately for each gene after pooling locations across the evaluated slides,

$$\text{PCC}_{\text{flat}} = \text{corr}\Big(\text{vec}(\mathbf{Y}), \text{vec}(\widehat{\mathbf{Y}})\Big), \quad \text{PCC}_{\text{gene}} = \frac{1}{|\mathcal{V}|}\sum_{g\in\mathcal{V}} \text{corr}\Big(\mathbf{Y}_{:g}, \widehat{\mathbf{Y}}_{:g}\Big),$$

where $\mathcal{V}$ contains genes for which the correlation is defined. Gene PCC measures gene-wise association across pooled locations and can reflect both between-slide and within-slide variation.

Because correlation does not measure the magnitude of prediction error, we additionally report R^2. Flat R^2 measures overall reconstruction, while Local R^2 measures reconstruction after removing the slide-wise mean,

$$R^2_{\text{flat}} = 1 - \frac{\|\widehat{\mathbf{Y}} - \mathbf{Y}\|_F^2}{\|\mathbf{Y} - \bar{y}\mathbf{1}_M\mathbf{1}_G^\top\|_F^2}, \quad R^2_{\text{local}} = 1 - \frac{\|\mathbf{C}(\widehat{\mathbf{Y}} - \mathbf{Y})\|_F^2}{\|\mathbf{C}\mathbf{Y}\|_F^2},$$

with $\bar{y} = (MG)^{-1}\sum_{i,g} Y_{ig}$. Each metric is computed separately for each cohort and then macro-averaged across the ten cohorts.

4.2 Aggregate prediction

We first examine what aggregate scores establish about the prediction task. MeanOnly assigns every location the training-set gene-mean vector, $\widehat{\mathbf{Y}}_{\text{MeanOnly}} = \mathbf{1}\boldsymbol{\mu}_Y^\top$. It uses no image information and contains no within-slide variation, so $R^2_{\text{local}} = 0$. Gene PCC is undefined because each gene has a constant prediction.

Table 1: Mean-only reference and learned predictors with UNI and $G = 500$. Scores are equal-cohort test averages over ten cohorts. Gene PCC is undefined for MeanOnly because its prediction is constant for every gene. Bold marks the highest mean in each column.

Method	Venue	PCC_{flat}	PCC_{gene}	R^2_{flat}	R^2_{local}
MeanOnly	–	0.6921	N.A.	0.4280	0.0000
Linear	–	0.7226	0.1354	0.4953	−0.0423
HiST	ICML'26	0.7363	0.1343	0.5067	0.0805
HistoPrism	ICLR'26	0.7372	0.1337	0.5032	0.0921
HEXST	ICML'26	0.7126	0.1560	0.4547	0.0497
STFlow	ICML'25	0.7053	0.1366	0.4236	−0.0565
BLEEP	NeurIPS'23	0.6766	0.1224	0.4187	−0.1171
SEPAL	ICCVW'23	0.7450	0.1385	0.5070	0.0717
CGL-Linear (Ours)	–	**0.7524**	**0.1656**	**0.5247**	**0.1137**

Table 1 shows that MeanOnly nevertheless obtains substantial Flat PCC and Flat R^2. These scores arise from gene-specific differences in average expression even though the prediction is identical at every spatial location. Linear improves both aggregate measures and produces nonzero gene-wise association, yet its macro Local R^2 remains negative. Strong aggregate performance can therefore coexist with poor reconstruction of variation within tissue slides.

The comparison also clarifies what a simple image-based readout already captures. Frozen image representations combined with a linear predictor recover substantial expression structure beyond the mean-only reference, while much less of the centered within-slide component is recovered. We next examine how these two error components behave during MSE training.

4.3 Optimization sensitivity

We examine whether the sensitivity difference characterized in Section 3.2 appears during learning. A one-hidden-layer MLP with 256 hidden units and GELU activation is trained on frozen DINOv2 (Oquab et al., 2024) and ResNet18 (He et al., 2016) representations with $G = 500$. Standard MSE and CGL use the same architecture, initialization, optimizer, and training budget. The control protocol is given in Appendix B.

Figure 2 shows a persistent sensitivity imbalance under standard MSE. The median κ_P/κ_C remains above one throughout training for both representations, and the same ordering appears on validation slides. Within-slide residuals therefore induce weaker parameter-space responses per unit residual energy in these controls. This empirical ordering identifies the within-slide component as the less responsive component in the setting studied here.

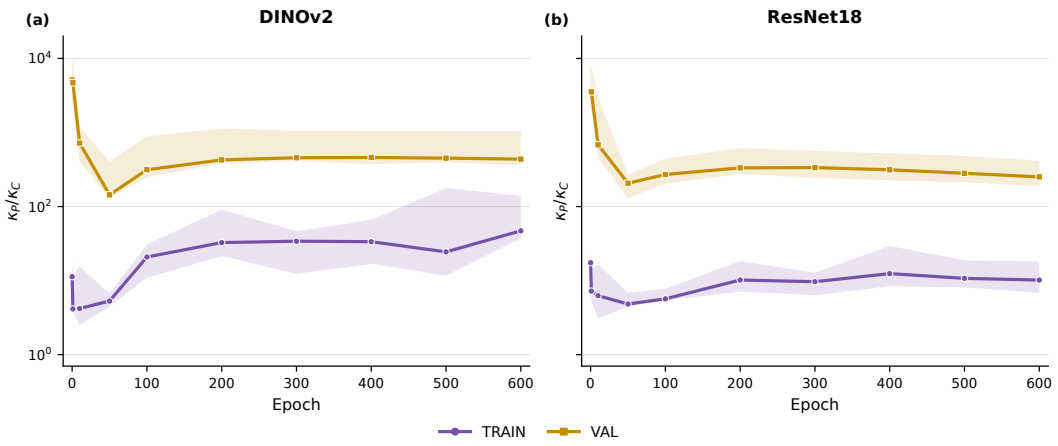

Figure 2: **Component sensitivity under standard MSE.** Curves show the cohort median of κ_P/κ_C for DINOv2 and ResNet18 on training and validation data. Shaded bands show cohort interquartile ranges. Values above one indicate greater residual-normalized sensitivity for the slide-level component.

We then train the same MLP with CGL. Figure 3 shows that the additional within-slide weight changes the allocation of fitting on the training data. CGL reduces $\mathcal{L}_C$ while allowing a larger $\mathcal{L}_P$. On validation slides, the mean difference favors CGL for the within-slide component and, over most of training, also for the slide-level component.

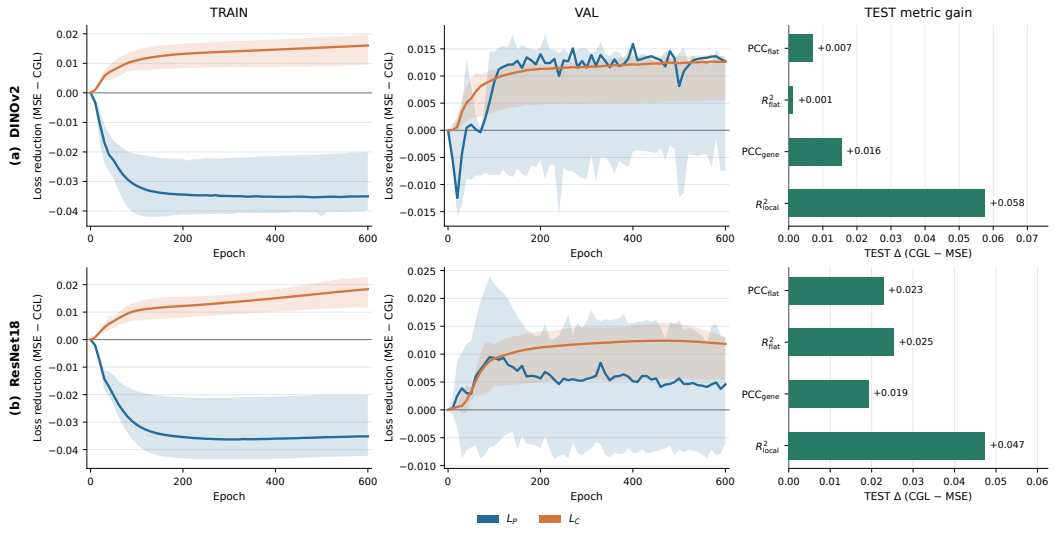

Figure 3: **Effect of CGL on the controlled MLP.** For DINOv2 (a) and ResNet18 (b), the left and middle columns show $\mathcal{L}_j^{\mathrm{MSE}} - \mathcal{L}_j^{\mathrm{CGL}}$ on training and validation data. Positive values indicate lower error under CGL. Lines show equal-cohort means and shaded regions show cohort interquartile ranges. The right column reports test metric changes from MSE to CGL.

The held-out metrics follow the same direction. CGL improves Local R^2 for both representations, directly reflecting better reconstruction of the component emphasized during training. Flat PCC, Gene PCC, and Flat R^2 also improve in these controls. The experiment therefore links the measured sensitivity imbalance to a simple change in supervision that improves held-out prediction.

4.4 BENCHMARK RESULTS

We evaluate CGL-Linear across the full benchmark in Figure 4. It achieves the highest macro $\mathrm{PCC}_{\mathrm{flat}}$ and R^2_{local} throughout the encoder–panel grid. It also leads R^2_{flat} in all but one configuration, where SEPAL (Mejia et al., 2023) is slightly higher. Gene PCC remains among the strongest results throughout the benchmark.

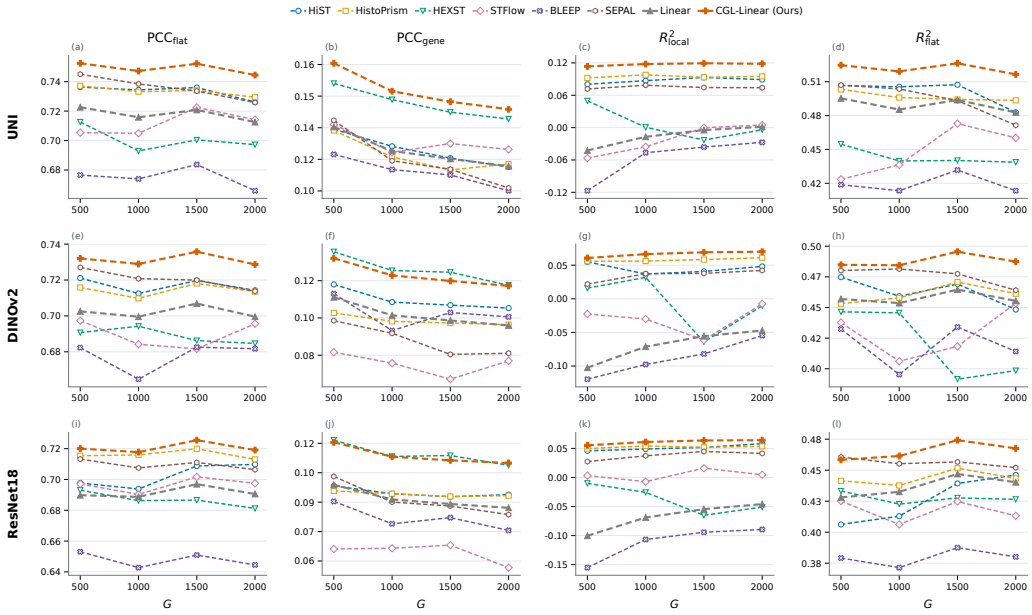

Figure 4: **Benchmark results across image representations and gene panels.** Each point is an equal-cohort macro test score over ten cohorts. Neural methods are averaged over three seeds, while Linear and CGL-Linear are deterministic. All methods use matched image representations, data splits, and targets. Higher is better for every metric.

The comparison with Linear makes the effect of CGL explicit. Linear uses the same affine predictor with the original MSE objective. Across the benchmark, CGL-Linear increases macro Local R^2 by 0.110–0.163 and makes it positive in every configuration. Flat PCC, Gene PCC, and Flat R^2 also improve throughout the grid. The strongest improvement is concentrated in within-slide reconstruction while overall predictive quality is maintained or improved.

The comparison across methods further illustrates the distinction between the evaluation views. HEXST (Byeon & Kwak, 2026) often obtains strong Gene PCC while showing lower Local R^2. SEPAL (Mejia et al., 2023) also performs strongly in several settings after explicitly modeling expression deviations. CGL-Linear remains competitive in gene-wise association while providing consistently stronger within-slide reconstruction. Cohort-level results in Appendix C report the corresponding tissue-specific performance and seed variation.

4.5 COMPUTATIONAL EFFICIENCY

We measure fitting and inference time from $10{,}000$ to $50{,}000$ spatial locations using frozen ResNet18 features with $G = 500$ on an NVIDIA A5000 GPU. CGL-Linear uses $256{,}000$ fitted coefficients and requires only 0.446–1.255 seconds for the complete closed-form fit and 0.039–0.078 seconds for inference across this range.

The computational gap increases with dataset size. At $50{,}000$ spots, CGL-Linear fits in 1.255 seconds, compared with total training times of 28.6 seconds for SEPAL (Mejia et al., 2023), 99.4 seconds for BLEEP (Xie et al., 2023), and 125.3 seconds for HistoPrism (Hu et al., 2026). HiST (Wu et al., 2026), HEXST (Byeon & Kwak, 2026), and STFlow (Huang et al., 2025) require substantially longer training. Inference for CGL-Linear remains below 0.1 seconds throughout the evaluated range.

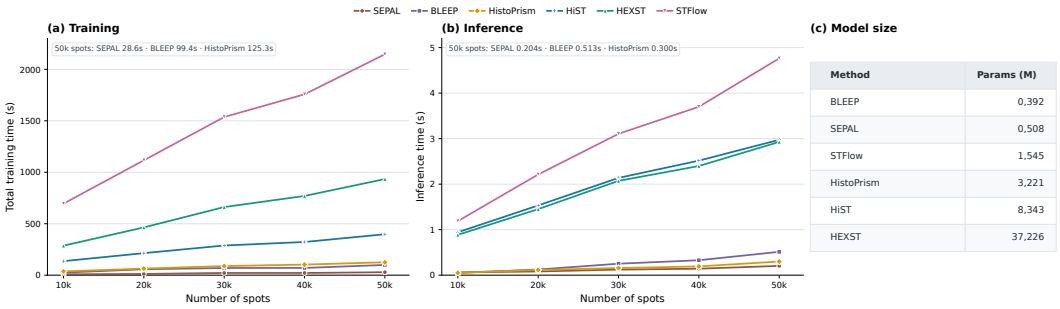

Figure 5: **Computational scaling.** Total training time, inference time, and model size are compared as the number of spatial locations increases.

Together with its strong predictive performance, this efficiency suggests that CGL-Linear could serve as a fast and lightweight reference baseline for H&E-to-spatial-transcriptomics research.

4.6 Transfer to existing models

We next examine whether within-slide supervision transfers to existing H&E-to-spatial-transcriptomics models whose training objectives include a squared-error regression component. We augment HiST (Wu et al., 2026), HistoPrism (Hu et al., 2026), HEXST (Byeon & Kwak, 2026), STFlow (Huang et al., 2025), and SEPAL (Mejia et al., 2023) using

$$\mathcal{L}_{\text{aug}} = \mathcal{L}_{\text{base}} + \lambda \mathcal{L}_C.$$

Each method retains its original architecture and full training objective, with $\mathcal{L}_C$ added as an auxiliary within-slide term. The component weight is selected using validation data for each method. Additional implementation details are given in Appendix B.1.

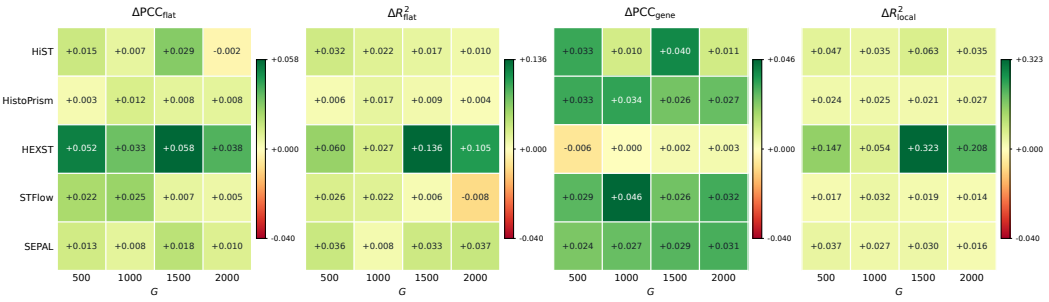

Figure 6: **Adding within-slide supervision to existing models.** Cells show equal-cohort macro test gains over the corresponding base method across ten cohorts using DINOv2 representations. Positive values favor the augmented objective.

Figure 6 shows positive Local R^2 gains for every method and gene panel evaluated. Most settings also improve the aggregate and gene-wise metrics, with only small regressions in a few comparisons. The effect therefore extends beyond the affine readout and the controlled MLP to substantially different model families.

The result is also informative for methods that already model expression variation. HEXST (Byeon & Kwak, 2026) shows large improvements in Local R^2 while Gene PCC changes little. SEPAL (Mejia et al., 2023) also improves despite explicitly predicting deviations from training-set gene means. Its deviation target removes a cohort-level gene baseline, whereas $\mathcal{L}_C$ measures prediction error around each slide mean. The two forms of supervision are therefore distinct.

Across hierarchical Transformers, graph-based predictors, and flow-based models, the same within-slide term can be incorporated without changing the underlying architecture or replacing the original

objective. These results demonstrate the practical flexibility of component-guided supervision across different H&E-to-spatial-transcriptomics model designs that already rely on squared-error learning.

4.7 Spatial predictions

We inspect two tissue-relevant genes to relate the quantitative results to spatial expression patterns. MUC2 encodes a gel-forming mucin secreted by intestinal goblet cells and is a major component of the colonic mucus layer (Johansson et al., 2011). MYH6 encodes the cardiac α-myosin heavy chain, a major contractile myosin isoform with predominant expression in human atrial myocardium (Reiser et al., 2001; Barrick & Greenberg, 2021). We examine MUC2 in held-out colon slide MISC71 and MYH6 in held-out heart slide MISC138.

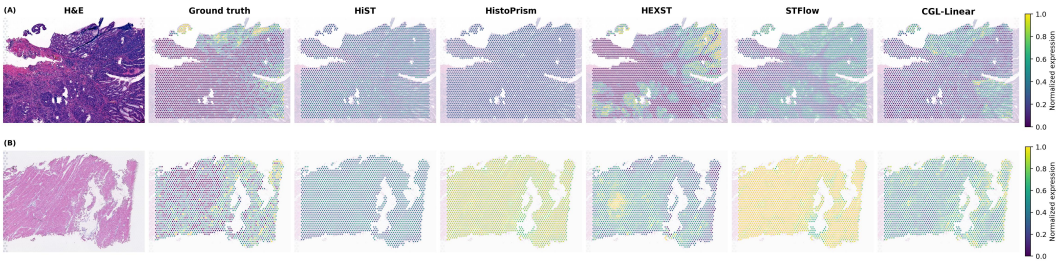

Figure 7: **Spatial expression predictions on held-out slides.** Row (A) shows MUC2 in colon slide MISC71, and row (B) shows MYH6 in heart slide MISC138. Each row contains the H&E image, measured expression, and predictions from the displayed methods.

For MUC2, the measured expression contains localized high-expression regions along the upper and right portions of the tissue. Several predictions capture the broad pattern but attenuate its contrast or place elevated expression in different regions. CGL-Linear follows the measured spatial arrangement more closely in this example and preserves the separation between these regions and the lower-expression interior.

MYH6 exhibits a broader expression pattern with marked regional variation across the heart section. Some methods produce relatively uniform predictions or shift the regions of high expression. CGL-Linear retains more of the measured contrast between lower-expression regions on the left and higher-expression regions toward the upper and right parts of the section. These examples complement the quantitative results by showing how differences in within-slide reconstruction appear spatially.

5 Conclusion

We studied H&E-to-spatial-transcriptomics prediction by separating mean-squared error into slide-level and within-slide components. This view reveals that strong aggregate prediction can coexist with weak within-slide recovery and, in controlled neural experiments, that the within-slide component is less responsive to shared parameters under standard MSE training. Component-Guided Loss directly increases supervision of this component. Its closed-form affine instantiation, CGL-Linear, substantially improves over ordinary linear regression, achieves overall state-of-the-art performance across HEST-1k cohorts, and could serve as a fast and lightweight reference baseline. The same within-slide supervision also improves several existing model families. We hope this perspective provides a basis for assessing how much additional predictive signal can be recovered by better exploiting the structure of the prediction problem before resorting to additional model complexity.

The current analysis is specific to squared-error prediction objectives. The exact decomposition and sensitivity characterization are derived for MSE, and the empirical sensitivity imbalance is established in controlled MSE-trained models. Although adding the within-slide term also improves models with more complex objectives, we do not establish that the same optimization behavior holds for arbitrary training losses. Extending the component-level analysis beyond squared-error objectives is therefore an important direction for future work.

REPRODUCIBILITY STATEMENT

Appendix A derives the error decomposition, sensitivity identities, and CGL-Linear estimator. Appendix B describes experimental configurations. Appendix C reports detailed cohort results.

AI USE STATEMENT

Generative AI tools were used to assist with literature retrieval and discovery, and to improve the clarity and readability of the manuscript. All AI-assisted writing and retrieved information were reviewed and verified by the authors. The authors take responsibility for the final content of this work.

REFERENCES

Samantha K. Barrick and Michael J. Greenberg. Cardiac myosin contraction and mechanotransduction in health and disease. *Journal of Biological Chemistry*, 297(5):101297, 2021. doi: 10.1016/j.jbc.2021.101297.

Keunho Byeon and Jin Tae Kwak. HEXST: Hexagonal shifted-window transformer for spatial transcriptomics gene expression prediction. In *Proceedings of the 43rd International Conference on Machine Learning*, 2026. URL https://arxiv.org/abs/2605.04682.

Richard J. Chen, Tong Ding, Ming Y. Lu, Drew F. K. Williamson, Guillaume Jaume, Andrew H. Song, Bowen Chen, Andrew Zhang, Daniel Shao, et al. Towards a general-purpose foundation model for computational pathology. *Nature Medicine*, 30, 2024. doi: 10.1038/s41591-024-02857-3. URL https://www.nature.com/articles/s41591-024-02857-3.

Rushin H. Gindra, Giovanni Palla, Mathias Nguyen, Sophia J. Wagner, Manuel Tran, Fabian J. Theis, Dieter Saur, Lorin Crawford, and Tingying Peng. A large-scale benchmark of cross-modal learning for histology and gene expression in spatial transcriptomics. In *Proceedings of the IEEE/CVF International Conference on Computer Vision Workshops*, pp. 1193–1203, 2025. doi: 10.1109/ICCVW69036.2025.00128. URL https://openaccess.thecvf.com/content/ICCV2025W/CVAMD/html/Gindra_A_Large-Scale_Benchmark_of_Cross-Modal_Learning_for_Histology_and_Gene_ICCVW_2025_paper.html.

Bryan He, Ludvig Bergenstråhle, Linnea Stenbeck, Abubakar Abid, Alma Andersson, Åke Borg, Jonas Maaskola, Joakim Lundeberg, and James Zou. Integrating spatial gene expression and breast tumour morphology via deep learning. *Nature Biomedical Engineering*, 4:827–834, 2020. doi: 10.1038/s41551-020-0578-x.

Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 770–778, 2016. URL https://openaccess.thecvf.com/content_cvpr_2016/html/He_Deep_Residual_Learning_CVPR_2016_paper.html.

Susu Hu, Qinghe Zeng, Nithya Bhasker, Jakob Nikolas Kather, and Stefanie Speidel. Histoprism: Unlocking functional pathway analysis from pan-cancer histology via gene expression prediction. In *The Fourteenth International Conference on Learning Representations*, 2026. URL https://openreview.net/forum?id=6dTHxb9JuA.

Tinglin Huang, Tianyu Liu, Mehrtash Babadi, Wengong Jin, and Rex Ying. Scalable generation of spatial transcriptomics from histology images via whole-slide flow matching. In *Forty-second International Conference on Machine Learning*, 2025. URL https://openreview.net/forum?id=Ossg1IbHDT.

Guillaume Jaume, Paul Doucet, Andrew H. Song, Ming Y. Lu, Cristina Almagro-Pérez, Sophia J. Wagner, Anurag J. Vaidya, Richard J. Chen, Drew F. K. Williamson, Ahrong Kim, and Faisal Mahmood. HEST-1k: A dataset for spatial transcriptomics and histology image analysis. In *Advances in Neural Information Processing Systems: Datasets and Benchmarks Track*, volume 37, 2024. doi: 10.52202/079017-1704. URL https://proceedings.neurips.cc/paper_files/paper/2024/hash/60a899cc31f763be0bde781a75e04458-Abstract-Datasets_and_Benchmarks_Track.html.

Malin E. V. Johansson, Jessica M. Holmén Larsson, and Gunnar C. Hansson. The two mucus layers of colon are organized by the muc2 mucin, whereas the outer layer is a legislator of host-microbial interactions. *Proceedings of the National Academy of Sciences*, 108(Suppl 1):4659–4665, 2011. doi: 10.1073/pnas.1006451107.

Gabriel Mejia, Paula Cárdenas, Daniela Ruiz, Angela Castillo, and Pablo Arbeláez. SEPAL: Spatial gene expression prediction from local graphs. In *Proceedings of the IEEE/CVF International Conference on Computer Vision Workshops*, pp. 2294–2303, 2023. URL https://openaccess.thecvf.com/content/ICCV2023W/CVAMD/

html/Mejia_SEPAL_Spatial_Gene_Expression_Prediction_from_Local_Graphs_ICCVW_2023_paper.html.

Maxime Oquab, Timothée Darcet, Théo Moutakanni, Huy V. Vo, Marc Szafraniec, Vasil Khalidov, Pierre Fernandez, Daniel Haziza, Francisco Massa, Alaaeldin El-Nouby, et al. DINOv2: Learning robust visual features without supervision. *Transactions on Machine Learning Research*, 2024. URL https://openreview.net/forum?id=a68SUt6zFt.

Shi Pan, Jianan Chen, and Maria Secrier. Teaching pathology foundation models to accurately predict gene expression with parameter efficient knowledge transfer. In *Medical Image Computing and Computer Assisted Intervention – MICCAI 2025*, volume 15966 of *Lecture Notes in Computer Science*, pp. 605–613. Springer, 2025. doi: 10.1007/978-3-032-04981-0_57.

Peter J. Reiser, Michael A. Portman, Xue-Han Ning, and Christine Schomisch Moravec. Human cardiac myosin heavy chain isoforms in fetal and failing adult atria and ventricles. *American Journal of Physiology-Heart and Circulatory Physiology*, 280(4):H1814–H1820, 2001. doi: 10.1152/ajpheart.2001.280.4.H1814.

Patrik L. Ståhl, Fredrik Salmén, Sanja Vickovic, Anna Lundmark, José Fernández Navarro, Jens Magnusson, Stefania Giacomello, Michaela Asp, Jakub O. Westholm, Mikael Huss, et al. Visualization and analysis of gene expression in tissue sections by spatial transcriptomics. *Science*, 353(6294): 78–82, 2016. doi: 10.1126/science.aaf2403.

Chuhan Wang, Adam S. Chan, Xiaohang Fu, Shila Ghazanfar, Jinman Kim, Ellis Patrick, and Jean Y. H. Yang. Benchmarking the translational potential of spatial gene expression prediction from histology. *Nature Communications*, 16:1544, 2025. doi: 10.1038/s41467-025-56618-y. URL https://www.nature.com/articles/s41467-025-56618-y.

Weiyi Wu, Xinwen Xu, Xingjian Diao, Siting Li, Zhi Wei, Alma Andersson, and Jiang Gui. HiST: A hierarchical sparse transformer for cross-modal spatial transcriptomics modeling. In *Proceedings of the 43rd International Conference on Machine Learning*, 2026. URL https://openreview.net/forum?id=ptbzlHzmEv.

Ronald Xie, Kuan Pang, Sai Chung, Catia Perciani, Sonya MacParland, Bo Wang, and Gary D. Bader. Spatially resolved gene expression prediction from histology images via bi-modal contrastive learning. In *Advances in Neural Information Processing Systems*, volume 36, 2023. URL https://proceedings.neurips.cc/paper_files/paper/2023/hash/df656d6ed77b565e8dcdfbf568aead0a-Abstract.html.

A MATHEMATICAL DETAILS

We use the notation of Section 3.1 and write $N = MG$. Observations and slide membership remain fixed during differentiation.

A.1 SLIDE-WISE PROJECTIONS

Each block $\mathbf{P}_s = n_s^{-1}\mathbf{1}_{n_s}\mathbf{1}_{n_s}^\top$ is symmetric, and

$$\mathbf{P}_s^2 = \frac{\mathbf{1}_{n_s}(\mathbf{1}_{n_s}^\top \mathbf{1}_{n_s})\mathbf{1}_{n_s}^\top}{n_s^2} = \mathbf{P}_s.$$

Consequently, $\mathbf{P}$ is an orthogonal projector. Its complement $\mathbf{C} = \mathbf{I} - \mathbf{P}$ satisfies $\mathbf{C}^\top = \mathbf{C}$, $\mathbf{C}^2 = \mathbf{C}$, and $\mathbf{PC} = \mathbf{CP} = \mathbf{0}$. For any matrix $\mathbf{Z}$, $\langle \mathbf{PZ}, \mathbf{CZ}\rangle_F = \mathrm{tr}(\mathbf{Z}^\top \mathbf{PCZ}) = 0$. Thus

$$\|\mathbf{Z}\|_F^2 = \|\mathbf{PZ}\|_F^2 + \|\mathbf{CZ}\|_F^2. \tag{9}$$

Taking $\mathbf{Z} = \mathbf{E}$ proves Eq. (2).

Let $\mathbf{e}_{si} = \widehat{\mathbf{y}}_{si} - \mathbf{y}_{si}$ and $\bar{\mathbf{e}}_s = n_s^{-1}\sum_i \mathbf{e}_{si}$. Then

$$\mathcal{L}_P = \frac{1}{N}\sum_s n_s \|\bar{\mathbf{e}}_s\|_2^2, \qquad \mathcal{L}_C = \frac{1}{N}\sum_s \sum_i \|\mathbf{e}_{si} - \bar{\mathbf{e}}_s\|_2^2. \tag{10}$$

These expressions separate error in slide-wise means from error in within-slide deviations. If $\mathbf{PD} = \mathbf{D}$, then $\mathbf{CD} = \mathbf{0}$, so $\mathbf{C}(\mathbf{Y} + \mathbf{D}) = \mathbf{CY}$.

For MeanOnly, $\mathbf{C}\widehat{\mathbf{Y}} = \mathbf{0}$. Therefore $\mathbf{C}(\widehat{\mathbf{Y}} - \mathbf{Y}) = -\mathbf{CY}$, giving $R^2_{\mathrm{local}} = 0$ whenever $\|\mathbf{CY}\|_F^2 > 0$. This identity is independent of how close the training gene means are to those in the evaluated split.

A.2 PARAMETER SENSITIVITY

Use column-wise vectorization and let $\mathbf{\Pi}_P = \mathbf{I}_G \otimes \mathbf{P}$, $\mathbf{\Pi}_C = \mathbf{I}_G \otimes \mathbf{C}$, and $\mathbf{e}_j = \mathbf{\Pi}_j \,\mathrm{vec}(\mathbf{E})$. With $\mathbf{J}_\theta = \partial\,\mathrm{vec}(\widehat{\mathbf{Y}}_\theta)/\partial\theta$, differentiation gives

$$\mathbf{g}_j = \nabla_\theta \mathcal{L}_j = \frac{2}{N}\mathbf{J}_\theta^\top \mathbf{e}_j, \qquad j \in \{P, C\}. \tag{11}$$

For $\mathcal{L}_j > 0$, the definition of κ_j implies

$$\|\mathbf{g}_j\|_2^2 = \frac{4}{N}\mathcal{L}_j \kappa_j, \qquad \frac{\|\mathbf{g}_P\|_2^2}{\|\mathbf{g}_C\|_2^2} = \frac{\mathcal{L}_P}{\mathcal{L}_C}\frac{\kappa_P}{\kappa_C}, \tag{12}$$

where the ratio requires both component errors to be nonzero and $\|\mathbf{g}_C\|_2 > 0$. If $\mathcal{L}_j = 0$, then $\mathbf{g}_j = \mathbf{0}$ and the normalized sensitivity is undefined. Orthogonal residuals need not yield orthogonal gradients because $\langle \mathbf{g}_P, \mathbf{g}_C\rangle = 4\mathbf{e}_P^\top \mathbf{J}_\theta \mathbf{J}_\theta^\top \mathbf{e}_C / N^2$.

The grouping gives a useful interpretation. Let $\mathbf{J}_{si} = \partial\widehat{\mathbf{y}}_{si}/\partial\theta$ and $\bar{\mathbf{J}}_s = n_s^{-1}\sum_i \mathbf{J}_{si}$. Since $\sum_i(\mathbf{e}_{si} - \bar{\mathbf{e}}_s) = \mathbf{0}$, Eq. (11) becomes

$$\mathbf{g}_P = \frac{2}{N}\sum_s n_s \bar{\mathbf{J}}_s^\top \bar{\mathbf{e}}_s, \tag{13}$$

$$\mathbf{g}_C = \frac{2}{N}\sum_s \sum_i (\mathbf{J}_{si} - \bar{\mathbf{J}}_s)^\top (\mathbf{e}_{si} - \bar{\mathbf{e}}_s). \tag{14}$$

Parameter responses common to a slide cancel from its contribution to $\mathbf{g}_C$. If responses are identical within every slide, $\mathbf{g}_C = \mathbf{0}$ even when within-slide error is nonzero. In general, the relative sensitivities depend on the response geometry and its alignment with the residuals. Slide membership alone does not determine their ordering.

Under standard MSE gradient flow, $\dot\theta = -(\mathbf{g}_P + \mathbf{g}_C)$. Applying the chain rule yields

$$\begin{aligned}\dot{\mathcal{L}}_P &= -\|\mathbf{g}_P\|_2^2 - \langle \mathbf{g}_P, \mathbf{g}_C\rangle,\\ \dot{\mathcal{L}}_C &= -\|\mathbf{g}_C\|_2^2 - \langle \mathbf{g}_P, \mathbf{g}_C\rangle.\end{aligned} \tag{15}$$

Hence $(-\dot{\mathcal{L}}_P) - (-\dot{\mathcal{L}}_C) = \|\mathbf{g}_P\|_2^2 - \|\mathbf{g}_C\|_2^2$. The sensitivity κ_j determines the self-descent contribution per unit component error. The full relative decrease also depends on $\langle \mathbf{g}_P, \mathbf{g}_C\rangle/\mathcal{L}_j$.

A.3 COMPONENT REWEIGHTING

At a fixed parameter state, let $\mathbf{v}_\lambda = -\mathbf{g}_P - (1+\lambda)\mathbf{g}_C$. Define $D_j(\lambda) = -\langle \mathbf{g}_j, \mathbf{v}_\lambda \rangle$, the instantaneous decrease rate of the unweighted component $\mathcal{L}_j$ along this direction. Direct subtraction gives

$$D_C(\lambda) - D_C(0) = \lambda \|\mathbf{g}_C\|_2^2, \tag{16}$$

$$D_P(\lambda) - D_P(0) = \lambda \langle \mathbf{g}_P, \mathbf{g}_C \rangle. \tag{17}$$

For $\lambda \geq 0$, CGL adds a nonnegative within-slide descent contribution. The change in slide-level progress depends on gradient alignment. This comparison holds at the same parameter state under gradient flow. It does not order complete training trajectories or held-out errors.

A.4 CGL-LINEAR

All fitting quantities refer to the training set. Write $\mathbf{W}_\lambda = \mathbf{I} + \lambda \mathbf{C}$ and $\mathbf{E}_{\mathbf{A},\mathbf{b}} = \mathbf{Q}\mathbf{A} + \mathbf{1}\mathbf{b}^\top - \mathbf{Y}$. The regularized affine problem is

$$\min_{\mathbf{A},\mathbf{b}} \langle \mathbf{E}_{\mathbf{A},\mathbf{b}}, \mathbf{W}_\lambda \mathbf{E}_{\mathbf{A},\mathbf{b}} \rangle_F + \alpha_{\text{raw}} \|\mathbf{A}\|_F^2. \tag{18}$$

The intercept is unregularized. Since $\mathbf{W}_\lambda \mathbf{1} = \mathbf{1}$, its first-order condition is $\mathbf{E}_{\mathbf{A},\mathbf{b}}^\top \mathbf{1} = \mathbf{0}$. This gives $\mathbf{b} = \boldsymbol{\mu}_Y - \mathbf{A}^\top \boldsymbol{\mu}_Q$, with $\boldsymbol{\mu}_Q = \mathbf{Q}^\top \mathbf{1}/M$ and $\boldsymbol{\mu}_Y = \mathbf{Y}^\top \mathbf{1}/M$.

Substituting $\mathbf{X} = \mathbf{Q} - \mathbf{1}\boldsymbol{\mu}_Q^\top$ and $\mathbf{Y}_c = \mathbf{Y} - \mathbf{1}\boldsymbol{\mu}_Y^\top$ reduces Eq. (18) to

$$\min_{\mathbf{A}} \|\mathbf{P}(\mathbf{X}\mathbf{A} - \mathbf{Y}_c)\|_F^2 + (1+\lambda)\|\mathbf{C}(\mathbf{X}\mathbf{A} - \mathbf{Y}_c)\|_F^2 + \alpha_{\text{raw}} \|\mathbf{A}\|_F^2. \tag{19}$$

Its stationarity condition is $(\mathbf{X}^\top \mathbf{W}_\lambda \mathbf{X} + \alpha_{\text{raw}} \mathbf{I})\mathbf{A} = \mathbf{X}^\top \mathbf{W}_\lambda \mathbf{Y}_c$, yielding

$$\mathbf{A}_{\lambda,\alpha} = \left[\mathbf{X}^\top (\mathbf{I} + \lambda \mathbf{C})\mathbf{X} + \alpha_{\text{raw}} \mathbf{I}\right]^{-1} \mathbf{X}^\top (\mathbf{I} + \lambda \mathbf{C}) \mathbf{Y}_c. \tag{20}$$

For $\lambda \geq 0$ and $\alpha_{\text{raw}} > 0$, the coefficient matrix is positive definite, so the solution is unique. The inverse notation represents a linear solve. Dividing Eq. (19) by N gives the normalized CGL objective plus $(\alpha_{\text{raw}}/N)\|\mathbf{A}\|_F^2$.

The implemented penalty convention is $\alpha_{\text{raw}} = \alpha \operatorname{tr}(\mathbf{X}^\top \mathbf{X})/d$, computed from unweighted training features independently of λ. At $\lambda = 0$, the estimator is ordinary affine ridge regression. For a new feature vector $\mathbf{q}_*$, prediction is $\widehat{\mathbf{y}}_* = \boldsymbol{\mu}_Y + \mathbf{A}_{\lambda,\alpha}^\top (\mathbf{q}_* - \boldsymbol{\mu}_Q)$. All fitted quantities come from training data; neither slide projection is required at inference.

B EXPERIMENTAL PROTOCOL

Data preparation The benchmark comprises ten organ-level cohorts derived from HEST-1k (Jaume et al., 2024). A separate model is fitted for each cohort. Fixed biological-unit-disjoint splits separate training, validation, and test data. Table 2 lists the aligned slides and locations.

Table 2: Cohort sizes in the fixed canonical split. Slide counts are recovered from the 500-gene evaluation support; spot counts are the aligned H&E–expression pairs used by the experiments.

Cohort	Slides (train / val / test)			Spots (train / val / test)		
Brain	9	2	3	23,780	4,891	4,219
Colon	33	4	4	94,871	6,359	15,138
Heart	26	7	8	70,802	14,100	14,195
Kidney	24	5	5	27,739	5,649	5,691
Liver	10	2	2	26,518	5,695	4,955
Lung	12	4	4	20,810	8,763	8,933
Prostate	23	2	2	62,481	7,898	7,784
Skin	42	7	10	36,585	7,984	7,488
Uterus	10	3	2	12,503	2,090	2,351
Breast	65	15	15	26,752	5,988	5,920
Total	254	51	55	402,841	69,417	76,674

Highly variable genes are selected from training data using `seurat_v3`, with slide identity as the batch key. The panels of 500, 1,000, 1,500, and 2,000 genes are nested. All compared methods use these same panels and splits. Scores therefore refer to this common benchmark, not the original HEST-Benchmark configuration.

Image representations Frozen UNI, DINOv2-Large, and ResNet18 representations have dimensions 1,024, 1,024, and 512, respectively. Features are extracted once per encoder and shared across the compared methods. Each method retains its downstream processing.

Model selection Model configurations are selected with DINOv2-Large using equal-cohort macro validation MSE, then held fixed for UNI and ResNet18. Checkpoints are selected using validation MSE. The main neural benchmark uses seeds $0, 1, 2$. Linear and CGL-Linear are deterministic once their hyperparameters have been selected. Test scores do not enter model selection.

Both affine estimators use the penalty convention in Appendix A.4. Linear has $\lambda = 0$; CGL-Linear uses a validation-selected $\lambda \geq 0$. Regularization is also selected on validation data. Table 3 summarizes the model configurations.

Optimization controls The controls use a single hidden layer with 256 units, GELU activation, and an affine output layer. Frozen DINOv2-Large and ResNet18 features predict $G = 500$ genes. The MSE and CGL runs start from identical parameters and use full-gradient Adam at learning rate 10^{-3} for 600 updates, with seed 0 and no parameter weight decay. The output bias is initialized with training gene means. CGL uses $\lambda = 30$, chosen on validation data.

Component losses use the normalization in Eq. (2). At each recorded state, the gradients of $\mathcal{L}_P$ and $\mathcal{L}_C$ are evaluated on the complete training or validation split. Validation gradients are diagnostic only and do not update the predictor. Sensitivity curves summarize cohort medians and interquartile ranges. Loss-difference curves summarize equal-cohort means and interquartile ranges. These bands describe variation across cohorts, not confidence intervals. The gradient-flow identities describe Euclidean gradient directions; the observed Adam trajectories are evaluated empirically.

Table 3: Configurations of the compared predictors. LR denotes learning rate and WD denotes weight decay.

Method	Epochs	Readout configuration	Optimization
HiST	300	Sparse hierarchy 111–21, window size 3, feed-forward ratio 4, three global tokens.	AdamW, LR 10^{-4}, WD 10^{-5}, cosine schedule, centered-target MSE.
HistoPrism	300	Hidden size 256, two Transformer layers, eight heads, dropout 0.1.	AdamW, LR 5×10^{-5}, WD 10^{-2}, linear decay with 5% warm-up, centered-target MSE.
HEXST	300	Hidden size 512, HexMSA/HexRoPE, 3,072-dimensional scFoundation alignment, dropout 0.2.	Adam, LR 4×10^{-5}, WD 10^{-4}, warm-up and cosine schedule, composite regression/correlation/deviation/alignment objective.
STFlow	300	Hidden size 128, four layers, four heads, eight neighbors, five sampling steps.	Adam, LR 5×10^{-4}, gradient clipping at 1.0, no scheduler, conditional flow-matching objective.
BLEEP	300	Dual 256-dimensional projections, dropout 0.1, temperature 1, top-50 retrieval.	AdamW, LR 10^{-3}, WD 10^{-3}, batch size 512, symmetric contrastive cross-entropy.
SEPAL	300	Image readout followed by a six-neighbor graph head, relative positional encoding, two GCN layers, SAGPooling.	SGD, LR 0.0099913, momentum 0.9, batch size 32, MSE in both stages, frozen first-stage predictor during graph fitting.
CGL-Linear	–	One affine readout.	Direct CGL solve, $\lambda = 5$ and $\alpha = 5$.

B.1 Neural augmentation

The transfer experiments use DINOv2-Large across the four gene panels. HiST, HistoPrism, HEXST, STFlow, and SEPAL retain their original architectures and training objectives, with the within-slide term added as

$$\mathcal{L}_{\text{aug}} = \mathcal{L}_{\text{base}} + \lambda \mathcal{L}_C.$$

For each method, we search $\lambda \in \{5, 10, 15, 20\}$ and select the value with the lowest equal-cohort macro validation MSE. Test performance is reported at the corresponding validation-selected checkpoint.

Table 4: Validation-selected component weights used in the neural augmentation experiments.

Method	λ
HiST	10
HistoPrism	15
HEXST	20
STFlow	10
SEPAL	5

The reported heatmap compares each augmented model with its corresponding base model across the four gene panels, yielding 20 paired method–panel comparisons.

C COHORT-LEVEL RESULTS

The following tables report the four evaluation metrics for all three encoders and four gene panels. Neural entries report mean and standard deviation over three seeds (0, 1, 2). Linear and CGL-Linear are deterministic. Macro rows average cohort means equally. Bold and underline identify the best and second-best unrounded means within each row. Display rounding can produce apparent ties. All values are generated from the accompanying benchmark CSV.

Table 5: Cohort-level test results with UNI and $G = 500$. Neural entries are mean ± seed standard deviation over three runs. Linear and CGL-Linear are deterministic. Bold and underlining mark the best and second-best unrounded means in each row. Macro rows average cohort means without weighting by cohort size. Higher is better.

Cohort	HiST	HistoPrism	HEXST	STFlow	BLEEP	SEPAL	Linear	CGL-Linear
$\mathrm{PCC}_{\mathrm{flat}}$ ↑								
Brain	0.775 ± 0.001	0.772 ± 0.004	0.764 ± 0.001	0.754 ± 0.010	0.702 ± 0.026	0.762 ± 0.008	0.735	**0.776**
Colon	0.744 ± 0.018	0.731 ± 0.009	0.719 ± 0.003	0.710 ± 0.002	0.622 ± 0.068	0.713 ± 0.011	0.713	**0.758**
Heart	0.789 ± 0.009	**0.804** ± 0.007	0.647 ± 0.002	0.791 ± 0.005	0.771 ± 0.009	0.768 ± 0.018	0.762	0.736
Kidney	0.734 ± 0.006	0.714 ± 0.000	0.719 ± 0.000	0.676 ± 0.004	0.702 ± 0.014	0.745 ± 0.011	0.727	**0.747**
Liver	0.864 ± 0.003	0.871 ± 0.002	0.858 ± 0.001	0.872 ± 0.000	0.858 ± 0.001	0.858 ± 0.004	0.862	**0.874**
Lung	0.622 ± 0.018	0.593 ± 0.037	0.655 ± 0.003	0.645 ± 0.043	0.564 ± 0.006	0.663 ± 0.005	**0.663**	0.658
Prostate	0.659 ± 0.003	0.685 ± 0.013	**0.726** ± 0.003	0.545 ± 0.068	0.476 ± 0.022	0.701 ± 0.023	0.591	0.716
Skin	0.680 ± 0.009	**0.710** ± 0.007	0.516 ± 0.008	0.659 ± 0.006	0.640 ± 0.005	0.667 ± 0.039	0.676	0.707
Uterus	0.827 ± 0.019	0.817 ± 0.011	0.849 ± 0.003	0.805 ± 0.028	0.807 ± 0.010	0.866 ± 0.006	0.835	**0.871**
Breast	0.669 ± 0.003	0.676 ± 0.006	0.674 ± 0.000	0.596 ± 0.016	0.624 ± 0.012	**0.707** ± 0.009	0.661	0.681
Macro	0.736	0.737	0.713	0.705	0.677	0.745	0.723	**0.752**
$\mathrm{PCC}_{\mathrm{gene}}$ ↑								
Brain	0.045 ± 0.004	0.018 ± 0.005	0.037 ± 0.013	0.040 ± 0.003	0.040 ± 0.004	0.010 ± 0.000	0.040	**0.051**
Colon	0.205 ± 0.004	0.198 ± 0.003	**0.248** ± 0.008	0.185 ± 0.002	0.142 ± 0.018	0.184 ± 0.022	0.195	0.227
Heart	0.053 ± 0.004	0.050 ± 0.006	0.042 ± 0.007	**0.069** ± 0.001	0.048 ± 0.001	0.036 ± 0.003	0.058	0.061
Kidney	**0.105** ± 0.010	0.088 ± 0.015	0.099 ± 0.006	0.077 ± 0.001	0.087 ± 0.009	0.081 ± 0.024	0.092	0.105
Liver	0.091 ± 0.001	0.082 ± 0.004	0.039 ± 0.002	0.104 ± 0.010	0.088 ± 0.003	0.062 ± 0.004	0.088	**0.107**
Lung	0.231 ± 0.011	0.235 ± 0.006	0.332 ± 0.024	0.313 ± 0.037	0.220 ± 0.004	0.296 ± 0.052	0.288	**0.347**
Prostate	0.045 ± 0.002	0.074 ± 0.018	0.071 ± 0.033	0.027 ± 0.021	0.040 ± 0.009	**0.075** ± 0.006	0.002	0.074
Skin	0.390 ± 0.001	0.416 ± 0.014	**0.453** ± 0.016	0.401 ± 0.012	0.397 ± 0.002	0.385 ± 0.046	0.409	0.444
Uterus	0.092 ± 0.022	0.078 ± 0.004	0.137 ± 0.003	0.095 ± 0.021	0.105 ± 0.008	0.095 ± 0.003	0.111	**0.155**
Breast	0.086 ± 0.004	0.099 ± 0.007	0.103 ± 0.004	0.054 ± 0.015	0.057 ± 0.011	**0.161** ± 0.027	0.070	0.085
Macro	0.134	0.134	0.156	0.137	0.122	0.138	0.135	**0.166**
R^2_{local} ↑								
Brain	−0.081 ± 0.015	−0.040 ± 0.009	**0.012** ± 0.005	−0.224 ± 0.087	−0.472 ± 0.027	0.002 ± 0.002	−0.290	−0.021
Colon	**0.158** ± 0.016	0.120 ± 0.015	0.094 ± 0.008	0.028 ± 0.002	−0.215 ± 0.034	0.028 ± 0.031	−0.138	0.153
Heart	−0.013 ± 0.002	**0.017** ± 0.004	0.007 ± 0.001	−0.152 ± 0.007	−0.219 ± 0.024	−0.093 ± 0.050	−0.213	−0.020
Kidney	0.051 ± 0.015	0.059 ± 0.010	0.052 ± 0.007	−0.174 ± 0.052	−0.122 ± 0.023	0.026 ± 0.016	−0.086	**0.076**
Liver	0.019 ± 0.031	0.030 ± 0.011	0.011 ± 0.012	**0.063** ± 0.008	−0.100 ± 0.027	0.031 ± 0.000	−0.049	0.055
Lung	0.194 ± 0.031	0.115 ± 0.015	0.119 ± 0.007	0.197 ± 0.006	0.096 ± 0.014	0.194 ± 0.050	0.161	**0.234**
Prostate	−0.019 ± 0.012	0.023 ± 0.018	0.008 ± 0.009	−0.121 ± 0.162	−0.252 ± 0.348	0.016 ± 0.007	−0.127	**0.023**
Skin	0.452 ± 0.005	**0.456** ± 0.014	0.103 ± 0.013	0.360 ± 0.007	0.325 ± 0.024	0.375 ± 0.055	0.380	0.430
Uterus	0.049 ± 0.010	0.080 ± 0.004	0.070 ± 0.008	−0.151 ± 0.002	−0.092 ± 0.063	0.088 ± 0.022	−0.006	**0.154**
Breast	−0.005 ± 0.002	**0.061** ± 0.001	0.020 ± 0.001	−0.393 ± 0.123	−0.121 ± 0.040	0.051 ± 0.005	−0.055	0.052
Macro	0.080	0.092	0.050	−0.057	−0.117	0.072	−0.042	**0.114**
R^2_{flat} ↑								
Brain	0.593 ± 0.004	0.593 ± 0.004	0.582 ± 0.002	0.548 ± 0.036	0.477 ± 0.033	0.580 ± 0.011	0.511	**0.599**
Colon	0.549 ± 0.031	0.506 ± 0.026	0.508 ± 0.004	0.460 ± 0.007	0.359 ± 0.105	0.502 ± 0.023	0.501	**0.568**
Heart	0.554 ± 0.005	**0.608** ± 0.008	0.205 ± 0.008	0.583 ± 0.015	0.574 ± 0.016	0.560 ± 0.022	0.505	0.458
Kidney	0.519 ± 0.012	0.477 ± 0.001	0.516 ± 0.000	0.413 ± 0.009	0.487 ± 0.022	0.547 ± 0.026	0.522	**0.557**
Liver	0.745 ± 0.006	**0.757** ± 0.005	0.727 ± 0.001	0.757 ± 0.005	0.732 ± 0.005	0.733 ± 0.007	0.740	0.757
Lung	0.371 ± 0.006	0.346 ± 0.050	0.250 ± 0.002	0.396 ± 0.075	0.166 ± 0.050	0.264 ± 0.100	**0.432**	0.353
Prostate	0.180 ± 0.021	0.134 ± 0.028	**0.346** ± 0.003	0.033 ± 0.031	−0.010 ± 0.168	0.220 ± 0.080	0.186	0.259
Skin	0.446 ± 0.013	0.500 ± 0.007	0.260 ± 0.008	0.405 ± 0.008	0.393 ± 0.014	0.433 ± 0.065	0.443	**0.500**
Uterus	0.679 ± 0.027	0.658 ± 0.022	0.701 ± 0.004	0.645 ± 0.046	0.635 ± 0.002	0.732 ± 0.013	0.693	**0.735**
Breast	0.431 ± 0.011	0.452 ± 0.009	0.454 ± 0.001	−0.004 ± 0.089	0.375 ± 0.017	**0.499** ± 0.013	0.421	0.460
Macro	0.507	0.503	0.455	0.424	0.419	0.507	0.495	**0.525**

Table 6: Cohort-level test results with UNI and $G = 1000$. Neural entries are mean ± seed standard deviation over three runs. Linear and CGL-Linear are deterministic. Bold and underlining mark the best and second-best unrounded means in each row. Macro rows average cohort means without weighting by cohort size. Higher is better.

Cohort	HiST	HistoPrism	HEXST	STFlow	BLEEP	SEPAL	Linear	CGL-Linear
$\mathrm{PCC}_{\mathrm{flat}}$ ↑								
Brain	0.807 ± 0.008	0.808 ± 0.004	0.800 ± 0.001	0.797 ± 0.029	0.748 ± 0.003	0.799 ± 0.007	0.779	**0.812**
Colon	0.746 ± 0.011	0.743 ± 0.004	0.716 ± 0.011	0.691 ± 0.044	0.659 ± 0.002	0.704 ± 0.003	0.717	**0.759**
Heart	0.776 ± 0.005	**0.787** ± 0.015	0.674 ± 0.001	0.776 ± 0.005	0.773 ± 0.004	0.760 ± 0.005	0.760	0.747
Kidney	0.720 ± 0.016	0.708 ± 0.001	0.709 ± 0.001	0.680 ± 0.015	0.712 ± 0.003	0.719 ± 0.020	0.714	**0.733**
Liver	0.861 ± 0.000	0.872 ± 0.005	0.856 ± 0.002	0.868 ± 0.002	0.860 ± 0.001	0.859 ± 0.001	0.865	**0.874**
Lung	0.605 ± 0.028	0.575 ± 0.022	0.589 ± 0.054	0.630 ± 0.027	0.557 ± 0.014	**0.643** ± 0.001	0.641	0.642
Prostate	0.673 ± 0.004	0.667 ± 0.002	**0.709** ± 0.004	0.573 ± 0.067	0.421 ± 0.022	0.702 ± 0.001	0.555	0.694
Skin	0.695 ± 0.011	0.710 ± 0.008	0.494 ± 0.010	0.658 ± 0.005	0.603 ± 0.032	0.680 ± 0.002	0.684	**0.712**
Uterus	0.824 ± 0.005	0.817 ± 0.017	0.751 ± 0.012	0.811 ± 0.007	0.809 ± 0.013	0.851 ± 0.002	0.823	**0.858**
Breast	0.638 ± 0.010	0.645 ± 0.009	0.632 ± 0.000	0.565 ± 0.005	0.597 ± 0.001	**0.667** ± 0.001	0.620	0.641
Macro	0.734	0.733	0.693	0.705	0.674	0.738	0.716	**0.747**
$\mathrm{PCC}_{\mathrm{gene}}$ ↑								
Brain	0.040 ± 0.007	0.013 ± 0.002	0.034 ± 0.007	0.033 ± 0.016	0.030 ± 0.004	0.010 ± 0.001	0.037	**0.050**
Colon	0.208 ± 0.000	0.194 ± 0.006	**0.259** ± 0.003	0.161 ± 0.026	0.156 ± 0.012	0.173 ± 0.018	0.193	0.224
Heart	0.043 ± 0.003	0.047 ± 0.012	0.040 ± 0.007	**0.053** ± 0.001	0.046 ± 0.005	0.033 ± 0.007	0.049	0.053
Kidney	0.086 ± 0.004	0.079 ± 0.009	0.080 ± 0.003	0.069 ± 0.004	0.084 ± 0.001	0.055 ± 0.020	0.078	**0.089**
Liver	0.083 ± 0.002	0.081 ± 0.009	0.051 ± 0.001	0.098 ± 0.002	0.095 ± 0.005	0.057 ± 0.002	0.091	**0.109**
Lung	0.231 ± 0.014	0.228 ± 0.019	**0.328** ± 0.005	0.282 ± 0.036	0.232 ± 0.062	0.253 ± 0.021	0.271	0.326
Prostate	0.032 ± 0.005	0.039 ± 0.019	0.061 ± 0.028	0.032 ± 0.011	0.019 ± 0.008	**0.064** ± 0.022	−0.015	0.052
Skin	0.355 ± 0.001	0.364 ± 0.000	0.412 ± 0.007	0.352 ± 0.014	0.340 ± 0.021	0.331 ± 0.000	0.377	**0.412**
Uterus	0.090 ± 0.002	0.070 ± 0.010	0.125 ± 0.006	0.093 ± 0.010	0.093 ± 0.000	0.074 ± 0.001	0.091	**0.127**
Breast	0.093 ± 0.028	0.097 ± 0.013	0.093 ± 0.008	0.062 ± 0.005	0.056 ± 0.002	**0.142** ± 0.004	0.067	0.082
Macro	0.126	0.121	0.148	0.123	0.115	0.119	0.124	**0.152**
R^2_{local} ↑								
Brain	−0.059 ± 0.011	−0.026 ± 0.010	**0.013** ± 0.002	−0.194 ± 0.202	−0.416 ± 0.067	−0.008 ± 0.019	−0.241	−0.010
Colon	0.193 ± 0.002	0.157 ± 0.005	0.098 ± 0.037	0.030 ± 0.061	−0.049 ± 0.004	0.114 ± 0.046	−0.058	**0.197**
Heart	−0.016 ± 0.019	**0.034** ± 0.009	0.004 ± 0.002	−0.120 ± 0.021	−0.120 ± 0.003	−0.035 ± 0.007	−0.145	−0.001
Kidney	0.042 ± 0.005	0.052 ± 0.004	0.035 ± 0.000	−0.101 ± 0.044	−0.080 ± 0.006	0.016 ± 0.017	−0.069	**0.069**
Liver	0.016 ± 0.017	0.056 ± 0.004	0.020 ± 0.002	0.068 ± 0.017	−0.070 ± 0.024	0.033 ± 0.016	−0.024	**0.068**
Lung	0.188 ± 0.033	0.125 ± 0.030	0.084 ± 0.004	0.206 ± 0.039	0.126 ± 0.061	0.163 ± 0.038	0.170	**0.225**
Prostate	−0.008 ± 0.002	0.000 ± 0.011	0.007 ± 0.006	−0.170 ± 0.116	−0.005 ± 0.013	**0.020** ± 0.022	−0.132	0.013
Skin	**0.437** ± 0.007	0.434 ± 0.001	0.065 ± 0.014	0.337 ± 0.004	0.261 ± 0.051	0.360 ± 0.008	0.375	0.425
Uterus	0.050 ± 0.010	0.094 ± 0.024	−0.332 ± 0.040	−0.041 ± 0.044	−0.052 ± 0.013	0.080 ± 0.008	0.002	**0.146**
Breast	0.028 ± 0.015	**0.053** ± 0.001	0.012 ± 0.000	−0.371 ± 0.065	−0.062 ± 0.006	0.043 ± 0.006	−0.048	0.047
Macro	0.087	0.098	0.001	−0.036	−0.047	0.079	−0.017	**0.118**
R^2_{flat} ↑								
Brain	0.649 ± 0.013	0.648 ± 0.005	0.639 ± 0.001	0.628 ± 0.056	0.552 ± 0.007	0.635 ± 0.009	0.590	**0.659**
Colon	0.540 ± 0.028	0.511 ± 0.001	0.505 ± 0.014	0.460 ± 0.069	0.398 ± 0.012	0.491 ± 0.007	0.505	**0.573**
Heart	0.495 ± 0.018	0.539 ± 0.040	0.246 ± 0.002	0.552 ± 0.003	**0.583** ± 0.015	0.552 ± 0.016	0.458	0.444
Kidney	0.507 ± 0.023	0.480 ± 0.003	0.501 ± 0.003	0.432 ± 0.036	0.502 ± 0.003	0.513 ± 0.033	0.504	**0.537**
Liver	0.739 ± 0.001	**0.759** ± 0.008	0.723 ± 0.003	0.752 ± 0.004	0.737 ± 0.002	0.733 ± 0.008	0.747	0.756
Lung	0.359 ± 0.034	0.324 ± 0.025	0.251 ± 0.042	0.380 ± 0.036	0.187 ± 0.029	0.281 ± 0.055	**0.398**	0.349
Prostate	0.235 ± 0.021	0.123 ± 0.002	**0.341** ± 0.001	0.163 ± 0.048	−0.155 ± 0.003	0.241 ± 0.055	0.159	0.245
Skin	0.461 ± 0.025	0.502 ± 0.012	0.243 ± 0.009	0.414 ± 0.003	0.336 ± 0.045	0.439 ± 0.014	0.452	**0.506**
Uterus	0.672 ± 0.014	0.664 ± 0.025	0.550 ± 0.017	0.651 ± 0.012	0.647 ± 0.032	0.708 ± 0.012	0.674	**0.715**
Breast	0.400 ± 0.013	0.411 ± 0.010	0.399 ± 0.000	−0.069 ± 0.093	0.346 ± 0.001	**0.443** ± 0.002	0.369	0.407
Macro	0.506	0.496	0.440	0.436	0.413	0.504	0.485	**0.519**

Table 7: Cohort-level test results with UNI and $G = 1500$. Neural entries are mean ± seed standard deviation over three runs. Linear and CGL-Linear are deterministic. Bold and underlining mark the best and second-best unrounded means in each row. Macro rows average cohort means without weighting by cohort size. Higher is better.

Cohort	HiST	HistoPrism	HEXST	STFlow	BLEEP	SEPAL	Linear	CGL-Linear
$\mathrm{PCC}_{\mathrm{flat}}$ ↑								
Brain	0.814 ± 0.005	0.817 ± 0.003	0.813 ± 0.001	0.806 ± 0.015	0.753 ± 0.014	0.817 ± 0.001	0.792	**0.823**
Colon	**0.768** ± 0.011	0.747 ± 0.008	0.716 ± 0.003	0.739 ± 0.014	0.688 ± 0.001	0.726 ± 0.016	0.727	0.767
Heart	0.783 ± 0.003	**0.789** ± 0.008	0.699 ± 0.000	0.786 ± 0.002	0.783 ± 0.000	0.770 ± 0.005	0.772	0.764
Kidney	0.758 ± 0.009	0.745 ± 0.002	0.747 ± 0.001	0.713 ± 0.016	0.750 ± 0.000	0.755 ± 0.016	0.753	**0.766**
Liver	0.873 ± 0.008	0.875 ± 0.003	0.866 ± 0.000	0.876 ± 0.003	0.872 ± 0.002	0.869 ± 0.008	0.870	**0.878**
Lung	0.590 ± 0.004	0.551 ± 0.002	0.619 ± 0.007	0.634 ± 0.010	0.521 ± 0.052	0.629 ± 0.026	0.637	**0.645**
Prostate	0.656 ± 0.007	0.657 ± 0.012	**0.703** ± 0.003	0.616 ± 0.010	0.434 ± 0.021	0.649 ± 0.031	0.542	0.684
Skin	0.693 ± 0.000	0.715 ± 0.004	0.511 ± 0.004	0.696 ± 0.032	0.651 ± 0.026	0.686 ± 0.013	0.698	**0.724**
Uterus	0.815 ± 0.009	0.820 ± 0.000	0.723 ± 0.006	0.819 ± 0.007	0.788 ± 0.021	0.787 ± 0.081	0.818	**0.854**
Breast	0.609 ± 0.002	0.623 ± 0.008	0.609 ± 0.000	0.541 ± 0.003	0.596 ± 0.007	**0.649** ± 0.009	0.599	0.618
Macro	0.736	0.734	0.700	0.722	0.684	0.734	0.721	**0.752**
$\mathrm{PCC}_{\mathrm{gene}}$ ↑								
Brain	0.037 ± 0.004	0.013 ± 0.000	0.034 ± 0.001	0.032 ± 0.005	0.023 ± 0.006	0.012 ± 0.000	0.033	**0.045**
Colon	0.218 ± 0.004	0.182 ± 0.014	**0.249** ± 0.006	0.193 ± 0.004	0.158 ± 0.006	0.177 ± 0.020	0.191	0.221
Heart	0.041 ± 0.001	0.040 ± 0.009	0.032 ± 0.010	**0.051** ± 0.001	0.048 ± 0.001	0.029 ± 0.004	0.047	0.051
Kidney	**0.084** ± 0.009	0.064 ± 0.003	0.072 ± 0.013	0.057 ± 0.009	0.077 ± 0.005	0.049 ± 0.019	0.073	0.083
Liver	0.096 ± 0.009	0.077 ± 0.003	0.068 ± 0.009	0.093 ± 0.001	0.115 ± 0.005	0.076 ± 0.000	0.098	**0.116**
Lung	0.211 ± 0.005	0.227 ± 0.006	0.307 ± 0.019	0.305 ± 0.003	0.182 ± 0.005	0.256 ± 0.029	0.266	**0.318**
Prostate	0.029 ± 0.006	0.028 ± 0.002	**0.047** ± 0.050	0.036 ± 0.036	0.011 ± 0.000	0.037 ± 0.022	−0.023	0.041
Skin	0.338 ± 0.000	0.348 ± 0.001	**0.404** ± 0.008	0.366 ± 0.040	0.352 ± 0.024	0.308 ± 0.019	0.366	0.400
Uterus	0.077 ± 0.005	0.074 ± 0.007	**0.117** ± 0.006	0.096 ± 0.005	0.080 ± 0.012	0.046 ± 0.042	0.083	0.115
Breast	0.074 ± 0.013	0.096 ± 0.013	0.094 ± 0.001	0.046 ± 0.007	0.081 ± 0.006	**0.162** ± 0.022	0.068	0.083
Macro	0.121	0.115	0.142	0.127	0.113	0.115	0.120	**0.147**
R^2_{local} ↑								
Brain	−0.081 ± 0.024	−0.020 ± 0.005	**0.010** ± 0.001	−0.204 ± 0.104	−0.423 ± 0.044	0.006 ± 0.001	−0.223	−0.009
Colon	**0.219** ± 0.022	0.158 ± 0.020	0.064 ± 0.012	0.106 ± 0.044	−0.077 ± 0.052	0.096 ± 0.057	−0.026	0.211
Heart	0.009 ± 0.007	**0.036** ± 0.001	0.003 ± 0.001	−0.086 ± 0.017	−0.082 ± 0.017	−0.014 ± 0.002	−0.115	0.008
Kidney	0.046 ± 0.014	0.043 ± 0.004	0.029 ± 0.012	−0.118 ± 0.085	−0.063 ± 0.018	0.021 ± 0.015	−0.057	**0.069**
Liver	0.065 ± 0.057	0.028 ± 0.039	0.026 ± 0.005	**0.083** ± 0.010	0.010 ± 0.012	0.037 ± 0.021	−0.008	0.077
Lung	0.170 ± 0.003	0.115 ± 0.001	0.072 ± 0.009	0.188 ± 0.020	0.046 ± 0.039	0.181 ± 0.044	0.180	**0.221**
Prostate	−0.005 ± 0.004	−0.006 ± 0.002	0.003 ± 0.011	−0.119 ± 0.167	−0.004 ± 0.008	0.000 ± 0.008	−0.131	**0.008**
Skin	0.427 ± 0.003	**0.436** ± 0.013	0.058 ± 0.005	0.377 ± 0.057	0.302 ± 0.042	0.335 ± 0.029	0.376	0.423
Uterus	0.058 ± 0.003	0.094 ± 0.004	−0.504 ± 0.134	0.015 ± 0.042	−0.027 ± 0.013	0.045 ± 0.055	0.005	**0.142**
Breast	0.022 ± 0.018	**0.048** ± 0.008	0.008 ± 0.000	−0.244 ± 0.002	−0.043 ± 0.011	0.038 ± 0.005	−0.045	0.042
Macro	0.093	0.093	−0.023	0.000	−0.036	0.074	−0.005	**0.119**
R^2_{flat} ↑								
Brain	0.661 ± 0.008	0.663 ± 0.004	0.659 ± 0.001	0.636 ± 0.042	0.558 ± 0.023	0.660 ± 0.001	0.614	**0.677**
Colon	**0.586** ± 0.013	0.483 ± 0.068	0.508 ± 0.004	0.531 ± 0.008	0.461 ± 0.008	0.513 ± 0.010	0.521	0.585
Heart	0.474 ± 0.002	0.533 ± 0.029	0.297 ± 0.001	0.563 ± 0.006	**0.596** ± 0.004	0.551 ± 0.021	0.460	0.464
Kidney	0.563 ± 0.019	0.545 ± 0.002	0.556 ± 0.001	0.484 ± 0.041	0.560 ± 0.001	0.570 ± 0.024	0.564	**0.587**
Liver	0.759 ± 0.016	0.761 ± 0.004	0.739 ± 0.001	**0.766** ± 0.005	0.759 ± 0.004	0.749 ± 0.022	0.756	0.760
Lung	0.340 ± 0.004	0.293 ± 0.007	0.207 ± 0.010	0.389 ± 0.008	0.164 ± 0.009	0.298 ± 0.020	**0.394**	0.350
Prostate	0.201 ± 0.012	0.121 ± 0.021	**0.324** ± 0.002	0.220 ± 0.070	−0.145 ± 0.031	0.149 ± 0.086	0.146	0.232
Skin	0.469 ± 0.003	0.506 ± 0.007	0.259 ± 0.004	0.471 ± 0.039	0.409 ± 0.029	0.447 ± 0.021	0.473	**0.523**
Uterus	0.660 ± 0.018	0.654 ± 0.005	0.483 ± 0.001	0.668 ± 0.014	0.610 ± 0.035	0.576 ± 0.157	0.667	**0.708**
Breast	0.360 ± 0.000	0.383 ± 0.010	0.369 ± 0.000	0.000 ± 0.035	0.345 ± 0.008	**0.419** ± 0.011	0.343	0.378
Macro	0.507	0.494	0.440	0.473	0.432	0.493	0.494	**0.526**

Table 8: Cohort-level test results with UNI and $G = 2000$. Neural entries are mean $\pm$ seed standard deviation over three runs. Linear and CGL-Linear are deterministic. Bold and underlining mark the best and second-best unrounded means in each row. Macro rows average cohort means without weighting by cohort size. Higher is better.

Cohort	HiST	HistoPrism	HEXST	STFlow	BLEEP	SEPAL	Linear	CGL-Linear
$\mathrm{PCC}_{\mathrm{flat}}$ ↑								
Brain	0.809 ± 0.003	0.810 ± 0.001	0.807 ± 0.001	0.802 ± 0.011	0.767 ± 0.019	0.809 ± 0.003	0.788	**0.818**
Colon	0.749 ± 0.035	0.740 ± 0.010	0.705 ± 0.002	0.718 ± 0.025	0.678 ± 0.007	0.715 ± 0.025	0.722	**0.763**
Heart	0.773 ± 0.000	**0.786** ± 0.008	0.709 ± 0.000	0.776 ± 0.002	0.772 ± 0.010	0.749 ± 0.006	0.771	0.766
Kidney	0.753 ± 0.002	0.747 ± 0.002	0.742 ± 0.001	0.726 ± 0.017	0.743 ± 0.007	0.750 ± 0.010	0.748	**0.762**
Liver	0.876 ± 0.002	0.872 ± 0.004	0.857 ± 0.006	**0.878** ± 0.007	0.870 ± 0.000	0.843 ± 0.012	0.868	0.875
Lung	0.585 ± 0.005	0.531 ± 0.014	0.604 ± 0.005	0.618 ± 0.019	0.460 ± 0.020	**0.636** ± 0.000	0.621	0.631
Prostate	0.639 ± 0.010	0.653 ± 0.002	**0.685** ± 0.005	0.593 ± 0.012	0.390 ± 0.091	0.658 ± 0.002	0.519	0.662
Skin	0.675 ± 0.008	**0.731** ± 0.029	0.507 ± 0.004	0.688 ± 0.027	0.614 ± 0.065	0.688 ± 0.008	0.700	0.726
Uterus	0.796 ± 0.015	0.818 ± 0.014	0.768 ± 0.080	0.821 ± 0.003	0.797 ± 0.025	0.783 ± 0.086	0.812	**0.846**
Breast	0.607 ± 0.010	0.606 ± 0.014	0.589 ± 0.000	0.522 ± 0.017	0.567 ± 0.009	**0.627** ± 0.006	0.577	0.597
Macro	0.726	0.729	0.697	0.714	0.666	0.726	0.713	**0.744**
$\mathrm{PCC}_{\mathrm{gene}}$ ↑								
Brain	0.038 ± 0.003	0.013 ± 0.001	0.033 ± 0.003	0.037 ± 0.002	0.038 ± 0.006	0.014 ± 0.001	0.032	**0.046**
Colon	0.190 ± 0.040	0.183 ± 0.013	**0.245** ± 0.003	0.178 ± 0.005	0.142 ± 0.012	0.167 ± 0.020	0.188	0.219
Heart	0.034 ± 0.001	0.038 ± 0.010	0.033 ± 0.014	0.044 ± 0.004	0.037 ± 0.009	0.022 ± 0.011	0.043	**0.048**
Kidney	0.072 ± 0.004	0.065 ± 0.007	0.067 ± 0.009	0.064 ± 0.011	0.069 ± 0.003	0.044 ± 0.015	0.067	**0.077**
Liver	0.110 ± 0.000	0.085 ± 0.011	0.073 ± 0.011	0.098 ± 0.022	0.122 ± 0.004	0.061 ± 0.022	0.105	**0.123**
Lung	0.203 ± 0.001	0.226 ± 0.001	0.304 ± 0.008	0.298 ± 0.020	0.190 ± 0.003	0.248 ± 0.006	0.262	**0.313**
Prostate	0.025 ± 0.011	0.032 ± 0.008	**0.047** ± 0.029	0.040 ± 0.005	0.012 ± 0.026	0.021 ± 0.023	−0.028	0.036
Skin	0.305 ± 0.006	0.357 ± 0.032	0.382 ± 0.003	0.341 ± 0.039	0.314 ± 0.038	0.296 ± 0.015	0.352	**0.386**
Uterus	0.069 ± 0.002	0.078 ± 0.008	0.104 ± 0.003	0.090 ± 0.005	0.071 ± 0.021	0.040 ± 0.049	0.078	**0.107**
Breast	0.118 ± 0.013	0.101 ± 0.033	0.103 ± 0.010	0.056 ± 0.013	0.056 ± 0.005	**0.151** ± 0.020	0.068	0.083
Macro	0.116	0.118	0.139	0.125	0.105	0.106	0.117	**0.144**
R^2_{local} ↑								
Brain	−0.093 ± 0.030	−0.011 ± 0.001	**0.011** ± 0.000	−0.187 ± 0.122	−0.358 ± 0.048	0.004 ± 0.000	−0.202	−0.005
Colon	0.199 ± 0.040	0.164 ± 0.009	0.055 ± 0.007	0.036 ± 0.105	0.018 ± 0.036	0.108 ± 0.070	−0.009	**0.219**
Heart	−0.010 ± 0.034	**0.032** ± 0.002	0.001 ± 0.000	−0.103 ± 0.024	−0.084 ± 0.032	0.000 ± 0.007	−0.101	0.012
Kidney	0.044 ± 0.009	0.047 ± 0.006	0.024 ± 0.011	−0.114 ± 0.001	−0.037 ± 0.005	0.018 ± 0.016	−0.053	**0.068**
Liver	**0.098** ± 0.005	0.024 ± 0.048	0.017 ± 0.002	0.090 ± 0.009	0.004 ± 0.010	0.030 ± 0.016	−0.002	0.077
Lung	0.169 ± 0.002	0.107 ± 0.001	0.066 ± 0.007	0.194 ± 0.029	0.037 ± 0.034	0.167 ± 0.005	0.179	**0.211**
Prostate	0.001 ± 0.014	0.002 ± 0.006	0.005 ± 0.004	−0.029 ± 0.018	−0.012 ± 0.030	−0.004 ± 0.015	−0.129	**0.006**
Skin	0.409 ± 0.012	**0.448** ± 0.028	0.049 ± 0.005	0.364 ± 0.039	0.267 ± 0.097	0.335 ± 0.018	0.376	0.422
Uterus	0.040 ± 0.023	0.093 ± 0.017	−0.272 ± 0.445	−0.006 ± 0.051	−0.054 ± 0.064	0.046 ± 0.064	0.003	**0.134**
Breast	0.035 ± 0.002	**0.045** ± 0.006	0.006 ± 0.000	−0.197 ± 0.022	−0.056 ± 0.001	0.036 ± 0.011	−0.047	0.039
Macro	0.089	0.095	−0.004	0.005	−0.027	0.074	0.001	**0.118**
R^2_{flat} ↑								
Brain	0.652 ± 0.004	0.653 ± 0.002	0.651 ± 0.001	0.633 ± 0.030	0.582 ± 0.034	0.647 ± 0.004	0.608	**0.668**
Colon	0.538 ± 0.084	0.517 ± 0.002	0.492 ± 0.002	0.505 ± 0.034	0.419 ± 0.015	0.488 ± 0.006	0.513	**0.578**
Heart	0.391 ± 0.049	0.512 ± 0.025	0.325 ± 0.001	**0.549** ± 0.022	0.543 ± 0.044	0.511 ± 0.053	0.451	0.464
Kidney	0.558 ± 0.010	0.553 ± 0.004	0.549 ± 0.001	0.502 ± 0.036	0.549 ± 0.013	0.561 ± 0.018	0.557	**0.580**
Liver	0.765 ± 0.007	0.757 ± 0.011	0.725 ± 0.005	**0.769** ± 0.013	0.752 ± 0.004	0.668 ± 0.056	0.753	0.754
Lung	0.326 ± 0.011	0.262 ± 0.015	0.189 ± 0.008	0.369 ± 0.028	0.144 ± 0.009	0.301 ± 0.012	**0.374**	0.334
Prostate	0.172 ± 0.015	0.136 ± 0.019	**0.296** ± 0.013	0.141 ± 0.064	−0.154 ± 0.020	0.191 ± 0.032	0.124	0.211
Skin	0.439 ± 0.006	**0.528** ± 0.038	0.256 ± 0.003	0.460 ± 0.023	0.363 ± 0.088	0.446 ± 0.007	0.475	0.526
Uterus	0.625 ± 0.035	0.655 ± 0.040	0.558 ± 0.154	0.673 ± 0.005	0.631 ± 0.047	0.512 ± 0.243	0.657	**0.698**
Breast	0.363 ± 0.014	0.363 ± 0.019	0.345 ± 0.000	0.002 ± 0.037	0.306 ± 0.006	**0.390** ± 0.008	0.316	0.353
Macro	0.483	0.494	0.439	0.460	0.413	0.472	0.483	**0.517**

Table 9: Cohort-level test results with DINOv2-Large and $G = 500$. Neural entries are mean $\pm$ seed standard deviation over three runs. Linear and CGL-Linear are deterministic. Bold and underlining mark the best and second-best unrounded means in each row. Macro rows average cohort means without weighting by cohort size. Higher is better.

Cohort	HiST	HistoPrism	HEXST	STFlow	BLEEP	SEPAL	Linear	CGL-Linear
$\mathrm{PCC}_{\mathrm{flat}}$ ↑								
Brain	0.752 ± 0.018	**0.777** ± 0.002	0.765 ± 0.001	0.770 ± 0.000	0.706 ± 0.010	0.771 ± 0.006	0.738	0.773
Colon	0.677 ± 0.009	0.699 ± 0.023	0.710 ± 0.002	0.687 ± 0.000	0.625 ± 0.021	0.694 ± 0.002	0.664	**0.731**
Heart	**0.765** ± 0.005	0.709 ± 0.009	0.646 ± 0.000	0.644 ± 0.013	0.748 ± 0.009	0.738 ± 0.039	0.711	0.684
Kidney	0.719 ± 0.020	0.684 ± 0.017	0.717 ± 0.000	0.707 ± 0.010	0.686 ± 0.010	**0.741** ± 0.000	0.720	0.730
Liver	0.863 ± 0.005	**0.871** ± 0.000	0.862 ± 0.000	0.859 ± 0.005	0.861 ± 0.001	0.857 ± 0.005	0.861	0.869
Lung	0.651 ± 0.006	0.603 ± 0.020	0.515 ± 0.023	**0.669** ± 0.016	0.522 ± 0.021	0.631 ± 0.005	0.591	0.639
Prostate	0.637 ± 0.049	0.653 ± 0.004	**0.718** ± 0.004	0.605 ± 0.028	0.615 ± 0.001	0.687 ± 0.059	0.622	0.710
Skin	**0.647** ± 0.011	0.640 ± 0.006	0.498 ± 0.009	0.541 ± 0.077	0.643 ± 0.021	0.598 ± 0.001	0.630	0.643
Uterus	0.822 ± 0.005	0.825 ± 0.001	0.803 ± 0.054	0.846 ± 0.004	0.785 ± 0.019	0.856 ± 0.004	0.827	**0.865**
Breast	0.678 ± 0.017	**0.699** ± 0.005	0.671 ± 0.001	0.646 ± 0.008	0.630 ± 0.017	0.697 ± 0.002	0.661	0.676
Macro	0.721	0.716	0.691	0.697	0.682	0.727	0.703	**0.732**
$\mathrm{PCC}_{\mathrm{gene}}$ ↑								
Brain	0.029 ± 0.010	0.014 ± 0.003	0.031 ± 0.002	0.027 ± 0.003	0.021 ± 0.001	0.006 ± 0.000	**0.037**	0.036
Colon	0.134 ± 0.008	0.118 ± 0.046	**0.219** ± 0.001	0.018 ± 0.002	0.135 ± 0.037	0.109 ± 0.027	0.130	0.180
Heart	0.031 ± 0.001	0.015 ± 0.003	0.028 ± 0.004	0.034 ± 0.002	0.028 ± 0.003	0.035 ± 0.001	0.028	**0.040**
Kidney	**0.091** ± 0.010	0.047 ± 0.002	0.075 ± 0.004	0.054 ± 0.008	0.068 ± 0.007	0.071 ± 0.001	0.077	0.078
Liver	0.088 ± 0.013	0.078 ± 0.001	0.049 ± 0.019	0.087 ± 0.013	**0.091** ± 0.000	0.070 ± 0.014	0.080	0.087
Lung	0.239 ± 0.011	0.198 ± 0.004	**0.344** ± 0.006	0.230 ± 0.009	0.223 ± 0.013	0.204 ± 0.003	0.246	0.285
Prostate	0.004 ± 0.036	0.019 ± 0.026	0.030 ± 0.016	−0.019 ± 0.005	**0.053** ± 0.003	0.018 ± 0.033	0.010	0.049
Skin	0.354 ± 0.000	0.314 ± 0.006	**0.419** ± 0.004	0.233 ± 0.089	0.378 ± 0.013	0.291 ± 0.000	0.347	0.365
Uterus	0.106 ± 0.012	0.090 ± 0.005	0.111 ± 0.003	0.071 ± 0.024	0.079 ± 0.002	0.047 ± 0.025	0.089	**0.126**
Breast	0.103 ± 0.029	0.134 ± 0.011	0.049 ± 0.011	0.081 ± 0.017	0.055 ± 0.013	**0.136** ± 0.009	0.068	0.074
Macro	0.118	0.103	**0.136**	0.082	0.113	0.099	0.111	0.132
R^2_{local} ↑								
Brain	−0.091 ± 0.034	−0.015 ± 0.005	**0.008** ± 0.001	−0.062 ± 0.026	−0.448 ± 0.003	−0.001 ± 0.007	−0.303	−0.042
Colon	0.070 ± 0.005	0.059 ± 0.006	0.060 ± 0.005	0.000 ± 0.000	−0.174 ± 0.043	0.036 ± 0.018	−0.142	**0.102**
Heart	−0.090 ± 0.006	−0.046 ± 0.006	**0.003** ± 0.000	−0.054 ± 0.006	−0.177 ± 0.014	−0.117 ± 0.061	−0.475	−0.071
Kidney	0.039 ± 0.003	0.021 ± 0.001	0.030 ± 0.007	−0.010 ± 0.001	−0.254 ± 0.028	−0.018 ± 0.001	−0.127	**0.042**
Liver	0.008 ± 0.034	−0.008 ± 0.002	0.013 ± 0.004	0.001 ± 0.005	−0.078 ± 0.060	0.004 ± 0.003	−0.103	**0.018**
Lung	**0.218** ± 0.010	0.087 ± 0.002	0.108 ± 0.089	0.141 ± 0.001	0.090 ± 0.018	0.069 ± 0.014	0.148	0.133
Prostate	−0.132 ± 0.070	−0.015 ± 0.010	**−0.014** ± 0.001	−0.258 ± 0.098	−0.232 ± 0.136	−0.063 ± 0.080	−0.177	−0.052
Skin	**0.396** ± 0.002	0.332 ± 0.003	0.079 ± 0.012	0.191 ± 0.090	0.296 ± 0.031	0.254 ± 0.004	0.292	0.323
Uterus	0.069 ± 0.000	0.080 ± 0.002	−0.143 ± 0.282	0.045 ± 0.020	−0.120 ± 0.021	0.044 ± 0.036	−0.087	**0.108**
Breast	**0.062** ± 0.016	0.062 ± 0.007	0.012 ± 0.005	−0.222 ± 0.022	−0.100 ± 0.022	0.009 ± 0.002	−0.049	0.046
Macro	0.055	0.056	0.016	−0.023	−0.120	0.022	−0.102	**0.061**
R^2_{flat} ↑								
Brain	0.556 ± 0.032	**0.599** ± 0.002	0.584 ± 0.001	0.590 ± 0.001	0.481 ± 0.005	0.593 ± 0.007	0.525	0.595
Colon	0.430 ± 0.015	0.470 ± 0.055	0.497 ± 0.002	0.464 ± 0.002	0.345 ± 0.060	0.462 ± 0.028	0.440	**0.525**
Heart	0.439 ± 0.008	0.342 ± 0.047	0.204 ± 0.001	0.300 ± 0.022	**0.553** ± 0.017	0.500 ± 0.091	0.359	0.303
Kidney	0.479 ± 0.049	0.453 ± 0.030	0.514 ± 0.000	0.494 ± 0.014	0.457 ± 0.004	**0.546** ± 0.001	0.514	0.533
Liver	0.741 ± 0.008	**0.755** ± 0.002	0.734 ± 0.000	0.731 ± 0.009	0.737 ± 0.001	0.733 ± 0.008	0.739	0.751
Lung	0.361 ± 0.021	0.295 ± 0.030	0.246 ± 0.028	**0.437** ± 0.020	0.155 ± 0.018	0.189 ± 0.053	0.279	0.253
Prostate	0.227 ± 0.099	0.077 ± 0.036	**0.358** ± 0.002	0.194 ± 0.002	0.247 ± 0.037	0.253 ± 0.089	0.219	0.297
Skin	0.399 ± 0.020	0.401 ± 0.007	0.244 ± 0.008	0.282 ± 0.090	0.381 ± 0.024	0.342 ± 0.005	0.391	**0.413**
Uterus	0.661 ± 0.011	0.658 ± 0.005	0.634 ± 0.078	0.713 ± 0.005	0.583 ± 0.024	0.704 ± 0.010	0.676	**0.724**
Breast	0.454 ± 0.025	0.481 ± 0.004	0.449 ± 0.002	0.172 ± 0.008	0.384 ± 0.024	**0.481** ± 0.005	0.429	0.455
Macro	0.475	0.453	0.446	0.438	0.432	0.480	0.457	**0.485**

Table 10: Cohort-level test results with DINOv2-Large and $G = 1000$. Neural entries are mean $\pm$ seed standard deviation over three runs. Linear and CGL-Linear are deterministic. Bold and underlining mark the best and second-best unrounded means in each row. Macro rows average cohort means without weighting by cohort size. Higher is better.

Cohort	HiST	HistoPrism	HEXST	STFlow	BLEEP	SEPAL	Linear	CGL-Linear
$\mathrm{PCC}_{\mathrm{flat}}$ ↑								
Brain	0.799 ± 0.002	0.810 ± 0.001	0.801 ± 0.001	**0.812** ± 0.000	0.747 ± 0.006	0.805 ± 0.003	0.780	0.808
Colon	0.731 ± 0.017	**0.737** ± 0.012	0.703 ± 0.005	0.683 ± 0.004	0.540 ± 0.066	0.697 ± 0.009	0.675	0.734
Heart	**0.748** ± 0.001	0.733 ± 0.002	0.674 ± 0.000	0.686 ± 0.008	0.727 ± 0.012	0.722 ± 0.000	0.722	0.710
Kidney	0.705 ± 0.013	0.673 ± 0.005	0.707 ± 0.001	0.679 ± 0.008	0.659 ± 0.006	0.710 ± 0.021	0.705	**0.719**
Liver	0.863 ± 0.001	0.866 ± 0.001	0.858 ± 0.001	0.865 ± 0.009	0.856 ± 0.010	0.863 ± 0.002	0.862	**0.867**
Lung	0.619 ± 0.012	0.543 ± 0.006	0.542 ± 0.095	0.516 ± 0.156	0.460 ± 0.065	0.619 ± 0.023	0.582	**0.621**
Prostate	0.538 ± 0.086	0.628 ± 0.011	0.707 ± 0.004	0.577 ± 0.025	0.610 ± 0.027	**0.723** ± 0.010	0.594	0.688
Skin	0.650 ± 0.013	0.647 ± 0.004	0.489 ± 0.006	0.576 ± 0.095	0.640 ± 0.043	0.592 ± 0.013	0.642	**0.654**
Uterus	0.821 ± 0.002	0.810 ± 0.002	0.831 ± 0.002	0.832 ± 0.004	0.804 ± 0.009	0.819 ± 0.043	0.814	**0.852**
Breast	0.651 ± 0.008	0.649 ± 0.008	0.631 ± 0.001	0.616 ± 0.007	0.603 ± 0.029	**0.658** ± 0.009	0.621	0.637
Macro	0.713	0.710	0.694	0.684	0.665	0.721	0.700	**0.729**
$\mathrm{PCC}_{\mathrm{gene}}$ ↑								
Brain	0.022 ± 0.000	0.010 ± 0.000	0.023 ± 0.004	0.027 ± 0.006	0.018 ± 0.003	0.003 ± 0.002	0.033	**0.034**
Colon	0.178 ± 0.005	0.176 ± 0.011	**0.218** ± 0.004	0.027 ± 0.007	0.091 ± 0.052	0.108 ± 0.009	0.131	0.183
Heart	0.023 ± 0.000	0.015 ± 0.003	0.029 ± 0.002	0.033 ± 0.002	0.025 ± 0.001	0.028 ± 0.003	0.024	**0.038**
Kidney	0.066 ± 0.001	0.042 ± 0.001	**0.068** ± 0.000	0.038 ± 0.010	0.051 ± 0.001	0.041 ± 0.028	0.063	0.068
Liver	0.085 ± 0.005	0.065 ± 0.002	0.059 ± 0.003	0.076 ± 0.010	0.084 ± 0.013	0.066 ± 0.007	0.083	**0.088**
Lung	0.217 ± 0.007	0.174 ± 0.004	**0.302** ± 0.003	0.175 ± 0.045	0.127 ± 0.017	0.194 ± 0.062	0.227	0.264
Prostate	−0.018 ± 0.019	−0.002 ± 0.012	0.023 ± 0.042	0.001 ± 0.014	0.047 ± 0.007	**0.048** ± 0.004	−0.011	0.034
Skin	0.303 ± 0.006	0.297 ± 0.006	**0.386** ± 0.004	0.242 ± 0.106	0.343 ± 0.030	0.260 ± 0.000	0.325	0.343
Uterus	0.091 ± 0.000	0.083 ± 0.000	0.097 ± 0.003	0.053 ± 0.015	0.078 ± 0.008	0.049 ± 0.031	0.074	**0.105**
Breast	0.120 ± 0.024	0.120 ± 0.009	0.051 ± 0.015	0.086 ± 0.004	0.070 ± 0.033	**0.121** ± 0.031	0.065	0.072
Macro	0.109	0.098	**0.125**	0.076	0.093	0.092	0.101	0.123
R^2_{local} ↑								
Brain	−0.118 ± 0.003	−0.013 ± 0.002	**0.007** ± 0.000	−0.031 ± 0.007	−0.374 ± 0.071	0.003 ± 0.006	−0.255	−0.036
Colon	0.135 ± 0.001	0.099 ± 0.002	0.050 ± 0.016	0.002 ± 0.003	−0.081 ± 0.058	0.038 ± 0.030	−0.072	**0.143**
Heart	−0.074 ± 0.011	−0.013 ± 0.003	**0.002** ± 0.000	−0.054 ± 0.018	−0.235 ± 0.126	−0.052 ± 0.046	−0.340	−0.046
Kidney	0.013 ± 0.000	0.024 ± 0.002	0.027 ± 0.003	−0.022 ± 0.013	−0.182 ± 0.012	0.002 ± 0.002	−0.107	**0.038**
Liver	−0.009 ± 0.001	−0.008 ± 0.004	**0.030** ± 0.008	0.006 ± 0.009	−0.144 ± 0.023	0.010 ± 0.005	−0.077	0.029
Lung	**0.190** ± 0.005	0.070 ± 0.001	0.086 ± 0.028	−0.089 ± 0.296	0.022 ± 0.020	0.081 ± 0.034	0.139	0.122
Prostate	−0.246 ± 0.084	−0.076 ± 0.028	−0.003 ± 0.013	−0.218 ± 0.015	−0.019 ± 0.053	**0.011** ± 0.011	−0.182	−0.056
Skin	**0.375** ± 0.011	0.336 ± 0.001	0.057 ± 0.007	0.218 ± 0.122	0.263 ± 0.052	0.231 ± 0.041	0.292	0.327
Uterus	0.055 ± 0.005	0.085 ± 0.000	0.053 ± 0.005	0.037 ± 0.011	−0.129 ± 0.038	0.036 ± 0.039	−0.069	**0.102**
Breast	0.048 ± 0.004	**0.058** ± 0.001	0.008 ± 0.001	−0.150 ± 0.060	−0.097 ± 0.003	0.013 ± 0.002	−0.041	0.042
Macro	0.037	0.056	0.032	−0.030	−0.098	0.037	−0.071	**0.066**
R^2_{flat} ↑								
Brain	0.629 ± 0.014	0.652 ± 0.001	0.641 ± 0.001	**0.660** ± 0.001	0.553 ± 0.011	0.644 ± 0.004	0.599	0.652
Colon	0.505 ± 0.031	0.523 ± 0.020	0.488 ± 0.007	0.459 ± 0.010	0.173 ± 0.161	0.484 ± 0.012	0.453	**0.534**
Heart	0.292 ± 0.016	0.341 ± 0.012	0.245 ± 0.001	0.373 ± 0.012	**0.491** ± 0.057	0.471 ± 0.023	0.336	0.330
Kidney	0.473 ± 0.010	0.448 ± 0.004	0.499 ± 0.002	0.443 ± 0.021	0.425 ± 0.015	0.499 ± 0.034	0.491	**0.517**
Liver	0.738 ± 0.005	0.745 ± 0.003	0.727 ± 0.000	0.745 ± 0.019	0.731 ± 0.018	0.744 ± 0.003	0.741	**0.748**
Lung	**0.323** ± 0.016	0.257 ± 0.011	0.201 ± 0.006	0.115 ± 0.382	0.093 ± 0.040	0.178 ± 0.023	0.279	0.248
Prostate	0.141 ± 0.085	0.135 ± 0.043	0.344 ± 0.007	0.175 ± 0.044	0.134 ± 0.086	**0.397** ± 0.081	0.199	0.281
Skin	0.404 ± 0.002	0.418 ± 0.006	0.238 ± 0.005	0.315 ± 0.112	0.377 ± 0.060	0.331 ± 0.001	0.403	**0.428**
Uterus	0.668 ± 0.008	0.642 ± 0.002	0.677 ± 0.003	0.691 ± 0.006	0.628 ± 0.034	0.642 ± 0.065	0.656	**0.704**
Breast	0.420 ± 0.013	0.418 ± 0.011	0.397 ± 0.001	0.086 ± 0.099	0.349 ± 0.032	**0.424** ± 0.014	0.378	0.402
Macro	0.459	0.458	0.446	0.406	0.395	0.481	0.454	**0.484**

Table 11: Cohort-level test results with DINOv2-Large and $G = 1500$. Neural entries are mean $\pm$ seed standard deviation over three runs. Linear and CGL-Linear are deterministic. Bold and underlining mark the best and second-best unrounded means in each row. Macro rows average cohort means without weighting by cohort size. Higher is better.

Cohort	HiST	HistoPrism	HEXST	STFlow	BLEEP	SEPAL	Linear	CGL-Linear
$\mathrm{PCC}_{\mathrm{flat}}$ ↑								
Brain	0.796 ± 0.006	0.819 ± 0.001	0.814 ± 0.001	**0.821** ± 0.001	0.773 ± 0.005	0.813 ± 0.002	0.793	0.818
Colon	0.715 ± 0.022	0.740 ± 0.014	0.713 ± 0.000	0.693 ± 0.002	0.619 ± 0.019	0.701 ± 0.003	0.690	**0.744**
Heart	0.742 ± 0.003	0.739 ± 0.002	0.658 ± 0.058	0.709 ± 0.002	**0.750** ± 0.004	0.732 ± 0.009	0.739	0.734
Kidney	0.738 ± 0.009	0.730 ± 0.007	0.745 ± 0.001	0.714 ± 0.002	0.718 ± 0.017	0.744 ± 0.003	0.741	**0.756**
Liver	0.863 ± 0.005	0.873 ± 0.001	0.864 ± 0.000	**0.875** ± 0.002	0.827 ± 0.032	0.857 ± 0.011	0.867	0.872
Lung	**0.632** ± 0.001	0.555 ± 0.044	0.504 ± 0.155	0.452 ± 0.094	0.559 ± 0.008	0.608 ± 0.001	0.589	0.623
Prostate	0.585 ± 0.022	0.615 ± 0.007	0.695 ± 0.001	0.519 ± 0.084	0.508 ± 0.110	**0.719** ± 0.008	0.581	0.679
Skin	**0.700** ± 0.007	0.646 ± 0.018	0.508 ± 0.009	0.613 ± 0.017	0.686 ± 0.003	0.620 ± 0.025	0.660	0.671
Uterus	0.820 ± 0.015	0.828 ± 0.002	0.753 ± 0.102	0.831 ± 0.003	0.796 ± 0.024	0.784 ± 0.086	0.809	**0.847**
Breast	0.606 ± 0.007	**0.635** ± 0.007	0.608 ± 0.001	0.589 ± 0.001	0.591 ± 0.009	0.620 ± 0.008	0.599	0.614
Macro	0.720	0.718	0.686	0.682	0.683	0.720	0.707	**0.736**
$\mathrm{PCC}_{\mathrm{gene}}$ ↑								
Brain	0.019 ± 0.005	0.008 ± 0.002	0.022 ± 0.001	0.023 ± 0.001	0.022 ± 0.005	0.003 ± 0.002	0.029	**0.030**
Colon	0.160 ± 0.021	0.172 ± 0.031	**0.225** ± 0.009	0.004 ± 0.003	0.129 ± 0.011	0.092 ± 0.001	0.132	0.183
Heart	0.014 ± 0.004	0.016 ± 0.004	0.035 ± 0.007	0.030 ± 0.001	0.027 ± 0.004	0.028 ± 0.001	0.024	**0.038**
Kidney	**0.067** ± 0.001	0.043 ± 0.002	0.059 ± 0.001	0.036 ± 0.004	0.056 ± 0.006	0.029 ± 0.009	0.059	0.065
Liver	0.078 ± 0.012	0.065 ± 0.001	0.062 ± 0.013	0.086 ± 0.007	0.049 ± 0.040	0.067 ± 0.015	0.089	**0.092**
Lung	0.222 ± 0.034	0.193 ± 0.019	**0.297** ± 0.032	0.152 ± 0.040	0.219 ± 0.013	0.159 ± 0.013	0.221	0.259
Prostate	−0.016 ± 0.006	−0.006 ± 0.010	0.003 ± 0.006	−0.045 ± 0.024	0.016 ± 0.023	**0.044** ± 0.007	−0.020	0.026
Skin	0.337 ± 0.000	0.281 ± 0.013	**0.391** ± 0.002	0.243 ± 0.022	0.362 ± 0.005	0.268 ± 0.015	0.320	0.338
Uterus	0.084 ± 0.004	0.076 ± 0.001	0.094 ± 0.000	0.057 ± 0.006	0.072 ± 0.008	0.040 ± 0.029	0.067	**0.095**
Breast	0.105 ± 0.031	**0.125** ± 0.019	0.059 ± 0.024	0.084 ± 0.004	0.078 ± 0.007	0.075 ± 0.024	0.066	0.074
Macro	0.107	0.097	**0.125**	0.067	0.103	0.080	0.099	0.120
R^2_{local} ↑								
Brain	−0.139 ± 0.015	−0.011 ± 0.002	**0.008** ± 0.000	−0.028 ± 0.005	−0.305 ± 0.016	0.006 ± 0.001	−0.234	−0.035
Colon	0.142 ± 0.014	0.109 ± 0.017	0.058 ± 0.002	0.000 ± 0.001	−0.068 ± 0.055	0.014 ± 0.002	−0.040	**0.157**
Heart	−0.107 ± 0.012	**−0.002** ± 0.004	−0.182 ± 0.260	−0.065 ± 0.020	−0.156 ± 0.016	−0.037 ± 0.042	−0.281	−0.035
Kidney	0.029 ± 0.009	0.024 ± 0.004	0.017 ± 0.001	−0.033 ± 0.004	−0.132 ± 0.033	−0.007 ± 0.003	−0.099	**0.039**
Liver	−0.010 ± 0.019	−0.004 ± 0.000	0.025 ± 0.017	0.011 ± 0.024	−0.229 ± 0.012	0.003 ± 0.014	−0.058	**0.035**
Lung	**0.191** ± 0.030	0.090 ± 0.017	−0.159 ± 0.329	−0.234 ± 0.281	0.118 ± 0.015	0.068 ± 0.001	0.140	0.120
Prostate	−0.211 ± 0.006	−0.078 ± 0.006	−0.010 ± 0.004	−0.433 ± 0.213	−0.122 ± 0.009	**0.015** ± 0.008	−0.181	−0.057
Skin	**0.415** ± 0.001	0.325 ± 0.014	0.054 ± 0.013	0.228 ± 0.023	0.301 ± 0.003	0.267 ± 0.007	0.300	0.332
Uterus	0.057 ± 0.004	0.079 ± 0.001	−0.458 ± 0.710	0.044 ± 0.003	−0.152 ± 0.039	0.031 ± 0.036	−0.065	**0.099**
Breast	0.042 ± 0.008	**0.052** ± 0.006	0.006 ± 0.002	−0.106 ± 0.024	−0.073 ± 0.003	0.017 ± 0.003	−0.038	0.037
Macro	0.041	0.058	−0.064	−0.062	−0.082	0.038	−0.056	**0.069**
R^2_{flat} ↑								
Brain	0.632 ± 0.008	0.666 ± 0.001	0.661 ± 0.002	**0.674** ± 0.001	0.596 ± 0.007	0.657 ± 0.002	0.622	0.669
Colon	0.476 ± 0.004	0.529 ± 0.037	0.503 ± 0.002	0.476 ± 0.003	0.332 ± 0.007	0.485 ± 0.003	0.475	**0.549**
Heart	0.251 ± 0.142	0.337 ± 0.020	−0.002 ± 0.424	0.424 ± 0.007	**0.552** ± 0.003	0.473 ± 0.009	0.355	0.372
Kidney	0.529 ± 0.022	0.528 ± 0.008	0.555 ± 0.001	0.492 ± 0.013	0.511 ± 0.025	0.549 ± 0.002	0.544	**0.571**
Liver	0.742 ± 0.008	0.755 ± 0.002	0.736 ± 0.001	**0.764** ± 0.005	0.664 ± 0.059	0.716 ± 0.040	0.751	0.755
Lung	**0.366** ± 0.033	0.279 ± 0.022	−0.005 ± 0.317	0.005 ± 0.282	0.220 ± 0.008	0.211 ± 0.024	0.284	0.248
Prostate	0.207 ± 0.031	0.124 ± 0.011	0.318 ± 0.002	0.167 ± 0.072	0.071 ± 0.152	**0.375** ± 0.054	0.185	0.269
Skin	**0.473** ± 0.006	0.414 ± 0.025	0.256 ± 0.009	0.356 ± 0.023	0.442 ± 0.001	0.343 ± 0.069	0.429	0.450
Uterus	0.658 ± 0.035	0.676 ± 0.005	0.525 ± 0.204	0.690 ± 0.004	0.617 ± 0.052	0.583 ± 0.165	0.649	**0.698**
Breast	0.355 ± 0.007	**0.401** ± 0.009	0.368 ± 0.002	0.134 ± 0.038	0.337 ± 0.007	0.382 ± 0.009	0.351	0.373
Macro	0.469	0.471	0.392	0.418	0.434	0.477	0.465	**0.496**

Table 12: Cohort-level test results with DINOv2-Large and $G = 2000$. Neural entries are mean $\pm$ seed standard deviation over three runs. Linear and CGL-Linear are deterministic. Bold and underlining mark the best and second-best unrounded means in each row. Macro rows average cohort means without weighting by cohort size. Higher is better.

Cohort	HiST	HistoPrism	HEXST	STFlow	BLEEP	SEPAL	Linear	CGL-Linear
$\mathrm{PCC}_{\mathrm{flat}}$ ↑								
Brain	0.797 ± 0.000	0.810 ± 0.000	0.809 ± 0.000	**0.815** ± 0.002	0.771 ± 0.005	0.809 ± 0.001	0.788	0.812
Colon	0.729 ± 0.013	0.731 ± 0.012	0.702 ± 0.003	0.688 ± 0.003	0.659 ± 0.013	0.693 ± 0.005	0.686	**0.739**
Heart	**0.748** ± 0.002	0.738 ± 0.002	0.668 ± 0.058	0.719 ± 0.008	0.739 ± 0.002	0.731 ± 0.000	0.742	0.740
Kidney	0.736 ± 0.014	0.732 ± 0.002	0.741 ± 0.000	0.706 ± 0.023	0.716 ± 0.021	0.740 ± 0.002	0.737	**0.752**
Liver	0.867 ± 0.004	0.866 ± 0.002	0.859 ± 0.000	**0.868** ± 0.004	0.859 ± 0.008	0.851 ± 0.018	0.865	0.868
Lung	0.605 ± 0.017	0.563 ± 0.029	0.482 ± 0.147	0.586 ± 0.060	0.522 ± 0.043	0.596 ± 0.018	0.577	**0.608**
Prostate	0.569 ± 0.027	0.595 ± 0.004	0.676 ± 0.000	0.551 ± 0.016	0.575 ± 0.031	**0.695** ± 0.002	0.558	0.658
Skin	**0.679** ± 0.017	0.658 ± 0.022	0.501 ± 0.016	0.641 ± 0.036	0.647 ± 0.060	0.650 ± 0.020	0.664	0.675
Uterus	0.802 ± 0.007	0.821 ± 0.007	0.821 ± 0.000	0.814 ± 0.017	0.783 ± 0.007	0.762 ± 0.092	0.802	**0.840**
Breast	0.611 ± 0.003	**0.621** ± 0.005	0.588 ± 0.000	0.567 ± 0.014	0.546 ± 0.009	0.610 ± 0.012	0.576	0.594
Macro	0.714	0.714	0.685	0.696	0.682	0.714	0.700	**0.729**
$\mathrm{PCC}_{\mathrm{gene}}$ ↑								
Brain	0.021 ± 0.003	0.006 ± 0.002	0.023 ± 0.004	0.027 ± 0.002	0.022 ± 0.003	0.004 ± 0.001	0.029	**0.031**
Colon	0.149 ± 0.008	0.157 ± 0.027	**0.222** ± 0.016	0.020 ± 0.034	0.146 ± 0.010	0.081 ± 0.014	0.132	0.182
Heart	0.015 ± 0.000	0.015 ± 0.001	0.029 ± 0.010	0.032 ± 0.001	0.023 ± 0.002	0.027 ± 0.003	0.022	**0.035**
Kidney	**0.062** ± 0.003	0.038 ± 0.004	0.052 ± 0.001	0.027 ± 0.003	0.052 ± 0.010	0.025 ± 0.002	0.054	0.060
Liver	0.095 ± 0.001	0.069 ± 0.003	0.063 ± 0.003	0.086 ± 0.010	**0.100** ± 0.016	0.063 ± 0.017	0.097	0.098
Lung	0.225 ± 0.010	0.212 ± 0.000	**0.277** ± 0.019	0.208 ± 0.040	0.198 ± 0.037	0.166 ± 0.050	0.216	0.255
Prostate	−0.021 ± 0.007	−0.017 ± 0.006	−0.002 ± 0.013	−0.016 ± 0.007	**0.036** ± 0.034	0.035 ± 0.003	−0.028	0.020
Skin	0.299 ± 0.002	0.279 ± 0.016	**0.362** ± 0.008	0.257 ± 0.025	0.308 ± 0.058	0.273 ± 0.004	0.310	0.328
Uterus	0.076 ± 0.013	0.073 ± 0.003	0.084 ± 0.004	0.061 ± 0.003	0.067 ± 0.009	0.032 ± 0.033	0.063	**0.089**
Breast	**0.133** ± 0.000	0.132 ± 0.014	0.066 ± 0.010	0.068 ± 0.012	0.054 ± 0.010	0.104 ± 0.043	0.066	0.074
Macro	0.105	0.096	**0.118**	0.077	0.101	0.081	0.096	0.117
R^2_{local} ↑								
Brain	−0.108 ± 0.001	−0.009 ± 0.001	**0.009** ± 0.001	−0.027 ± 0.000	−0.285 ± 0.010	0.007 ± 0.000	−0.212	−0.030
Colon	0.148 ± 0.007	0.106 ± 0.007	0.040 ± 0.013	0.013 ± 0.017	0.011 ± 0.023	0.024 ± 0.022	−0.022	**0.164**
Heart	−0.090 ± 0.043	**−0.009** ± 0.006	−0.212 ± 0.301	−0.067 ± 0.032	−0.189 ± 0.005	−0.010 ± 0.005	−0.253	−0.030
Kidney	0.024 ± 0.000	0.023 ± 0.001	0.012 ± 0.001	−0.038 ± 0.026	−0.118 ± 0.011	−0.004 ± 0.009	−0.094	**0.039**
Liver	0.033 ± 0.004	−0.006 ± 0.002	0.013 ± 0.001	0.000 ± 0.006	−0.094 ± 0.034	0.010 ± 0.001	−0.050	**0.037**
Lung	**0.189** ± 0.013	0.113 ± 0.006	−0.042 ± 0.130	0.081 ± 0.087	0.094 ± 0.036	0.069 ± 0.027	0.134	0.113
Prostate	−0.200 ± 0.033	−0.076 ± 0.023	−0.005 ± 0.003	−0.209 ± 0.028	−0.017 ± 0.067	**0.004** ± 0.001	−0.176	−0.056
Skin	**0.387** ± 0.002	0.334 ± 0.009	0.040 ± 0.021	0.253 ± 0.054	0.272 ± 0.074	0.284 ± 0.018	0.303	0.334
Uterus	0.052 ± 0.017	0.081 ± 0.005	0.036 ± 0.003	0.004 ± 0.064	−0.058 ± 0.005	0.024 ± 0.034	−0.062	**0.093**
Breast	0.047 ± 0.011	**0.054** ± 0.002	0.006 ± 0.001	−0.085 ± 0.045	−0.163 ± 0.157	0.013 ± 0.006	−0.040	0.036
Macro	0.048	0.061	−0.010	−0.007	−0.055	0.042	−0.047	**0.070**
R^2_{flat} ↑								
Brain	0.632 ± 0.000	0.654 ± 0.000	0.653 ± 0.001	**0.663** ± 0.004	0.591 ± 0.009	0.649 ± 0.004	0.615	0.660
Colon	0.525 ± 0.012	0.499 ± 0.045	0.487 ± 0.004	0.468 ± 0.002	0.375 ± 0.016	0.475 ± 0.002	0.470	**0.543**
Heart	0.086 ± 0.170	0.327 ± 0.001	−0.009 ± 0.472	0.455 ± 0.015	**0.510** ± 0.016	0.482 ± 0.009	0.358	0.386
Kidney	0.533 ± 0.031	0.530 ± 0.009	0.548 ± 0.000	0.481 ± 0.042	0.503 ± 0.036	0.547 ± 0.004	0.537	**0.566**
Liver	0.749 ± 0.006	0.741 ± 0.002	0.726 ± 0.001	**0.751** ± 0.009	0.731 ± 0.022	0.688 ± 0.080	0.748	0.748
Lung	**0.338** ± 0.053	0.282 ± 0.013	0.032 ± 0.212	0.329 ± 0.085	0.145 ± 0.080	0.193 ± 0.053	0.270	0.232
Prostate	0.194 ± 0.059	0.101 ± 0.018	0.290 ± 0.000	0.164 ± 0.037	0.097 ± 0.006	**0.352** ± 0.028	0.162	0.248
Skin	0.437 ± 0.044	0.430 ± 0.029	0.250 ± 0.015	0.397 ± 0.060	0.394 ± 0.066	0.401 ± 0.028	0.434	**0.455**
Uterus	0.628 ± 0.004	0.665 ± 0.013	0.663 ± 0.002	0.644 ± 0.054	0.592 ± 0.012	0.487 ± 0.240	0.639	**0.688**
Breast	0.361 ± 0.015	**0.384** ± 0.006	0.344 ± 0.000	0.185 ± 0.087	0.203 ± 0.127	0.368 ± 0.011	0.322	0.348
Macro	0.448	0.461	0.399	0.453	0.414	0.464	0.455	**0.487**

Table 13: Cohort-level test results with ResNet18 and $G = 500$. Neural entries are mean ± seed standard deviation over three runs. Linear and CGL-Linear are deterministic. Bold and underlining mark the best and second-best unrounded means in each row. Macro rows average cohort means without weighting by cohort size. Higher is better.

Cohort	HiST	HistoPrism	HEXST	STFlow	BLEEP	SEPAL	Linear	CGL-Linear
$\mathrm{PCC}_{\mathrm{flat}}$ ↑								
Brain	0.761 ± 0.004	0.772 ± 0.001	0.763 ± 0.002	**0.778** ± 0.002	0.704 ± 0.045	0.774 ± 0.002	0.755	0.774
Colon	0.706 ± 0.024	0.717 ± 0.005	0.709 ± 0.008	0.688 ± 0.002	0.511 ± 0.037	0.686 ± 0.014	0.693	**0.720**
Heart	0.710 ± 0.014	0.719 ± 0.003	0.647 ± 0.000	0.654 ± 0.000	**0.727** ± 0.005	0.644 ± 0.006	0.642	0.661
Kidney	0.671 ± 0.013	0.679 ± 0.007	0.711 ± 0.006	0.708 ± 0.001	0.687 ± 0.001	0.710 ± 0.030	0.716	**0.725**
Liver	0.869 ± 0.002	0.875 ± 0.000	0.856 ± 0.006	0.860 ± 0.003	0.852 ± 0.009	0.860 ± 0.001	0.853	**0.877**
Lung	0.598 ± 0.013	0.638 ± 0.010	0.586 ± 0.035	0.556 ± 0.040	0.473 ± 0.002	**0.649** ± 0.008	0.578	0.619
Prostate	0.600 ± 0.014	0.658 ± 0.035	**0.726** ± 0.006	0.656 ± 0.032	0.555 ± 0.044	0.690 ± 0.027	0.540	0.678
Skin	0.611 ± 0.029	0.612 ± 0.030	0.485 ± 0.009	0.583 ± 0.017	0.610 ± 0.024	0.553 ± 0.095	**0.620**	0.608
Uterus	0.790 ± 0.015	0.792 ± 0.009	0.776 ± 0.015	0.844 ± 0.001	0.798 ± 0.003	**0.864** ± 0.001	0.836	0.861
Breast	0.661 ± 0.012	0.692 ± 0.004	0.672 ± 0.001	0.644 ± 0.002	0.614 ± 0.009	**0.704** ± 0.009	0.665	0.678
Macro	0.698	0.715	0.693	0.697	0.653	0.713	0.690	**0.720**
$\mathrm{PCC}_{\mathrm{gene}}$ ↑								
Brain	0.024 ± 0.007	0.016 ± 0.004	0.031 ± 0.004	0.024 ± 0.002	0.026 ± 0.003	0.018 ± 0.003	0.036	**0.038**
Colon	0.155 ± 0.014	0.121 ± 0.019	**0.207** ± 0.010	−0.023 ± 0.011	0.073 ± 0.029	0.094 ± 0.022	0.129	0.152
Heart	0.017 ± 0.005	0.026 ± 0.001	0.025 ± 0.009	−0.002 ± 0.002	0.022 ± 0.005	0.018 ± 0.001	0.022	**0.041**
Kidney	0.059 ± 0.001	0.042 ± 0.008	0.060 ± 0.000	0.005 ± 0.006	0.066 ± 0.011	0.040 ± 0.019	0.064	**0.068**
Liver	0.087 ± 0.000	0.076 ± 0.004	0.024 ± 0.013	0.085 ± 0.010	0.080 ± 0.009	0.061 ± 0.005	0.068	**0.104**
Lung	0.151 ± 0.004	0.173 ± 0.014	0.197 ± 0.039	0.138 ± 0.043	0.136 ± 0.003	**0.233** ± 0.007	0.167	0.206
Prostate	0.033 ± 0.001	0.043 ± 0.008	**0.106** ± 0.009	0.055 ± 0.018	0.069 ± 0.020	0.057 ± 0.016	0.008	0.074
Skin	0.322 ± 0.028	0.292 ± 0.024	**0.363** ± 0.009	0.278 ± 0.012	0.335 ± 0.023	0.267 ± 0.052	0.334	0.322
Uterus	0.062 ± 0.004	0.048 ± 0.000	**0.125** ± 0.003	0.029 ± 0.014	0.062 ± 0.000	0.091 ± 0.007	0.088	0.114
Breast	0.071 ± 0.008	0.121 ± 0.008	0.078 ± 0.012	0.072 ± 0.001	0.035 ± 0.012	**0.152** ± 0.018	0.072	0.086
Macro	0.098	0.096	**0.122**	0.066	0.090	0.103	0.099	0.121
R^2_{local} ↑								
Brain	−0.052 ± 0.011	−0.007 ± 0.002	**0.004** ± 0.002	−0.015 ± 0.002	−0.551 ± 0.350	0.004 ± 0.008	−0.224	−0.010
Colon	**0.090** ± 0.002	0.029 ± 0.009	0.059 ± 0.032	0.000 ± 0.000	−0.220 ± 0.134	0.008 ± 0.011	−0.105	0.079
Heart	−0.169 ± 0.045	−0.051 ± 0.020	**0.005** ± 0.001	0.000 ± 0.000	−0.155 ± 0.019	−0.178 ± 0.067	−0.568	−0.092
Kidney	0.005 ± 0.004	0.010 ± 0.005	0.021 ± 0.002	0.000 ± 0.000	−0.192 ± 0.075	−0.033 ± 0.002	−0.102	**0.028**
Liver	−0.029 ± 0.032	0.001 ± 0.005	−0.013 ± 0.021	−0.045 ± 0.044	−0.257 ± 0.184	0.014 ± 0.002	−0.164	**0.019**
Lung	**0.141** ± 0.014	0.098 ± 0.001	0.074 ± 0.042	0.025 ± 0.064	0.038 ± 0.001	0.127 ± 0.001	0.097	0.106
Prostate	0.002 ± 0.006	0.013 ± 0.003	**0.030** ± 0.014	0.006 ± 0.031	−0.027 ± 0.009	0.001 ± 0.036	−0.110	0.029
Skin	**0.342** ± 0.030	0.288 ± 0.027	0.065 ± 0.012	0.227 ± 0.021	0.224 ± 0.043	0.235 ± 0.059	0.257	0.257
Uterus	0.081 ± 0.004	0.065 ± 0.000	−0.360 ± 0.082	0.008 ± 0.004	−0.195 ± 0.082	0.064 ± 0.019	−0.060	**0.095**
Breast	0.048 ± 0.012	**0.052** ± 0.003	0.014 ± 0.000	−0.173 ± 0.001	−0.217 ± 0.032	0.031 ± 0.006	−0.025	0.044
Macro	0.046	0.050	−0.010	0.003	−0.155	0.027	−0.100	**0.055**
R^2_{flat} ↑								
Brain	0.576 ± 0.005	0.595 ± 0.003	0.581 ± 0.003	**0.605** ± 0.004	0.455 ± 0.114	0.598 ± 0.002	0.554	0.593
Colon	0.432 ± 0.058	0.431 ± 0.032	0.495 ± 0.012	0.466 ± 0.003	0.174 ± 0.094	0.465 ± 0.020	0.480	**0.512**
Heart	0.230 ± 0.122	0.422 ± 0.004	0.208 ± 0.000	0.250 ± 0.001	**0.523** ± 0.010	0.319 ± 0.019	0.183	0.236
Kidney	0.420 ± 0.001	0.426 ± 0.023	0.505 ± 0.007	0.500 ± 0.002	0.462 ± 0.004	0.501 ± 0.046	0.509	**0.525**
Liver	0.754 ± 0.003	**0.763** ± 0.000	0.721 ± 0.012	0.737 ± 0.004	0.722 ± 0.014	0.738 ± 0.002	0.725	0.763
Lung	0.210 ± 0.041	0.210 ± 0.008	0.202 ± 0.027	**0.284** ± 0.034	0.036 ± 0.008	0.253 ± 0.028	0.234	0.211
Prostate	0.036 ± 0.016	0.120 ± 0.036	**0.353** ± 0.025	0.181 ± 0.042	0.120 ± 0.063	0.262 ± 0.050	0.088	0.207
Skin	0.367 ± 0.035	0.371 ± 0.041	0.231 ± 0.008	0.322 ± 0.026	0.348 ± 0.009	0.249 ± 0.155	**0.380**	0.366
Uterus	0.612 ± 0.026	0.603 ± 0.018	0.584 ± 0.040	0.706 ± 0.001	0.625 ± 0.011	**0.729** ± 0.009	0.690	0.713
Breast	0.426 ± 0.019	0.474 ± 0.010	0.451 ± 0.002	0.202 ± 0.045	0.326 ± 0.038	**0.492** ± 0.011	0.436	0.457
Macro	0.406	0.441	0.433	0.425	0.379	**0.461**	0.428	0.458

Table 14: Cohort-level test results with ResNet18 and $G = 1000$. Neural entries are mean ± seed standard deviation over three runs. Linear and CGL-Linear are deterministic. Bold and underlining mark the best and second-best unrounded means in each row. Macro rows average cohort means without weighting by cohort size. Higher is better.

Cohort	HiST	HistoPrism	HEXST	STFlow	BLEEP	SEPAL	Linear	CGL-Linear
$\mathrm{PCC}_{\mathrm{flat}}$ ↑								
Brain	0.801 ± 0.006	0.808 ± 0.002	0.796 ± 0.003	**0.818** ± 0.003	0.752 ± 0.036	0.807 ± 0.001	0.793	0.808
Colon	**0.736** ± 0.012	0.710 ± 0.001	0.705 ± 0.005	0.685 ± 0.001	0.516 ± 0.096	0.682 ± 0.020	0.694	0.718
Heart	0.721 ± 0.021	**0.724** ± 0.031	0.674 ± 0.001	0.675 ± 0.011	0.713 ± 0.008	0.685 ± 0.007	0.679	0.695
Kidney	0.701 ± 0.004	0.690 ± 0.001	0.704 ± 0.000	0.699 ± 0.001	0.689 ± 0.008	**0.718** ± 0.014	0.706	0.716
Liver	0.865 ± 0.003	0.871 ± 0.004	0.852 ± 0.006	0.862 ± 0.006	0.841 ± 0.022	0.863 ± 0.007	0.854	**0.874**
Lung	0.495 ± 0.043	0.613 ± 0.006	0.582 ± 0.002	0.526 ± 0.004	0.449 ± 0.044	**0.626** ± 0.004	0.559	0.597
Prostate	0.571 ± 0.037	0.687 ± 0.014	**0.705** ± 0.004	0.612 ± 0.077	0.550 ± 0.076	0.683 ± 0.052	0.522	0.662
Skin	**0.654** ± 0.024	0.603 ± 0.001	0.475 ± 0.000	0.651 ± 0.014	0.562 ± 0.056	0.539 ± 0.040	0.632	0.621
Uterus	0.767 ± 0.006	0.796 ± 0.003	0.739 ± 0.011	0.817 ± 0.024	0.790 ± 0.003	0.807 ± 0.059	0.823	**0.848**
Breast	0.627 ± 0.009	0.658 ± 0.002	0.631 ± 0.000	0.557 ± 0.002	0.565 ± 0.017	**0.665** ± 0.004	0.625	0.638
Macro	0.694	0.716	0.686	0.690	0.643	0.708	0.689	**0.718**
$\mathrm{PCC}_{\mathrm{gene}}$ ↑								
Brain	0.021 ± 0.004	0.017 ± 0.001	0.023 ± 0.011	0.031 ± 0.007	0.014 ± 0.008	0.020 ± 0.004	0.033	**0.035**
Colon	0.169 ± 0.010	0.133 ± 0.004	**0.208** ± 0.015	0.008 ± 0.017	0.076 ± 0.047	0.089 ± 0.040	0.130	0.149
Heart	0.018 ± 0.006	0.026 ± 0.007	0.038 ± 0.002	0.016 ± 0.022	0.018 ± 0.004	0.011 ± 0.015	0.022	**0.040**
Kidney	**0.065** ± 0.001	0.038 ± 0.002	0.047 ± 0.005	0.003 ± 0.001	0.058 ± 0.002	0.043 ± 0.010	0.058	0.061
Liver	0.078 ± 0.000	0.077 ± 0.002	0.040 ± 0.019	0.083 ± 0.019	0.068 ± 0.025	0.061 ± 0.001	0.068	**0.104**
Lung	0.112 ± 0.024	0.157 ± 0.003	0.167 ± 0.021	0.045 ± 0.078	0.122 ± 0.005	**0.209** ± 0.020	0.150	0.188
Prostate	0.047 ± 0.002	0.072 ± 0.001	**0.094** ± 0.008	0.045 ± 0.029	0.051 ± 0.038	0.070 ± 0.004	−0.005	0.070
Skin	0.309 ± 0.016	0.254 ± 0.003	**0.331** ± 0.004	0.309 ± 0.025	0.289 ± 0.008	0.206 ± 0.012	0.315	0.305
Uterus	0.058 ± 0.005	0.048 ± 0.001	**0.111** ± 0.003	0.056 ± 0.015	0.064 ± 0.008	0.050 ± 0.034	0.075	0.095
Breast	0.065 ± 0.014	0.122 ± 0.001	0.074 ± 0.007	0.069 ± 0.012	0.031 ± 0.014	**0.141** ± 0.019	0.070	0.084
Macro	0.094	0.094	**0.113**	0.066	0.079	0.090	0.091	0.113
R^2_{local} ↑								
Brain	−0.051 ± 0.008	−0.001 ± 0.001	0.008 ± 0.004	0.004 ± 0.008	−0.332 ± 0.063	**0.014** ± 0.001	−0.187	−0.004
Colon	**0.125** ± 0.014	0.060 ± 0.001	0.045 ± 0.015	0.000 ± 0.000	−0.118 ± 0.136	0.038 ± 0.032	−0.044	0.110
Heart	−0.095 ± 0.037	−0.031 ± 0.036	**0.004** ± 0.002	−0.039 ± 0.055	−0.224 ± 0.044	−0.078 ± 0.054	−0.403	−0.061
Kidney	0.019 ± 0.010	0.010 ± 0.002	0.016 ± 0.005	0.000 ± 0.000	−0.114 ± 0.011	−0.013 ± 0.026	−0.084	**0.027**
Liver	−0.084 ± 0.078	0.007 ± 0.001	−0.014 ± 0.029	0.013 ± 0.005	−0.218 ± 0.158	0.012 ± 0.009	−0.150	**0.026**
Lung	0.086 ± 0.017	0.079 ± 0.008	0.033 ± 0.002	0.009 ± 0.013	0.041 ± 0.002	**0.118** ± 0.024	0.088	0.093
Prostate	0.018 ± 0.004	**0.040** ± 0.003	0.022 ± 0.009	−0.026 ± 0.056	−0.034 ± 0.032	0.034 ± 0.002	−0.106	0.027
Skin	**0.361** ± 0.013	0.266 ± 0.004	0.040 ± 0.000	0.296 ± 0.017	0.194 ± 0.006	0.190 ± 0.000	0.258	0.262
Uterus	0.063 ± 0.004	0.065 ± 0.005	−0.419 ± 0.046	−0.052 ± 0.104	−0.137 ± 0.034	0.025 ± 0.027	−0.041	**0.090**
Breast	**0.049** ± 0.005	0.047 ± 0.002	0.008 ± 0.002	−0.277 ± 0.012	−0.126 ± 0.053	0.031 ± 0.007	−0.019	0.040
Macro	0.049	0.054	−0.026	−0.007	−0.107	0.037	−0.069	**0.061**
R^2_{flat} ↑								
Brain	0.636 ± 0.014	0.649 ± 0.003	0.633 ± 0.004	**0.669** ± 0.004	0.558 ± 0.060	0.649 ± 0.000	0.622	0.652
Colon	**0.523** ± 0.027	0.462 ± 0.011	0.491 ± 0.007	0.466 ± 0.002	0.139 ± 0.230	0.424 ± 0.084	0.481	0.513
Heart	0.220 ± 0.084	0.380 ± 0.086	0.248 ± 0.001	0.303 ± 0.002	**0.500** ± 0.017	0.394 ± 0.035	0.233	0.288
Kidney	0.479 ± 0.011	0.460 ± 0.003	0.495 ± 0.000	0.488 ± 0.001	0.464 ± 0.015	**0.515** ± 0.020	0.495	0.513
Liver	0.747 ± 0.004	0.755 ± 0.009	0.719 ± 0.007	0.741 ± 0.010	0.695 ± 0.039	0.744 ± 0.012	0.727	**0.759**
Lung	0.081 ± 0.077	0.171 ± 0.023	0.170 ± 0.006	0.208 ± 0.078	0.046 ± 0.057	**0.271** ± 0.018	0.230	0.206
Prostate	0.060 ± 0.043	0.101 ± 0.027	**0.329** ± 0.009	0.140 ± 0.118	0.116 ± 0.122	0.272 ± 0.117	0.087	0.203
Skin	**0.421** ± 0.032	0.358 ± 0.003	0.225 ± 0.000	0.412 ± 0.012	0.288 ± 0.062	0.229 ± 0.092	0.396	0.384
Uterus	0.578 ± 0.008	0.609 ± 0.011	0.519 ± 0.024	0.656 ± 0.049	0.604 ± 0.012	0.615 ± 0.107	0.670	**0.693**
Breast	0.383 ± 0.015	0.431 ± 0.002	0.398 ± 0.001	−0.020 ± 0.001	0.303 ± 0.023	**0.440** ± 0.006	0.384	0.404
Macro	0.413	0.438	0.423	0.406	0.371	0.455	0.433	**0.461**

Table 15: Cohort-level test results with ResNet18 and $G = 1500$. Neural entries are mean ± seed standard deviation over three runs. Linear and CGL-Linear are deterministic. Bold and underlining mark the best and second-best unrounded means in each row. Macro rows average cohort means without weighting by cohort size. Higher is better.

Cohort	HiST	HistoPrism	HEXST	STFlow	BLEEP	SEPAL	Linear	CGL-Linear
$\mathrm{PCC}_{\mathrm{flat}}$ ↑								
Brain	0.813 ± 0.001	0.818 ± 0.001	0.809 ± 0.002	0.817 ± 0.007	0.766 ± 0.007	0.817 ± 0.000	0.805	**0.819**
Colon	0.721 ± 0.001	0.724 ± 0.004	0.711 ± 0.005	0.694 ± 0.001	0.558 ± 0.008	0.706 ± 0.016	0.705	**0.727**
Heart	**0.748** ± 0.012	0.740 ± 0.021	0.699 ± 0.000	0.707 ± 0.001	0.741 ± 0.001	0.711 ± 0.001	0.707	0.722
Kidney	0.741 ± 0.019	0.738 ± 0.005	0.744 ± 0.000	0.739 ± 0.001	0.723 ± 0.018	0.745 ± 0.011	0.745	**0.755**
Liver	0.869 ± 0.006	**0.880** ± 0.001	0.846 ± 0.025	0.867 ± 0.007	0.857 ± 0.013	0.855 ± 0.008	0.857	0.880
Lung	0.559 ± 0.001	0.599 ± 0.004	0.558 ± 0.028	0.535 ± 0.016	0.478 ± 0.031	**0.615** ± 0.007	0.562	0.597
Prostate	0.573 ± 0.046	0.648 ± 0.046	**0.698** ± 0.001	0.577 ± 0.124	0.436 ± 0.007	0.686 ± 0.036	0.516	0.658
Skin	0.679 ± 0.015	0.603 ± 0.019	0.487 ± 0.005	**0.687** ± 0.001	0.585 ± 0.012	0.552 ± 0.013	0.649	0.638
Uterus	0.772 ± 0.002	0.810 ± 0.000	0.704 ± 0.003	0.811 ± 0.027	0.810 ± 0.020	0.779 ± 0.092	0.820	**0.844**
Breast	0.612 ± 0.019	0.640 ± 0.001	0.609 ± 0.000	0.583 ± 0.017	0.556 ± 0.006	**0.645** ± 0.012	0.603	0.616
Macro	0.709	0.720	0.687	0.702	0.651	0.711	0.697	**0.726**
$\mathrm{PCC}_{\mathrm{gene}}$ ↑								
Brain	0.022 ± 0.004	0.017 ± 0.000	0.021 ± 0.003	0.022 ± 0.004	0.009 ± 0.006	0.018 ± 0.000	0.030	**0.031**
Colon	0.148 ± 0.002	0.130 ± 0.004	**0.220** ± 0.005	0.020 ± 0.033	0.078 ± 0.025	0.091 ± 0.021	0.129	0.148
Heart	0.021 ± 0.006	0.024 ± 0.009	0.036 ± 0.002	0.001 ± 0.001	0.030 ± 0.001	0.031 ± 0.008	0.022	**0.040**
Kidney	**0.061** ± 0.018	0.041 ± 0.006	0.046 ± 0.002	0.001 ± 0.000	0.057 ± 0.005	0.037 ± 0.013	0.056	0.059
Liver	0.077 ± 0.012	0.086 ± 0.000	0.026 ± 0.021	0.085 ± 0.028	0.082 ± 0.018	0.049 ± 0.034	0.070	**0.111**
Lung	0.126 ± 0.014	0.169 ± 0.000	**0.185** ± 0.056	0.071 ± 0.103	0.144 ± 0.022	0.183 ± 0.020	0.146	0.184
Prostate	0.034 ± 0.014	0.045 ± 0.011	**0.090** ± 0.010	0.031 ± 0.049	0.012 ± 0.017	0.067 ± 0.009	−0.011	0.066
Skin	0.310 ± 0.022	0.233 ± 0.009	**0.325** ± 0.012	0.323 ± 0.004	0.305 ± 0.001	0.218 ± 0.037	0.309	0.301
Uterus	0.053 ± 0.005	0.055 ± 0.000	**0.102** ± 0.007	0.049 ± 0.010	0.065 ± 0.004	0.040 ± 0.032	0.068	0.087
Breast	0.075 ± 0.038	0.130 ± 0.006	0.087 ± 0.006	0.079 ± 0.020	0.039 ± 0.011	**0.147** ± 0.028	0.071	0.085
Macro	0.093	0.093	**0.114**	0.068	0.082	0.088	0.089	0.111
R^2_{local} ↑								
Brain	−0.071 ± 0.003	0.001 ± 0.002	0.006 ± 0.002	0.001 ± 0.019	−0.355 ± 0.118	**0.011** ± 0.000	−0.170	−0.003
Colon	**0.132** ± 0.003	0.055 ± 0.004	0.042 ± 0.016	0.000 ± 0.000	−0.126 ± 0.033	0.041 ± 0.020	−0.020	0.122
Heart	−0.090 ± 0.028	−0.036 ± 0.053	**0.002** ± 0.001	0.000 ± 0.000	−0.206 ± 0.015	−0.018 ± 0.008	−0.332	−0.048
Kidney	0.013 ± 0.005	0.015 ± 0.004	0.015 ± 0.001	0.000 ± 0.000	−0.145 ± 0.018	0.005 ± 0.004	−0.080	**0.029**
Liver	−0.033 ± 0.008	0.017 ± 0.004	−0.195 ± 0.296	0.014 ± 0.015	−0.203 ± 0.170	0.022 ± 0.007	−0.136	**0.034**
Lung	**0.104** ± 0.007	0.086 ± 0.001	0.061 ± 0.043	0.019 ± 0.027	0.055 ± 0.017	0.086 ± 0.012	0.087	0.090
Prostate	0.009 ± 0.004	0.019 ± 0.009	0.016 ± 0.002	−0.017 ± 0.070	−0.017 ± 0.005	**0.036** ± 0.013	−0.101	0.025
Skin	**0.359** ± 0.020	0.248 ± 0.011	0.028 ± 0.006	0.323 ± 0.004	0.227 ± 0.002	0.212 ± 0.066	0.263	0.267
Uterus	0.057 ± 0.002	0.068 ± 0.001	−0.632 ± 0.104	−0.037 ± 0.086	−0.080 ± 0.000	0.029 ± 0.038	−0.039	**0.087**
Breast	0.038 ± 0.008	**0.045** ± 0.001	0.007 ± 0.000	−0.143 ± 0.077	−0.093 ± 0.007	0.025 ± 0.005	−0.017	0.036
Macro	0.052	0.052	−0.065	0.016	−0.094	0.045	−0.055	**0.064**
R^2_{flat} ↑								
Brain	0.657 ± 0.002	0.666 ± 0.001	0.654 ± 0.003	0.667 ± 0.013	0.579 ± 0.019	0.664 ± 0.000	0.643	**0.670**
Colon	0.478 ± 0.015	0.450 ± 0.017	0.501 ± 0.006	0.477 ± 0.002	0.225 ± 0.039	0.483 ± 0.009	0.497	**0.526**
Heart	0.328 ± 0.073	0.380 ± 0.087	0.298 ± 0.000	0.339 ± 0.003	**0.534** ± 0.003	0.463 ± 0.012	0.290	0.342
Kidney	0.533 ± 0.038	0.535 ± 0.009	0.553 ± 0.000	0.545 ± 0.001	0.518 ± 0.029	0.554 ± 0.017	0.553	**0.569**
Liver	0.752 ± 0.009	**0.770** ± 0.004	0.710 ± 0.035	0.747 ± 0.010	0.727 ± 0.017	0.708 ± 0.028	0.733	0.767
Lung	0.197 ± 0.000	0.226 ± 0.049	0.193 ± 0.036	0.165 ± 0.022	0.088 ± 0.007	0.217 ± 0.044	**0.230**	0.202
Prostate	0.055 ± 0.043	0.096 ± 0.023	**0.316** ± 0.001	0.068 ± 0.132	−0.057 ± 0.023	0.283 ± 0.087	0.084	0.197
Skin	0.441 ± 0.006	0.354 ± 0.023	0.237 ± 0.004	**0.454** ± 0.001	0.322 ± 0.020	0.243 ± 0.051	0.418	0.406
Uterus	0.589 ± 0.004	0.631 ± 0.000	0.446 ± 0.002	0.646 ± 0.056	0.645 ± 0.044	0.541 ± 0.184	0.665	**0.689**
Breast	0.362 ± 0.029	0.408 ± 0.003	0.369 ± 0.000	0.138 ± 0.056	0.294 ± 0.007	**0.411** ± 0.012	0.357	0.375
Macro	0.439	0.452	0.428	0.425	0.387	0.457	0.447	**0.474**

Table 16: Cohort-level test results with ResNet18 and $G = 2000$. Neural entries are mean ± seed standard deviation over three runs. Linear and CGL-Linear are deterministic. Bold and underlining mark the best and second-best unrounded means in each row. Macro rows average cohort means without weighting by cohort size. Higher is better.

Cohort	HiST	HistoPrism	HEXST	STFlow	BLEEP	SEPAL	Linear	CGL-Linear
$\mathrm{PCC}_{\mathrm{flat}}$ ↑								
Brain	0.810 ± 0.000	0.812 ± 0.000	0.806 ± 0.002	0.812 ± 0.000	0.763 ± 0.010	0.809 ± 0.001	0.801	**0.813**
Colon	**0.730** ± 0.004	0.719 ± 0.003	0.699 ± 0.001	0.686 ± 0.000	0.579 ± 0.029	0.700 ± 0.012	0.699	0.721
Heart	**0.753** ± 0.006	0.748 ± 0.022	0.709 ± 0.000	0.705 ± 0.017	0.741 ± 0.004	0.718 ± 0.019	0.718	0.732
Kidney	0.740 ± 0.001	0.731 ± 0.001	0.739 ± 0.000	0.731 ± 0.000	0.723 ± 0.005	0.739 ± 0.004	0.741	**0.751**
Liver	0.866 ± 0.003	0.873 ± 0.002	0.833 ± 0.035	0.865 ± 0.012	0.854 ± 0.011	0.850 ± 0.024	0.855	**0.877**
Lung	0.546 ± 0.007	0.578 ± 0.009	0.510 ± 0.067	0.510 ± 0.001	0.447 ± 0.004	**0.600** ± 0.017	0.547	0.581
Prostate	0.606 ± 0.034	0.614 ± 0.041	**0.681** ± 0.001	0.667 ± 0.022	0.493 ± 0.052	0.664 ± 0.016	0.500	0.642
Skin	**0.679** ± 0.001	0.635 ± 0.001	0.493 ± 0.003	0.648 ± 0.004	0.517 ± 0.211	0.598 ± 0.026	0.653	0.642
Uterus	0.785 ± 0.019	0.805 ± 0.004	0.754 ± 0.093	0.815 ± 0.005	0.803 ± 0.004	0.765 ± 0.096	0.813	**0.837**
Breast	0.583 ± 0.011	0.615 ± 0.005	0.589 ± 0.000	0.536 ± 0.010	0.525 ± 0.013	**0.621** ± 0.013	0.581	0.595
Macro	0.710	0.713	0.681	0.698	0.645	0.706	0.691	**0.719**
$\mathrm{PCC}_{\mathrm{gene}}$ ↑								
Brain	0.027 ± 0.003	0.018 ± 0.000	0.026 ± 0.002	0.012 ± 0.018	0.018 ± 0.010	0.019 ± 0.001	0.031	**0.032**
Colon	0.155 ± 0.007	0.134 ± 0.013	**0.211** ± 0.000	−0.004 ± 0.014	0.101 ± 0.037	0.075 ± 0.006	0.129	0.147
Heart	0.019 ± 0.000	0.026 ± 0.009	0.035 ± 0.003	0.017 ± 0.024	0.028 ± 0.002	0.029 ± 0.008	0.021	**0.039**
Kidney	**0.060** ± 0.002	0.035 ± 0.006	0.040 ± 0.003	0.000 ± 0.000	0.051 ± 0.004	0.025 ± 0.007	0.052	0.056
Liver	0.084 ± 0.001	0.091 ± 0.001	0.014 ± 0.016	0.098 ± 0.021	0.093 ± 0.020	0.050 ± 0.044	0.076	**0.119**
Lung	0.122 ± 0.008	0.160 ± 0.007	0.171 ± 0.051	−0.001 ± 0.000	0.100 ± 0.005	**0.185** ± 0.043	0.142	0.180
Prostate	0.047 ± 0.043	0.039 ± 0.004	**0.087** ± 0.023	0.067 ± 0.038	0.027 ± 0.022	0.054 ± 0.025	−0.014	0.067
Skin	0.292 ± 0.005	0.248 ± 0.003	**0.322** ± 0.017	0.267 ± 0.008	0.240 ± 0.108	0.227 ± 0.018	0.300	0.293
Uterus	0.058 ± 0.006	0.055 ± 0.002	**0.083** ± 0.015	0.040 ± 0.032	0.064 ± 0.006	0.039 ± 0.037	0.064	0.082
Breast	0.075 ± 0.011	0.127 ± 0.010	0.098 ± 0.008	0.070 ± 0.010	0.033 ± 0.009	**0.133** ± 0.040	0.071	0.085
Macro	0.094	0.093	0.109	0.057	0.076	0.084	0.087	**0.110**
R^2_{local} ↑								
Brain	−0.059 ± 0.014	0.003 ± 0.002	0.007 ± 0.002	**0.012** ± 0.017	−0.311 ± 0.011	0.005 ± 0.006	−0.151	0.000
Colon	**0.143** ± 0.016	0.063 ± 0.004	0.028 ± 0.003	0.000 ± 0.000	−0.149 ± 0.085	0.030 ± 0.013	−0.008	0.128
Heart	−0.071 ± 0.005	−0.028 ± 0.040	**0.001** ± 0.000	−0.059 ± 0.083	−0.162 ± 0.005	−0.015 ± 0.005	−0.295	−0.042
Kidney	0.026 ± 0.008	0.014 ± 0.002	0.010 ± 0.002	0.000 ± 0.000	−0.108 ± 0.022	0.003 ± 0.001	−0.075	**0.029**
Liver	−0.049 ± 0.034	0.008 ± 0.002	−0.305 ± 0.440	−0.003 ± 0.001	−0.186 ± 0.124	0.017 ± 0.004	−0.127	**0.034**
Lung	**0.101** ± 0.003	0.075 ± 0.000	0.054 ± 0.049	0.000 ± 0.000	0.023 ± 0.008	0.091 ± 0.039	0.082	0.083
Prostate	0.024 ± 0.049	0.016 ± 0.001	0.011 ± 0.005	**0.041** ± 0.039	−0.017 ± 0.001	0.018 ± 0.030	−0.094	0.024
Skin	**0.371** ± 0.009	0.276 ± 0.011	0.029 ± 0.003	0.256 ± 0.023	0.170 ± 0.142	0.218 ± 0.023	0.265	0.270
Uterus	0.060 ± 0.011	0.066 ± 0.002	−0.350 ± 0.533	0.017 ± 0.010	−0.045 ± 0.018	0.025 ± 0.036	−0.035	**0.082**
Breast	0.036 ± 0.008	**0.043** ± 0.004	0.006 ± 0.000	−0.218 ± 0.019	−0.110 ± 0.047	0.023 ± 0.006	−0.020	0.034
Macro	0.058	0.054	−0.051	0.005	−0.090	0.041	−0.046	**0.064**
R^2_{flat} ↑								
Brain	0.652 ± 0.004	0.656 ± 0.000	0.649 ± 0.003	0.657 ± 0.001	0.574 ± 0.017	0.651 ± 0.003	0.636	**0.661**
Colon	0.501 ± 0.019	0.445 ± 0.001	0.484 ± 0.001	0.468 ± 0.000	0.273 ± 0.083	0.459 ± 0.024	0.488	**0.517**
Heart	0.356 ± 0.085	0.422 ± 0.070	0.326 ± 0.000	0.399 ± 0.060	**0.534** ± 0.010	0.482 ± 0.040	0.311	0.365
Kidney	0.538 ± 0.004	0.524 ± 0.005	0.546 ± 0.001	0.534 ± 0.001	0.519 ± 0.010	0.546 ± 0.007	0.547	**0.564**
Liver	0.749 ± 0.005	0.754 ± 0.005	0.684 ± 0.055	0.748 ± 0.021	0.726 ± 0.017	0.685 ± 0.093	0.728	**0.760**
Lung	0.175 ± 0.013	0.182 ± 0.010	0.163 ± 0.036	0.141 ± 0.002	0.014 ± 0.002	**0.221** ± 0.053	0.214	0.187
Prostate	0.104 ± 0.024	0.056 ± 0.001	**0.293** ± 0.003	0.150 ± 0.091	0.022 ± 0.081	0.266 ± 0.063	0.072	0.182
Skin	**0.451** ± 0.007	0.399 ± 0.005	0.242 ± 0.002	0.397 ± 0.029	0.259 ± 0.243	0.332 ± 0.040	0.423	0.411
Uterus	0.605 ± 0.033	0.624 ± 0.005	0.532 ± 0.185	0.638 ± 0.042	0.630 ± 0.000	0.495 ± 0.235	0.654	**0.678**
Breast	0.330 ± 0.017	0.372 ± 0.003	0.345 ± 0.000	0.000 ± 0.036	0.251 ± 0.028	**0.384** ± 0.017	0.329	0.350
Macro	0.446	0.443	0.426	0.413	0.380	0.452	0.440	**0.468**