

Improving tissue-specific expression prediction of sequence-to-function models

Introduction

- ▶ Many multitask sequence-to-expression models (e.g., Enformer, Borzoi) are trained to predict gene expression from DNA sequence across numerous biological contexts (e.g., cell lines, tissues, disease states)
- ▶ These models often struggle to predict context-specific expression or regulatory features (Kathail et al. 2024)
 - Approaches to improve prediction of context-specific regulatory features (e.g., chromatin accessibility) include increasing model capacity and including an additional cross-context loss
- ▶ Decima (Lal et al. 2024) aimed to improve prediction of context-specific expression by:
 - **More data:** Training on pseudobulked single-cell RNA-seq data from additional contexts
 - **Cross-context loss:** Incorporating an additional cross-context multinomial term into the loss that focuses on predicting a gene's *relative* expression profile across contexts, rather than absolute expression
- ▶ In this work, we aim to better understand the impact of a cross-tissue loss on predicting tissue-specific expression. We
 - Fine-tune Enformer on bulk RNA-seq data from 218 tissues
 - Incorporate a **cross-tissue Pearson loss** term into a hybrid loss function

Hybrid loss function

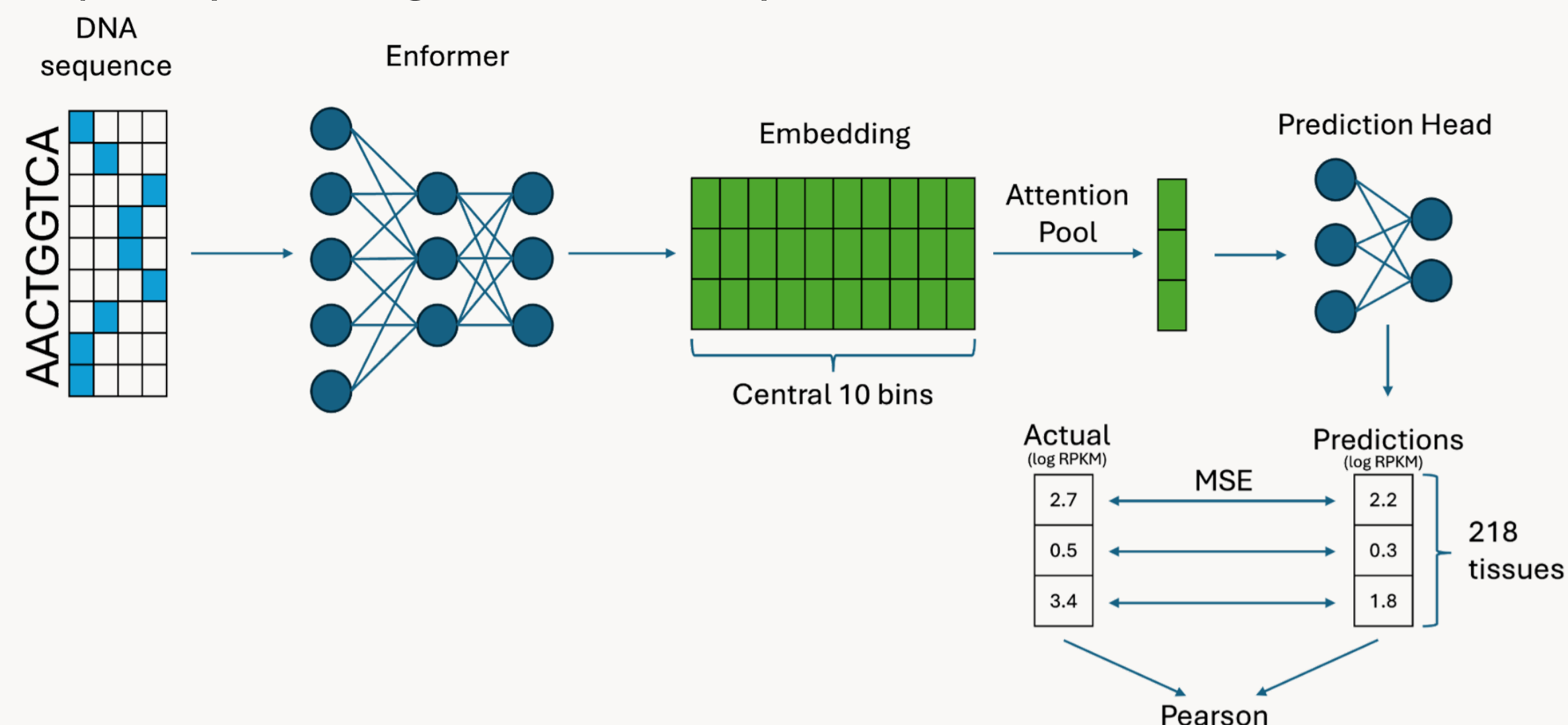
- ▶ Poisson negative log likelihood (NLL) loss used in Enformer may not be best choice for optimal performance on cross-tissue task
 - Poisson NLL loss aims to output values close to corresponding ground truth
 - Expression values for a single gene across tissues are closer together than different genes in the same tissue = harder to learn ordering of a gene across tissues!
- ▶ We want our loss function to more explicitly penalize poor cross-tissue predictions
- ▶ We introduce a hybrid loss function that is a weighted combination of Mean Squared Error loss (MSE) and cross-tissue Pearson loss

$$L(y, \hat{y}) = \alpha \cdot (1 - \text{Pearson}(y, \hat{y})) + \beta \cdot \text{MSE}(y, \hat{y})$$

where $y, \hat{y} \in \mathbb{R}^{218}$ and $\alpha, \beta \in [0, 1]$

Methods

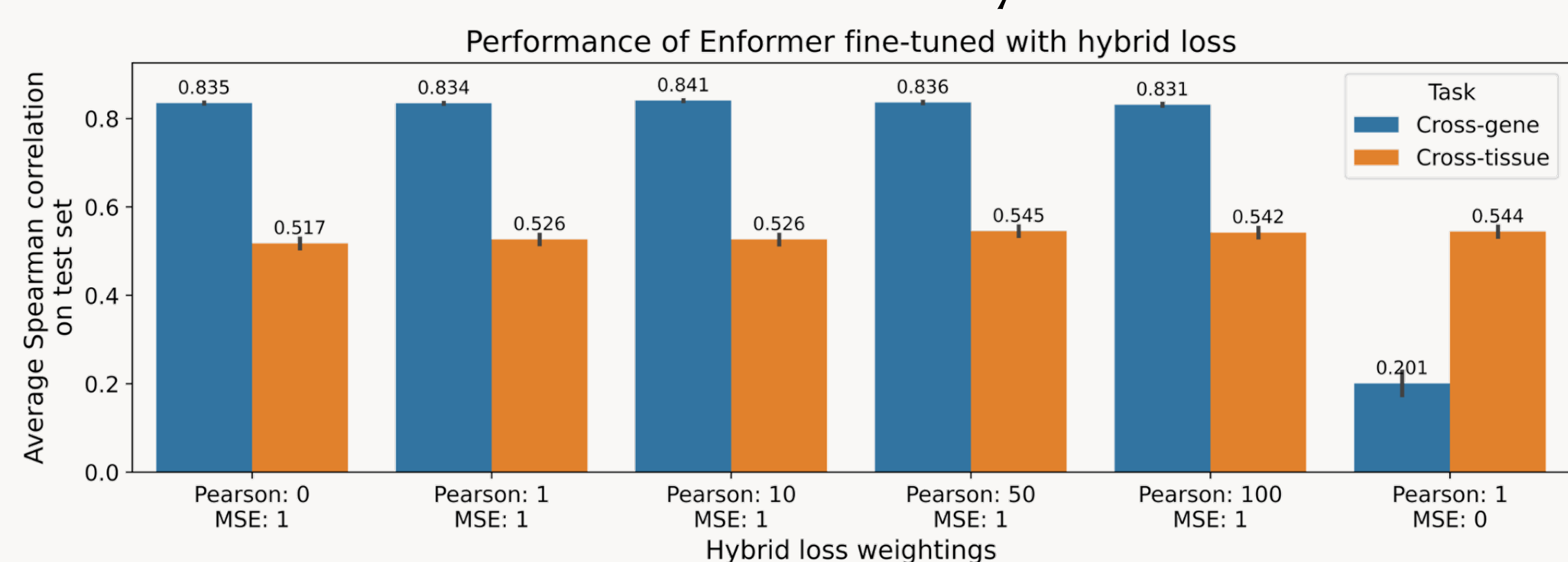
- ▶ Used bulk RNA-seq data (log RPKM) from 218 distinct tissue samples across >20,000 genes (excluding rRNA + genes on sex chromosomes)
 - Test set is chr14 (Enformer is not trained on any genes from chr14)
- ▶ Fine-tuned Enformer with hybrid loss function
 - Pass ~50kb sequence centered at TSS into model and extract embeddings before output heads
 - Attention pooling to aggregate information across central 10 bins
 - Linear layer to predict log RPKM RNA-seq values across 218 tissues



- ▶ Used hyperparameters largely inspired by Rastogi et al. 2024.

Performance with hybrid (MSE + cross-tissue Pearson) loss

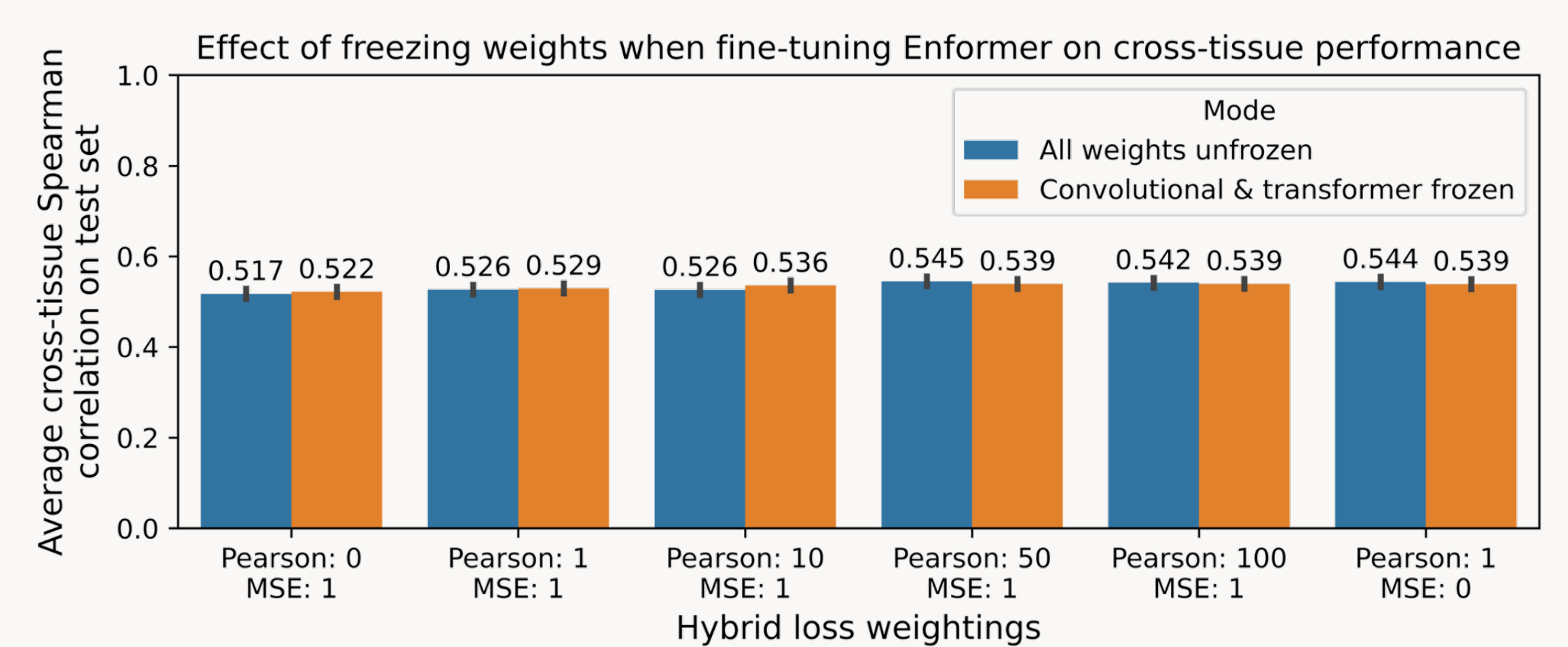
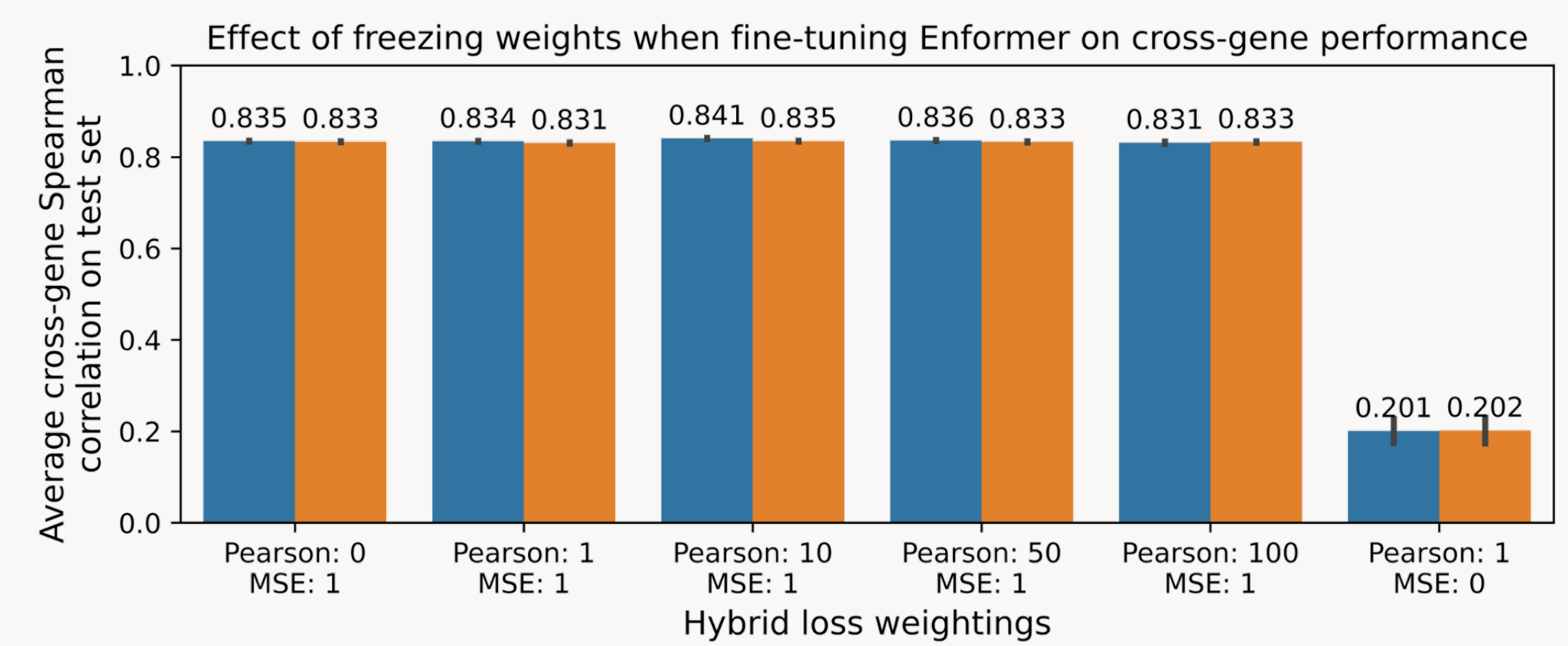
- ▶ Evaluated Enformer models fine-tuned with 6 hybrid loss function weightings



- ▶ As we increase weight of cross-tissue Pearson loss, cross-tissue performance increases (by 5.4% from $\alpha = 0$ to $\alpha = 50$) with minimal change in cross-gene performance
 - Omitting MSE loss causes dramatic drop in cross-gene performance

Freezing weights when fine-tuning Enformer

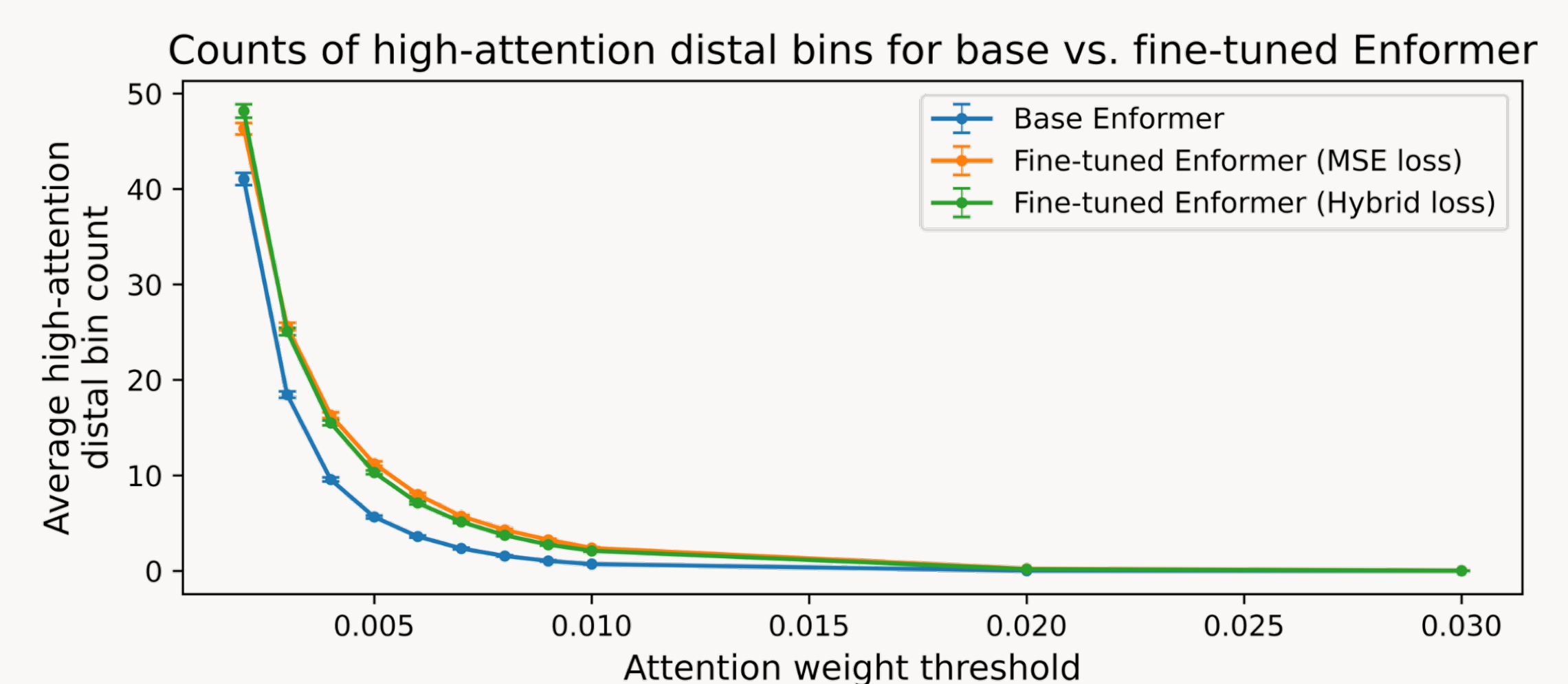
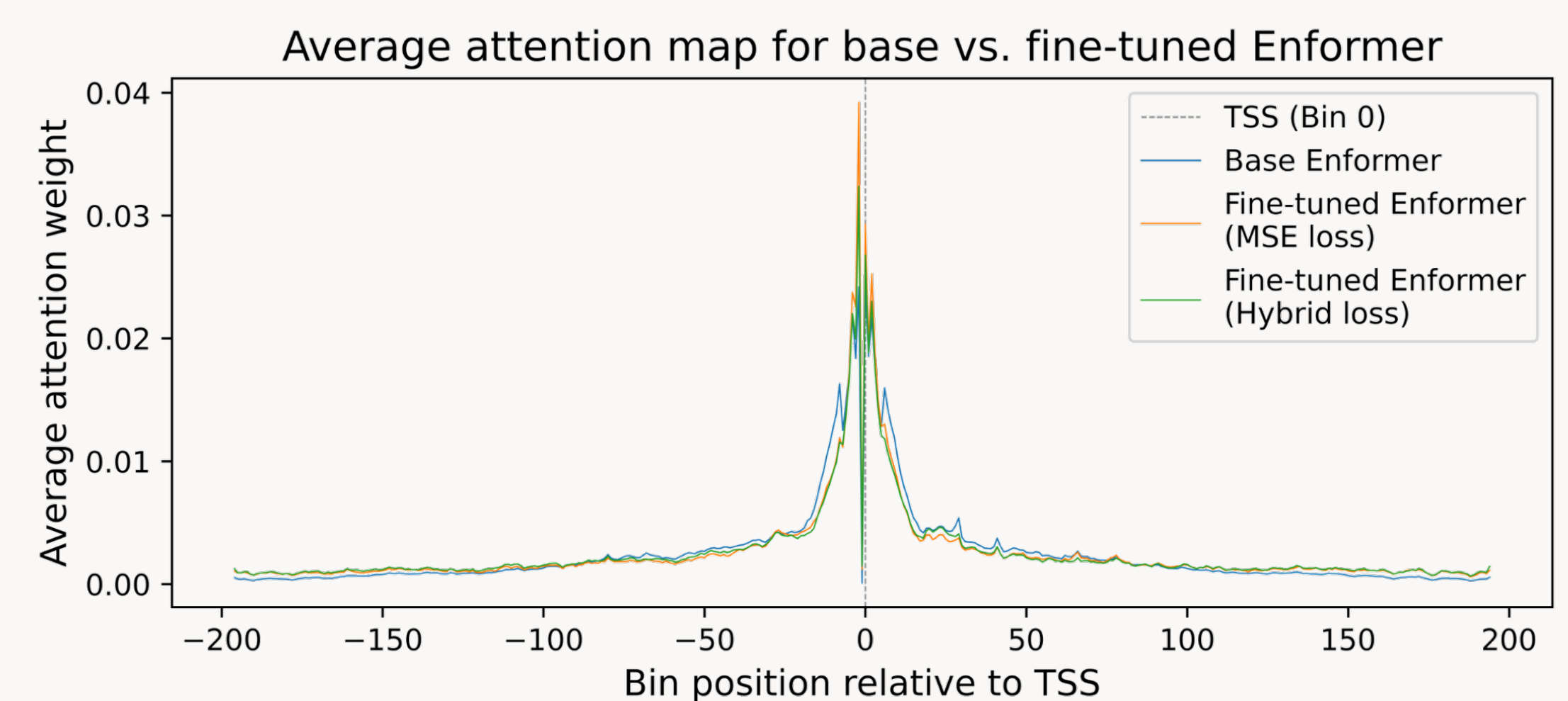
- ▶ Fine-tuned Enformer model with weights of convolutional and transformer modules frozen
- ▶ Freezing reduces computational cost with negligible decay in performance



- ▶ Hypothesis: Cross-tissue loss should help Enformer better capture signal from distal enhancers that modulate tissue-specific expression
 - Frozen attention layers: unable to learn to capture distal signal
 - Unfrozen attention layers: can learn to capture distal signal, performance should rise
- ▶ Found that performance does not rise when unfreezing attention layers

Effect of cross-tissue loss on Enformer attention maps

- ▶ Computed attention weights between TSS bin and all other bins (averaged across layers and heads) for genes in test set



- ▶ Fine-tuning with MSE loss versus hybrid loss produce highly similar attention patterns, including in distal regions
 - Cross-tissue loss term doesn't help model pay more attention to distal bins (>50 bins from TSS)
- ▶ Fine-tuning with either loss increases attention to distal regions
 - Distal enhancers known to modulate tissue-specific expression

Conclusion

- ▶ Introduced a hybrid loss function with a cross-tissue Pearson loss component to penalize poor cross-tissue expression predictions
- ▶ Fine-tuning Enformer with hybrid loss improves cross-tissue performance
 - Does not significantly alter attention maps compared to MSE loss
- ▶ Future work may look at use of cross-context losses throughout model training, on tracks besides gene expression, etc. and design of loss functions to better integrate distal signal

Kanav Mittal, Ruchir Rastogi, Nilah Ioannidis

kanavmittal@berkeley.edu, nilah@berkeley.edu

UC Berkeley