Bridging Morphology and Molecular Signatures: Multi-Task Deep Learning for Multi-Omics Prediction from Histopathology

Anonymous CVPR submission

Paper ID *****

Abstract

Whole slide images (WSIs) capture intricate morphological features that correlate with molecular profiles in tumors, making them valuable for non-invasive molecular profiling. While previous work in computational pathology has focused on predicting gene expression, protein levels, and mutation status from WSIs, we extend this by introducing a deep learning framework that predicts pathway activity and microRNA (miRNA) expression in addition to these commonly studied molecular layers. By employing multi-task learning, our model efficiently captures shared histological patterns across different molecular modalities, enhancing prediction accuracy. We show that pathway activity is the most reliably predicted feature, followed by protein expression, with gene expression and miRNA predictions being more challenging. These findings highlight the necessity of incorporating pathway and miRNA data for a more comprehensive and biologically relevant understanding of tumor biology. Our approach demonstrates significant potential for improving cancer diagnostics and biomarker discovery, offering a more comprehensive alternative to traditional molecular assays.

1. Introduction

Hematoxylin and eosin (H&E)-stained histopathology slides have long been fundamental in cancer diagnosis, offering intricate insights into tissue architecture and cellular organization. Recent advances, however, have revealed that these slides also contain rich molecular information that correlates with alterations at the transcriptomic, proteomic, pathway, and microRNA levels [4–6]. Harnessing this latent information to predict omics profiles directly from WSIs presents a promising, rapid, and cost-effective alternative to traditional molecular profiling methods, which are often labor-intensive and expensive [8].

Deep learning techniques have revolutionized computational pathology, enabling the extraction of biologically meaningful features from WSIs [14, 17, 21]. While molecular profiling typically relies on specialized laboratory workflows, recent advancements in convolutional neural networks (CNNs), multiple-instance learning (MIL), and transformer-based architectures have demonstrated that histopathological patterns in WSIs can predict underlying molecular phenotypes [2, 7, 10, 11, 19]. These techniques have successfully been applied to predict gene expression, protein abundance, and other molecular signatures from histology, setting the stage for AI-driven virtual multi-omics profiling [8].

Related works to predict omics level from WSI

Recent advances in deep learning have enabled the prediction of molecular profiles (mainly gene expression, mutation status and protein expression profiles directly from WSIs offering a non-invasive alternative for molecular profiling in histopathology. Existing approaches typically follow a shared pipeline with variations in architectural components. The standard workflow involves preprocessing WSIs by segmenting and tiling them into square patches. A subset of N patches is randomly sampled, and patch-level embeddings are generated using a pre-trained feature extractor. The WSI-level representation is then obtained by aggregating these patch embeddings into a single latent vector $W \in R^{1\times d}$ which is passed to a downstream predictor often a shallow linear layer to map embeddings to molecular profiles. For brevity, we refer to the combined tiling, feature extraction, and aggregation modules as the encoder part.

In WSI-based omics prediction, model architectures typically consist of two core components: feature extraction and patch aggregation. The feature extractor generates patch-level embeddings, while the aggregator combines these embeddings into a WSI-level representation for downstream omics prediction.

For feature extraction, recent approaches leverage both CNN-based and transformer-based models, pre-trained on large-scale histopathology datasets. Commonly used models include CTransPath[20] , which uses a hierarchical transformer architecture to capture local and global tissue

patterns, and UNI[3], a unified framework trained in multiple histology datasets for robust feature extraction. These foundation models effectively encode WSI patches into expressive feature representations, capturing morphological variations relevant to molecular profiles.

For patch aggregation, various strategies are employed to combine patch embeddings into a single WSI-level vector. MLP-based aggregators, as seen in HE2RNA[17] , apply fully connected layers to directly merge patch features. Transformer-based aggregators, such as tRNAsformer[1], use self-attention to capture inter-patch relationships, enabling the model to account for spatial dependencies. More recent methods, like SEQUOIA[16] adopt SummaryMixing[15], a linearized transformer variant, to enhance efficiency while preserving representational power.

While prior research has primarily concentrated on predicting gene expression or mutation status from WSIs, with some extending to protein expression, the inference of pathway activity and microRNA (miRNA) expression has largely been unexplored. Pathway activity provides a more interpretable representation of molecular states compared to individual gene expression, capturing the collective behavior of genes involved in biological processes. This makes pathway-based inference not only more robust but also more biologically meaningful for downstream analysis. Similarly, miRNAs, which regulate gene expression post-transcriptionally, are vital biomarkers for cancer detection and prognosis, yet their prediction from WSIs remains largely unexplored. Moreover, most existing models focus on predicting individual molecular modalities, such as gene expression or proteomics, independently. However, the molecular processes in cells are intricately interconnected, with gene expression, protein regulation, and pathway activity influencing one another. By ignoring these interdependencies, current models may fail to capture biologically meaningful cross-omics relationships, thereby limiting their overall predictive accuracy.

In this work, we expand the scope of WSI-based molecular inference to include multiple omics layers, including gene expression, protein expression, pathway activity, and miRNA levels. Through both single-task and multi-task learning frameworks, we demonstrate that pathway and miRNA predictions not only enhance interpretability but also provide valuable prognostic insights. Our findings suggest that inferring higher-level molecular features, such as pathway activity, offers greater robustness and clinical relevance, thus advancing the integration of histology and molecular data in cancer research.
By incorporating *pathway* and *microRNA* expression levels into WSI-based omics prediction, the framework moves beyond gene-level inference, providing a more stable, interpretable, and biologically relevant representation of tumor biology. This expanded approach enhances the utility of WSI-based profiling for both research and clinical applications, enabling more reliable biomarker discovery and precision medicine strategies.

Key contributions: i) The first study to predict both pathway activity and microRNA expression levels directly from WSIs. ii) A multi-task learning framework that leverages shared and complementary information across omics modalities, enhancing predictive performance. iii) The first comprehensive comparative analysis of omics prediction accuracy from WSIs, covering pathway activity, protein expression, gene expression, and microRNA expression.

2. Method

Here we present our framework for predicting multi-omics profile from WSI slides Figure 1. Our framework consists of two main steps. i) A vision encoder that embeds the patches using a vision transformer, followed by an aggregator module for learning a slide-level embedding. ii) The omics predictor module processes slide embeddings to predict.

2.1. Slide encoder

Feature extractor: we use a similar architecture to the TANGLE framework, selecting UNI [3] as the feature extractor due to its superior consistency in prior benchmarks [16]. The workflow tessellates the slide into small patches, extracts patch embeddings using the pre-trained UNI encoder, and aggregates them into a slide-level representation [13].

Aggregator To aggregate the patch embeddings obtained from the pre-trained vision encoder, we train the widely used attention-based multiple instance learning (ABMIL) model, which learns patch-level attention weights to effectively pool the embeddings into a slide-level representation [12]

2.2. Predictors

Single task learning: our framework predicts multi-omics profiles, including gene expression, microRNA expression, pathway activity, and protein expression levels, from WSI embeddings. The WSI embeddings are obtained from the first stage of our framework (Figure 1), which employs TANGLE, a transformer-based model designed to learn hierarchical histopathology representations. Given an input WSI, TANGLE generates a d-dimensional embedding $z \in R^d$, capturing spatial and morphological features relevant to downstream molecular characterization. We model the prediction of each omics modality as an independent high-dimensional regression task, where a function f learns to map the extracted WSI embedding z to a target omics profile y corresponding to one specific molecular layer. Formally, for gene expression prediction, the network approximates:

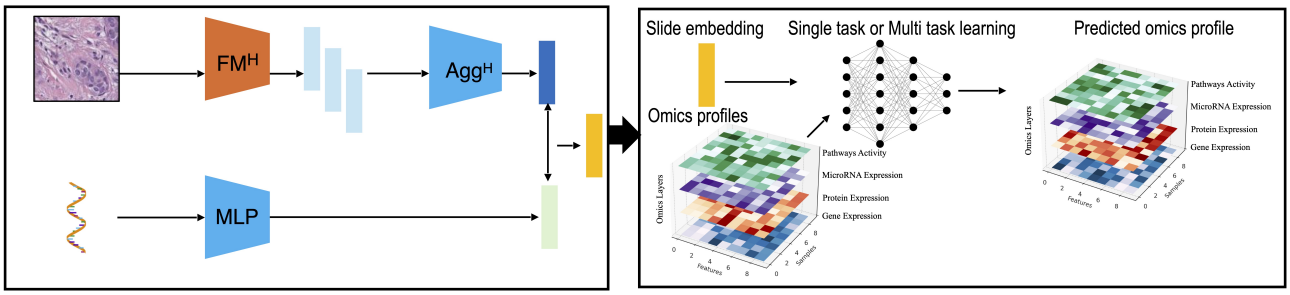

Figure 1. A histology slide is divided into patches and encoded using a pre-trained UNI encoder (FM^H) . The patch embeddings are aggregated by an ABMIL module (Agg^H) into a slide-level representation. In parallel, gene expression data is encoded with an MLP. A SymCLR objective aligns the embeddings. The learned slide embeddings are then used to train deep regression models in single-task or multi-task settings to predict multi-omics profiles, including gene, protein, pathway, and microRNA expression.

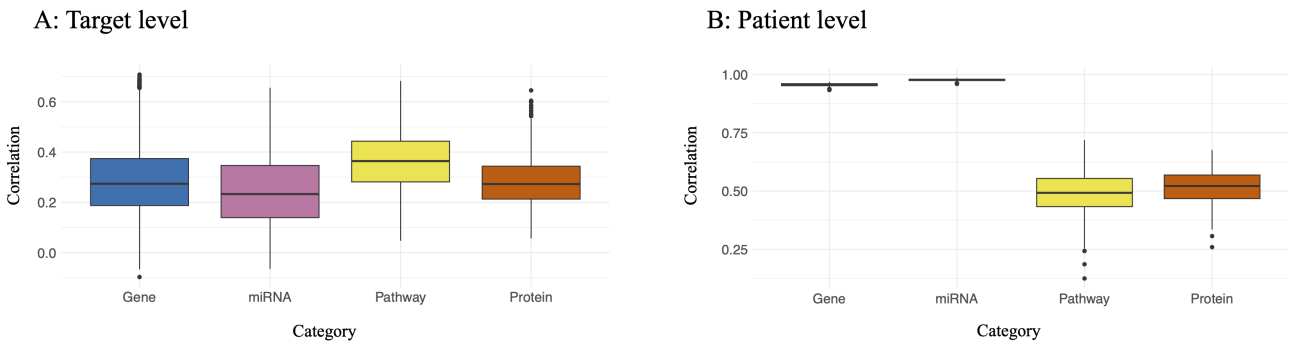

Figure 2. Model performance comparison across omics layers. (A) Correlation between predicted and actual omics profiles across patients. (B) Correlation of model performance at the individual patient level.

$$f_g : R^d \mapsto R^{m_g}$$

where $m_g$denotes the number of genes. Similarly, for pathway activity and protein expression prediction, we define:

$$f_p : R^d \mapsto R^{m_p}, f_k : R^d \mapsto R^{m_k} f_l : R^d \mapsto R^{m_l}$$

where m_p and m_k represent the number of pathways and proteins, respectively. Each function f is parameterized as a fully connected neural network with a single hidden layer of 512 neurons, activated by ReLU, followed by an output layer tailored to the corresponding omics dimension. The model is trained independently for each omics type using the *Adam* optimizer and a mean squared error (MSE) loss function, ensuring accurate reconstruction of omics profiles from histopathological features. Model selection is performed using 5-fold cross-validation, where the model achieving the lowest validation MSE is selected for final evaluation.

Multi-task learning Building on our single-task framework, we extend our approach to multi-task learning (MTL) to jointly predict gene expression, pathway activity, and protein expression from WSI embeddings. Rather than training independent models for each molecular layer, we introduce a multi-task architecture that learns shared representations while allowing task-specific adaptations. This enables the model to leverage common histopathological features while maintaining flexibility to capture distinct patterns relevant to each omics type. Given a WSI embedding $z \in R^d$ extracted using TANGLE (Fig.1), we define a joint function:

$$f_\theta : R^d \mapsto (R^{m_g}, R^{m_p}, R^{m_k}, R^{m_m})$$

where m_g, m_p, m_k, m_m correspond to the number of genes, pathways, proteins, and microRNAs respectively.

Each omics prediction task is modeled as a high-dimensional regression problem, where a shared feature extraction backbone is followed by task specific output layers. The shared backbone consists of a fully connected (FC) layer with 512 neurons and ReLU activation, capturing a common histological representation of tumor morphology. The task-specific heads, one for each omics modality, consist of independent linear layers, each mapping the features

	Top 50		Top 500	
	Single Task	Multi Task	Single Task	Multi Task
Gene	0.6778 ± 0.0226	0.6741 ± 0.0281	0.5573 ± 0.0244	0.5935 ± 0.0194
Protein	0.4694 ± 0.0218	0.4924 ± 0.0088	0.2897 ± 0.0161	0.2958 ± 0.0127
microRNA	0.5295 ± 0.0278	0.5338 ± 0.0170	0.2624 ± 0.0243	0.3259 ± 0.0099
Pathways	0.5970 ± 0.0071	0.6449 ± 0.0263	0.4902 ± 0.0207	0.5031 ± 0.0187

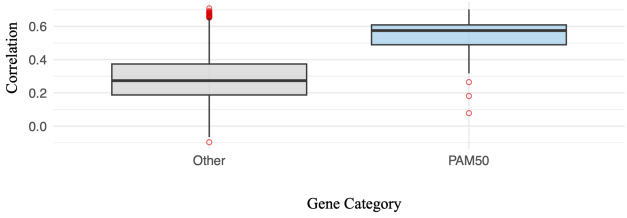

Figure 3. Comparison of single task vs. multi task learning and the prediction performance of breast cancer-relevant genes (PAM50). The left table presents quantitative results (mean Pearson correlation ± standard deviation) for the top 50 and top 500 molecular targets across different omics modalities, while the right panel visualizes model performance on PAM50 gene signatures.

to the corresponding output dimension.

$$f_g(z) = W_g \cdot z + b_g, \quad f_p(z) = W_p \cdot z + b_p,$$
$$f_k(z) = W_k \cdot z + b_k, \quad f_m(z) = W_m \cdot z + b_m$$

where W and b are task-specific weight matrices and biases.

$$\mathcal{L} = \frac{1}{N} \sum_{i=1}^{N} \sum_{t \in \{g,p,k,m\}} M_t^i \cdot ||y_t^i - f_t(z^i)||^2$$

To address missing omics measurements across patients, we employ task-specific masking, where a binary mask is applied during loss computation to exclude missing targets. The model is trained using a partial supervision strategy, ensuring that only available omics labels contribute to the gradient update. Optimization is performed using the Adam optimizer, with the loss function defined as the masked mean squared error (MSSE) across all available omics tasks. Model selection is conducted via 5-fold cross-validation, selecting the model with the lowest validation MSE for final evaluation.

3. Experiments and results

3.1. Dataset

This study utilizes publicly available multimodal data from the Xena browser for breast cancer (BRCA)[9]. We collected a dataset comprising matched gene expression profiles, WSIs, protein expression levels, microRNA expression, and pathway activity measurements for a cohort of 1,024 BRCA patients.

Histology slides: Matched histopathology slides for BRCA were obtained from the Xena browser and processed using the TANGLE framework. This approach generated patch embeddings, enabling the model to learn meaningful representations of tissue morphology for downstream predictive tasks.

Gene expression data: We obtained matched gene expression data with FPKM values for protein-coding genes from the Xena browser. To improve the interpretability of the model, we excluded genes with a median expression of zero, retaining $17,759$ protein-coding genes. To mitigate the dominance of highly expressed genes in the regression analysis, we applied a $log10(1 + a)$ transformation to expression values.

microRNA expression data: We obtained matched microRNA expression data with FPKM values for 1800 microRNA from the Xena browser. To enhance the interpretability of the model, we excluded microRNA with a median expression of zero, retaining 751 microRNA genes. To mitigate the dominance of highly expressed genes in the regression analysis, we applied a $log10(1+a)$ transformation to expression values.

Protein expression data: We obtained proteomics data for breast cancer across 980 samples from the Xena Browser, measured using reverse-phase protein array (RPA), which includes expression levels for 480 proteins. Missing values were imputed using the median expression of the respective protein across all samples, ensuring data completeness. To stabilize variance and improve downstream analysis, the entire dataset was subsequently logarithmically transformed.

Pathway activity data: We obtained pathway activity data for breast cancer from the Xena Browser, using the PARADIGM algorithm for inference[18]. PARADIGM integrates pathway information with gene expression and copy number variation data to estimate the activation status of pathway components within a unified pathway network. This network structure consolidates $1,387$ constituent pathways derived from three major pathway databases: NCI-PID, BioCarta, and Reactome, providing a comprehensive framework for pathway activity analysis. Furthermore, we evaluated TANGLE in both multimodal mode (contrastive alignment) and unimodal mode. However, we did not observe a significant improvement in omics layer prediction between the two settings, suggesting that multimodal alignment did not substantially enhance predictive performance in this context.

3.2. Single task learning of multi-omics

We performed omics prediction at both the target and patient levels and evaluated the model's performance across different molecular layers. At the target level, we assessed the accuracy of predicting individual gene, protein, pathway, and microRNA expression levels, while at the patient level, we predicted the entire omics profile for each patient. The predictive performance of our model varied between different omics layers inferred from WSI of breast cancer tissues Figure 2. Pathway activity exhibited the highest correlation between predicted and hold-out data, followed by protein expression, while miRNA and gene expression showed comparatively lower correlations (Figure 2).

Pathway activity prediction: The superior performance in pathway prediction is likely due to the aggregation of gene expression into functional networks. By capturing collective biological activity rather than individual gene fluctuations, this approach reduces transcriptional noise and stochastic variability, leading to more stable and interpretable predictions. The top predicted pathways that were strongly linked to cancer progression and relevant biological processes include Targets of *MYC* transcriptional activation, *Aurora B signaling*, and *Cyclin B2* mediated events. These pathways play pivotal roles in cell cycle regulation, tumor invasion, and cellular growth. *C-MYC*, a key oncogene, activates several transcriptional programs that drive tumor genesis and progression, including genes involved in cellular metabolism, proliferation, and survival. The model's ability to predict pathway activity associated with *C-MYC* transcriptional activation implies that the morphological features of WSI reflect key signaling events that drive the pathology of breast cancer. *Aurora B signaling* is involved in the regulation of mitosis and chromosome segregation, crucial for proper cell division and genomic stability. These pathway predictions suggest that WSI captures cellular processes that underlie tumor proliferation and metastasis, providing clinically relevant insights into cancer progression at the tissue level.

Protein expression prediction: showed stronger correlation compared to microRNA and gene expression. Protein expression prediction from WSI is inherently more accurate than gene expression inference due to the stronger morphological imprint of proteins on tissue architecture. Unlike mRNA, which reflects transient transcriptional states, protein abundance integrates post-transcriptional and translational regulation, resulting in more stable and visually discernible patterns. Proteins directly influence cellular morphology, adhesion, and staining intensity, creating distinct histological signatures that deep learning models can effectively capture. In contrast, gene expression exhibits higher variability due to transcriptional noise, cell-type heterogeneity, and dynamic regulation, making it less reliably linked to tissue morphology. Additionally, the lower dimensionality of proteomics data compared to transcriptomics reduces prediction complexity, further improving model accuracy. These factors collectively make protein expression more amenable to robust inference from WSI in computational pathology.

miRNA Prediction: despite its moderate accuracy, the model's ability to predict miRNA expression from WSI reveals that breast cancer tissue morphology encodes subtle yet detectable imprints of miRNA activity. miRNAs are critical regulators of post-transcriptional gene expression, influencing key cancer processes such as cell proliferation, apoptosis, and metastasis. Their successful inference from histological data suggests that the morphological features of tumor tissues capture molecular dysregulation associated with miRNA activity, providing a potential new biomarker for histology-based cancer profiling. However, the relatively lower accuracy of miRNA prediction compared to protein and pathway activity can be attributed to the complexity of miRNA regulation, which is highly tissue-specific and often present at lower abundance. These factors make it more challenging to directly capture miRNA expression from morphological patterns alone. Nonetheless, the ability to predict miRNA expression from WSI opens up exciting possibilities for incorporating small RNA data into histopathological analyses, enhancing the predictive capabilities of cancer diagnostics and enabling more comprehensive molecular profiling from tissue slides.

Gene expression prediction: Gene expression prediction from WSI showed the lowest correlation among the evaluated omics layers. This lower performance can be attributed to several challenges, including technical noise, cell-type heterogeneity, and regulatory mechanisms that extend beyond the morphological features captured in WSI. Since gene expression is influenced by complex, tissue-specific regulatory networks, predicting expression for all genes from histological data alone is inherently more difficult. The morphological features of tissue sections provide only indirect insights into the transcriptomic landscape, leading to reduced accuracy compared to protein and pathway predictions, where functional relationships are more directly reflected in tissue morphology. This highlights the limitations of using WSI for gene expression-based biomarker inference, which may be suboptimal and a limiting factor. This underscores the need to model other omics layers, such as pathways, proteins, or small RNA, to infer more accurate and comprehensive predictions from WSI.

Patient level prediction: beyond target-level predictions, our model demonstrates superior performance in accurately predicting omics profiles at the patient level. This can likely be attributed to the inherent correlations within each omics modality—genes within the transcriptomics layer are correlated with each other, and similarly, pro-

teins within the proteomics layer exhibit strong correlations. When the model performs at the patient level, it can leverage these correlation structures, which helps improve predictive accuracy. In contrast, at the target level, the model struggles to utilize these correlations since the focus is on individual targets like specific genes or proteins, rather than on the collective relationships between them. This lack of structural correlation at the target level limits the model's ability to make accurate predictions across modalities.

3.3. Multi task learning of multi-omics

Using the architecture described in the Methods section, we conducted multi-task learning (MTL) across multiple settings. Initially, we trained a unified model to predict 17759 genes, 778 microRNAs, 1857 pathways, and 457 proteins, following the STL baseline. However, we observed no significant performance improvements with MTL. A key factor could be the imbalance in task dimensionality, where large output spaces dominate optimization, leading to suboptimal representation learning for smaller tasks. To mitigate this, we constrained MTL to the top 500 genes, pathways, microRNAs, and proteins, aligning task dimensions and reducing the risk of unequal loss scaling. The 480-protein constraint naturally set 500 as a threshold, harmonizing the learning dynamics across tasks.

To further refine task balance, we trained MTL on the top 50 genes, pathways, and proteins. This was motivated by the hypothesis that reducing dimensionality while retaining biologically relevant targets would enhance feature sharing and optimization stability. From an AI/ML perspective, high-dimensional MTL can suffer from negative transfer, where conflicting gradients degrade task-specific learning. By restricting MTL to a compact, high-signal subset, we aimed to improve cross-task knowledge transfer, stabilize shared representations, and prevent task interference.

Table 3 demonstrates the effectiveness of multi-task learning (MTL) across different omics modalities. While MTL and single-task learning (STL) showed similar performance for the top 50 features, MTL consistently outperformed STL in the top 500 setting, particularly for genes, microRNAs, and pathways. Notably, MTL improved gene prediction by $6.5\%(0.5573 \rightarrow 0.5935)$ and microRNA prediction by $24\%(0.2624 \rightarrow 0.3259)$, highlighting its ability to leverage shared representations across omics layers. Pathway prediction also benefited from MTL, showing an increase from 0.4902 to 0.5031, indicating that incorporating cross-task dependencies enhances predictive power. While protein prediction showed only a marginal improvement $(0.2897\text{ß}0.2958)$, this suggests that further architectural refinements may be needed to better capture protein interactions in the MTL framework. Overall, these results demonstrate that MTL effectively leverages shared histological representations to improve multi-omics prediction when task dimensionalities are aligned. The significant performance gains in gene and microRNA prediction suggest that incorporating structured multi-task objectives facilitates cross-task knowledge transfer, particularly for omics layers with overlapping regulatory mechanisms. These findings reinforce the potential of MTL for enhancing histology-based biomarker discovery by integrating complementary molecular signals.

3.4. Survival analysis

To evaluate the clinical relevance of the inferred omics profiles, we examined whether the top predicted targets were associated with patient survival outcomes. Specifically, we performed survival analysis on AGR, FOXM1, and HSA-miR-190b expression levels in breast cancer (BRCA) patients. The expression of these biomarkers was accurately predicted from WSI of BRCA tumors using our deep learning model (Fig. 4). Kaplan-Meier survival curves revealed significant associations between the predicted expression levels and patient survival.

AGR (Activator of G-protein signaling) was strongly linked to poorer overall survival (OS), suggesting its potential role in driving BRCA progression. Similarly, FOXM1 (Forkhead Box M1), a key regulator of cell cycle and proliferation, was associated with reduced survival, highlighting its value as a prognostic marker. In contrast, HSA-miR-190b, a tumor-suppressive microRNA, showed a protective effect, with higher expression correlating with improved survival outcomes. These results demonstrate that WSI-derived omics predictions can capture clinically relevant biomarkers, offering potential for non-invasive risk stratification and personalized treatment planning.

4. Discussion

In this study, we demonstrated the ability to predict multiple omics layers directly from WSI of breast cancer tissues using a deep learning framework. The predictive performance of our model varied across the different molecular modalities, with pathway activity emerging as the most predictable feature, followed by protein expression. miRNA and gene expression predictions showed lower correlations, highlighting the challenges and opportunities in histology-driven molecular inference.

The superior performance of pathway-level predictions suggests that aggregating gene expression into functional networks enhances predictive robustness. Biological pathways, by their nature, integrate signals from multiple genes that contribute collectively to a functional process. This aggregation helps to reduce transcriptional noise and stochastic variability, which are often associated with individual gene expression measurements. As a result, pathway-level inference offers more stable and interpretable predictions,

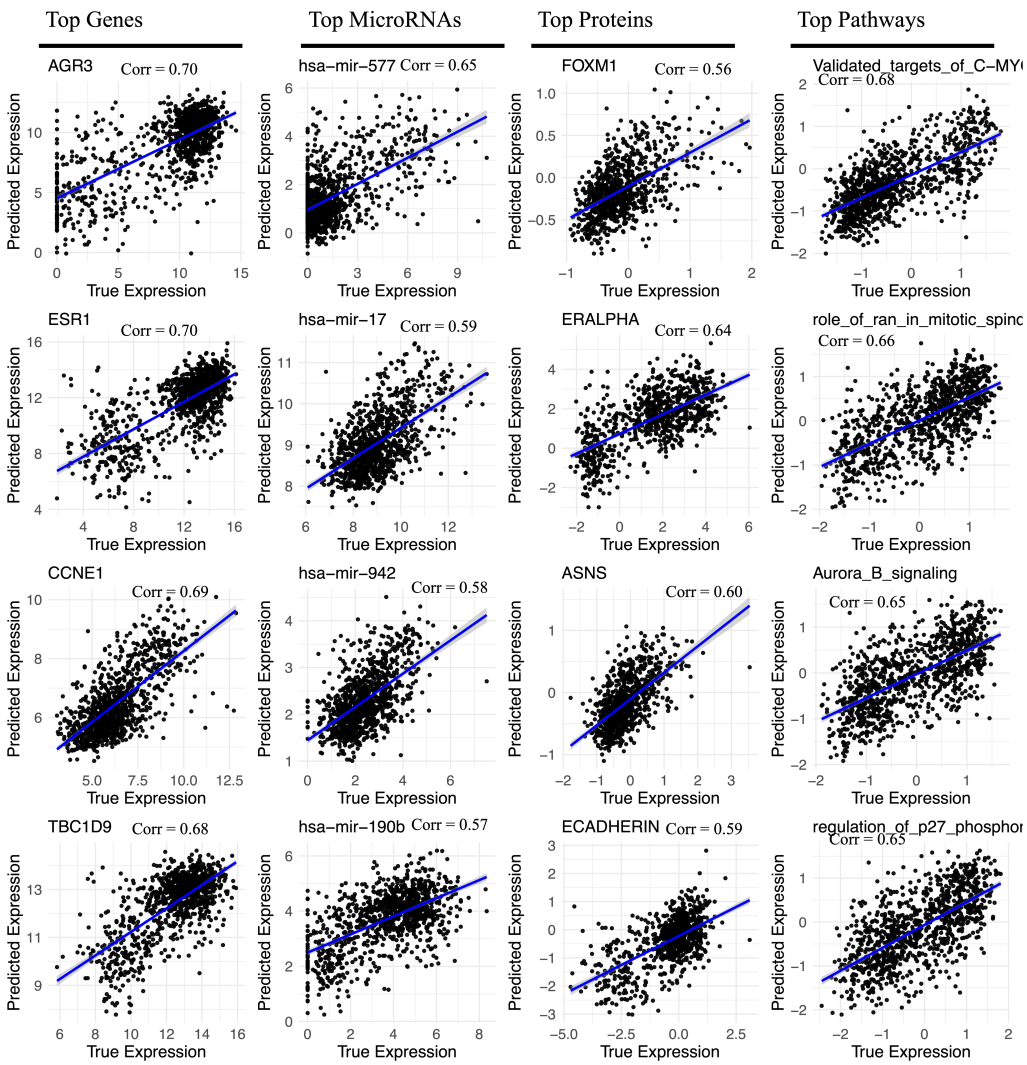

Figure 4. Correlation plot showing predicted vs. ground truth expression or activity values for the top 4 genes, proteins, microRNAs, and pathways.

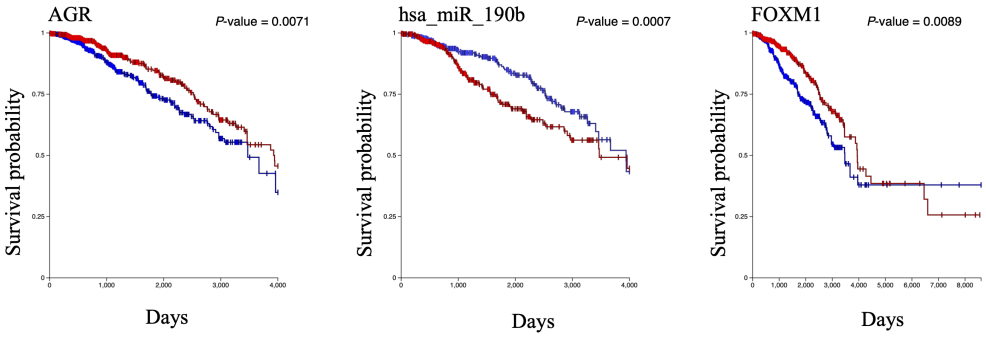

Figure 5. Survival analysis of AGR, FOXM1, and HSA-miR-190b expression levels in breast cancer (BRCA) patients, highlighting their prognostic significance.

effectively mitigating the challenges of capturing transcriptional variability.

Protein expression prediction from WSI also demonstrated a higher correlation compared to miRNA and gene expression. This finding indicates that histological features, which reflect tissue morphology, are better at capturing aspects of post-transcriptional regulation and protein abundance than transcriptomic data alone. Since protein expression levels are more directly influenced by cellular architecture and histological features, it is not surprising that these predictions were more accurate than gene expression predictions. The intermediate performance of miRNA prediction underscores the potential for leveraging histological patterns to infer regulatory elements beyond mRNA and protein levels. While the prediction of miRNA expression from WSI is still a relatively novel approach, the model's moderate success in this area suggests that miRNA-associated regulatory activity leaves a morphological imprint in cancer tissues, offering a promising avenue for future exploration.

In contrast, gene expression prediction exhibited the lowest performance in our study. This may be attributed to several factors, including the greater variability of transcript levels due to technical noise, cell-type heterogeneity, and complex regulatory mechanisms that are not fully captured by histological features alone. The relatively lower correlation for miRNA prediction compared to protein and pathway predictions further suggests that miRNA expression, despite its biological relevance, is more difficult to infer from histology. miRNA regulation is highly tissue-specific and influenced by a complex network of post-transcriptional processes, making it more challenging to detect directly from morphological patterns. Nevertheless, the ability to predict miRNA expression from WSI represents a significant advancement in integrating small RNA profiles into histopathological analysis, contributing to a more comprehensive molecular understanding of cancer biology.

Beyond predictive accuracy, our study explored the biological relevance of the model's predictions by focusing on the PAM50 gene signature, a widely used panel for molecular subtyping in breast cancer. Our findings reveal that the model achieved significantly higher prediction accuracy for the PAM50 genes compared to the rest of the transcriptome3. This result underscores the model's ability to capture clinically relevant signals from histology slides, as the PAM50 genes are tightly associated with specific breast cancer subtypes that exhibit distinct histopathological patterns. The high prediction accuracy for these genes highlights the potential of using WSI-based models for molecular subtyping, which could enhance clinical decision-making in personalized treatment strategies.

We conducted survival analysis on the top predicted genes, proteins, and miRNAs (AGR, FOXM1, and hsa-miR-190b) to evaluate their prognostic significance. Kaplan-Meier survival curves revealed strong associations between their expression levels and patient outcomes. AGR, a regulator of G-protein signaling, correlated with poorer overall survival, suggesting its role in cancer progression. Similarly, FOXM1, a key driver of cell cycle regulation and proliferation, was linked to reduced survival, reinforcing its potential as a prognostic marker. Conversely, lower expression of hsa-miR-190b, a tumor-suppressive microRNA, was associated with worse survival outcomes, underscoring its protective role. These findings highlight the clinical relevance of histology-derived molecular predictions and their potential for risk stratification and personalized treatment strategies in breast cancer. Our study highlights the advantage of pathway-based inference and multi-omics integration in improving predictive accuracy and clinical relevance. Pathway-level predictions offer a more stable and interpretable representation of molecular states, while protein and miRNA predictions provide valuable insights into post-transcriptional regulation. The ability to predict these features directly from WSI opens new possibilities for integrating histological data with molecular profiling, offering a more accessible and cost-effective alternative to traditional molecular assays.

5. Conclusion

This work extends molecular inference from WSIs beyond gene and protein expression to include pathway activity and miRNA expression, providing a more holistic view of tumor biology. By leveraging MTL, we demonstrate that multi-omics integration enhances prediction accuracy, particularly for pathway activity and microRNAs, which are critical in cancer progression. These findings highlight the potential for deep learning frameworks to transform histology-based molecular profiling into a non-invasive, scalable approach for biomarker discovery and precision oncology.

6. Code Availability

The implementation of our framework, including data preprocessing and model training scripts, is available upon request.

References

[1] Areej Alsaafin, Amir Safarpoor, Milad Sikaroudi, Jason D Hipp, and HR Tizhoosh. Learning to predict rna sequence expressions from whole slide images with applications for search and classification. *Communications Biology*, 6(1): 304, 2023. 2

[2] Salim Arslan, Julian Schmidt, Cher Bass, Debapriya Mehrotra, Andre Geraldes, Shikha Singhal, Julius Hense, Xiusi Li, Pandu Raharja-Liu, Oscar Maiques, et al. A systematic pan-cancer study on deep learning-based prediction of multi-

omic biomarkers from routine pathology images. *Communications Medicine*, 4(1):48, 2024. 1

[3] Richard J Chen, Tong Ding, Ming Y Lu, Drew FK Williamson, Guillaume Jaume, Andrew H Song, Bowen Chen, Andrew Zhang, Daniel Shao, Muhammad Shaban, et al. Towards a general-purpose foundation model for computational pathology. *Nature Medicine*, 30(3):850–862, 2024. 2

[4] Xingjian Chen, Jiecong Lin, Yuchen Wang, Weitong Zhang, Weidun Xie, Zetian Zheng, and Ka-Chun Wong. He2gene: image-to-rna translation via multi-task learning for spatial transcriptomics data. *Bioinformatics*, 40(6):btae343, 2024. 1

[5] Nicolas Coudray, Paolo Santiago Ocampo, Theodore Sakellaropoulos, Navneet Narula, Matija Snuderl, David Fenyö, Andre L Moreira, Narges Razavian, and Aristotelis Tsirigos. Classification and mutation prediction from non–small cell lung cancer histopathology images using deep learning. *Nature medicine*, 24(10):1559–1567, 2018.

[6] Muhammad Dawood, Mark Eastwood, Mostafa Jahanifar, Lawrence Young, Asa Ben-Hur, Kim Branson, Louise Jones, Nasir Rajpoot, et al. Cross-linking breast tumor transcriptomic states and tissue histology. *Cell Reports Medicine*, 4 (12), 2023. 1

[7] Andreas Ekholm, Yinxi Wang, Johan Vallon-Christersson, Constance Boissin, and Mattias Rantalainen. Prediction of gene expression-based breast cancer proliferation scores from histopathology whole slide images using deep learning. *BMC cancer*, 24(1):1–13, 2024. 1

[8] Omar SM El Nahhas, Chiara ML Loeffler, Zunamys I Carrero, Marko van Treeck, Fiona R Kolbinger, Katherine J Hewitt, Hannah S Muti, Mara Graziani, Qinghe Zeng, Julien Calderaro, et al. Regression-based deep-learning predicts molecular biomarkers from pathology slides. *nature communications*, 15(1):1253, 2024. 1

[9] Mary J Goldman, Brian Craft, Mim Hastie, Kristupas Repečka, Fran McDade, Akhil Kamath, Ayan Banerjee, Yunhai Luo, Dave Rogers, Angela N Brooks, et al. Visualizing and interpreting cancer genomics data via the xena platform. *Nature biotechnology*, 38(6):675–678, 2020. 4

[10] Max Hallemeesch, Marija Pizurica, Paloma Rabaey, Olivier Gevaert, Thomas Demeester, and Kathleen Marchal. Prior knowledge injection into deep learning models predicting gene expression from whole slide images. *arXiv preprint arXiv:2501.14056*, 2025. 1

[11] Danh-Tai Hoang, Gal Dinstag, Eldad D Shulman, Leandro C Hermida, Doreen S Ben-Zvi, Efrat Elis, Katherine Caley, Stephen-John Sammut, Sanju Sinha, Neelam Sinha, et al. A deep-learning framework to predict cancer treatment response from histopathology images through imputed transcriptomics. *Nature cancer*, 5(9):1305–1317, 2024. 1

[12] Maximilian Ilse, Jakub Tomczak, and Max Welling. Attention-based deep multiple instance learning. pages 2127–2136, 2018. 2

[13] Guillaume Jaume, Lukas Oldenburg, Anurag Vaidya, Richard J Chen, Drew FK Williamson, Thomas Peeters, Andrew H Song, and Faisal Mahmood. Transcriptomics-guided slide representation learning in computational pathology. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 9632–9644, 2024. 2

[14] Hui Liu, Xiaodong Xie, and Bin Wang. Deep learning infers clinically relevant protein levels and drug response in breast cancer from unannotated pathology images. *NPJ Breast Cancer*, 10(1):18, 2024. 1

[15] Titouan Parcollet, Rogier van Dalen, Shucong Zhang, and Sourav Bhattacharya. Summarymixing: A linear-complexity alternative to self-attention for speech recognition and understanding. *arXiv preprint arXiv:2307.07421*, 2023. 2

[16] Marija Pizurica, Yuanning Zheng, Francisco Carrillo-Perez, Humaira Noor, Wei Yao, Christian Wohlfart, Antoaneta Vladimirova, Kathleen Marchal, and Olivier Gevaert. Digital profiling of gene expression from histology images with linearized attention. *Nature Communications*, 15(1):9886, 2024. 2

[17] Benoît Schmauch, Alberto Romagnoni, Elodie Pronier, Charlie Saillard, Pascale Maillé, Julien Calderaro, Aurélie Kamoun, Meriem Sefta, Sylvain Toldo, Mikhail Zaslavskiy, et al. A deep learning model to predict rna-seq expression of tumours from whole slide images. *Nature communications*, 11(1):3877, 2020. 1, 2

[18] Charles J Vaske, Stephen C Benz, J Zachary Sanborn, Dent Earl, Christopher Szeto, Jingchun Zhu, David Haussler, and Joshua M Stuart. Inference of patient-specific pathway activities from multi-dimensional cancer genomics data using paradigm. *Bioinformatics*, 26(12):i237–i245, 2010. 4

[19] Joshua M Wang, Runyu Hong, Elizabeth G Demicco, Jimin Tan, Rossana Lazcano, Andre L Moreira, Yize Li, Anna Calinawan, Narges Razavian, Tobias Schraink, et al. Deep learning integrates histopathology and proteogenomics at a pan-cancer level. *Cell Reports Medicine*, 4(9), 2023. 1

[20] Xiyue Wang, Sen Yang, Jun Zhang, Minghui Wang, Jing Zhang, Wei Yang, Junzhou Huang, and Xiao Han. Transformer-based unsupervised contrastive learning for histopathological image classification. *Medical image analysis*, 81:102559, 2022. 1

[21] Yinxi Wang, Kimmo Kartasalo, Philippe Weitz, Balazs Acs, Masi Valkonen, Christer Larsson, Pekka Ruusuvuori, Johan Hartman, and Mattias Rantalainen. Predicting molecular phenotypes from histopathology images: a transcriptome-wide expression–morphology analysis in breast cancer. *Cancer research*, 81(19):5115–5126, 2021. 1

Bridging Morphology and Molecular Signatures: Multi-Task Deep Learning for Multi-Omics Prediction from Histopathology

Supplementary Material

7. Top predicted targets

Top Genes	Pearson Correlation	Top Proteins	Pearson Correlation	Top microRNA	Pearson Correlation
AGR3	0.707977	ERALPHA	0.645216	hsa-mir-18a	0.655745
ESR1	0.701862	ASNS	0.604106	hsa-mir-577	0.597472
CCNE1	0.693196	ECADHERIN	0.600070	hsa-mir-17	0.591355
TBC1D9	0.693071	PLK1	0.597685	hsa-mir-942	0.585266
MLPH	0.687592	CYCLINB1	0.586213	hsa-mir-190b	0.579965
THSD4	0.685178	GATA3	0.583716	hsa-mir-210	0.574278
C6orf97	0.684455	CDK1_pT14	0.575290	hsa-mir-505	0.573169
FOXA1	0.682475	FOXM1	0.563957	hsa-mir-590	0.572802
SCUBE2	0.681247	BETACATENIN	0.563820	hsa-mir-130b	0.550578
PSAT1	0.677510	CASPASE7CLEAVEDD198	0.553280	hsa-mir-19a	0.550189
ORC6L	0.676083	IDO	0.545630	hsa-mir-301b	0.548072
CENPA	0.675721	BCL2	0.542713	hsa-mir-301a	0.547313
KCNJ11	0.674130	Aurora-A	0.532306	hsa-mir-934	0.547292
CENPN	0.668963	INPP4B	0.528886	hsa-mir-1307	0.540670
GATA3	0.667477	MSH6	0.527805	hsa-mir-106b	0.538964
FAM54A	0.667091	Mnk1	0.501646	hsa-mir-197	0.535915
DNALI1	0.665293	CYCLINE1	0.500675	hsa-mir-1306	0.535651
CDC25A	0.664781	DJ1	0.488249	hsa-mir-1301	0.533062
DEGS2	0.663996	MCT4	0.488224	hsa-mir-877	0.522733
ZMYND10	0.661833	Puma	0.488050	hsa-mir-16-2	0.519643

Table 1. List of top predicted genes, proteins, microRNA, and pathways.