Network-Informed Functional Kinase Activity Profiling in a BRAF-Mutant Colorectal Cancer Patient-Derived Xenograft Model

UNIVERSITY OF SAN FRANCISCO | MS in Data Science and Artificial Intelligence

Borna Karimi[1,2], Cody Carroll [1], Hui Lin[2], and Jean-Philippé Coppe[2]

[1]Data Institute, University of San Francisco, [2]UCSF Health, Department of Radiation Oncology

Background and Problem Overview

Kinases and Kinase-Inhibitors

A **kinase** is an enzyme which regulates cellular functions by trasnfering a phosphate (PO_4^{3-}) group from ATP (Adenine Triphosphate) to substrates in a process known as phosphorylation. This is necessary for cell growth, division and signaling cascades.

Kinase inhibitors are a class of targeted drugs which can block kinases from phosphorylating substrates. These are largely used in cancer therapies to disrupt the pathways that lead to tumor growth.

Therapeutic Resistance

When a kinase is inhibited, tumors often reroute signaling through parallel pathways or feedback loops that restore growth. This adaptive reprogramming is the main obstacle to sustained responses from targeted therapies, making it essential to identify which kinases become active after treatment.

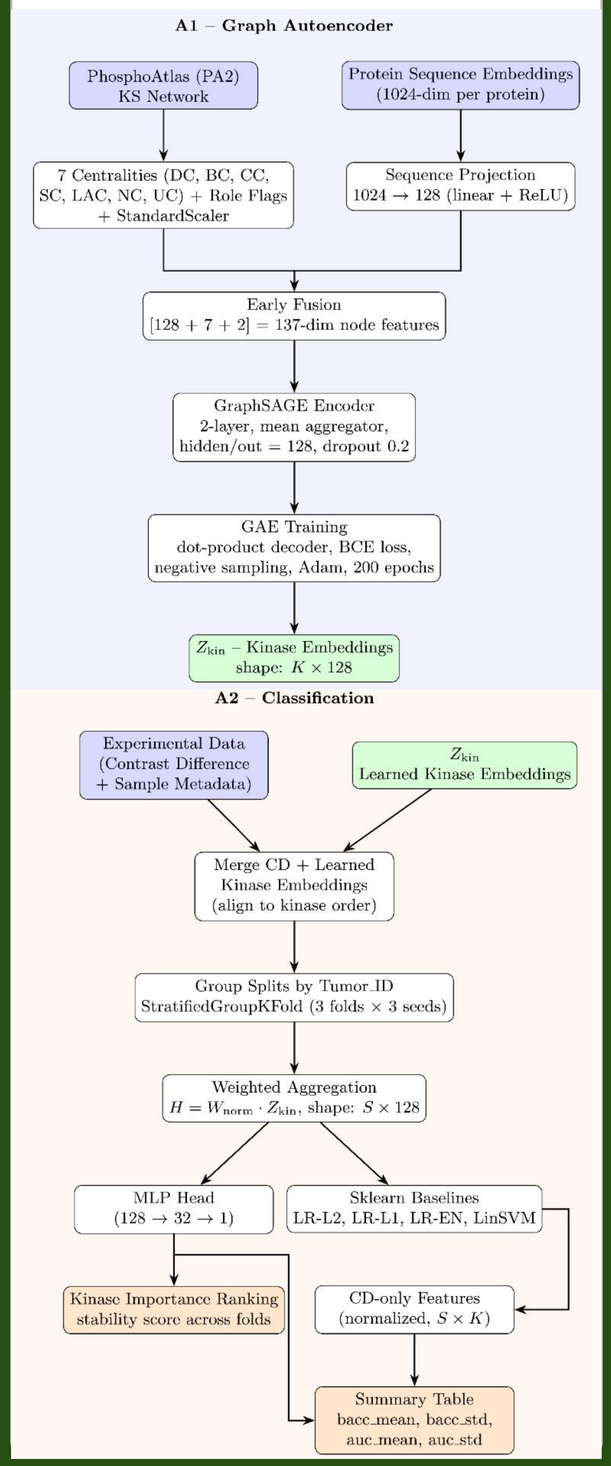
Data Description

The dataset includes 44 tumor samples from a BRAF-mutant colorectal cancer Patient-Derived Xenograft model. Each sample came from a different mouse, split between two treatment groups: combined BRAF + EGFR inhibition or vehicle control.

Samples were profiled using High-Throughput Kinase Activity Mapping (HT-KAM), a functional assay that measures kinase activity by tracking ATP consumption across short peptide sequences taken from known kinase substrates. ATP depletion at each peptide reflects the activity of the kinase that targets it, yielding direct activity measurements for 192 kinases per sample.

The kinase signaling graph used downstream came from PhosphoAtlas, a curated map of experimentally validated human kinase-substrate interactions. Treatment status is the binary label, and leave-one-mouse-out cross-validation is used for evaluation.

System Architecture Diagram



System Architecture

The pipeline has two stages. A Graph Autoencoder first learns 128-dim kinase embeddings from PhosphoAtlas, using a 2-layer GraphSAGE encoder over node features that combine projected proteinBERT sequence embeddings, graph centralities, and role flags. These embeddings are then merged with experimental contrast differences through weighted aggregation and passed to an MLP head and linear baselines, with CD-only models used for comparison and StratifiedGroupKFold splits by Tumor ID.

Results

We compared three classifiers (Ridge Logistic Regression, ElasticNet Logistic Regression, and a Multi-Layer Perceptron) trained on either direct activity measurements (DAMs) alone or DAMs augmented with network-projected features, evaluated with leave-one-mouse-out cross-validation.

	Direct Activity Measurements only		Network Projected + DAMs	
Model	BACC	AUC	BACC	AUC
LR-Ridge	0.74 ± 0.33	0.88 ± 0.20	0.78 ± 0.20	0.80 ± 0.20
LR-EN	0.75 ± 0.30	0.86 ± 0.20	0.80 ± 0.20	0.80 ± 0.20
MLP	0.56 ± 0.16	0.56 ± 0.32	0.82 ± 0.16	0.81 ± 0.22

The headline finding is not higher mean accuracy but more stable accuracy. Mean performance is comparable across feature sets for the linear models, but variance differs sharply: DAM-only models swing from near-random to near-perfect across folds, while network-projected features cut standard deviations roughly in half. The MLP shows the pattern most clearly, jumping from chance-level performance with DAMs alone to the best overall results in the table once network features are added.

Network-projected features also produced more stable kinase rankings across folds, consistently flagging the same kinases as treatment-responsive regardless of which mouse was held out. Stable differentiators included CDK1, PRKACA, SRC, AKT1, and MAPK pathway members. This ranking stability, more than the raw metrics, is what makes the approach useful for prioritizing combination therapy candidates.

Discussion

This work shows that augmenting direct kinase activity measurements with signaling network topology produces more reproducible identification of treatment-responsive kinases. While mean classifier performance was comparable between feature sets, network-projected features cut variance roughly in half and stabilized kinase rankings across folds. This stability matters more than raw accuracy: a high-variance ranking cannot reliably guide which kinases to pursue as combination therapy targets.

The kinases consistently flagged as treatment-responsive (CDK1, PRKACA, SRC, AKT1, and MAPK pathway members) are plausible mediators of resistance to BRAF + EGFR inhibition. Stable identification of these targets is the first step toward prioritizing them for combination therapy testing, where the goal is to block the escape routes tumors use to restore growth after initial treatment.

Limitations include the small sample size (44 mice) and the use of a single PDX model, which constrains how broadly the findings generalize. The approach should be validated on additional PDX models, other cancer types, and ideally patient samples before clinical claims can be made. Future work will extend this framework to other drug combinations and explore whether the same network projection improves kinase identification in non-BRAF-mutant contexts.

Acknowledgements

This work was supported by UCSF Health, Department of Radiation Oncology. Data was provided by the lab of Dr. Jean-Philippe Coppe.

References

1. Coppé, J.-P., Mori, M., Pan, B., et al. (2019). Mapping phospho-catalytic dependencies of therapy-resistant tumours reveals actionable vulnerabilities. *Nature Cell Biology* 21, 778–790.
2. Hunter, T. (2000). Signaling—2000 and beyond. *Cell* 100, 113–127.
3. Olow, A., Chen, Z., Niedner, R. H., et al. (2016). An atlas of the human kinome reveals the mutational landscape underlying dysregulated phosphorylation cascades in cancer. *Cancer Research* 76(7), 1733–1745.
4. Zhou, J., Cui, G., Hu, S., et al. (2020). Graph neural networks: A review of methods and applications. *AI Open* 1, 57–81.