ICML 2026

RNA-FM

Flow-Matching Generative Model for Genome-wide RNA-Seq Prediction

Yaxuan Song[1], Jianan Fan[1], Tianyi Wang[1], Qiuyue Hu[1], Hang Chang[2], Heng Huang[3], Weidong Cai[1]

[1]The University of Sydney · [2]Lawrence Berkeley National Laboratory · [3]University of Maryland College Park

github.com/YXSong000/RNA-FM

Why predict RNA-seq from histology?

Whole-Slide Images (WSI)

- Routinely acquired & low-cost
- Rich tissue morphology
- But: no direct molecular readout

RNA Sequencing

- Genome-wide: >20,000 genes
- Molecular ground truth
- But: costly & resource-intensive

The opportunity

Learn to predict gene expression directly from a slide: enriching routine diagnosis with genome-wide molecular insight, in silico, at near-zero extra cost.

WSI → RNA-seq

The limitation of existing methods

Prior work treats this as **a deterministic regression**, resulting in a single one-to-one mapping to a point estimate per gene.

Ignores heterogeneity

Collapses inter-patient and intra-tumoral variability to one answer.

No uncertainty

Point estimates look equally confident — no calibrated risk.

Hard to scale

Modeling >20,000 genes as independent tokens is infeasible.

Our stance Model the full **conditional distribution** of expression given morphology: probabilistic & generative.

Pathway-aware flow-matching model

Flow Matching

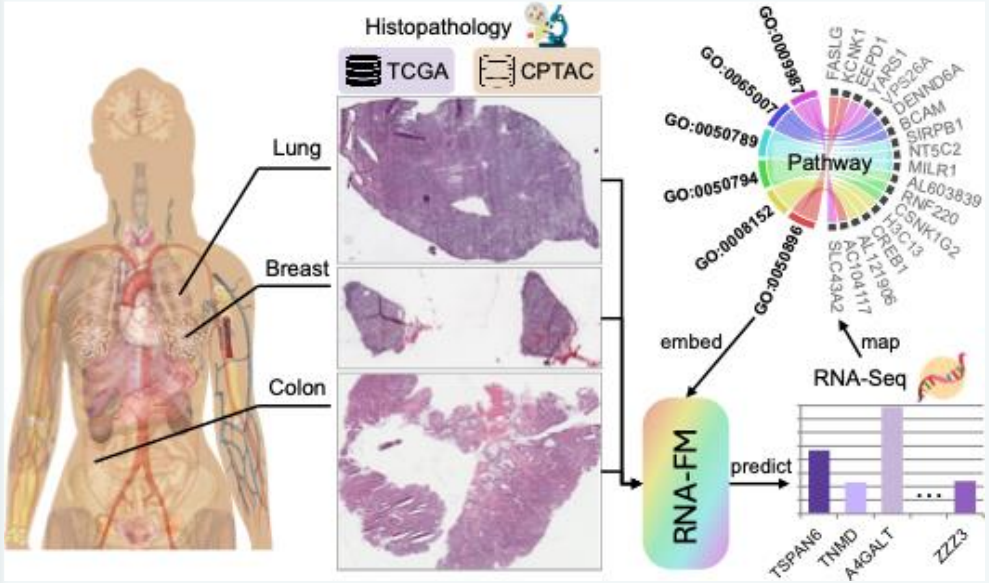
Learn a continuous-time velocity field that transports a simple Gaussian prior to the gene-expression distribution conditioned on the slide.

- **No iterative denoising / likelihood**
- **Stable training, fast ODE sampling**
- **One-to-many, uncertainty-aware**

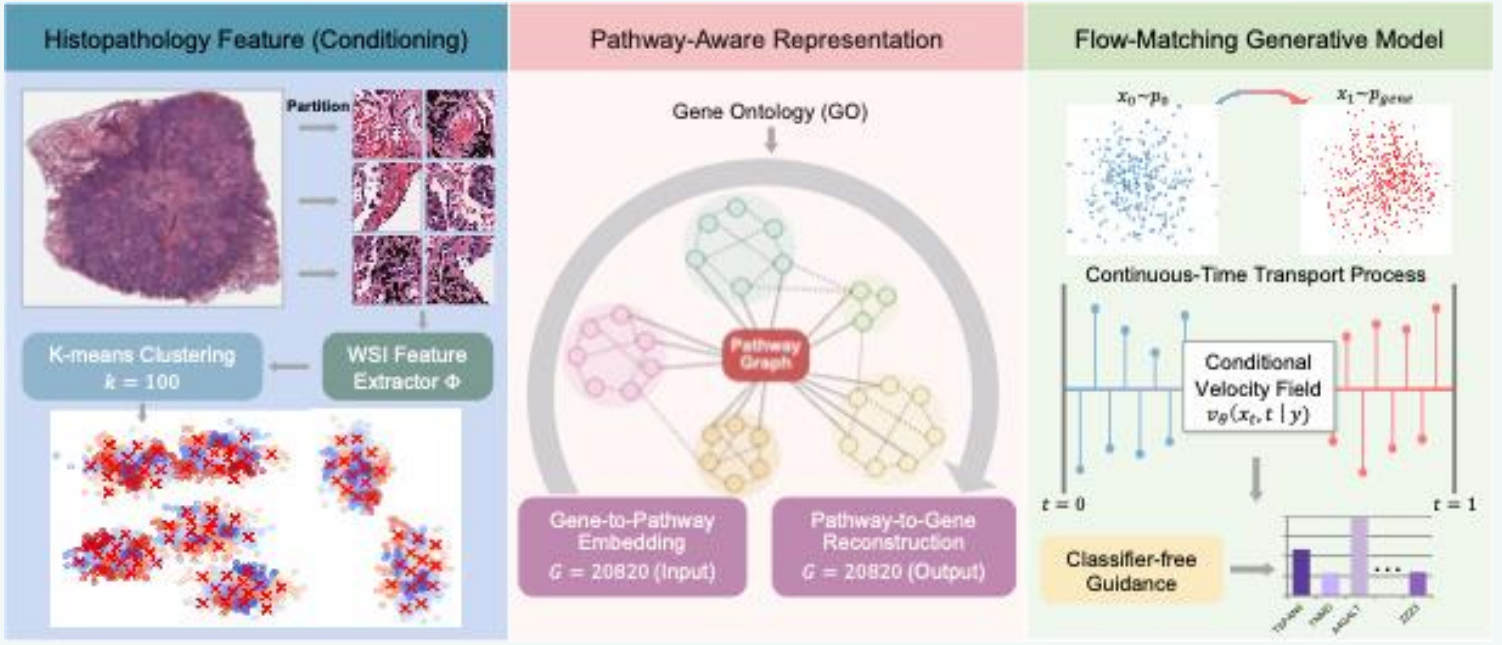
Pathway-Aware Tokens

Group 20,820 genes into Gene Ontology biological-process tokens, with a graph linking pathways that share genes.

- **Scalable to the full genome**
- **Biologically interpretable units**
- **Coherent cross-pathway information**



Framework overview



1 WSI conditioning
Tile → encode → k-means slide tokens

2 Pathway representation
GO tokens + inter-pathway graph attention

3 Flow-matching core
Conditional velocity field + CFG sampling

Continuous-time conditional transport

Linear interpolation (prior → target)

$$x_t = \alpha(t) \cdot x_1 + \sigma(t) \cdot x_0\,, \qquad t \in [0, 1]$$

Probability-flow ODE / velocity field

$$v^\star(x_t, t) = \frac{dx_t}{dt} = v_\theta(x_t, t \mid y)$$

Flow-matching objective

$$\mathcal{L} = \mathbb{E}_{x_0, x_1, t}[\| v_\theta(x_t, t|y) - v^\star(x_t, t) \|^2]$$

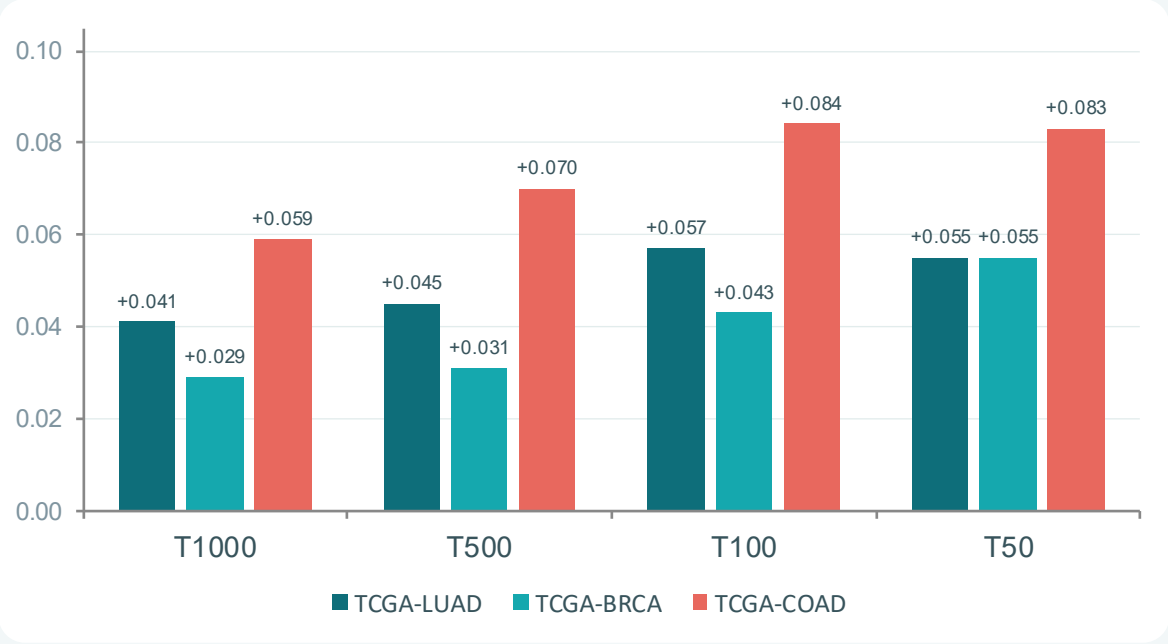
Classifier-free guidance (sampling)

$$v_\theta^{cfg} = v_\theta(\cdot\,|\emptyset) + s \cdot (v_\theta(\cdot\,|y) - v_\theta(\cdot\,|\emptyset))$$

$x_0 \sim \mathcal{N}(0, I)$ prior · $x_1 \sim p(x \mid y)$ gene expression · inference: Euler ODE over 20 steps

How much does RNA-FM gain?

Pearson Correlation Coefficient (PCC) improvement of RNA-FM (with UNI) over the strongest baseline.



Δ PCC vs. best baseline

COAD · top-100

0.81 → 0.92

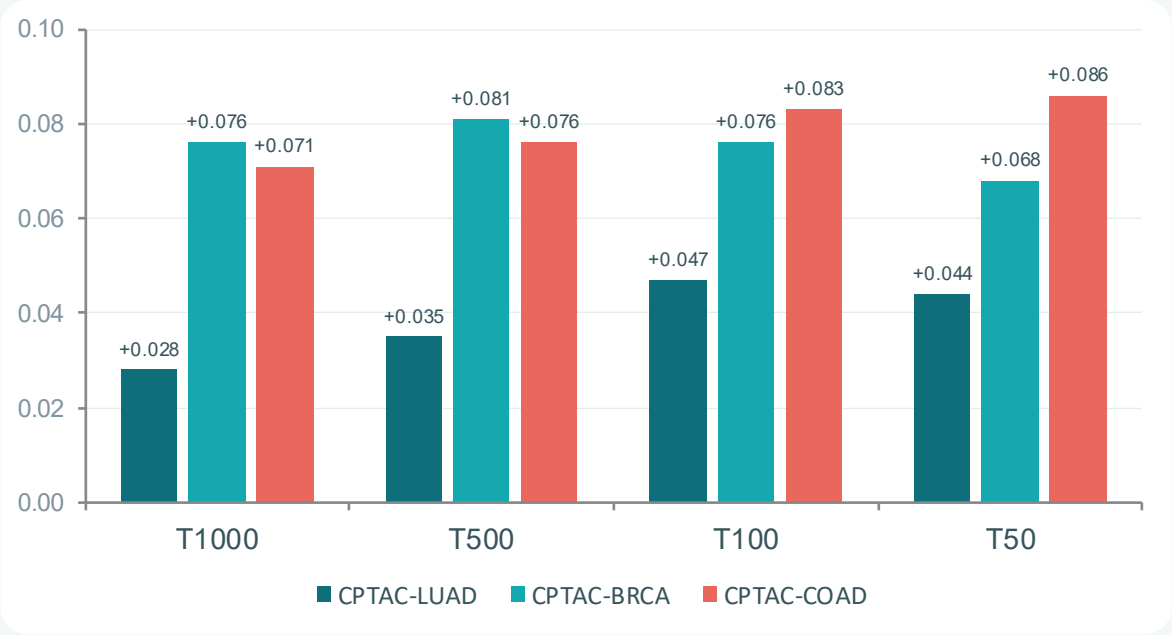
PCC vs. best baseline

Best in every setting

- All gene subsets (T1000→T20)
- ResNet-50 & UNI backbones
- Largest gains on smaller gene sets

Generalization

PCC improvement of RNA-FM over the strongest baseline on CPTAC.



Δ PCC vs. best baseline (CPTAC)

COAD · top-50

+0.086

largest single gain

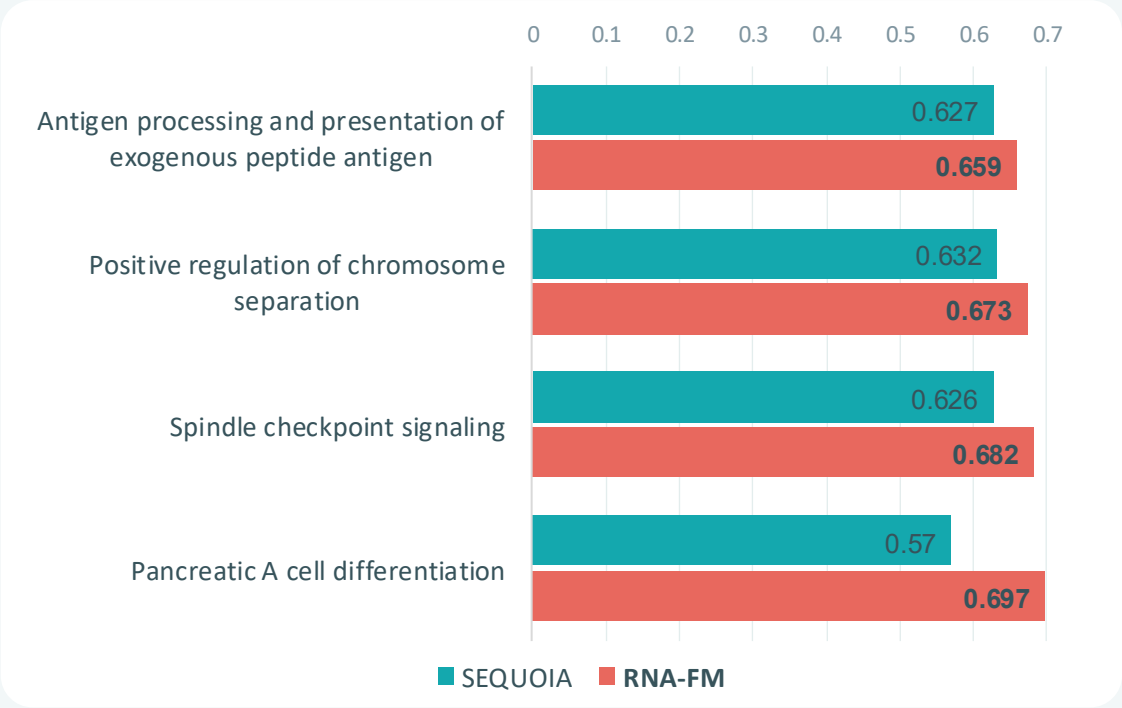
BRCA · top-500

+0.081

gains hold under shift

Pathway-Level Analysis

PCC comparison between RNA-FM with the strongest baseline (SEQUOIA) on representative pathways.



Takeaways

First of its kind — A flow-matching generative framework for genome-wide bulk RNA-seq prediction from whole-slide images.

Scalable & interpretable — Pathway-aware Gene Ontology tokens turn >20,000 genes into biologically meaningful, tractable units.

SOTA + reliable — Best accuracy on TCGA & CPTAC, calibrated uncertainty, and strong cross-cohort generalization.

github.com/YXSong000/RNA-FM

Thank you · Questions welcome

yson2999@uni.sydney.edu.au