ingegneria sismica
International Journal of Earthquake Engineering
Anno XLIII - Num. 1 - 2026

Construction of deep learning tumor prognosis risk assessment model based on genomic and clinical data fusion

Yuzhu Fang[1] *and Yan Zhang*[1,*]

[1] MOE Key Laboratory of Metabolism and Molecular Medicine, Department of Biochemistry and Molecular Biology, School of Basic Medical Sciences, Fudan University, Shanghai, 200032, China

SUMMARY: *In this study, a deep learning framework for tumor prognostic risk assessment was constructed around the joint modeling of dynamic genomic signals associated with target genes and structured clinical features. The genome input was GRO-seq data from Huh7 Mock cells and ADAMTS5e overexpressing cells, and the sequencing quality was maintained above Q38 overall and close to Q40 locally. groHMM identification results show that the density of the transcript part reaches 0.7197, and the FivePrimeFP signal is 0.4125 near the transcription start site. Differential analysis detected 312 upregulated and 145 downregulated de novo transcripts. Functional enrichment mainly involved apoptosis regulation, cell junction assembly, adhesion junction organization and receptor inhibitory activity. Site level analysis further confirmed that RUNX1 and LOC100507412 had definite changes, while the global RNA Pol II pause index was generally stable. The above genomic descriptors were co-coded with clinical variables such as age, stage, pathological index, treatment record and follow-up outcome, and the prognostic risk score was output after cross-modal fusion. This framework emphasizes dynamic feature compression, joint representation learning and reproducible calculation process, which provides a computational basis for subsequent tumor prognosis stratification and model expansion.*

KEYWORDS: *Genomic data fusion; Clinical feature representation; Deep learning; Risk assessment of tumor prognosis*

1 Introduction

Tumor prognosis assessment is shifting from single clinical judgment to collaborative modeling of multi-source heterogeneous information. Clinical staging, grading, treatment records, and follow-up outcomes can provide risk clues, but their characterization of transcriptional remodeling is limited. Genomic data can reflect the state of tumor activity, but it has the characteristics of high dimensionality and strong noise. Incorporating the two types of data into a unified computing framework has become a feasible path to improve the accuracy of risk representation and enhance the ability of model transfer. Schneider et al. [1] reviewed the integration of deep learning image analysis and genomic information in tumor pathology. Stahlschmidt et al. [2] summarized the biomedical multimodal deep learning fusion process. Wang et al. [3] used gene expression images and clinical data for long-term survival prediction of lung cancer. Ma et al. [4] proposed an improved XGBoost survival prediction model. Bashiri et al. [5] summarized the main paths for machine learning to process cancer gene expression

---

*zhangyan08@fudan.edu.cn
https://doi.org/10.65102/is2026227

data.

Multi-omics research around tumor prognosis modeling is advancing towards more fine-grained fusion strategies. Li et al. [6] proposed a hierarchical bilinear fusion model HFBSurv to compress heterogeneous survival information into a unified risk space. Zhang et al. [7] constructed a deep latent space fusion method for multi-omics adaptive representation. Zhang and Kiryu[8] proposed MODEC unsupervised integration framework for identifying tumor subtypes. Hu et al. [9] used digital pathology and deep learning to jointly predict the clinicopathological characteristics and prognosis of breast cancer. Li et al. [10] established a skin melanoma risk stratification model based on multi-omics integration. Existing results have proved the clear value of multimodal learning, but dynamic genomic signals rarely enter the clinical combined prognosis model, and the model input still stays at the static expression matrix level.

Target gene-related de novo transcript data provide a direct entry point for dynamic genomic feature construction. Previous experimental results showed that Huh7 cells had established Mock and target gene overexpression systems, and captured nascent RNA transcription events by GRO-seq. The sequencing quality was stable at about Q40 in each cycle. groHMM identification results showed that the transcription density was significantly enriched in the gene body region, with a TP of 0.7197, and a significant FivePrimeFP signal was still retained near the transcription start site with a value of 0.4125. A total of 312 up-regulated genes and 145 down-regulated genes were detected by differential analysis, and the functional enrichment was concentrated in the pathways of apoptosis, cell junction and adhesion junction organization. The track signal of chr21 representative locus LOC100507412 and RUNX1 showed clear changes, while the global RNA Pol II pause index remained stable. These results indicate that ADAMTS5e overexpression forms a structured set of dynamic transcriptional signatures that can be used to describe the genomic activity status associated with tumor progression.

Based on this, this paper focuses on the construction of a deep learning tumor prognosis risk assessment model combining genomic and clinical data. The de novo transcription signals related to the target genes are introduced as the input of the genome branch, and the clinical variables are encoded as structured feature vectors to complete cross-modal alignment, feature fusion and risk score output. In this paper, TSS neighborhood signals, differential transcripts, functional enrichment features, key site activity and pause index are used to form dynamic molecular representations, while age, stage, pathological indicators and follow-up information are mapped into trainable representations, and then the joint risk space is learned through a deep fusion module. This pathway not only retains the experimental basis in the previous experimental data, but also transforms the genomic signal into a high information density input that can be used for computational prediction, so as to provide a research starting point for tumor prognosis stratification and subsequent technology implementation.

2 Literature Review

The joint modeling of multi-omics and clinical data is pushing the research of tumor prognosis from single-source statistical analysis to cross-modal representation learning. Wen et al. proposed FGCNSurv dual fusion graph convolutional network, which maps multi-omics node relationship and survival information into graph structure space together, and enhances the correlation modeling ability of risk representation [11]. Salimy et al. constructed a deep learning framework integrating multi-omics and clinical data for colon cancer survival related group prediction, and proved that the joint input had higher hierarchical stability than the single-modal input [12]. Ballard et al. systematically summarized the multi-omics integration path driven by deep learning, and pointed out that the encoding method, alignment strategy and fusion layer structure directly affect the feature transfer effect [13]. Steyaert et al. applied multimodal deep learning to the prognosis prediction of brain tumors in adults and children, demonstrating the practical feasibility of collaborative modeling of heterogeneous medical data in a unified representation space [14].

In terms of joint modeling of imaging, clinical and molecular information, Wang et al. used multi-instance learning to fuse pathological full-field images and clinical features to complete the prognosis prediction of HER2-positive breast cancer patients [15]. Wang et al. further proposed a multi-modal data integration framework for breast cancer disease-free survival prediction, indicating that deep feature interaction can improve the consistency of risk scores [16]. Yang et al. compared the performance of multiple machine learning models in triple-negative breast cancer classification, and used deep learning-based multi-omics integration to improve the accuracy of subtype recognition [17]. Lan et al. proposed the DeepKEGG framework to complete cancer recurrence prediction and marker discovery under the constraints of biological pathway knowledge, so that feature learning has more explicit biological semantics [18]. Tanvir et al. constructed the MOGAT graph attention framework for cancer subtype prediction, which strengthened the weight allocation ability between multi-omics nodes [19]. Wang et al. proposed DFASGCNS dual-fusion channel stacked graph convolution model, which reflects the advantages of graph structure integration in ovarian cancer prognosis prediction [20].

Existing research has formed a basic technical chain from feature coding, representation alignment to risk output. However, most methods are mainly based on static expression matrix, pathological images or routine clinical variables, and the absorption of real-time transcription activity features is still weak. The previous experimental results showed that Huh7 cells completed GRO-seq sequencing under the conditions of Mock and target gene overexpression, and the data quality maintained at Q40 on each cycle. groHMM identified 312 up-regulated genes and 145 down-regulated genes. LOC100507412 and RUNX1 on chr21 showed clear differential coverage, and the RNA Pol II pausing index remained stable as a whole. These results indicate that ADAMTS5E-related de novo transcriptional signals have structural coding value and can be used as dynamic genomic input in tumor prognosis modeling. As shown in Table 1.

Table 1: Summary of related studies and the results of this paper

Study	Data Type	Method	Result Characteristics	Gap Relative to This Study
Wen et al. [11]	Multi-omics	Dual-Fusion GCN	Emphasizes graph-structured survival prediction	Does not incorporate dynamic transcriptional signals with joint clinical encoding
Salimy et al. [12]	Multi-omics + Clinical Data	Deep Learning Fusion Framework	Supports survival subgroup prediction	Lacks real-time transcription-layer feature construction
Steyaert et al. [14]	Multimodal Medical Data	Deep Learning Prediction	Applicable to brain tumor prognosis	Large differences in tumor type and input structure
Lan et al. [18]	Multi-omics + Pathway Knowledge	DeepKEGG	Strengthens biological interpretability	Strong pathway constraints with insufficient dynamic regulatory modeling
Wang et al. [20]	Molecular Features	Dual-Channel Graph Convolution	Improves ovarian cancer prognosis prediction	Limited depth of clinical feature fusion
This Study	Genomic Data + Clinical Data	Deep Learning Risk Assessment Model	Introduces nascent transcription features related to target genes	Emphasizes dynamic feature encoding and cross-modal risk representation

In addition, the change of chr21 representative sites in the previous experiments is not an isolated phenomenon, and it forms a computational correspondence with the direction of TSS neighborhood signal enhancement, differential transcript up-regulation and functional enrichment, which provides a direct basis and interface basic support for the subsequent construction of risk feature compression and cross-modal mapping for patient outcome.

From the perspective of computational modeling, the focus of future work is not to continue to increase the input sources, but to establish a unified representation mechanism that can accommodate dynamic genomic signals and structured clinical variables at the same time. The TSS neighborhood signals, differential transcripts, functional enrichment results, key site activity and pause index under ADAMTS5e overexpression condition are encoded and compressible, which are suitable to be sent into the deep learning framework together with age, stage, treatment record and follow-up outcome. Based on this understanding, based on the existing multimodal fusion research, this paper extracts the dynamic features of de novo transcription related to target genes, and further constructs a joint representation model for tumor prognostic risk assessment, so as to strengthen the consistency, interpretability and computational adaptability of risk stratification.

3 Materials and Methods

3.1 Data source and sample composition

The data source of this study consists of two parts: genomic branch and clinical branch. In the genome branch, Huh7 liver cancer cells were used as the experimental objects, and the Mock control group and ADAMTS5e overexpression group were set up, and the CRISPR/dCas9-SAM system was used to establish a stable overexpression model. The clinical branch is oriented to the subsequent prognostic risk assessment task, taking on structured variables such as patient age, stage, pathological indicators, treatment records and follow-up outcomes. The two types of data are functionally separated by design: the former is used to extract the dynamic features of ADAMTS5e related nascent transcription, and the latter is used to characterize the differences in disease status and outcome at the patient level, and finally enters the multi-modal deep learning framework under a unified sample index. The composition of the sample branches and their functional positioning in the model are shown in Table 2.

Table 2: Sample branches and functional localization

Data Branch	Samples / Subjects	Main Content	Role in the Model
Genomic Branch	Huh7 Mock, Huh7 OE-ADAMTS5e	GRO-seq, ChIP-seq, site coverage, pausing index	Provides dynamic transcription features
Clinical Branch	Patient-level structured records	Age, stage, pathology, treatment, follow-up outcomes	Provides prognostic phenotype constraints
Fusion Layer	Unified sample index	Aligned multimodal inputs	Forms a joint risk representation

The Huh7 cell system of the same source was used for the experimental samples. The Mock group and the OE group were cultured under the same conditions, and the medium was DMEM containing 10% fetal bovine serum and 1% penicilin-streptomycin, and the environment was maintained at 37 ℃ and 5%CO2. The GRO-seq process is completed in four days: On the second day, the nuclear run-on reaction containing Br-UTP was performed. On the third day, RNA purification, dnase digestion, fragmentation and anti-brdu magnetic beads immunoprecipitation were performed. On the fourth day, directed library construction was completed and sequenced on Illumina platform. This flow directly captures nascent Rnas, bringing the input data closer to the immediate transcriptional regulatory state. The histone modification branch calls H3K27ac, H3K4me1 and H3K4me3 ChIP-seq data from GEO database and uniformly maps to hg38 reference genome. The genomic data sources and experimental compositions are shown in Table 3.

Table 3: Genomic data sources and experimental composition

Data Type	Source / Platform	Key Settings	Output Object
GRO-seq	Huh7 Mock and OE	Br-UTP run-on, Illumina sequencing	Nascent RNA reads
ChIP-seq	Public GEO dataset	H3K27ac, H3K4me1, H3K4me3	Chromatin activity tracks
Reference Genome	hg38	Unified coordinate system	Interval-level mapping results

The annotations required for computational analysis are determined synchronously with the filtering criteria during the data entry phase. Only unique alignment reads were retained after GRO-seq alignment. Pol II pause index was calculated from 50 bp upstream to 200 bp downstream of TSS in promoter region and from 200 bp downstream to 5000 bp downstream of TSS in gene body region. To ensure reliability, genes with total reads lower than 20 in the gene body and with non-finite values of the pause index were removed. For differential transcription analysis, $|\log2FC|>1$ and $P<0.05$ were used as the screening criteria. groHMM was used to identify active transcription units, and DESeq2, GenomicRanges, GenomicAlignments and Rsamtools were combined to complete statistics and interval mapping. The preprocessing rules and calculation entry are shown in Table 4.

Table 4: Preprocessing rules and computation entry

Module	Rule / Parameter	Purpose
Reads Filtering	Retain only uniquely mapped reads	Reduce mapping ambiguity
Pausing Index	Promoter: −50 to +200 bp; Gene body: +200 to 5000 bp	Quantify Pol II pausing
Coverage Filtering	Total gene body reads ≥ 20	Improve statistical reliability
Differential Screening	$|\log2FC| > 1$, $P < 0.05$	Generate candidate transcription features
Analysis Environment	R 4.2.2 + Python	Support reproducible computation

Before entering the model, the data is organized into three layers of objects: sample-level indexes are responsible for linking experimental conditions with clinical records, site-specific and transcription-level features are responsible for carrying TSS neighborhood density, differential expression, critical site coverage, and pause index, and function-level summary features are used to integrate GO enrichment results with histone modification background. Through this hierarchical structure, the genomic branch outputs an input tensor with local trajectory information and global statistical constraints, and the clinical branch maintains a low-dimensional, structured and completable feature representation. Therefore, the data source and sample composition directly determine the input boundary, alignment granularity and computational stability of the subsequent fusion model, and have the ability to reproduce experiments.

3.2 Genomic Data encoding and Clinical feature representation

Before fusion modeling, genomic branches and clinical branches need to be transformed into trainable objects in the same numerical space. Instead of directly using raw reads on the genome side, the de novo transcriptional changes formed around ADAMTS5e overexpression, The TSS neighborhood density, gene body coverage, FivePrimeFP, Post-TTS, TUA, expression amplitude of differential transcripts, GO functional tags and local activity of key sites were collated into a multi-layer feature matrix. On the clinical side, variables such as age, tumor stage, pathological grade, treatment record, and follow-up outcome were mapped into structured vectors. Previous experimental results have given Huh7 Mock/OE system, Q40 sequencing quality, TP=0.7197, FivePrimeFP=0.4125, 312 up-regulated genes and 145 down-regulated genes. Therefore, this section does not focus on repeated statistics. Instead, we need to condense these results into standard input suitable for reading by deep learning prognostic models.

In order to achieve location-level compression while preserving local regulatory information, the following feature mapping relationship is constructed for the g transcription unit, and the promoter region and the gene body region are spliced after normalization respectively:

$$u_g=[\phi_1(\rho_g^{tss}),\phi_2(\rho_g^{body}),\phi_3(\pi_g),\phi_4(\delta_g),\phi_5(\omega_g)] \tag{1}$$

Here, ρ_g^{tss} represents the coverage density of the g gene in the tss neighborhood, ρ_g^{body} represents the coverage density of the gene body region, π_g represents the pause index, δ_g represents the differential expression strength, ω_g represents the functional layer summary vector, and ϕ_1~ϕ_5 represent the nonlinear compression mapping that acts on different source features respectively. The function of this formula is to project genomic indicators with obvious differences in distribution scales into a fixed dimension.

In order to avoid the local high coverage sites dominating the overall input numerically, we further use robust normalization based on quantile truncation and logarithmic compression to transform the original genomic features into stable distribution values:

$$\tilde{x}_{g,j}=\frac{\log(1+\max(x_{g,j}-q_{0.01,j},0))}{\log(1+q_{0.99,j}-q_{0.01,j})+\epsilon} \tag{2}$$

Here, $x_{g,j}$ represents the original value of the g sample on the j genomic index, $q_{0.01,j}$ and $q_{0.99,j}$ represent the 1% and 99% quantile of the index, respectively, and ϵ is a stability constant to prevent the denominator from being zero. This formula is used to suppress long-tail coverage values and improve the comparability between representative sites such as RUNX1 and LOC100507412 and general sites in the same input space.

Instead of simple concatenation, clinical feature representations were treated separately according to continuous variables, categorical variables, and time outcomes. Continuous variables were standardized by intervals, categorical variables were mapped by embedding, and temporal outcomes were represented by auxiliary masks with censored survival goals, resulting in the following mixed clinical representation:

$$v_i=[\psi(c_i^{num}) ; \sum_{r=1}^{R} E_r [z_{i,r}] ; m_i\tau_i] \tag{3}$$

Here, c_i^{num} is the continuous clinical variable of the i patient, $\psi(\cdot)$ represents the standardized mapping, E_r is the embedding matrix of the r category field, $z_{i,r}$ is the category value index, m_i is the censoring indicator, and τ_i is the follow-up time. This formula collates the clinical end information into vector inputs that are equally trainable as the genomic end.

After finishing the original variables, it is necessary to further clarify the field composition, representation mode and output dimension of the two types of inputs in the encoding stage. The specific Settings are shown in Table 5.

Table 5: Genomic coding and clinical representation field composition

Branch	Input Object	Encoding Method	Output Dimension	Description
Genomic Branch	TSS density, gene body density	Quantile normalization + nonlinear mapping	64	Preserves differences between promoter and elongation regions
Genomic Branch	Pausing index, FivePrimeFP	Standardization + rank compression	16	Describes Pol II dynamics
Genomic Branch	Differential transcripts, GO summary	Sparse screening + functional encoding	48	Characterizes ADAMTS5e-induced transcriptional remodeling
Clinical Branch	Age, tumor stage, pathological grade	Continuous normalization + categorical embedding	32	Forms the patient state vector
Clinical Branch	Treatment records, follow-up outcomes	Temporal masking + censoring representation	16	Connects to the prognostic target

Through the above encoding, the genomic branch retains the local dynamics of ADAMTS5E-related de novo transcription, and the clinical branch retains the outcome constraint at the patient level, both of which already have a unified numerical shape before entering the fusion layer.

3.3 Multi-modal feature fusion method

After the genome and clinical branches have been independently encoded, the next step is not to directly sum or simply concatenate, but to establish fusion structures that can simultaneously preserve local transcriptional dynamics and patient outcome semantics. Considering that the genomic features associated with ADAMTS5e contain both dense numerical values and summary information from GO enrichment and orbit maps, a "two-layer interaction-gated writeback" fusion method is adopted in this paper: The first layer is used to learn cross-modal correlations, and the second layer is used to invert the clinical context to constrain the genome representation, avoiding overfitting of high-dimensional omics signals in the small-sample survival task.

To establish an explicit association between the two modalities, we first define the genomic representation matrix $G \in \mathbb{R}^{n\times d_g}$ and the clinical representation matrix $C \in \mathbb{R}^{n\times d_c}$, and then calculate the cross-modal coupling strength by bilinear correlation mapping:

$$A = \mathrm{softmax}\left(\frac{GW_{gc}C^{\top}}{\sqrt{\kappa}}\right) \tag{4}$$

Here, W_{gc} is the trainable correlation matrix, κ is the scaling factor, and A is the cross-modal attention weight. This formula is used to characterize the degree of coupling between a certain omics model and a certain clinical state.

Subsequently, the genome representation is contextually rewritten using clinical semantics to obtain the genome enhancement vector under clinical conditions:

$$\hat{g}_i = g_i + \lambda_i \sum_{j=1}^{n} a_{ij}\ Uc_j,\quad \lambda_i = \sigma(r^{\top}g_i) \tag{5}$$

Here, g_i and c_j denote the genomic representation of the i sample and the clinical representation of the j sample, respectively, U is the linear transformation matrix, a_{ij} is the element in the weight matrix A, $\sigma(\cdot)$ is the Sigmoid function, and λ_i is the gating coefficient. This formula enables the clinical structure of different patients to form conditional constraints on similar omics patterns.

In order to control the proportion imbalance of two types of modes after fusion, this paper designs an adaptive counterweight based on energy difference:

$$\eta_i = \frac{\exp(\|\hat{g}_i\|_2)}{\exp(\|\hat{g}_i\|_2) + \exp(\|c_i\|_2)},\quad h_i = \eta_i\,\hat{g}_i + (1-\eta_i)Vc_i \tag{6}$$

Here, h_i is the final fusion representation and V is the clinical branch projection matrix. This formula adaptively adjusts the contribution ratio through the modal energy to avoid the high-dimensional genomic input completely suppressing the low-dimensional clinical information.

Considering the obvious local structure of ADAMTS5E-related features, we further introduce a consistency constraint to keep the fused representations from similar prognostic groups locally compact in the latent space:

$$\mathcal{R}_{\mathrm{coh}} = \frac{1}{|\Omega|}\sum_{(p,q)\in\Omega} \omega_{pq}\ \|h_p - h_q\|_2^2 \tag{7}$$

Here, Ω is the set of sample pairs and ω_{pq} is the neighboring weights generated based on prognostic labels and local omics similarity. The effect of the term is to enhance the intra-class consistency of the fusion representation, making the arrangement of 312 up-regulated genes, 145 down-regulated genes and their functional tags more stable in the risk space.

The key to this step is not to simply stack omics and clinical, but to let the clinical end participate in the interpretation of the genomic end, and let the omics end complement the resolution of the clinical end. For this task, TSS neighborhood density, differential transcription, GO enrichment, and RUNX1 and LOC100507412 activity constitute the primary information on the dynamic molecular side, while stage, treatment response, and follow-up outcome provide external risk constraints. Through the above fusion method, the input received by the subsequent risk model is no longer the result of loose splicing, but a joint representation with clear cross-modal structure.

3.4 Construction of deep learning tumor prognosis risk assessment model

After the joint representation construction, this paper further establishes a deep learning tumor prognostic risk assessment model for censored survival data. The overall model adopts a bi-objective structure of "shared trunk road - risk head - sorting head", in which the shared trunk road is used to receive the fused multimodal features and complete the unified representation, the risk head is responsible for output continuous risk scores, and the sorting head is used to strengthen the ability to distinguish the prognosis order between samples. This structure is mainly used to take into account both absolute risk estimation and relative order discrimination. If only a single survival risk output module is used, the local ranking information is easy to be weakened. If only relying on ranking constraints, the risk scale is difficult to maintain stability. Therefore, the bi-objective structure is more suitable for the prognostic modeling task of fusing genomic and clinical data addressed in this paper.

In order to make the fusion representation enter the unified survival risk space, this paper first defines the shared trunk road output as follows.

$$s_i = \varphi(B_2\,\zeta(B_1 h_i + b_1) + b_2) \tag{8}$$

Here, B_1,B_2 are the shared trunk parameters, b_1,b_2 are the bias terms, $\zeta(\cdot)$ is the GELU activation, and $\varphi(\cdot)$ is the layer normalization and residual combination.

On this basis, the risk head adopts the survival risk output with temperature control:

$$r_i = \frac{q^\top s_i}{\tau_r}, \quad \underset{i}{\hat{\lambda}} = \exp(r_i) \tag{9}$$

Here, q is the risk projection vector, τ_r is the temperature scaling parameter, r_i is the log risk score, and $\underset{i}{\hat{\lambda}}$ is the predicted hazard rate. This equation is used to produce prognostic risk outputs with comparable scales.

To simultaneously preserve temporal order consistency, the sorting head establishes pairwise ordering constraints for any pair of samples (i,j) as follows.

$$\mathcal{L}_{\mathrm{ord}} = \frac{1}{|\Pi|}\sum_{(i,j)\in\Pi} \log(1 + \exp[-\chi_{ij}(r_i - r_j)]) \tag{10}$$

Here, Π is the set of comparable sample pairs, $\chi_{ij} = 1$ means that the risk of i should be higher than$j$, otherwise it is taken as -1. This term constrains the model to learn a clearer boundary between high and low risk stratification.

In order to facilitate the illustration of the structural positioning and output relationship of each module in the overall network, the core configuration of the model is arranged in Table 6.

Table 6: Structural configuration of deep learning prognostic risk model

Module	Structure	Output Dimension	Function
Shared Backbone	Two fully connected layers + GELU + LayerNorm	64	Receives the fused representation and compresses risk semantics
Risk Head	Linear projection + temperature scaling	1	Outputs a continuous risk score
Ranking Head	Pairwise difference mapping	1	Enhances the relative prognostic ordering
Regularization Layer	Dropout + weight decay	—	Controls overfitting in small-sample settings

Finally, the overall model goal is written as follows:

$$\mathcal{J} = \mathcal{L}_{\mathrm{cox}} + \alpha \mathcal{L}_{\mathrm{ord}} + \beta \mathcal{R}_{\mathrm{coh}} \tag{11}$$

Here $\mathcal{L}_{\mathrm{cox}}$ is the partial likelihood loss, and α and β are the weight coefficients. The role of the overall objective is to make the model have censored survival modeling ability, hierarchical ability, and cross-modal consistency.

The model structure matches the results of Chapter IV: TSS enrichment, differential transcription, GO enrichment, critical site activity and pause index are not regarded as dispersed evidence, but as unified omics input into the shared trunk road; Clinical variables then provide outcome constraints through the fusion layer. The risk model thus established both inherits the dynamic characteristics of ADAMTS5E-related de novo transcripts and retains the patient-level discriminative power required for tumor prognostic tasks.

3.5 Model Training strategy and parameter Settings

In the training phase, a three-stage strategy of "branch warming-joint fine-tuning-stable convergence" is adopted. Since the dimension of genomic branches is significantly higher than that of clinical branches, if trained directly end-to-end, it is easy to cause the early gradient to be dominated by the omics side, and the clinical embedding is quickly compressed. Therefore, in this paper, the two branches are warmed up independently, and then jointly fine-tuned at the fusion layer, and the sample reweighting and exponential moving average parameter update are introduced in the later stage to ensure that the survival risk score remains stable in the mini-batch training.

In the early stage of joint training, the high-dimensional input of the genomic branch is more likely to produce large gradient responses, while the update amplitude of the clinical branch is usually smaller due to the lower dimension. In order to avoid a single branch to dominate parameter iteration too fast, this paper adopts an adaptive learning rate adjustment strategy based on gradient energy ratio, which is formally defined as follows:

$$\ell_{\mathrm{g}}^{(\mathrm{t})} = \ell_0 \left(1 + \beta \frac{\|\nabla \mathcal{J}_{\mathrm{g}}^{(\mathrm{t})}\|_2}{\|\nabla \mathcal{J}_{\mathrm{c}}^{(\mathrm{t})}\|_2 + \varepsilon}\right)^{-1}, \quad \ell_{\mathrm{c}}^{(\mathrm{t})} = \ell_0 \left(1 + \beta \frac{\|\nabla \mathcal{J}_{\mathrm{c}}^{(\mathrm{t})}\|_2}{\|\nabla \mathcal{J}_{\mathrm{g}}^{(\mathrm{t})}\|_2 + \varepsilon}\right)^{-1} \tag{12}$$

Here, $\ell_{\mathrm{g}}^{(\mathrm{t})}$ and $\ell_{\mathrm{c}}^{(\mathrm{t})}$ denote the adaptive learning rate of genomic branch and clinical branch in the t round of training, ℓ_0 denotes the basic learning rate, $\nabla \mathcal{J}_{\mathrm{g}}^{(\mathrm{t})}$ denotes the gradient vector of the objective function corresponding to the genomic branch in the current round, $\nabla\mathcal{J}_c^{(t)}$ denotes the gradient vector of the objective function corresponding to the clinical branch, and $\|\cdot\|_2$ denotes the two-norm. Let β denote the regulation coefficient and ε denote the stability constant that prevents the denominator from being zero. The function of this formula is to dynamically scale the learning rate according to the gradient energy difference of the two branches, so that the high-response branch automatically slows down, and the low-response branch obtains a more stable update space, so as to alleviate the convergence imbalance caused by dimension asymmetry in the fusion training stage.

Considering the uneven distribution of censored samples and event samples, this paper introduces reweighting based on event density:

$$\upsilon_i = 1 + \rho \frac{\mathbb{I}(e_i = 1)}{\sum_{m=1}^{N} \mathbb{I}\,(e_m = 1) + \epsilon_e} \tag{13}$$

Here, υ_i is the sample weight, e_i is the event indicator, ρ is the reinforcement coefficient, and ϵ_e is the smoothing term. This formula is used to improve the contribution of real event samples in training and reduce the risk head shift caused by the dominance of censored data.

Parameter update uses exponential moving average to form a stable inference model:

$$\overline{\Theta}^{(t)} = \mu\,\overline{\Theta}^{(t-1)} + (1-\mu)\Theta^{(t)} \tag{14}$$

Here, $\Theta^{(t)}$ is the current round parameter, $\overline{\Theta}^{(t)}$ is the smoothing parameter, and μ is the attenuation coefficient. This formula can reduce the impact of mini-batch fluctuations on the final risk score, which is especially suitable for the training scenario such as this one where dynamic omics features and structured clinical variables are simultaneously integrated.

During training, the genomic branch was warmed up for 30 rounds and the clinical branch was warmed up for 20 rounds, after which the joint training was performed and early stopping was implemented on the validation set. Since the number of up-regulated genes is significantly higher than that of down-regulated genes, it is more necessary to suppress the unilateral drive of omics side to risk head for model training, so the adaptive learning rate and sample reweighting are necessary configurations. The training strategy ultimately serves the goal of stably transforming these omics signals into patient-level prognostic risk scores without disrupting the transcriptional dynamics of ADAMTS5e.

3.6 Evaluation indicators and statistical analysis methods

The evaluation part simultaneously covers three types of objectives: the first is the prognostic discrimination ability, the second is the consistency of risk ranking, and the third is the stability of the fusion representation. Since this paper is not a separate differential analysis, but to put ADAMTS5E-related omics features and clinical variables into the survival model together, a single accuracy or general classification AUC is not enough. The evaluation caliber needs to simultaneously reflect the ranking effect under censored data, the time-dependent discrimination ability, and the robustness of the model output to omics perturbations.

At the level of prognosis discrimination, in order to simultaneously consider the bias caused by censored samples and the consistency of risk ranking, this paper uses the inverse probability censored weighted consistency index as the core evaluation metric, which is defined as follows:

$$C_{IPCW} = \frac{\sum_{i=1}^{N} \sum_{j=1}^{N} \mathbb{I}\,(t_i < t_j)\, \mathbb{I}(e_i = 1)\, \dfrac{\mathbb{I}(r_i > r_j) + \frac{1}{2}\mathbb{I}(r_i = r_j)}{\widehat{G}\,(t_i)^2}}{\sum_{i=1}^{N} \sum_{j=1}^{N} \mathbb{I}\,(t_i < t_j)\, \mathbb{I}(e_i = 1)\, \dfrac{1}{\widehat{G}\,(t_i)^2}} \tag{15}$$

Here, t_i represents the follow-up time of the i sample, $\mathbb{I}(\cdot)$ is the indicative function, and $\widehat{G}\,t_i)$ represents the Kaplan-Meier estimate of the censored distribution at time t_i. The index imposes an inverse probability censoring weight on comparable sample pairs, weakens the interference of the censoring mechanism on the risk ranking evaluation, and makes the prognosis discrimination results of the model after fusing genomic and clinical data more stable and comparable.

In the time-dependent evaluation level, this paper retains the dynamic discrimination ability under a specific time cut-off to supplement the overall ranking index:

$$AUC(\tau) = Pr(r_a > r_b \mid t_a \leq \tau, e_a = 1, t_b > \tau) \tag{16}$$

Here, τ denotes the preset follow-up time cut-off, a denotes the samples whose events occurred before τ, and b denotes the samples that are still in the risk set after τ time. This metric is used to judge the discriminative power of the model in the short, medium and relatively long-term prognostic intervals.

In summary, the above evaluation system does not only interpret the results for a single risk score, but synchronously constrain the model output from three levels: ranking consistency, time cut-off discrimination ability and statistical robustness. In this way, it can not only test the prognostic discrimination effect after fusing genomic and clinical data, but also measure the actual contribution and stability of omics features in the joint modeling process.

4 Results and analysis

4.1 Experimental process and sample acquisition overview

In this study, Huh7 liver cancer cells were used as the source of genomic signal, two groups of Mock and ADAMTS5e overexpression samples were set, and subsequent clinical data were mapped into independent structured branches to form a dual input framework of "dynamic transcription feature-clinical variables". As shown in Fig. 1, the nuclei were extracted and stored at −80 ° C on the first day, the nuclear run-on reaction was performed at 30 ° C for 7 min on the second day to capture BR-UTP-labeled newborn RNA, and RNA purification, dnase treatment and anti-brdu magnetic beads immunoprecipitation were performed on the third day. On day 4, directed library construction was completed and entered into Illumina sequencing.

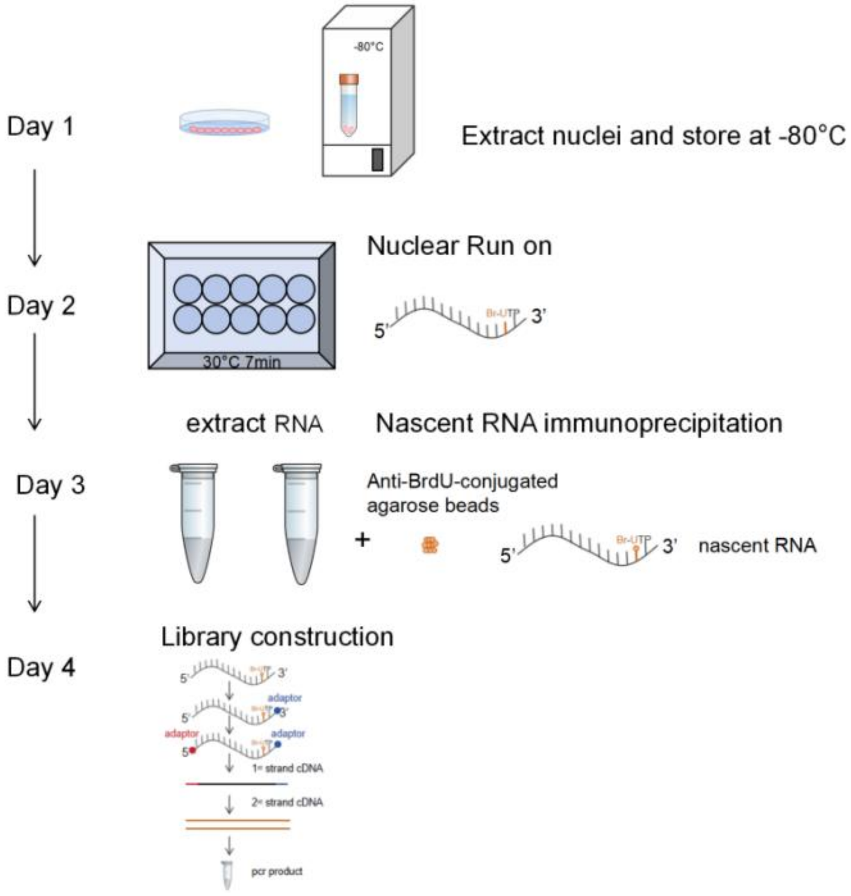

Figure 1: Schematic diagram of the experimental flow of ADAMTS5e related GRO-seq

The above process converts the raw reads generated on the experimental side into a computable timing transcription signal, which is then aligned, filtered and annotated to generate a candidate genomic feature matrix. Different from using only static expression, this pathway retains information related to transcription initiation, elongation and pause, providing more fine-grained dynamic input for subsequent risk scoring. After sample acquisition, the genomic branch mainly constructs representations for indicators such as TSS neighborhood density, differential transcripts, functional enrichment characteristics, key site coverage and pause index, and the clinical branch carries on patient staging, pathological indicators, treatment records and follow-up outcomes. The two enter the fusion model under a unified numbering system.

The data base thus formed not only maintains the biological authenticity of ADAMTS5e experimental results, but also meets the requirements of standardized, reproducible and scalable input structures for deep learning prognostic modeling. At the same time, this process provides a unified data starting point and input interface for subsequent sequencing quality assessment, groHMM transcription unit identification and cross-modal feature alignment, and facilitates the stable reproduction of experiments.

4.2 Quality assessment of sequencing data

This section presents a quality assessment around the reliability of genomic branch inputs. Based on the GRO-seq results of Huh7 Mock and ADAMTS5e overexpressing cells, the average quality distribution of the original reads on each sequencing cycle was calculated, and the quality curve was used as the basis for the first screening before subsequent feature coding. This step not only serves the usability judgment at the experimental level, but also directly relates to the stability of transcription start site signal, differential transcript identification and pause index calculation. If the base quality fluctuates too much, sequence alignment errors, coverage shifts, and local enrichment distortions are further amplified in the multimodal fusion stage, thus weakening the alignment accuracy between clinical variables and genomic signals.

As shown in Fig. 2, the average mass value of Mock samples in the full read range basically maintained between Q37.4 and Q39.3, and the first 20 base positions were Q38.8, then slowly decreased to Q37.7 around the 140/150 position, and briefly rose to about Q39.1 after crossing the boundary. The end of the read length remains above Q37.4. The overall fluctuation range of the curve was 1.9 mass units, and there was no sudden collapse or continuous attenuation, indicating that the quality of the control group library was stable, which could be used as an effective reference for subsequent difference comparison.

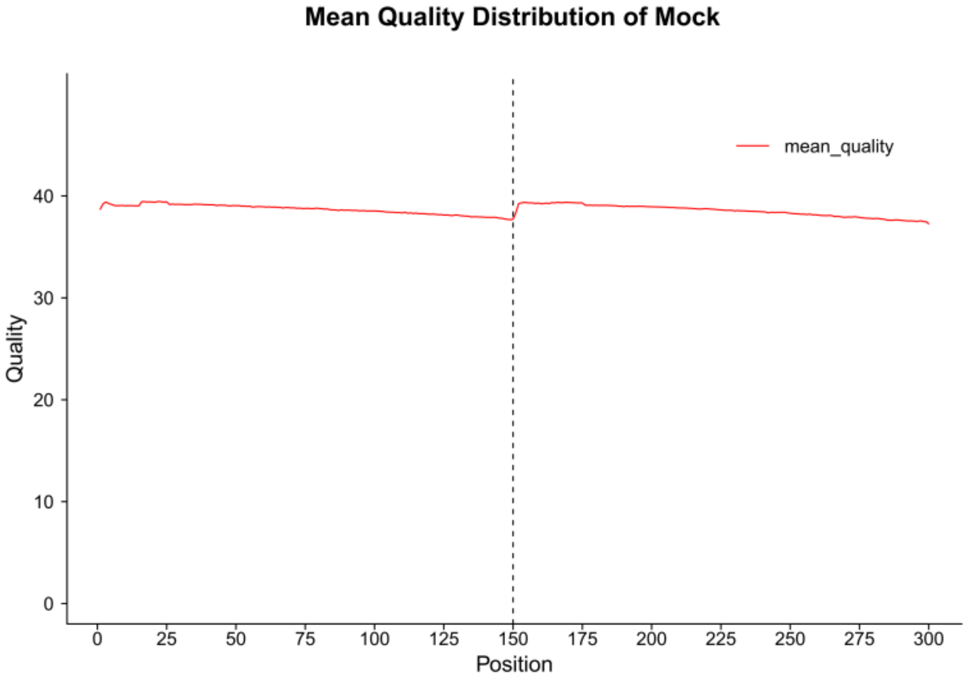

Figure 2: Sequencing quality profile of Mock samples

As shown in Fig. 3, the average mass distribution of OE samples is highly close to that of the Mock group, which is in the range Q37.3 to Q39.2 as a whole. The mass value of the starting region was Q38.9, decreased to Q37.8 near the 150th site, then recovered to around Q39.0, and maintained at Q37.3 at the end. The difference between the two sets of curves at key positions was mostly less than 0.2, indicating that ADAMTS5e overexpression did not introduce systematic bias caused by sample preparation or sequencing processes. For computational modeling, this consistency means that the subsequent changes of transcript unit boundaries, TSS neighborhood density and gene body coverage identified by groHMM are more comparable, and the dynamic genomic features in the model input can be encoded under a uniform noise level.

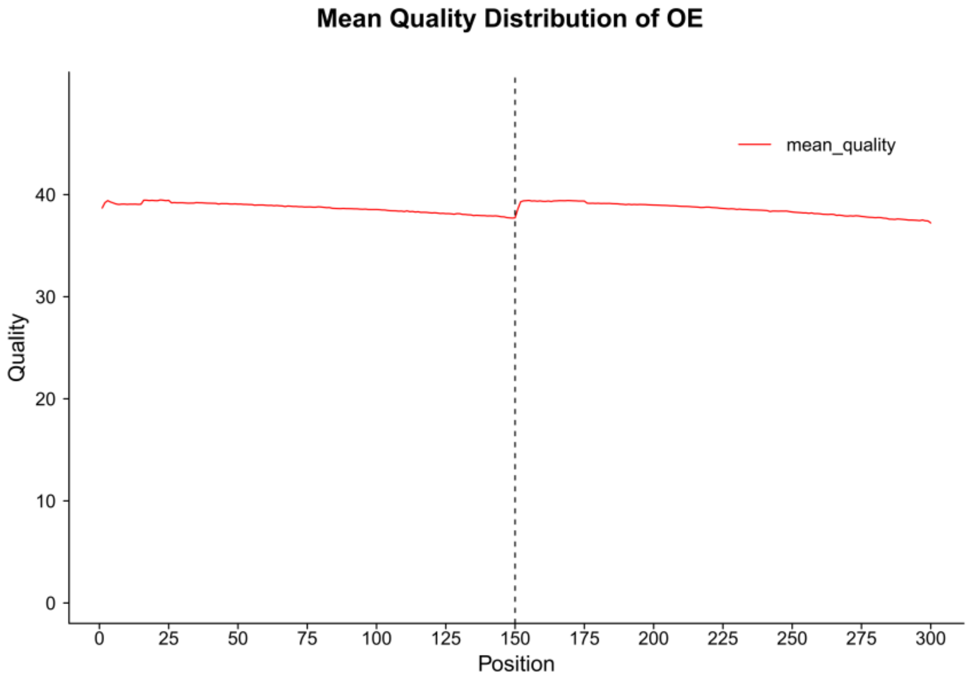

Figure 3: Sequencing quality profile of OE samples

Combining the two sets of results, the current GRO-seq data has met the input quality requirements of the subsequent deep learning prognostic risk assessment framework. The average quality of Mock and OE samples is stable at about Q38. Although the middle and posterior segments show a slight decrease, they do not enter the low confidence interval, and the two groups of curves are consistent in shape and amplitude. This result provides a reliable starting point for sequence alignment, signal normalization, differential feature screening, and cross-modal fusion. It also ensures that the genomic branch has a reproducible data basis when jointly training with clinical features, and reduces the error accumulation in the cross-modal coding stage.

4.3 Analysis of signal distribution characteristics near the transcription start site

This section centers the analysis around the signal distribution near the transcription start site and uses it as an important criterion for the reliability of genomic branch input. GRO-seq can directly record the status of polymerase occupancy during nascent transcription, so the density change of TSS neighborhood not only reflects the transcriptional activity of the proximal promoter, but also determines whether the subsequent feature coding can stably distinguish the initiation, elongation and pause events. For the fusion model, if the signals near the TSS lack boundary and hierarchy, the joint representation between clinical variables and genomic variables will be dragged by noise, and the risk score output is difficult to form an interpretable hierarchical structure.

As shown in Fig. 4, the transcript density identified by groHMM was significantly higher in the 0 to 1000 bp interval downstream of the TSS than in the upstream region, indicating that the main signal was concentrated in the gene body initiation segment rather than disordered diffusion. In the figure, the ideal framework defined 0 to 1000 bp as the Transcriptional Part, and the corresponding density main peak was maintained above 0.70, and the TP value recorded by the experiment was 0.7197. A weak but continuous FivePrimeFP signal was retained in the upstream region of the TSS with a statistical value of 0.4125, indicating that there was still identifiable promoter proximal polymerase occupancy under the ADAMTS5e overexpression system. At the same time, the Post-TTS and TUA regions reached 0.4872 and 0.5972, respectively, suggesting that the nascent transcription data did not just stay at the local peak, but maintained the spatial continuity consistent with the real transcription elongation process.

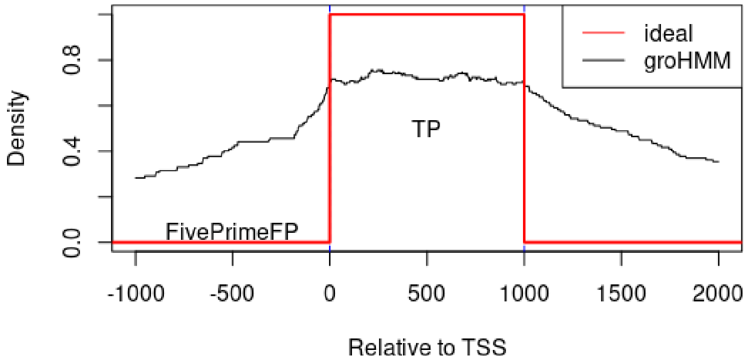

Figure 4: Map of signal density around transcription start sites

From a computational modeling perspective, this result has two implications. First, a clear density gradient is formed between the TSS neighborhood and the gene body region, which indicates that the promoter signal, gene body signal and upstream background signal can be encoded as independent feature channels to reduce the information aliasing caused by simple cover summation. Second, the FivePrimeFP did not increase abnormally, while the TP region remained stable and enriched, indicating that ADAMTS5E-related de - crept transcription is closer to structural remodeling than random perturbation, which provides higher confidence for the risk model to extract dynamic genomic representation. Combined with the above Q40-level sequencing quality results, the current sample has a computational foundation to support TSS neighborhood modeling, differential transcript screening, and joint input of clinical features, and provides a stable data interface and training constraints for subsequent cross-modal alignment in deep learning cancer prognostic risk assessment.

4.4 Analysis of expression changes of differential transcripts

This section provides a unified assessment of de novo transcriptional differences between Mock and ADAMTS5e overexpression samples based on the transcription units identified by groHMM, and the results are used as genomic candidate input for subsequent tumor prognostic risk models. The MA plot takes the average expression logCPM as the horizontal axis and log2 fold change as the vertical axis, which can simultaneously observe the coupling relationship between expression intensity and change direction. For fusion modeling, this step is not simply screening significant genes, but is used to identify dynamic molecular features that can stably characterize ADAMTS5E-driven transcriptional remodeling, avoiding sending low coverage, low consistency or random perturbation signals directly into the deep learning branch, which affects the quality of joint representation between clinical variables and genomic variables.

As shown in Fig. 5, the differential transcripts were mainly concentrated in the logCPM −4 to 10 range, and most genes were distributed around the zero line, but a clear up-down differentiation occurred in the low to moderate expression regions. A total of 312 significantly up-regulated genes and 145 significantly down-regulated genes were detected in the experimental records, and the number of up-regulated genes was 2.15 times that of down-regulated genes, indicating that the immediate response after ADAMTS5e overexpression was dominated by transcriptional activation. In the figure, LINC02685, HCN4, MAGI2-AS3 and

LINC02171 were located in the up-regulated region, and the log2 fold change was mostly higher than 1. FGF13, LINC01942, and MIR548M were located in the down-regulated region, and some transcripts showed a decrease close to −6. LOC100507412 was in the middle and high expression band and showed a significant upward shift, indicating that this site change was not a local isolated phenomenon, but an integral part of ADAMTS5E-related transcriptional reprogramming.

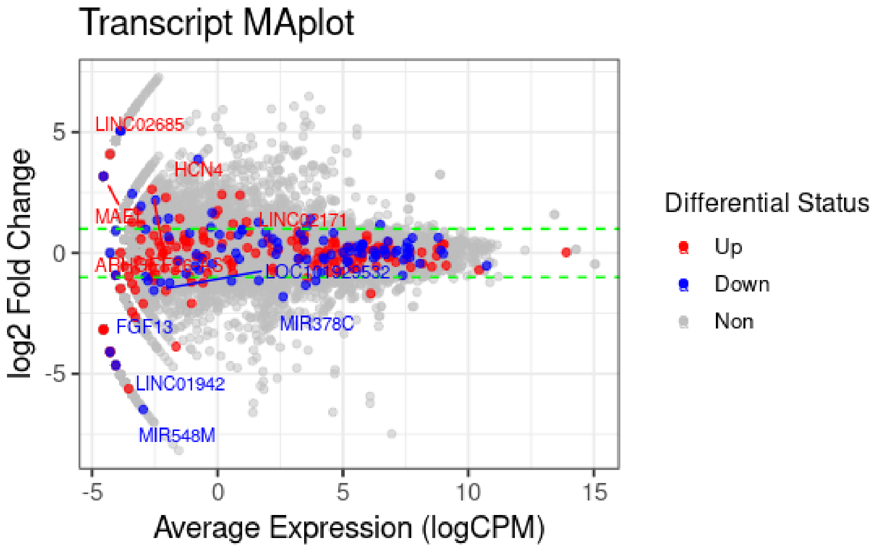

Figure 5: MA plot of differential transcripts

From the computational perspective, the MA distribution presents two characteristics. One is that significant differences mainly occur in the set of transcripts with a certain expression basis, indicating that the subsequent feature matrix can preserve statistical significance while maintaining numerical stability. Second, upregulated transcripts were dominant in both number and magnitude, suggesting that genome branches are more suitable for direction-sensitive coding rather than compression by absolute change only. Based on this result, the expression intensity, change direction and significance of differential transcripts will be jointly mapped into the joint input of the subsequent model to enhance the responsiveness of the tumor prognostic risk score to dynamic transcriptional remodeling and provide a candidate set for subsequent functional enrichment analysis and key site validation.

4.5 Functional enrichment analysis of up-regulated genes

In this section, functional enrichment analysis of up-regulated genes was performed to determine the functional range corresponding to de novo transcriptional activation after ADAMTS5e overexpression and to provide compressible functional layer features for subsequent tumor prognostic risk assessment models. GO enrichment can map discrete gene changes into functional modules, thereby reducing the redundancy of high-dimensional genomic input and enhancing the interpretability of the model in joint learning with clinical variables. The previous experimental results showed that among the 312 significantly up-regulated genes, the enrichment results mainly focused on the direction of apoptosis regulation, cell junction remodeling and receptor inhibitory activity.

As shown in Fig. 6, The top entries at the BP level include negative regulation of execution phase of apoptosis, execution phase of apoptosis, and regulation of execution phase of apoptosis and cell-cell junction assembly, etc., indicate that ADAMTS5E-related transcriptional activation is not disordered enhancement, but directional rearrangement around the process of apoptosis execution and intercellular adhesion structure.

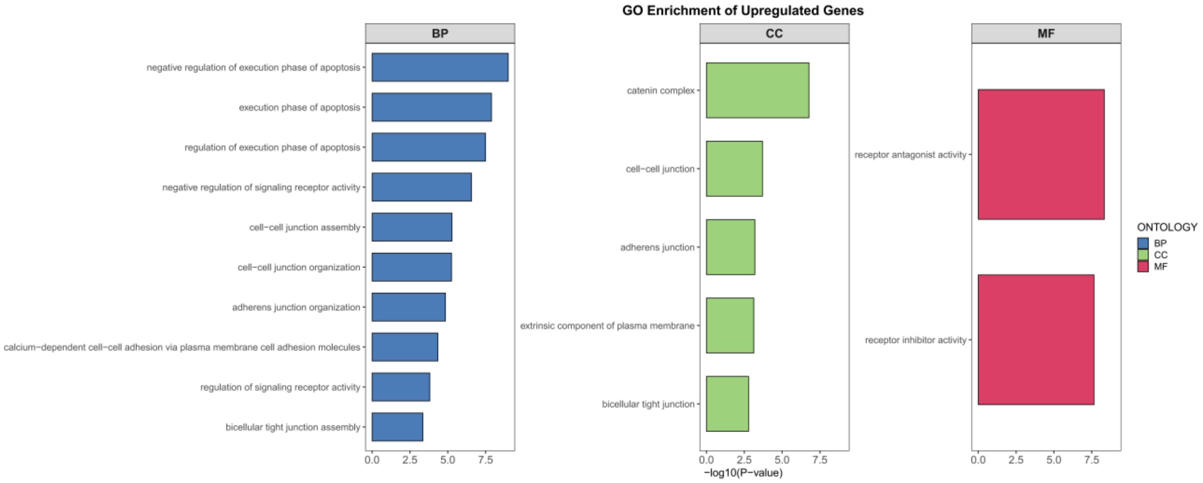

Figure 6: GO enrichment analysis of up-regulated genes

At the CC level, catenin complex, cell-cell junction, adherens junction, extrinsic component of plasma membrane, and bicellular tight All junction showed enrichment, showing that the up-regulated genes mainly fell in the membrane structure, junctional complex and adhesion interface regions. Two significant entries of receptor antagonist activity and receptor inhibitor activity were retained at the MF level, indicating that receptor inhibitory regulation was the concentrated output direction in this transcriptional remodeling. Combined with the −log10(P-value) distribution in the figure, it can be seen that the significance of BP and MF entries is higher as a whole, some entries are close to or more than 7, and CC entries are mostly distributed between 2 and 6, indicating that the current set of up-regulated genes has a clear hierarchical structure in functional output, rather than evenly scattered in multiple weakly correlated pathways.

From the perspective of computational modeling, this set of GO results can encode three functional labels of apoptosis regulation, cell connectivity and receptor inhibition as prior features, which together with TSS neighborhood density, differential transcripts and key site activity constitute the genome branch input. At the same time, it provides a controllable molecular interpretation interface for stage, metastasis status, and treatment response variables in the clinical branch, and provides a more stable computational basis for risk weight assignment in subsequent cross-modal fusion. Therefore, the functional enrichment results are not only a supplement at the biological interpretation level, but also an important intermediate bridge information for deep learning prognostic risk assessment models to perform feature compression, semantic clustering and joint representation learning.

4.6 Analysis of transcriptional activity changes at key gene loci

This section focuses on the analysis of changes in transcriptional activity at representative differential loci of chr21 and considers local track signals as key evidence to verify the reliability of genomic candidate features. Different from the global difference results provided by the MA map, the IGV tracks were able to simultaneously present the GRO-seq coverage changes under Mock and OE conditions at the single gene scale, and were compared with active chromatin marks such as H3K27ac, H3K4me1 and H3K4me3 in parallel. Therefore, it is more suitable to determine whether the immediate transcriptional remodeling after ADAMTS5e overexpression is site-specific. The previous experimental results showed that LOC100507412 on chr21 was a representative up-regulated gene, and RUNX1 was a representative down-regulated gene.

As shown in Fig. 7, the RUNX1 window is roughly between 34.80 Mb and 35.10 Mb, covering an interval of 300 kb. In this region, the main signal peaks of the positive strand track of the Mock group were concentrated near the RUNX1 promoter and the front segment of the gene body, and the peak value was close to the upper limit of the track 100. The positive chain signal at the corresponding position of the OE group was significantly weakened, and the local peak height was only maintained between 40 and 60, with a decrease of nearly half. The overall fluctuation of the reverse strand track was low and no continuous enrichment bands comparable to the positive strand were formed, indicating that the main changes at this site were focused on weakened transcriptional output rather than strand switching. At the same time, the range of H3K27ac, H3K4me1 and H3K4me3 histone modification tracks were set from 0 to 30, the main peak of H3K4me3 was located at the proximal end of the promoter, and H3K27ac and H3K4me1 were scattered, indicating that RUNX1 retained the promoter active mark after the overexpression of ADAMTS5e. However, the de novo transcriptional coverage has been substantially downregulated.

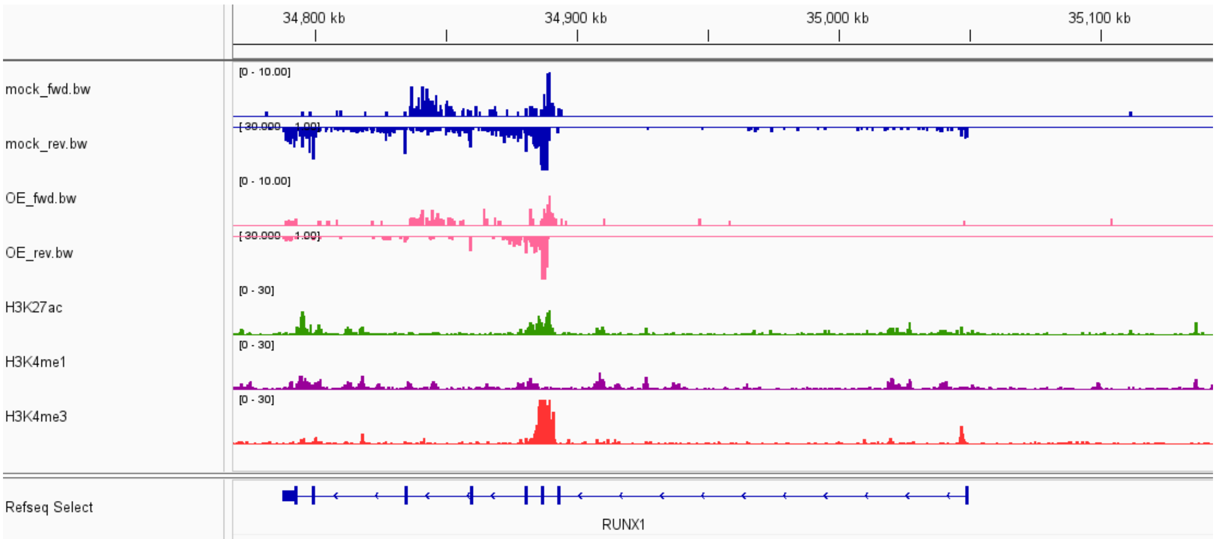

Figure 7: Track diagram of transcriptional activity at the RUNX1 locus

As shown in Fig. 8, the window of LOC100507412 ranges from 8.38 Mb to 8.48 Mb, covering an interval of 100 kb, and the signal change is more concentrated than that of RUNX1. In this locus, the positive strand track of the OE group formed a continuous enhancement band, and multiple peaks were close to the upper limit of 500, covering from the vicinity of the promoter to the inside of the gene body. The corresponding signal of the Mock group is weaker, with most segments below 200. The reverse chain track also shows a more pronounced negative coverage in the OE condition, with amplitude roughly reaching −200 to −300, while the fluctuations in the Mock group are relatively limited. The range of histone modification tracks was set from 0 to 80, in which H3K27ac and H3K4me1 maintained a high background in the region near LOC100507412, and H3K4me3 formed a clear peak near the promoter, indicating that this site had the characteristics of active promoter and enhancer at the same time. The previous analysis also pointed out that the GRO-seq signal of LOC100507412 was significantly enhanced near the TSS and in the gene body region in OE cells, while the overall coverage of RUNX1 was significantly reduced in OE cells.

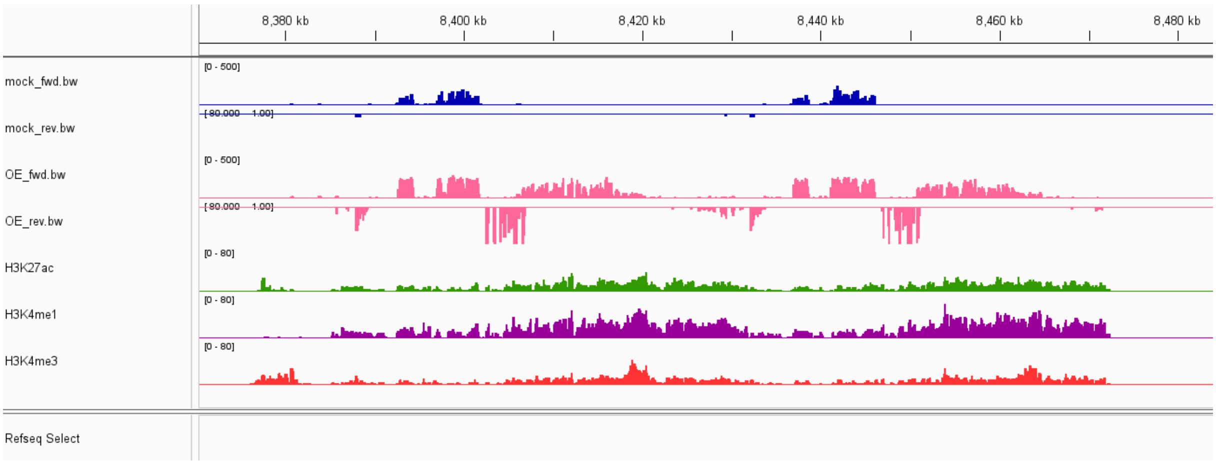

Figure 8: Track diagram of transcriptional activity at LOC100507412

These two types of local site changes in opposite directions have direct value from a computational modeling perspective. The track level evidence shows that the differential genes are not only from statistical threshold screening, but structural transcriptional changes with visual support, which can be used as high confidence nodes in subsequent feature selection. The spatial synergistic relationship between GRO-seq coverage and histone modification signals also provides local priors for constructing the joint representation of "transcription intensity-chromatin activity-clinical phenotype". Therefore, RUNX1 and LOC100507412 not only assume the function of biological validation, but also provide interpretable key site inputs for deep learning prognostic risk assessment models.

4.7 Analysis of changes in RNA Pol II pause index

This section focuses on the analysis of RNA Pol II pause index changes and uses them as an important basis for determining whether ADAMTS5e overexpression changes the transcriptional dynamics at the proximal end of the promoter. The pause index is expressed as the ratio of reads density between the promoter region and the gene body region, which can describe the relative state of the polymerase transition from pause to elongation after transcription initiation. For prognostic modeling, this index is used to determine whether ADAMTS5E-related de novo transcriptional remodeling is represented by global pause rearrangement or directional regulation at the local site and differential transcript level. Therefore, when the overall distribution remains stable, the pause index is more suitable as an auxiliary constraint feature to enter the fusion model together with the TSS neighborhood density, differential transcripts, and key site signals.

As shown in Fig. 9, the main body of log10(pausing index) distribution of Mock group and OE group were both in the negative range, and the two violin profiles overlapped around the median, indicating that the relative proportion of RNA Pol II between proximal promoter pause and gene body extension did not systematically shift after ADAMTS5e overexpression. In the figure, there are 54 genes with high pause and 80 genes with low pause, and the scatter is mainly concentrated between −1.2 and 0.5, but both groups retain a small number of outliers with high values, the highest being close to 2.0. In terms of box line position, the median of the Mock group was −0.78, and that of the OE group was −0.86. The interquartile range difference was very small, and the lower boundary of the OE group slightly extended to the lower value, suggesting that the pause degree of some genes was weakened under the overexpression condition, but this change was not enough to promote the overall migration of the global distribution. The previous experimental analysis pointed out that the pausing index of genes in

OE cells did not change significantly compared with Mock cells.

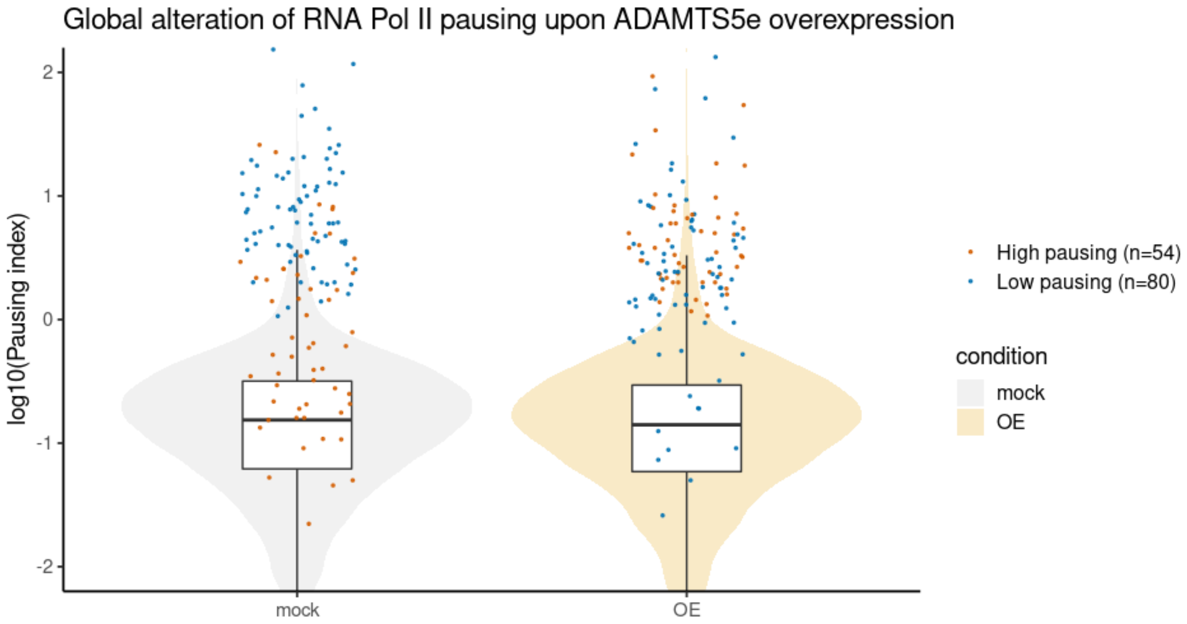

Figure 9: Comparison plot of the exponential distribution of RNA Pol II pauses

This result makes sense from a computational modeling point of view. The extensive transcriptional changes induced by ADAMTS5e overexpression were reflected by increased up-regulated genes, maintained enrichment of TSS neighborhood signals, and differential coverage of critical sites, rather than driving transcriptional output through comprehensive alterations in the pausing release mechanism. Therefore, in the deep learning tumor prognosis risk assessment model, the pause index is more suitable as a steady-state constraint input, rather than as a dominant branch to amplify alone. This conclusion is helpful to maintain the balance of feature weights within the genome branches and improve the representation stability when the model is jointly learned with clinical variables.

5 Conclusion

This paper focuses on the construction of a deep learning tumor prognostic risk assessment model that fuses genomic and clinical data, and completes the complete technical link from dynamic transcription feature extraction, clinical variable representation, multimodal fusion to risk score output. After the Mock and overexpression system based on Huh7 cells obtained the de vivo transcript data, the overall sequencing quality was maintained at Q40. The transcript unit identification results showed that TP was 0.7197, FivePrimeFP was 0.4125, and 312 up-regulated genes and 145 down-regulated genes were detected. Local coverage changes at key sites such as RUNX1 and LOC100507412 provide interpretable high-confidence features for model inputs, while the global pause index remains stable, indicating that the risk-related signal mainly results from directed transcriptional remodeling rather than global pause imbalance. In the computational implementation, the hierarchical coding uniformly maps the TSS neighborhood density, differential transcripts, functional enrichment labels, key site activity and clinical stage, and treatment records into a trainable representation space, so that omics signals can participate in patient-level prognosis modeling in a stable and low redundancy manner, and retain the subsequent replication and deployment interface. The limitations of this paper are that the sample source is still focused on a single cell system, the clinical branch still lacks external verification of large-scale follow-up cohorts, and the ability to migrate across cancer species and time-dependent calibration still need to be enhanced. Future studies will introduce multi-center clinical samples, pathological phenotypes and treatment response information, and combine graph representation learning, domain adaptation and survival calibration strategies to further improve the generalization ability, interpretability and deployment efficiency of the model.

References

[1] Schneider L, Laiouar-Pedari S, Kuntz S, et al. Integration of deep learning-based image analysis and genomic data in cancer pathology: A systematic review[J]. European journal of cancer, 2022, 160: 80-91.

[2] Stahlschmidt S R, Ulfenborg B, Synnergren J. Multimodal deep learning for biomedical data fusion: a review[J]. Briefings in bioinformatics, 2022, 23(2): bbab569.

[3] Wang S, Zhang H, Liu Z, et al. A novel deep learning method to predict lung cancer long-term survival with biological knowledge incorporated gene expression images and clinical data[J]. Frontiers in Genetics, 2022, 13: 800853.

[4] Ma B, Yan G, Chai B, et al. XGBLC: an improved survival prediction model based on XGBoost[J]. Bioinformatics, 2022, 38(2): 410-418.

[5] Bashiri A, Ghazisaeedi M, Safdari R, et al. Improving the prediction of survival in cancer patients by using machine learning techniques: experience of gene expression data: a narrative review[J]. Iranian journal of public health, 2017, 46(2): 165.

[6] Li R, Wu X, Li A, et al. HFBSurv: hierarchical multimodal fusion with factorized bilinear models for cancer survival prediction[J]. Bioinformatics, 2022, 38(9): 2587-2594.

[7] Zhang C, Chen Y, Zeng T, et al. Deep latent space fusion for adaptive representation of heterogeneous multi-omics data[J]. Briefings in Bioinformatics, 2022, 23(2): bbab600.

[8] Zhang Y, Kiryu H. MODEC: an unsupervised clustering method integrating omics data for identifying cancer subtypes[J]. Briefings in Bioinformatics, 2022, 23(6): bbac372.

[9] Hu J, Lv H, Zhao S, et al. Prediction of clinicopathological features, multi-omics events and prognosis based on digital pathology and deep learning in HR+/HER2− breast cancer[J]. Journal of thoracic disease, 2023, 15(5): 2528.

[10] Li W, Huang Q, Peng Y, et al. A deep learning approach based on multi-omics data integration to construct a risk stratification prediction model for skin cutaneous melanoma[J]. Journal of Cancer Research and Clinical Oncology, 2023, 149(17): 15923-15938.

[11] Wen G, Li L. FGCNSurv: dually fused graph convolutional network for multi-omics survival prediction[J]. Bioinformatics, 2023, 39(8): btad472.

[12] Salimy S, Lanjanian H, Abbasi K, et al. A deep learning-based framework for predicting

survival-associated groups in colon cancer by integrating multi-omics and clinical data[J]. Heliyon, 2023, 9(7).

[13] Ballard J L, Wang Z, Li W, et al. Deep learning-based approaches for multi-omics data integration and analysis[J]. BioData Mining, 2024, 17(1): 38.

[14] Steyaert S, Qiu Y L, Zheng Y, et al. Multimodal deep learning to predict prognosis in adult and pediatric brain tumors[J]. Communications Medicine, 2023, 3(1): 44.

[15] Wang Y, Zhang L, Li Y, et al. Predicting the prognosis of HER2-positive breast cancer patients by fusing pathological whole slide images and clinical features using multiple instance learning[J]. Math. Biosci. Eng, 2023, 20(6): 11196-11211.

[16] Wang Z, Lin R, Li Y, et al. Deep learning-based multi-modal data integration enhancing breast cancer disease-free survival prediction[J]. Precision clinical medicine, 2024, 7(2): pbae012.

[17] Yang S, Wang Z, Wang C, et al. Comparative evaluation of machine learning models for subtyping triple-negative breast cancer: a deep learning-based multi-omics data integration approach[J]. Journal of Cancer, 2024, 15(12): 3943.

[18] Lan W, Liao H, Chen Q, et al. DeepKEGG: a multi-omics data integration framework with biological insights for cancer recurrence prediction and biomarker discovery[J]. Briefings in bioinformatics, 2024, 25(3): bbae185.

[19] Tanvir R B, Islam M M, Sobhan M, et al. MOGAT: a multi-omics integration framework using graph attention networks for cancer subtype prediction[J]. International Journal of Molecular Sciences, 2024, 25(5): 2788.

[20] Wang H, Han X, Niu S, et al. DFASGCNS: A prognostic model for ovarian cancer prediction based on dual fusion channels and stacked graph convolution[J]. Plos one, 2024, 19(12): e0315924.