

RESEARCH ARTICLE

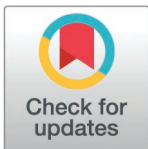
Structure-aware deep learning enhances m6A prediction and reveals cell type-associated RNA structural signatures

Mingze Sun¹, Di Zhang²*, Zhiyuan Li³*, Yihan Lin^{2,3,4}*

1 Yingcai Honors College, University of Electronic Science and Technology of China, Chengdu, Sichuan, China, **2** Peking University Chengdu Academy for Advanced Interdisciplinary Biotechnologies, Chengdu, Sichuan, China, **3** Center for Quantitative Biology and Peking-Tsinghua Center for Life Sciences, Academy for Advanced Interdisciplinary Studies, Peking University, Beijing, China, **4** The MOE Key Laboratory of Cell Proliferation and Differentiation, School of Life Sciences, Peking University, Beijing, China

* These authors contributed equally to this work.

* zhangdi@aaib.pku.edu.cn (DZ), zhiyuanli@pku.edu.cn (ZL), yihan.lin@pku.edu.cn (YL)



OPEN ACCESS

Citation: Sun M, Zhang D, Li Z, Lin Y (2026) Structure-aware deep learning enhances m6A prediction and reveals cell type-associated RNA structural signatures. PLoS Comput Biol 22(8): e1014649. <https://doi.org/10.1371/journal.pcbi.1014649>

Editor: Marc R Birtwistle, Clemson University, UNITED STATES OF AMERICA

Received: March 10, 2026

Accepted: July 31, 2026

Published: August 14, 2026

Peer Review History: PLOS recognizes the benefits of transparency in the peer review process; therefore, we enable the publication of all of the content of peer review and author responses alongside final, published articles. The editorial history of this article is available here: <https://doi.org/10.1371/journal.pcbi.1014649>

Copyright: © 2026 Sun et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution,

Abstract

N6-methyladenosine (m6A), the most abundant mRNA modification in eukaryotes, plays essential roles in gene regulation and disease pathogenesis. Computational prediction of m6A sites offers a scalable alternative to costly experimental approaches, yet current methods rely predominantly on linear sequence features. This overlooks potentially informative RNA structural context, which is associated with local methylation patterns and may provide complementary predictive information beyond linear sequence motifs. To incorporate this complementary information, we propose SMART-m6A (Sequence–structure Multifeature Attention RNA Transformer for m6A), a deep learning framework that integrates sequence and structural information through parallel convolutional feature extraction and structure-guided attention for multifeature fusion. SMART-m6A achieves superior predictive performance compared to existing methods, with particularly clear advantages in sequence-ambiguous candidates. Beyond prediction accuracy, learned attention patterns reveal strong concordance with experimentally validated m6A-binding protein recognition sites and identify potentially novel regulatory motifs. Through systematic ablation studies and targeted structural-input perturbation analyses, we show that sequence and structure provide complementary predictive information, and that sites with greater prediction sensitivity to structural perturbation exhibit distinct local structural profiles between cell lines. Collectively, this work demonstrates the predictive value of sequence-derived structural features in m6A modeling and provides a multifeature deep learning framework for accurate and interpretable structure-aware epitranscriptomic prediction.

and reproduction in any medium, provided the original author and source are credited.

Data availability statement: This study did not generate any new experimental data. All datasets analyzed in this study are publicly available from previously published sources, as described in the manuscript. The processed data and materials necessary to reproduce the results are available at OSF: <https://osf.io/qd9xh/>. The source code is available at GitHub: <https://github.com/mzsun-ai/SMART-m6A>.

Funding: This work was supported by Sichuan Science and Technology Program (2025ZNSFSC0993 to D.Z.), the National Natural Science Foundation of China (T2325002, 32088101, T2321001, 32550512, and 92570207 to Y.L.), the National Key R&D Program of China (2020YFA0906900 and 2018YFA0900703 to Y.L.; 2024YFA0919500 to Z.L.), and the Ministry of Education of China (JYB2025XDXM502 to Z.L.). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: I have read the journal's policy and the authors of this manuscript have the following competing interests: The authors declare that a patent application related to this work has been filed (application number: 2026103796665).

Author summary

Chemical modifications of RNA provide a crucial regulatory layer beyond the genetic code. Among them, N6-methyladenosine (m6A) is the most abundant internal modification in eukaryotic messenger RNA and has been implicated in diverse biological processes, including development, immune regulation, and tumorigenesis. Understanding how m6A is deposited across transcripts is therefore essential for deciphering epitranscriptomic regulation and its roles in health and disease. Although high-throughput technologies can map m6A sites genome-wide, these experiments remain expensive and technically demanding, motivating the development of computational prediction methods. However, most existing algorithms rely primarily on linear RNA sequence features and largely ignore RNA secondary structure, which has been associated with RNA-binding protein recognition and m6A-related regulatory contexts. Here, we introduce SMART-m6A, a structure-aware deep learning framework that integrates RNA sequence and secondary structure through a multifeature neural architecture with structure-guided attention. By jointly modeling these complementary sources of information, SMART-m6A achieves improved predictive accuracy across multiple benchmark datasets. Importantly, the model is not only predictive but also biologically informative: it identifies structural features associated with m6A deposition, reveals that cell type-specific m6A sites exhibit greater structural variability, and uncovers sequence motifs linked to known RNA-binding proteins. Together, our results indicate that sequence-derived structural features are informative for m6A prediction and highlight the value of structure-aware machine learning approaches for understanding epitranscriptomic mechanisms.

Introduction

Chemical modifications of RNA represent a crucial layer of post-transcriptional regulation, with more than 160 distinct types identified to date [1–3]. Among them, N6-methyladenosine (m6A) is a modification of mammalian mRNA with essential biological functions, playing pivotal roles in post-transcriptional control [4–9]. The deposition of m6A is dynamically reversible [10]: it is installed by the methyltransferase complex (“writers,” e.g., METTL3/METTL14 [11]/WTAP [12]), removed by demethylases (“erasers,” e.g., FTO [5] and ALKBH5 [13]), and interpreted by a variety of binding proteins (“readers,” including YTH family proteins [14] and hnRNPs [15]). Through these effectors, m6A influences gene expression across multiple steps from transcription to translation, including splicing, nuclear export, translation, and RNA stability [16]. Functionally, m6A serves as a key regulatory node in diverse biological processes and disease phenotypes, such as embryonic development, immune regulation, and tumorigenesis [17]. Advances in next-generation sequencing (NGS) technologies have enabled transcriptome-wide detection of m6A through multiple approaches, including antibody-based methods such as MeRIP-seq [18] and miCLIP

[19], and enzyme- or chemistry-based strategies such as MAZTER-seq [20] and m6A-SEAL [21]. In addition, third-generation sequencing methods, such as Nanopore direct RNA sequencing, allow the direct identification of RNA modifications from electrical current signals [22]. More recently, Sun *et al.* introduced the GLORI method, which integrates mild chemical reactions with NGS readouts to achieve absolute quantification of m6A under low-input conditions, thereby opening new possibilities for studies in scarce samples [23]. Despite these advances, the widespread application of m6A profiling remains constrained by the high cost and technical complexity of existing methods.

Meanwhile, computational prediction methods have increasingly emerged as a valuable complement to experimental detection, enabling large-scale predictions within a short timeframe. Early approaches were largely based on traditional machine learning [24]; however, given the intricate interactions within RNA sequences, handcrafted features may not fully capture the underlying rules, which can limit their robustness across datasets. With the advent of deep learning, a growing number of powerful neural architectures have been developed, providing effective tools for modeling RNA sequences. For example, Zhang *et al.* proposed the DeepM6ASeq framework, which integrates CNNs [25] with BiLSTMs [26] to simultaneously capture local sequence motifs and contextual dependencies [27]. Xiong *et al.* developed the MASS framework, leveraging multitask learning and attention mechanisms to enable cross-species extraction and interpretation of m6A-related features [28]. Luo *et al.* introduced iM6A, which incorporates a ResNet [29] architecture to enhance feature representation [30]. More recently, Fan *et al.* presented deepSRAMP, which combines Transformer [31] and BiGRU [32] modules while integrating genomic positional information, thereby improving m6A prediction accuracy at the isoform level [33]. While these methods have significantly improved m6A prediction accuracy, current approaches face complementary challenges: sequence-only models may lack transcript-level regulatory context, whereas model incorporating positional distribution features, though effective on similar datasets, may inadvertently encode population-level statistics rather than causal sequence properties, potentially limiting robust generalization to novel transcripts or altered cellular contexts.

Recent studies indicate that m6A deposition is not solely determined by canonical sequence motifs such as DRACH [34], but is intricately modulated by the mutual influence between RNA secondary structure and m6A modification itself, likely through the regulation of RNA-binding protein accessibility and binding affinity [35,36]. Shachar *et al.* demonstrated *in vitro* that when RNA adopts a stable double-stranded conformation, the METTL3–METTL14 writer complex has limited accessibility for methylation, whereas relaxed single-stranded regions are more readily modified [37]. Wang *et al.* also reported that Pol II pausing can shape RNA secondary structure, thereby modulating the accessibility of m6A motifs to the methyltransferase complex and influencing subsequent methylation events [38]. Conversely, the m6A-switch model shows that m6A itself can remodel local RNA structure and alter the accessibility of RNA-binding protein motifs, providing evidence that m6A modification can also act upstream of local structural changes [39]. In addition, De La Cruz *et al.* showed that structurally flexible regions not only facilitate entry of writer complexes (such as METTL3–METTL14–WTAP) but also enhance the binding of reader proteins including YTHDF1/2/3 and regulatory RBPs such as IGF2 BPs and HuD [36]. Collectively, these findings suggest that sequence information alone is insufficient to fully explain the distribution of m6A. Incorporating RNA structural context into computational models may therefore improve m6A prediction and facilitate the identification of structure-associated patterns potentially related to RNA–protein interactions.

Here, we present SMART-m6A, a multifeature deep learning framework that explicitly integrates RNA secondary structure to advance m6A site prediction and uncover structure-associated regulatory signatures in epitranscriptomic patterns. By combining parallel convolutional feature extraction with structure-guided attention, SMART-m6A achieves consistently strong performance across multiple benchmarks and remains applicable when genomic positional annotations are unavailable, addressing a practical limitation in cross-transcript or less-annotated settings. Model interpretability reveals learned attention patterns that recapitulate experimentally validated m6A motifs and identify novel structural signatures enriched at modification sites. Comparative analysis of HEK293T and HeLa m6A landscapes further shows that cell line-specific sites exhibit greater local structural divergence than sites shared between the two cell lines. Model-based structural-input perturbation analysis identifies m6A candidates with differential prediction sensitivity to local structural

changes, with prediction-sensitive sites exhibiting greater structural variation in matched icSHAPE profiles. Together with the successful extension to quantitative modification-intensity prediction, these findings support the predictive and interpretive value of sequence-derived structural features and position SMART-m6A as a structure-aware framework for modeling m6A-associated patterns across cellular contexts.

Results

Overview of SMART-m6A

SMART-m6A enables joint modeling of RNA sequences and secondary structures, combined with a two-stage training strategy to enhance predictive performance and generalization. The model first employs a dual-branch convolutional module to capture sequence-level k-mer patterns and local structural neighborhood features, which are then integrated within the StructureTransformer encoder. Here, the attention mechanism learns long-range dependencies while explicitly incorporating structural biases, ensuring that global representations retain essential structural constraints. Through the pretraining–supervised learning paradigm, SMART-m6A not only outperforms existing methods across multiple benchmark datasets with accurate site-level identification, but also models the coupling between m6A and secondary structure across cell lines.

SMART-m6A accurately predicts m6A sites

We trained and evaluated SMART-m6A across six benchmark datasets, including four publicly available antibody-based datasets (SRAMP, DeepM6ASeq, m6ATCPred-Cerebellum7, and m6ATCPred-Lung4) and two datasets constructed in this study from publicly available GLORI-derived data (SMART-m6A and SMART-m6A-Position). Antibody-dependent detection methods are limited in site-level resolution and quantitative precision, whereas GLORI enables single-nucleotide resolution, bias-free, and absolute quantification of m6A via chemical reactions. Evaluating SMART-m6A on such high-quality data provides a stringent test of its predictive power and extends its potential applications to quantitative m6A analyses.

We first performed global feature analyses on five sequence-only datasets, including SRAMP, DeepM6ASeq, m6ATCPred-Cerebellum7, m6ATCPred-Lung4, and SMART-m6A. Minimum free energy (MFE) distributions calculated by ViennaRNA revealed no clear differences between m6A-containing and non-m6A fragments, and the distributional profiles were largely consistent across datasets (Fig 1a). Embeddings extracted from the pretrained model and visualized by t-SNE (Fig 1b) similarly showed no distinct clustering boundaries between positive and negative samples, nor substantial inter-dataset differences. These results indicate that common biophysical metrics or unsupervised learning alone are insufficient to distinguish m6A features, underscoring the necessity of supervised modeling.

To comprehensively assess performance, we compared SMART-m6A against seven representative methods, including four deep learning models (MASS, iM6A, DeepM6ASeq-RNN, and DeepM6ASeq-CNN) and three traditional machine learning models (XGBoost, RandomForest, and SVM). Across all five benchmark datasets, SMART-m6A achieved the best overall performance (Fig 1c and 1d and Tables 1–5). On the GLORI-based SMART-m6A dataset, our model reached an Accuracy of 0.7824, F1-score of 0.7949, AUROC of 0.8646, and AUPRC of 0.8521, outperforming competing approaches. Compared with sequence-only methods, SMART-m6A leveraged secondary structure constraints to resolve sites where sequence context alone was ambiguous, thereby reducing misclassification rates.

Previous studies have shown that m6A exhibits non-random transcriptome-wide distribution, typically enriched in the 3' UTR and near stop codons [2]. Thus, access to genomic positional information can further refine modification predictions. To evaluate this, we benchmarked SMART-m6A against deepSRAMP on the SMART-m6A-Position dataset (Fig 1e and Table 6). When genomic positional information is included, SMART-m6A and deepSRAMP achieve comparable performance. However, in the absence of genomic positional information, both methods

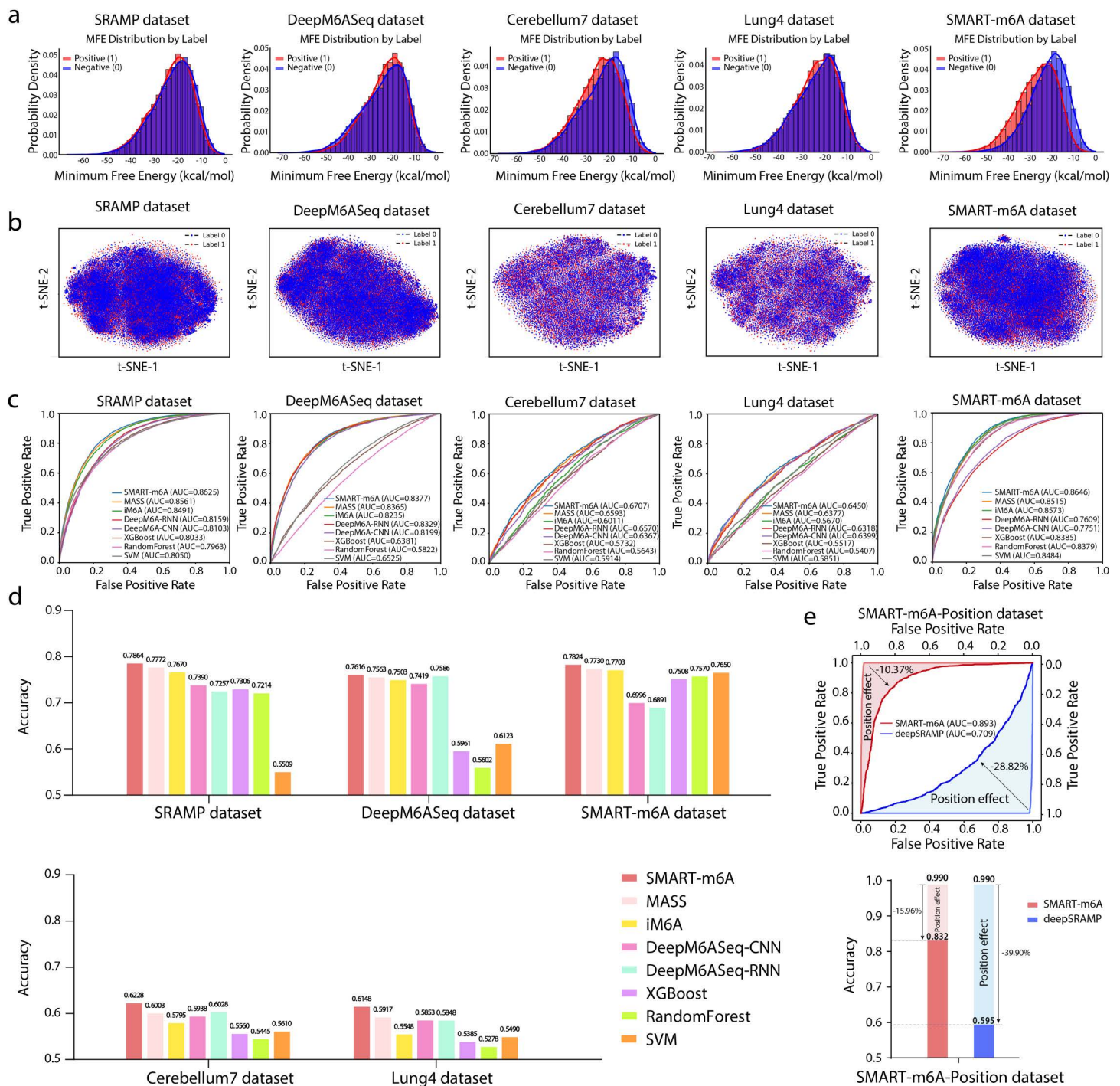


Fig 1. SMART-m6A leverages structural features and achieves superior performance across multiple benchmarks. (a) Distribution of minimum free energy (MFE) values calculated by ViennaRNA for positive (with m6A) and negative (without m6A) samples across five benchmark datasets (DeepM6ASeq, SRAMP, m6ATCPrd-Cerebellum7, m6ATCPrd-Lung4, and SMART-m6A). (b) t-SNE visualization of embeddings extracted from the pre-trained model across five benchmark datasets. (c) ROC curves comparing SMART-m6A against seven representative methods, including MASS, iM6A, DeepM6ASeq-CNN, DeepM6ASeq-RNN, XGBoost, RandomForest, and SVM, on the five benchmark datasets. (d) Accuracy comparison of SMART-m6A and baseline models across the five datasets. (e) SMART-m6A remains superior to deepSRAMP when positional information is unavailable.

<https://doi.org/10.1371/journal.pcbi.1014649.g001>

Table 1. Accuracy of 8 algorithms across 5 datasets.

Dataset	SMART-m6A	MASS	iM6A	DeepM6ASeq-CNN	DeepM6ASeq-RNN	XGBoost	RandomForest	SVM
SRAMP	0.7864	0.7772	0.7670	0.7390	0.7257	0.7306	0.7214	0.5509
DeepM6ASeq	0.7616	0.7563	0.7503	0.7419	0.7586	0.5961	0.5602	0.6123
Cerebellum7	0.6228	0.6003	0.5795	0.5938	0.6028	0.5560	0.5445	0.5610
Lung4	0.6148	0.5917	0.5548	0.5853	0.5848	0.5385	0.5278	0.5490
SMART-m6A	0.7824	0.7730	0.7703	0.6996	0.6891	0.7508	0.7570	0.7650

<https://doi.org/10.1371/journal.pcbi.1014649.t001>

Table 2. Recall of 8 algorithms across 5 datasets.

Dataset	SMART-m6A	MASS	iM6A	DeepM6ASeq-CNN	DeepM6ASeq-RNN	XGBoost	RandomForest	SVM
SRAMP	0.8124	0.8352	0.8419	0.7487	0.6511	0.7436	0.7523	0.1062
DeepM6ASeq	0.7969	0.7862	0.7785	0.6984	0.7975	0.6000	0.5860	0.6421
Cerebellum7	0.5745	0.4811	0.7175	0.7795	0.7525	0.5550	0.5870	0.5945
Lung4	0.5785	0.6880	0.5425	0.7455	0.7335	0.5540	0.5530	0.5990
SMART-m6A	0.8433	0.8664	0.7730	0.6955	0.7403	0.7785	0.8062	0.7820

<https://doi.org/10.1371/journal.pcbi.1014649.t002>

Table 3. F1-score of 8 algorithms across 5 datasets.

Dataset	SMART-m6A	MASS	iM6A	DeepM6ASeq-CNN	DeepM6ASeq-RNN	XGBoost	RandomForest	SVM
SRAMP	0.7918	0.7894	0.7833	0.7415	0.7036	0.7340	0.7298	0.1912
DeepM6ASeq	0.7697	0.7634	0.7572	0.7302	0.7677	0.5977	0.5713	0.6236
Cerebellum7	0.6036	0.5462	0.6305	0.6574	0.6545	0.5556	0.5631	0.5752
Lung4	0.6003	0.6276	0.5492	0.6425	0.6385	0.5455	0.5394	0.5705
SMART-m6A	0.7949	0.7924	0.7709	0.6984	0.7042	0.7575	0.7684	0.7689

<https://doi.org/10.1371/journal.pcbi.1014649.t003>

Table 4. AUROC of 8 algorithms across 5 datasets.

Dataset	SMART-m6A	MASS	iM6A	DeepM6ASeq-CNN	DeepM6ASeq-RNN	XGBoost	RandomForest	SVM
SRAMP	0.8625	0.8561	0.8491	0.8103	0.8159	0.8033	0.7963	0.8050
DeepM6ASeq	0.8377	0.8365	0.8235	0.8199	0.8329	0.6381	0.5822	0.6525
Cerebellum7	0.6707	0.6593	0.6011	0.6367	0.6570	0.5732	0.5643	0.5914
Lung4	0.6450	0.6377	0.5670	0.6399	0.6318	0.5517	0.5407	0.5851
SMART-m6A	0.8646	0.8515	0.8573	0.7751	0.7609	0.8385	0.8379	0.8484

<https://doi.org/10.1371/journal.pcbi.1014649.t004>

Table 5. AUPRC of 8 algorithms across 5 datasets.

Dataset	SMART-m6A	MASS	iM6A	DeepM6ASeq-CNN	DeepM6ASeq-RNN	XGBoost	RandomForest	SVM
SRAMP	0.8444	0.8382	0.8284	0.7774	0.7863	0.7837	0.7845	0.8058
DeepM6ASeq	0.8251	0.8243	0.7987	0.8009	0.8201	0.6206	0.5592	0.6219
Cerebellum7	0.6505	0.6483	0.5741	0.6202	0.6442	0.5569	0.5455	0.5654
Lung4	0.6468	0.6343	0.5558	0.6456	0.6282	0.5487	0.5383	0.5795
SMART-m6A	0.8521	0.8259	0.8428	0.7534	0.7401	0.8269	0.8185	0.8272

<https://doi.org/10.1371/journal.pcbi.1014649.t005>

Table 6. Performance on the SMART-m6A-Position dataset.

Model	Accuracy	Precision	Recall	F1-score	AUROC	AUPRC
SMART-m6A-Full	0.990	0.988	0.995	0.991	0.996	0.995
deepSRAMP-Full	0.990	0.984	0.999	0.991	0.996	0.996
SMART-m6A	0.832	0.823	0.900	0.860	0.893	0.889
deepSRAMP-single	0.595	0.588	0.983	0.736	0.709	0.744

Table 1–5 Comparison of classification performance metrics (Accuracy, Recall, F1-score, AUROC, AUPRC) of SMART-m6A and baseline models across multiple datasets. **Table 6** Comparative performance of SMART-m6A and deepSRAMP variants on the SMART-m6A-Position dataset across multiple evaluation metrics (Accuracy, Precision, Recall, F1-score, AUROC, and AUPRC).

<https://doi.org/10.1371/journal.pcbi.1014649.t006>

exhibit a degree of performance decline, with SMART-m6A showing a much smaller drop than deepSRAMP, demonstrating a clear advantage. Specifically, SMART-m6A experiences a 15.96% reduction in Accuracy (from 0.990 to 0.832) and a 10.34% reduction in AUROC (from 0.996 to 0.893), whereas deepSRAMP suffers a 39.90% reduction in Accuracy (from 0.990 to 0.595) and a 28.82% reduction in AUROC (from 0.996 to 0.709). Even under these conditions, SMART-m6A consistently outperforms deepSRAMP, achieving higher Accuracy (+0.237), Precision (+0.235), F1-score (+0.124), AUROC (+0.183), and AUPRC (+0.145). Considering that genomic positional information is often difficult to obtain in practice, whereas RNA secondary structure features can be computationally inferred from known sequences, the superior performance of SMART-m6A in the no-position setting further highlights its robustness and practical value in real-world applications.

Ablation experiments support the predictive contribution of structure-derived features

To systematically assess the relative contributions of sequence and structural features in SMART-m6A, we conducted ablation experiments in two settings: structural ablation and sequence ablation (**Fig 2a**). In the structural ablation setting, RNA sequences were preserved while loop annotations, dot-bracket structures, and base-pair probability (BPP) matrices were uniformly replaced with those from a fixed reference sample. In the sequence ablation setting, all input sequences were replaced with an identical reference sequence, while structural features were kept intact.

The results revealed that structural ablation led to a modest performance drop, with Accuracy decreasing by 1.4–4.1% and AUROC decreasing by 0.68–3.48% (**Fig 2b** and **2c**). By contrast, sequence ablation caused a substantial degradation in performance, with Accuracy decreasing by 10.2–24.9% to an average of 52.8%, and AUROC decreasing by 12.56–29.40% to an average of 0.5507. As an additional control, we also tested structure-misaligned and random-structure inputs while keeping the sequence inputs and labels unchanged (**S1 Fig**). Both controls performed similarly to the sequence-only model, whereas the full SMART-m6A model achieved higher ACC and AUROC. These findings indicate that sequence information remains the dominant predictive component, while correctly matched structural features provide complementary information that contributes to the overall predictive performance of SMART-m6A. We reasoned that the contribution of structural features might be underestimated when evaluated on the full test set, because many m6A candidates can be classified primarily from canonical sequence patterns. Therefore, we examined whether structural information provides greater benefit for samples with less confident sequence-only predictions. We defined sequence-only low-confidence samples as the 30% of test samples with the smallest prediction margins under the sequence-only model, representing candidates that were less clearly resolved by sequence information alone. In this subset, the benefit of adding structural features was more pronounced than that observed across all samples (**S2 Fig**). In the pooled analysis, the Accuracy gain increased from 0.032 in all samples to 0.095 in sequence-only low-confidence samples, and the AUROC gain increased from 0.055 to 0.123. This result indicates that structural features provide clearer additional value for sequence-ambiguous m6A candidates.

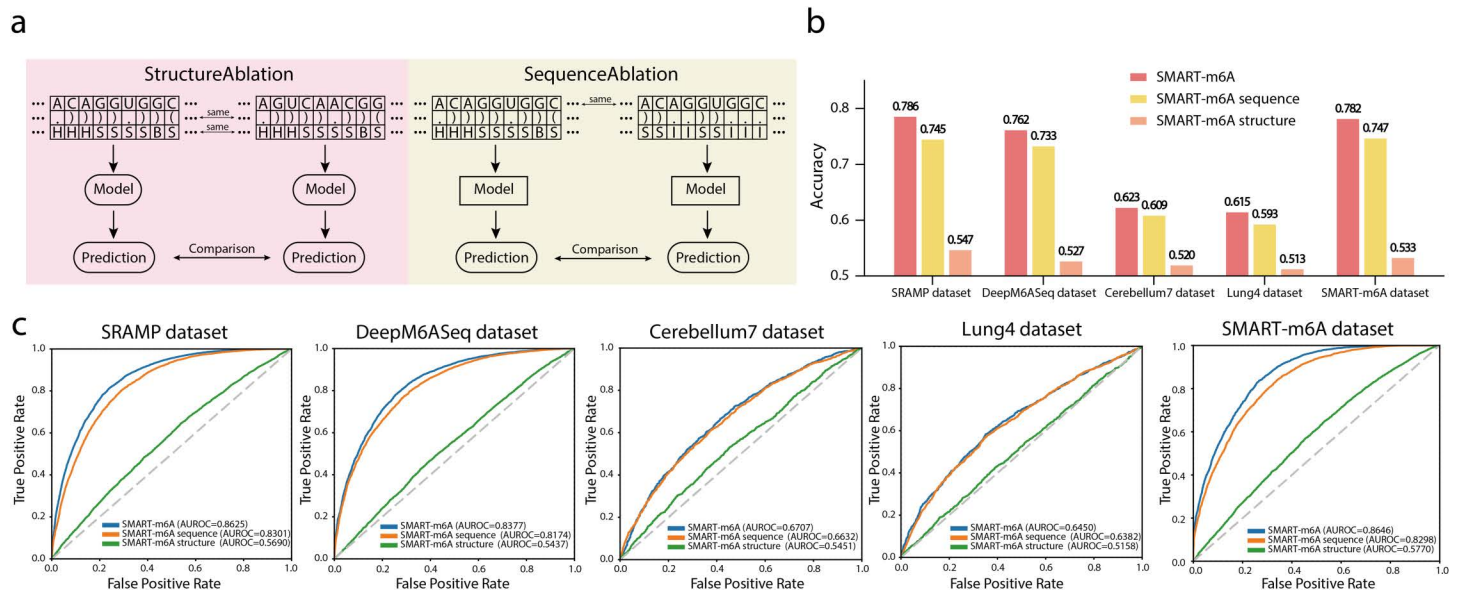


Fig 2. Sequence and structure features provide complementary information for m6A prediction. (a) Illustration of ablation settings. In the *StructureAblation* setting, structural features are removed while sequence information is retained; in the *SequenceAblation* setting, sequence features are removed while structural information is retained. Predictions from the ablated models are compared with those of the full SMART-m6A model. (b) Accuracy comparison of ablated vs. non-ablated models across five benchmark datasets (SMART-m6A, SRAMP, DeepM6ASeq, m6ATCPred-Cerebellum7, m6ATCPred-Lung4). (c) AUROC comparison of ablated vs. non-ablated models across five benchmark datasets.

<https://doi.org/10.1371/journal.pcbi.1014649.g002>

SMART-m6A identifies structurally sensitive m6A sites and models modification levels across cell lines

m6A modifications are reversible and exhibit cell state specificity. Previous studies have reported systematic differences between HeLa and HEK293T m6A landscapes [22]. Given that RNA secondary structures also undergo dynamic changes across cellular contexts and are associated with m6A-related regulatory features, we hypothesized that cell line-specific m6A sites may show more pronounced local structural differences. Moreover, if SMART-m6A indeed captures structure-associated predictive signals, it should be capable of quantifying modification intensities across different cellular states and predicting cell-specific m6A modifications. To test this, we integrated GLORI-based m6A profiles from HeLa and HEK293T cells (including both site locations and modification intensities) with icSHAPE-derived structural data [40] for comparative analysis.

We first summarized the icSHAPE-based comparison workflow (Fig 3a) and quantified cross-cell-line structural divergence between shared and cell line-specific m6A sites using cumulative distributions and a two-sample Kolmogorov-Smirnov (KS) test (Fig 3b). A complementary boxplot comparison is provided in S3 Fig. This analysis revealed a rightward shift for cell line-specific sites, indicating larger icSHAPE changes relative to shared sites (KS, $p = 4.621 \times 10^{-53}$). Based on this finding, we further examined whether SMART-m6A could recognize sensitivity to structural input changes. To this end, we designed and implemented a structure-perturbation assay (Fig 3c). Specifically, during the inference stage, we perturbed the 15 consecutive nucleotides with the highest attention scores in each input dot-bracket structural string by swapping the symbols “.” and “()”, and then re-evaluated the model predictions. Sites whose predicted outcomes changed after perturbation were defined as structurally sensitive m6A sites, whereas those remaining unchanged were defined as structurally insensitive m6A sites. We then compared the icSHAPE-based structural variation between these two categories of sites (Fig 3d). The results showed that, in both HEK293T and HeLa cells, structure-sensitive sites exhibited significantly greater structural fluctuations (HEK293T: KS test $p = 3.713 \times 10^{-2}$; HeLa: KS test $p = 4.335 \times 10^{-2}$). These findings

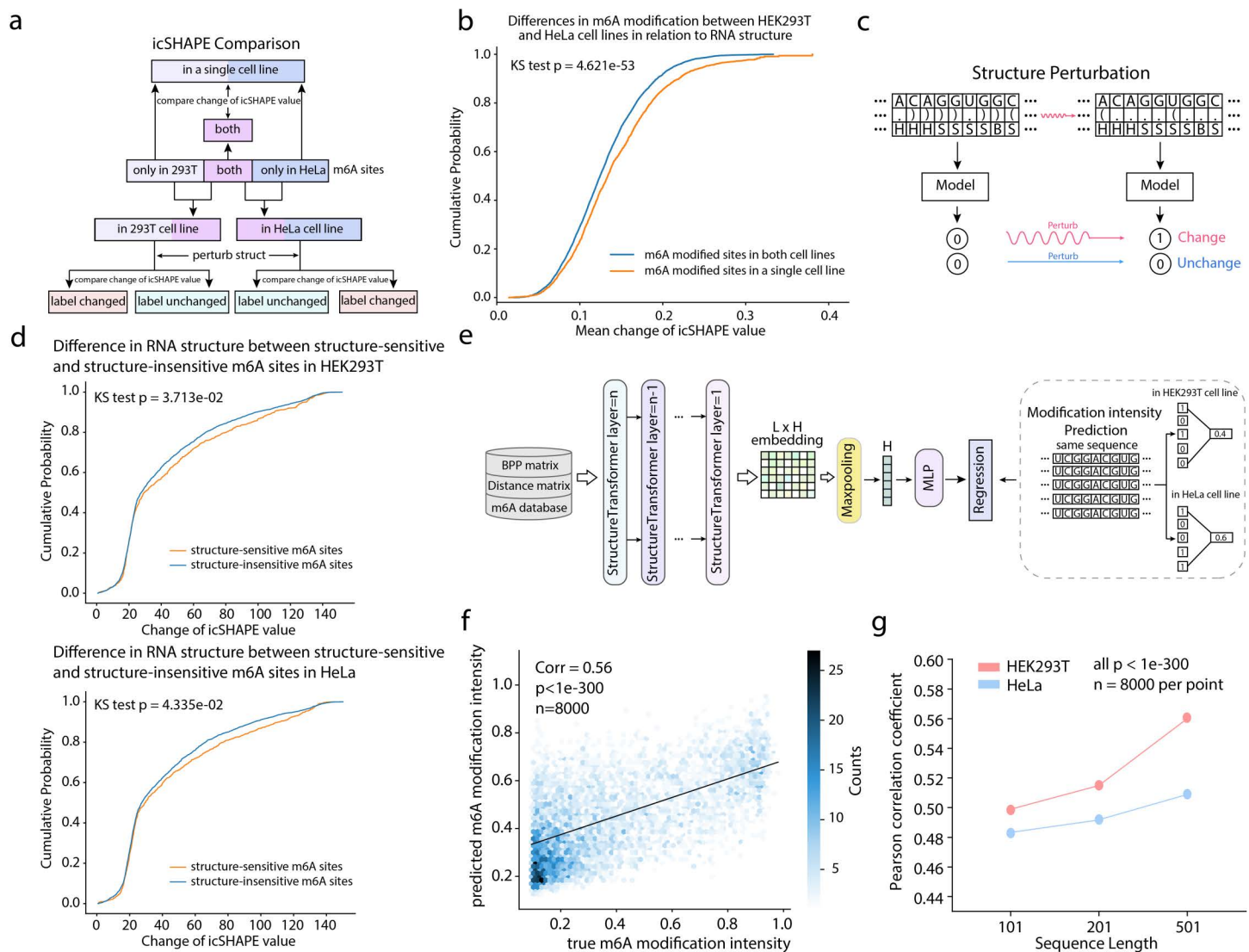


Fig 3. SMART-m6A identifies structure-associated differences in cell line-specific m6A sites and predicts modification intensity. (a) Workflow of icSHAPE-based comparison between shared and cell line-specific m6A sites. (b) Cumulative distribution of RNA structural probe values for differential versus shared m6A sites between the two cell lines. (c) Illustration of the structural perturbation strategy by modifying dot-bracket annotations. (d) RNA structural differences between structure-sensitive and structure-insensitive m6A sites in HEK293T and HeLa cells. (e) Overview of the regression framework for quantitative prediction of m6A modification intensity. (f) Scatter plot showing correlation between predicted and true modification intensity in HEK293T. (g) Pearson correlation coefficients under different input sequence lengths (101 nt, 201 nt, 501 nt).

<https://doi.org/10.1371/journal.pcbi.1014649.g003>

indicate that SMART-m6A can distinguish m6A sites with differential prediction sensitivity to structural perturbation, and that prediction-sensitive sites are associated with greater local structural variation in experimental icSHAPE profiles.

Building on this, we extended SMART-m6A to quantitative prediction of modification intensities (Fig 3e). Considering that the same transcript can exhibit different modification levels at the same adenosine across cell types, we trained models for HeLa and HEK293T separately. As input sequence length increased from 101 nt to 201 nt and 501 nt, correlations between predicted and observed intensities progressively improved (Fig 3f). In HEK293T, the Pearson correlation coefficient reached 0.56, representing a 12.5% increase over the 101-nt model (Fig 3g). In HeLa, correlation reached 0.51,

corresponding to a 5.6% increase. Each correlation was calculated using 8,000 paired samples, and all two-sided Pearson correlation tests were highly significant, with $p < 1 \times 10^{-300}$. These results demonstrate that longer input sequences provide richer contextual information, thereby enhancing prediction accuracy for modification strength. Moreover, we conducted additional experiments to assess the performance of other algorithms across different sequence lengths (S4 Fig).

Taken together, SMART-m6A captures structure-associated differences in cell line-specific m6A sites and extends to quantitative prediction of modification intensity across cellular contexts. These results highlight the value of integrating structural features for improving model interpretability and for identifying m6A candidates linked to distinct local structural contexts and cell type-dependent modification patterns.

Motif-based interpretation of SMART-m6A

To explore the interpretability of SMART-m6A, we further analyzed the multi-head attention distribution of the last layer of the StructureTransformer (Fig 4a). Fig 4b presents two representative attention heads, while the results of the remaining heads are provided in S5 Fig. The overall distribution revealed that different attention heads exhibited distinct focus patterns, with each head attending to different sequence regions, suggesting that the multi-head mechanism enables the model to capture information across multiple scales and regions (Fig 4b). Meanwhile, we also performed loop-type statistics on the high-scoring regions attended by different attention heads in the SMART-m6A dataset (S6 Fig). For local attention visualization, we selected a representative attention head and representative positive samples to illustrate sample-to-sample heterogeneity in attention patterns. Information content analysis showed a significant enrichment of DRACH motifs, with the highest bit score at the modification center, consistent with the known sequence conservation of m6A sites. Notably, the attention focus of this head varied across sequences, which aligns with biological intuition: key motifs in different m6A sites may be distributed across different regions. This indicates that the model can capture motif-specific differences among m6A-containing sequences (Fig 4c). To systematically identify motifs learned by the attention mechanism, we applied MEME to high-attention regions in positive test samples, yielding 16 high-scoring motifs. These were compared against the human CIS-BP database using TOMTOM (Fig 4d). Several motifs matched known RNA-binding proteins (RBPs), including hnRNPK, which has been reported to interact with METTL3 as a cofactor to influence m6A deposition and post-transcriptional regulation [41], and FUS, which interacts with m6A-modified RNAs to regulate RNA metabolism and contribute to neurodegenerative disease processes [42]. More matched motifs are provided in S7 Fig.

Given the structure-aware design of SMART-m6A, we further mapped top-attended regions from each attention head to their corresponding RNA structural contexts, including stem, multiloop, bulge, internal loop, hairpin loop, and exterior/unpaired regions (Fig 4e). This analysis revealed distinct structural-context preferences across attention heads, indicating that SMART-m6A attention captures not only sequence motifs but also structure-associated signals relevant to model prediction. To further examine whether these attention-selected structural regions carried predictive information, we extracted structural motif patterns from top-attended windows and compared them with matched background windows (S8a Fig). Several structural motifs were enriched in high-attention regions (S8b Fig). Sequence motifs recovered from windows containing these enriched structural motifs matched DRACH-family m6A motifs, indicating that the enriched structural contexts are coupled with known m6A sequence signals (S8c Fig). In addition, perturbing attention-selected structural windows caused larger prediction changes than perturbing matched high-attention random windows (S8d Fig). These results suggest that the structural regions highlighted by attention are not merely descriptive, but are also functionally relevant to SMART-m6A prediction.

Together, these findings confirm that SMART-m6A successfully captures biologically relevant signals associated with m6A regulation from both sequence and structural contexts. Importantly, we also identified four motifs without significant matches in existing databases. Although their functions remain unclear, these motifs may represent previously uncharacterized regulatory elements, providing potential leads for future experimental validation.

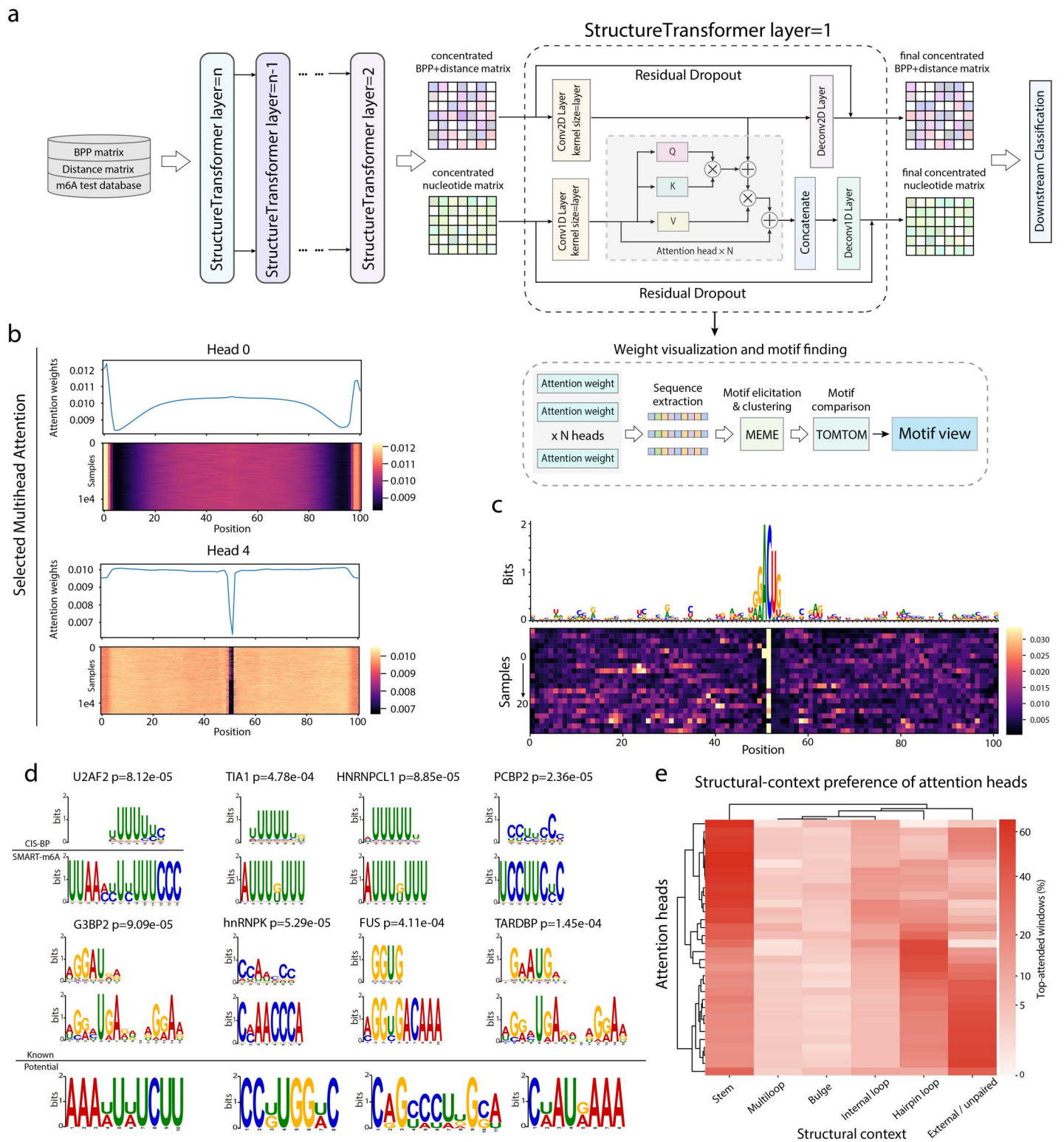


Fig 4. Attention-based interpretability reveals sequence motifs and structural-context preferences associated with m6A prediction. (a) Framework of attention weight extraction and motif discovery from the final StructureTransformer layer. **(b)** Global attention distributions across different

attention heads. **(c)** Local attention visualization and information content analysis for representative samples. **(d)** Motifs identified from high-attention regions and matched to the CIS-BP human database using MEME–TOMTOM. **(e)** Distribution of top-attended regions across RNA structural contexts for different attention heads.

<https://doi.org/10.1371/journal.pcbi.1014649.g004>

Discussion

In this study, we developed SMART-m6A, a multifeature deep learning framework that integrates RNA sequence information with sequence-derived structural representations to predict m6A modification sites and modification intensities. Across multiple benchmark datasets, SMART-m6A achieved consistently strong predictive performance, with structural features providing particularly clear additional value for sequence-ambiguous candidates. Attention-based analyses and structural-input perturbation experiments further identified prediction-relevant local sequence–structure patterns. In parallel, cross-cell line analyses showed that cell line-specific m6A sites are associated with greater local structural divergence than shared sites. Together, these results support SMART-m6A as an interpretable structure-aware prediction framework and provide testable hypotheses regarding the relationship between local RNA structural context and cell line-associated m6A patterns.

Despite these advantages, several limitations remain. First, the predictive accuracy of SMART-m6A is inevitably influenced by the quality and resolution of available data. For instance, current icSHAPE profiles and m6A annotations still contain experimental noise and variability, which may constrain the model's ability to capture subtle local structural features associated with m6A patterns. Second, although our framework effectively integrates sequence and structural features, it does not yet account for additional regulatory factors, such as RNA-binding protein dynamics, cellular stress responses, or epitranscriptomic crosstalk. These factors likely contribute to the cell-specific deposition of m6A. In addition, SMART-m6A uses sequence-derived local structural proxies from isolated mature mRNA fragments, rather than native co-transcriptional RNA structures *in vivo*.

Looking forward, our results suggest that structure-informed modeling holds strong potential for understanding cell type–specific m6A regulation. Future work could benefit from integrating higher-resolution structural datasets, multi-omics measurements, or single-cell m6A maps to disentangle the interplay between sequence context, secondary structure, and cellular environment. While our study represents a step toward more interpretable and biologically grounded prediction models, it also raises important questions that remain to be addressed, paving the way for deeper exploration of the principles governing epitranscriptomic regulation.

Materials and methods

Datasets

We evaluated SMART-m6A across six benchmark datasets, including four publicly available resources (SRAMP [43], DeepM6ASeq [27], m6ATCPred-Cerebellum7 [44], and m6ATCPred-Lung4 [44]) and two in-house datasets (SMART-m6A and SMART-m6A-Position). The dataset sources, detection technologies, training and test sizes, and availability of positional information are summarized in [S1 Table](#). The two datasets constructed in this study were newly generated using the recently developed GLORI technology. We downloaded single-nucleotide resolution m6A modification sites identified using the GLORI method and mapped them to their corresponding genomic coordinates. During dataset construction, negative samples were selected to ensure that the central nucleotide was adenosine (A) and conformed to the DRACH motif, thereby increasing the stringency and difficulty of prediction. All datasets consisted of 101-nt mammalian mRNA fragments. In SRAMP, Cerebellum7, Lung4, and SMART-m6A, sequences were extracted by centering on candidate adenosines, whereas DeepM6ASeq imposed no central nucleotide restriction. Training and test sets were strictly separated for each dataset, and positive and negative samples were approximately balanced in all benchmark datasets. For structural feature data, we used ViennaRNA [45] to generate each sequence's base-pairing probability matrix and dot-bracket

pairing annotation, and employed bpRNA [46] to derive the corresponding loop-domain (loop type) structures. All structural features used in SMART-m6A were computed from isolated mature mRNA fragments. In this setting, the structural inputs represented sequence-derived local structural proxies, and no modification-aware folding model or m6A-aware folding parameters were introduced during model training or cross-validation. In addition, a three-channel relative distance matrix was custom-computed, where each element is inversely proportional to the absolute nucleotide distance $|i-j|^k$ for three channels ($k = 1, 2, 3$), with non-zero values only when $|i-j| > 0$. For all datasets, model training and evaluation were conducted under a tenfold cross-validation scheme.

Dual-branch convolutional feature extraction

To jointly capture local sequence patterns and neighborhood structural relationships, SMART-m6A (Fig 5a) employs parallel convolutional modules for the sequence and structural branches at the input stage.

In the sequence branch, each RNA fragment of length L is described by three parallel symbolic channels representing its nucleotide identity (A, C, G, U), dot-bracket structural annotation (paired “()” or unpaired “.”), and loop-context category

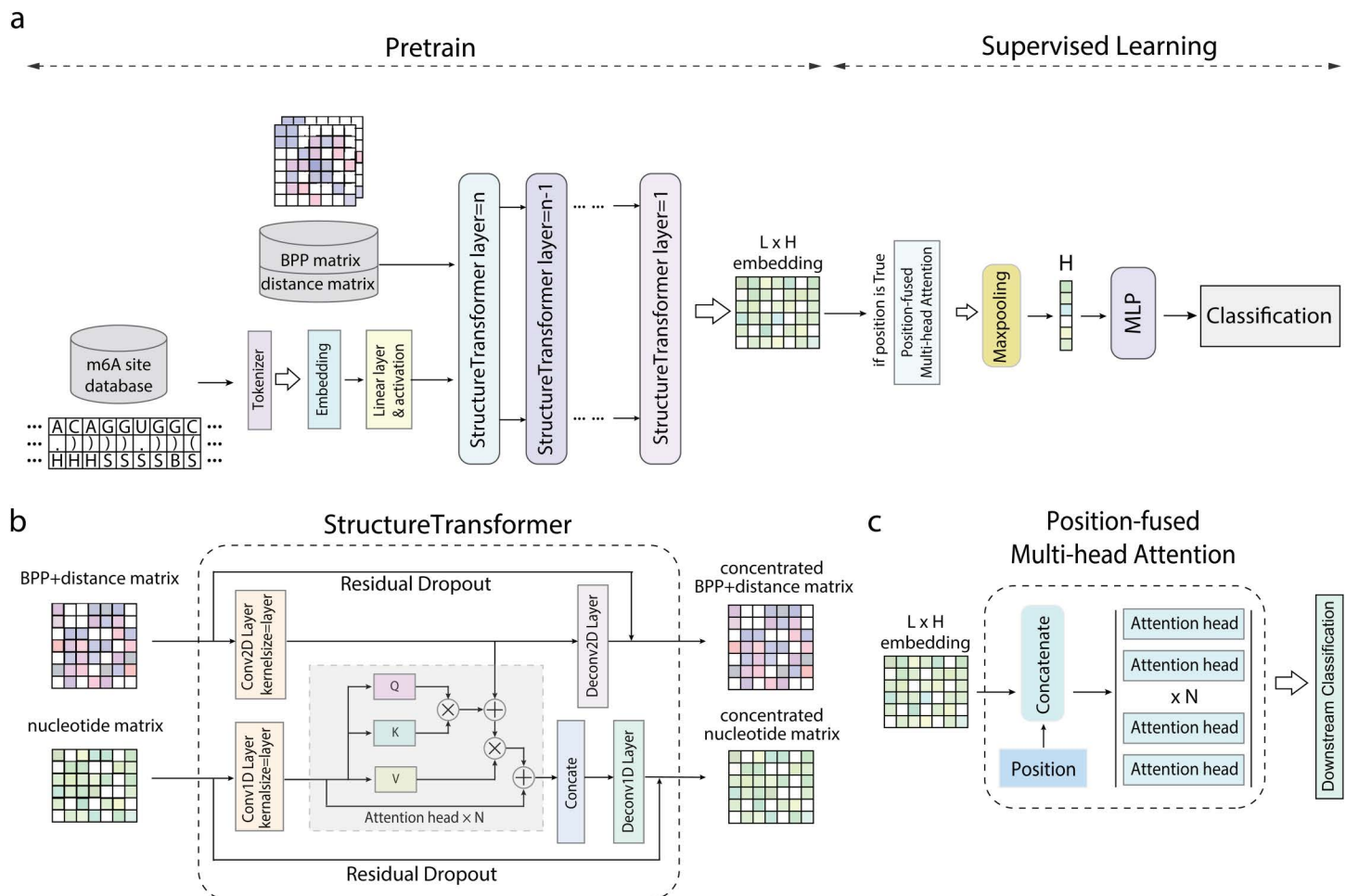


Fig 5. Architecture and workflow of SMART-m6A for multifeature sequence-structure integration. (a) Overall workflow of SMART-m6A, from sequence and structure input to m6A site classification. (b) StructureTransformer architecture integrating nucleotide and structural features. (c) Position-fused multi-head attention incorporating positional information into embeddings.

<https://doi.org/10.1371/journal.pcbi.1014649.g005>

such as hairpin, internal, bulge, and other types of loops. Each of these channels is first embedded into a dense vector representation through an independent embedding matrix, yielding $E^{(seq)}$, $E^{(dot)}$, and $E^{(loop)}$ in $\mathbb{R}^{L \times d_{model}}$. The three embeddings are then concatenated along the last dimension and linearly projected into a unified feature space through a learnable projection matrix $W_p \in \mathbb{R}^{3d_{model} \times d_{model}}$, forming

$$X = [E^{(seq)} \parallel E^{(dot)} \parallel E^{(loop)}] W_p. \quad (1)$$

Here, $[\cdot \parallel \cdot]$ denotes vector concatenation, d_{model} is the embedding dimension, and the resulting tensor $X \in \mathbb{R}^{B \times L \times d_{model}}$ integrates base-composition and structural semantics for each nucleotide, where B is the batch size.

To capture local contextual dependencies, SMART-m6A applies a one-dimensional convolution across the sequence dimension, such that the output at position p is computed as

$$O_p = \sum_{t=0}^{k-1} W_t \cdot X_{p+t}, \quad p \in [1, L - k + 1], \quad (2)$$

where $W_t \in \mathbb{R}^{d_{model} \times d_{model}}$ denotes the kernel weight at offset t , and X_{p+t} is the projected embedding at position $p + t$. The resulting vector $O_p \in \mathbb{R}^{d_{model}}$ serves as a structure-aware k-mer embedding centered at position p . Owing to the convolution operation, the sequence length is reduced from L to $L - k + 1$, yielding an overall output tensor $O \in \mathbb{R}^{B \times (L-k+1) \times d_{model}}$ that preserves the channel dimension while aggregating local context over k-mer windows. By progressively reducing the kernel size k across convolutional layers, the model first learns short-range conserved motifs in shallow layers and gradually expands its receptive field to capture long-range dependencies in deeper ones.

In parallel, the structural branch processes a four-channel tensor $S_{original} \in \mathbb{R}^{4 \times L \times L}$ that combines the base-pairing probability matrix (BPP) with three relative-distance matrices describing the spatial proximity between nucleotide pairs. This tensor is first projected through a 1×1 convolution and then passed through a series of two-dimensional convolutions to extract local neighborhood correlations, producing

$$S = \text{Conv2D}(S_{original}). \quad (3)$$

The resulting tensor $S \in \mathbb{R}^{n_{head} \times L \times L}$ represents head-specific structural feature maps, where n_{head} is the number of attention heads. The 2D convolutions slide receptive fields across each $L \times L$ structural matrix, capturing spatial dependencies between nucleotide pairs, while transposed convolutions and residual connections refine and stabilize the propagation of structural constraints through deeper layers.

Ultimately, the sequence branch yields structure-aware k-mer representations $O \in \mathbb{R}^{B \times (L-k+1) \times d_{model}}$, and the structural branch produces multi-head spatial feature maps $S \in \mathbb{R}^{n_{head} \times L \times L}$. These complementary feature streams are subsequently fused and fed into the Transformer encoder, enabling integrated modeling of sequence and structural information across multiple spatial scales.

Transformer encoder with structural integration

Following convolutional feature extraction, SMART-m6A employs a structure-integrated Transformer encoder to jointly model global dependencies and secondary-structure constraints. The dual-branch convolutional backbone together with the structure-aware attention mechanism constitutes what we term the StructureTransformer encoder (Fig 5b).

In this framework, the output tensor from the sequence branch is first linearly projected into the query, key, and value matrices $Q, K, V \in \mathbb{R}^{B \times L \times d_k}$ through independent learnable transformations:

$$Q = XW_Q, \quad K = XW_K, \quad V = XW_V, \quad (4)$$

where $W_Q, W_K, W_V \in \mathbb{R}^{d_{model} \times d_k}$.

In parallel, the structural branch provides two-dimensional convolutional feature maps $S \in \mathbb{R}^{n_{head} \times L \times L}$, which serve as head-specific structural biases injected into the attention scoring function. The scaled dot-product attention is defined as:

$$\text{Attention}(Q, K, V, S) = \text{Softmax} \left(\frac{QK^T}{\sqrt{d_k}} + \lambda \text{Conv2D}(S) \right) V, \quad (5)$$

where d_k denotes the dimensionality of the key vectors, and λ is a learnable balancing coefficient controlling the contribution of structure-aware biases relative to sequence-based similarity. Here, $\text{Conv2D}(S)$ represents an additional two-dimensional convolution applied to the structural feature maps, which reinforces local base-pair correlations and neighborhood dependencies so that the attention weights are simultaneously constrained by both sequence context and RNA secondary structure.

In implementation, the attention outputs are normalized and refined through residual connections and layer normalization, followed by a feed-forward network. A deconvolutional module is then applied to restore the structural resolution, enabling the stacking of multiple encoder layers. Across successive layers, the model begins with broader convolutional kernels to capture coarse contextual patterns and progressively reduces kernel size to focus on fine-grained local dependencies. At each layer, structure-based biases are explicitly incorporated into attention computation, ensuring that sequence–structure coupling remains consistent and is further reinforced during global representation learning.

Downstream classification

After five layers of StructureTransformer encoding, the resulting joint sequence–structure representation was used as input to the downstream classification module. For datasets without genomic positional information, the encoded embeddings were directly subjected to max pooling, followed by a multilayer perceptron (MLP) for site prediction. For datasets with positional annotations, we first incorporated a multi-head attention module prior to pooling, enabling effective integration of structural features with positional cues and thereby enhancing the model's ability to capture position-dependent patterns (Fig 5c). The fused representation was then processed by max pooling and an MLP to generate classification outputs, achieving accurate prediction of m6A modification sites. This design ensures that the model retains flexibility and adaptability across diverse data scenarios.

Training strategy

During training, SMART-m6A adopted a two-stage pretraining–supervised learning paradigm. In the pretraining stage, we designed reconstruction tasks by randomly masking or mutating nucleotides, structural states (dot-brackets), or loop types. An autoregressive loss was applied to optimize the model, requiring it to accurately recover the original information through softmax prediction. This procedure enabled the model to capture the coupling between RNA sequences and their underlying biophysical properties. In the subsequent supervised learning stage, pretrained weights were loaded on the same dataset, and the model was fine-tuned using available labels. Here, a cross-entropy loss was employed to optimize site-level predictions. The multi-layer embeddings produced by the StructureTransformer were aggregated through global pooling and passed to a multilayer perceptron (MLP), enabling accurate identification of m6A modification sites.

Quantitative regression of m6A modification intensity

For quantitative prediction of m6A modification intensity, we adapted the downstream prediction module of SMART-m6A from binary classification to regression. The StructureTransformer encoder was kept unchanged, while the final classification layer was replaced with a single-neuron regression head. GLORI-derived modification levels were used as quantitative labels.

Because GLORI provides modification stoichiometry as a continuous value, labels were represented on a normalized scale between 0 and 1. The regression model was optimized using mean squared error loss between predicted and observed modification levels. Models were trained separately for HEK293T and HeLa datasets to account for cell line-specific modification profiles. To evaluate whether longer sequence contexts improve quantitative prediction, we trained models using 101-nt, 201-nt, and 501-nt input windows centered on candidate m6A sites. Predictive performance was evaluated using Pearson correlation coefficient between predicted and observed modification levels, with corresponding p-values and sample sizes.

Identifying prediction sensitivity to structural-input perturbation

Some m6A sites may be closely associated with local RNA structural contexts, where pronounced conformational differences are observed, whereas others appear less coupled to structural variation. We employed a structural perturbation assay to distinguish these two types of sites. In this assay, the nucleotide sequence was kept unchanged while the dot-bracket annotations representing RNA secondary structures were systematically perturbed. For each RNA fragment in the test set, consecutive nucleotides with the highest attention scores from the final layer of the StructureTransformer were identified as structurally critical regions, and their structural states were inverted by swapping paired “()” and unpaired “.” symbols, mimicking the local opening or closing of base-paired segments. Sites whose predicted labels changed after perturbation were defined as structure-sensitive sites, whereas those remaining unchanged were defined as structure-insensitive sites. To verify that SMART-m6A correctly distinguishes between these two categories, we compared icSHAPE-derived structural reactivity changes and found that structure-sensitive sites exhibited significantly greater structural variation than structure-insensitive ones (KS test, $p < 0.05$). This analysis evaluates the sensitivity of SMART-m6A predictions to local structural-input changes and does not establish a causal effect of RNA structure on m6A deposition.

Motif extraction for model interpretation

To probe the biological signals captured by the model, we extracted the attention matrices from the final layer of the StructureTransformer and averaged them along the vertical dimension to obtain a base-level attention profile across the input sequence. This distribution reflects the relative contribution of each nucleotide to the overall prediction. We then applied the model to positive samples in the test set and selected subsequences with high attention scores for motif discovery using MEME. The identified motifs were subsequently compared against known databases via TOMTOM. Motifs with successful matches were interpreted as previously established biological signals recognized by the model, whereas unmatched motifs may represent putative novel regulatory elements.

Supporting information

S1 Fig. Structure-misaligned and random-structure control analyses. Comparison of ACC and AUROC among the full SMART-m6A model, sequence plus misaligned structure, sequence plus randomly shuffled structure, and sequence-only SMART-m6A on the SMART-m6A dataset.

(TIF)

S2 Fig. Structural-feature gains in all samples and sequence-ambiguous samples. (a) Accuracy gain from structural features in all samples and sequence-ambiguous samples across pooled and individual datasets. (b) AUROC gain from structural features in all samples and sequence-ambiguous samples across pooled and individual datasets.

(TIF)

S3 Fig. Local structural divergence of shared and cell line-specific m6A sites. (a) Boxplot comparison of icSHAPE differences between shared and cell line-specific m6A sites. (b) Representative cell line-specific m6A sites with MFE and icSHAPE-guided RNA structures.

(TIF)

S4 Fig. Quantitative regression performance across different input sequence lengths. Regression performance of SMART-m6A, Transformer, and CNN using different input sequence lengths in HEK293T and HeLa cell lines.
(TIF)

S5 Fig. Attention distributions across all heads in the final StructureTransformer layer. Attention distribution of 32 heads in the last layer of the StructureTransformer on the SMART-m6A test set.
(TIF)

S6 Fig. Structural-context composition of top-attended windows. Distribution of loop structure types derived from the top 300 high-attention 16-nt segments in each of the 32 heads of the last StructureTransformer layer on the SMART-m6A test set.
(TIF)

S7 Fig. Additional RBP motif matches identified by SMART-m6A. Motifs identified by SMART-m6A that additionally matched entries in the CIS-BP database.
(TIF)

S8 Fig. Attention-guided structural motif analysis. (a) Workflow of attention-guided structural motif analysis. (b) Structural motifs enriched in top-attended windows compared with matched background windows, shown as log2 odds ratios. (c) Recovered sequence motifs from enriched structural motifs matched DRACH-family m6A motifs. (d) Perturbation of attention-selected structural motifs caused larger prediction changes than matched random perturbations.
(TIF)

S1 Table. Summary of datasets used for model evaluation. Summary of dataset sources, detection technologies, training sizes, test sizes, and availability of positional information for all datasets used in this study.
(XLSX)

S1 File. Graphical abstract.
(DOCX)

Author contributions

Data curation: Mingze Sun, Di Zhang.

Formal analysis: Mingze Sun, Di Zhang.

Methodology: Mingze Sun, Di Zhang, Zhiyuan Li, Yihan Lin.

Supervision: Di Zhang, Zhiyuan Li, Yihan Lin.

Visualization: Mingze Sun.

Writing – original draft: Mingze Sun, Di Zhang.

Writing – review & editing: Di Zhang, Zhiyuan Li, Yihan Lin.

References

1. Delaunay S, Helm M, Frye M. RNA modifications in physiology and disease: towards clinical applications. *Nat Rev Genet.* 2024;25(2):104–22. <https://doi.org/10.1038/s41576-023-00645-2> PMID: [37714958](https://pubmed.ncbi.nlm.nih.gov/37714958/)
2. Sun H, Li K, Liu C, Yi C. Regulation and functions of non-m6A mRNA modifications. *Nat Rev Mol Cell Biol.* 2023;24(10):714–31. <https://doi.org/10.1038/s41580-023-00622-x>
3. Boccaletto P, Stefaniak F, Ray A, Cappannini A, Mukherjee S, Purta E, et al. MODOMICS: a database of RNA modification pathways. 2021 update. *Nucleic Acids Res.* 2022;50(D1):D231–5. <https://doi.org/10.1093/nar/gkab1083> PMID: [34893873](https://pubmed.ncbi.nlm.nih.gov/34893873/)

4. Yang C, Hu Y, Zhou B, Bao Y, Li Z, Gong C, et al. The role of m6A modification in physiology and disease. *Cell Death Dis.* 2020;11(11):960. <https://doi.org/10.1038/s41419-020-03143-z> PMID: 33162550
5. Jia G, Fu Y, Zhao X, Dai Q, Zheng G, Yang Y, et al. N6-methyladenosine in nuclear RNA is a major substrate of the obesity-associated FTO. *Nat Chem Biol.* 2011;7(12):885–7. <https://doi.org/10.1038/nchembio.687> PMID: 22002720
6. Wang X, Lu Z, Gomez A, Hon GC, Yue Y, Han D, et al. N6-methyladenosine-dependent regulation of messenger RNA stability. *Nature.* 2014;505(7481):117–20. <https://doi.org/10.1038/nature12730> PMID: 24284625
7. Roost C, Lynch SR, Batista PJ, Qu K, Chang HY, Kool ET. Structure and thermodynamics of N6-methyladenosine in RNA: a spring-loaded base modification. *J Am Chem Soc.* 2015;137(5):2107–15. <https://doi.org/10.1021/ja513080v> PMID: 25611135
8. Deng X, Su R, Feng X, Wei M, Chen J. Role of N6-methyladenosine modification in cancer. *Curr Opin Genet Dev.* 2018;48:1–7. <https://doi.org/10.1016/j.gde.2017.04.005> PMID: 29040886
9. Roignant J-Y, Soller M. m6A in mRNA: an ancient mechanism for fine-tuning gene expression. *Trends Genet.* 2017;33(6):380–90. <https://doi.org/10.1016/j.tig.2017.04.003> PMID: 28499622
10. Hong J, Xu K, Lee JH. Biological roles of the RNA m6A modification and its implications in cancer. *Exp Mol Med.* 2022;54(11):1822–32. <https://doi.org/10.1038/s12276-022-00897-8> PMID: 36446846
11. Qin Y, Li L, Luo E, Hou J, Yan G, Wang D, et al. Role of m6A RNA methylation in cardiovascular disease (Review). *Int J Mol Med.* 2020;46(6):1958–72. <https://doi.org/10.3892/ijmm.2020.4746> PMID: 33125109
12. Liu J, Yue Y, Han D, Wang X, Fu Y, Zhang L, et al. A METTL3-METTL14 complex mediates mammalian nuclear RNA N6-adenosine methylation. *Nat Chem Biol.* 2014;10(2):93–5. <https://doi.org/10.1038/nchembio.1432> PMID: 24316715
13. Zheng G, Dahl JA, Niu Y, Fedorcsak P, Huang C-M, Li CJ, et al. ALKBH5 is a mammalian RNA demethylase that impacts RNA metabolism and mouse fertility. *Mol Cell.* 2013;49(1):18–29. <https://doi.org/10.1016/j.molcel.2012.10.015> PMID: 23177736
14. Shi H, Wei J, He C. Where, when, and how: context-dependent functions of RNA methylation writers, readers, and erasers. *Mol Cell.* 2019;74(4):640–50. <https://doi.org/10.1016/j.molcel.2019.04.025> PMID: 31100245
15. Alarcón CR, Goodarzi H, Lee H, Liu X, Tavazoie S, Tavazoie SF. HNRNPA2B1 is a mediator of m(6)A-dependent nuclear RNA processing events. *Cell.* 2015;162(6):1299–308. <https://doi.org/10.1016/j.cell.2015.08.011> PMID: 26321680
16. Meyer KD, Jaffrey SR. Rethinking m6A readers, writers, and erasers. *Annu Rev Cell Dev Biol.* 2017;33(1):319–42.
17. Zhang C, Chen Y, Sun B, Wang L, Yang Y, Ma D, et al. m6A modulates haematopoietic stem and progenitor cell specification. *Nature.* 2017;549(7671):273–6. <https://doi.org/10.1038/nature23883>
18. Dominissini D, Moshitch-Moshkovitz S, Schwartz S, Salmon-Divon M, Ungar L, Osenberg S, et al. Topology of the human and mouse m6A RNA methylomes revealed by m6A-seq. *Nature.* 2012;485(7397):201–6. <https://doi.org/10.1038/nature11112> PMID: 22575960
19. Chen K, et al. High-resolution N6-methyladenosine (m6A) map using photo-crosslinking-assisted m6A sequencing. *Angew Chem.* 2015;127(5):1607–10.
20. Garcia-Campos MA, Edelheit S, Toth U, Safra M, Shachar R, Viukov S, et al. Deciphering the “m6A Code” via antibody-independent quantitative profiling. *Cell.* 2019;178(3):731–747.e16. <https://doi.org/10.1016/j.cell.2019.06.013> PMID: 31257032
21. Zhang Z, Chen L-Q, Zhao Y-L, Yang C-G, Roundtree IA, Zhang Z, et al. Single-base mapping of m6A by an antibody-independent method. *Sci Adv.* 2019;5(7):eaax0250. <https://doi.org/10.1126/sciadv.aax0250> PMID: 31281898
22. Garalde DR, Snell EA, Jachimowicz D, Sipos B, Lloyd JH, Bruce M, et al. Highly parallel direct RNA sequencing on an array of nanopores. *Nat Methods.* 2018;15(3):201–6. <https://doi.org/10.1038/nmeth.4577> PMID: 29334379
23. Sun H, Lu B, Zhang Z, Xiao Y, Zhou Z, Xi L, et al. Mild and ultrafast GLORI enables absolute quantification of m6A methylome from low-input samples. *Nat Methods.* 2025;22(6):1226–36. <https://doi.org/10.1038/s41592-025-02680-9> PMID: 40325216
24. Chen W, Feng P, Ding H, Lin H, Chou K-C. iRNA-Methyl: Identifying N6-methyladenosine sites using pseudo nucleotide composition. *Anal Biochem.* 2015;490:26–33. <https://doi.org/10.1016/j.ab.2015.08.021>
25. Krizhevsky A, Sutskever I, Hinton GE. Imagenet classification with deep convolutional neural networks[J]. *Advances in neural information processing systems*, 2012.
26. Schuster M, Paliwal KK. Bidirectional recurrent neural networks. *IEEE Trans Signal Process.* 1997;45(11):2673–81. <https://doi.org/10.1109/78.650093>
27. Zhang Y, Hamada M. DeepM6ASeq: prediction and characterization of m6A-containing sequences using deep learning. *BMC Bioinform.* 2018;19(Suppl 19):524. <https://doi.org/10.1186/s12859-018-2516-4> PMID: 30598068
28. Xiong Y, He X, Zhao D, Tian T, Hong L, Jiang T, et al. Modeling multi-species RNA modification through multi-task curriculum learning. *Nucleic Acids Res.* 2021;49(7):3719–34. <https://doi.org/10.1093/nar/gkab124> PMID: 33744973
29. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*; 2016. p. 770–8
30. Luo Z, Zhang J, Fei J, Ke S. Deep learning modeling m6A deposition reveals the importance of downstream cis-element sequences. *Nat Commun.* 2022;13(1):2720. <https://doi.org/10.1038/s41467-022-30209-7> PMID: 35581216

31. Vaswani A, et al. Attention is all you need. *Adv Neural Inf Process Syst*. 2017;30.
32. Cho K, et al. Learning phrase representations using RNN encoder-decoder for statistical machine translation. *arXiv:1406.1078 [Preprint]*. 2014.
33. Fan R, Cui C, Kang B, Chang Z, Wang G, Cui Q. A combined deep learning framework for mammalian m6A site prediction. *Cell Genom*. 2024;4(12):100697. <https://doi.org/10.1016/j.xgen.2024.100697> PMID: [39571573](https://pubmed.ncbi.nlm.nih.gov/39571573/)
34. Dubin DT, Taylor RH. The methylation state of poly A-containing messenger RNA from cultured hamster cells. *Nucleic Acids Res*. 1975;2(10):1653–68. <https://doi.org/10.1093/nar/2.10.1653> PMID: [1187339](https://pubmed.ncbi.nlm.nih.gov/1187339/)
35. Höfler S, Duss O. Interconnections between m(6)A RNA modification, RNA structure, and protein-RNA complex assembly. *Life Sci Alliance*. 2024;7(1).
36. De La Cruz BM, Darsinou M, Riccio A. From form to function: m6A methylation links mRNA structure to metabolism. *Adv Biol Regul*. 2023;87:100926.
37. Shachar R, Dierks D, Garcia-Campos MA, Uzonyi A, Toth U, Rossmanith W, et al. Dissecting the sequence and structural determinants guiding m6A deposition and evolution via inter- and intra-species hybrids. *Genome Biol*. 2024;25(1):48. <https://doi.org/10.1186/s13059-024-03182-1> PMID: [38360609](https://pubmed.ncbi.nlm.nih.gov/38360609/)
38. Wang Y, Wang S, Meng Z, Liu X-M, Mao Y. Determinant of m6A regional preference by transcriptional dynamics. *Nucleic Acids Res*. 2024;52(7):3510–21. <https://doi.org/10.1093/nar/gkae169> PMID: [38452220](https://pubmed.ncbi.nlm.nih.gov/38452220/)
39. Liu N, Dai Q, Zheng G, He C, Parisien M, Pan T. N(6)-methyladenosine-dependent RNA structural switches regulate RNA-protein interactions. *Nature*. 2015;518(7540):560–4. <https://doi.org/10.1038/nature14234> PMID: [25719671](https://pubmed.ncbi.nlm.nih.gov/25719671/)
40. Sun L, Xu K, Huang W, Yang YT, Li P, Tang L, et al. Predicting dynamic cellular protein-RNA interactions by deep learning using in vivo RNA structures. *Cell Res*. 2021;31(5):495–516. <https://doi.org/10.1038/s41422-021-00476-y> PMID: [33623109](https://pubmed.ncbi.nlm.nih.gov/33623109/)
41. Barbieri I, Tzelepis K, Pandolfini L, Shi J, Millán-Zambrano G, Robson SC, et al. Promoter-bound METTL3 maintains myeloid leukaemia by m6A-dependent translation control. *Nature*. 2017;552(7683):126–31. <https://doi.org/10.1038/nature24678> PMID: [29186125](https://pubmed.ncbi.nlm.nih.gov/29186125/)
42. Chen J, Zhang Y-C, Huang C, Shen H, Sun B, Cheng X, et al. m6A regulates neurogenesis and neuronal development by modulating histone methyltransferase Ezh2. *Genom Proteom Bioinform*. 2019;17(2):154–68. <https://doi.org/10.1016/j.gpb.2018.12.007> PMID: [31154015](https://pubmed.ncbi.nlm.nih.gov/31154015/)
43. Zhou Y, Zeng P, Li Y-H, Zhang Z, Cui Q. SRAMP: prediction of mammalian N6-methyladenosine (m6A) sites based on sequence-derived features. *Nucleic Acids Res*. 2016;44(10):e91. <https://doi.org/10.1093/nar/gkw104> PMID: [26896799](https://pubmed.ncbi.nlm.nih.gov/26896799/)
44. Tu G, Wang X, Xia R, Song B. m6A-TCPred: a web server to predict tissue-conserved human m6A sites using machine learning approach. *BMC Bioinform*. 2024;25(1):127. <https://doi.org/10.1186/s12859-024-05738-1> PMID: [38528499](https://pubmed.ncbi.nlm.nih.gov/38528499/)
45. Lorenz R. ViennaRNA package 2.0. *Algorithms Mol Biol*. 2011;6(1):26.
46. Danaee P, Rouches M, Wiley M, Deng D, Huang L, Hendrix D. bpRNA: large-scale automated annotation and analysis of RNA secondary structure. *Nucleic Acids Res*. 2018;46(11):5381–94. <https://doi.org/10.1093/nar/gky285> PMID: [29746666](https://pubmed.ncbi.nlm.nih.gov/29746666/)