SWAMamba: A Sliding Window Attention Mamba Framework for Predicting Translation Elongation Rates

Xi Zeng[1], Fei Ni[2], Shaoqing Jiao[1], Dazhi Lu[1], Jianye Hao[2*], Jiajie Peng[1, 3, 4*]

[1]AI for Science Interdisciplinary Research Center, School of Computer Science, Northwestern Polytechnical University
[2]College of Intelligence and Computing, Tianjin University
[3]School of Computer Science, Research and Development Institute of Northwestern Polytechnical University in Shenzhen
[4]Key Laboratory of Big Data Storage and Management, Northwestern Polytechnical University, Ministry of Industry and Information Technology
{xizeng,jiaosq22,dzlu}@mail.nwpu.edu.cn, {fei_ni,jianye.hao}@tju.edu.cn, jiajiepeng@nwpu.edu.cn

Abstract

Translation elongation is essential for cellular proteostasis and is implicated in cancer and neurodegeneration. Accurately predicting the rate of ribosome elongation in each codon (also called ribosomal A site) on mRNA is important for understanding and modulating protein synthesis. However, predicting elongation rates is challenging due to the trade-off between capturing distal codon interactions and focusing on proximal codon effects at the A site. Approaches capturing distal codon interactions in the coding sequences (CDS) of mRNA fail to effectively differentiate critical regions (codons near the A site) due to insufficient effective mechanisms for focusing on these regions. Conversely, due to the limitations of models when handling long mRNA sequences, some methods simplify inputs by conditioning solely on proximal codons surrounding the A site, leading to the loss of important information from distal codons. To address this issue, we leverage Mamba's success in capturing long-range dependencies to enable the consideration of distant codons' impact on the A site. Additionally, we introduce a sliding window attention mechanism to emphasize the proximal codons around the A site during ribosome elongation. Building on these advancements, we present Sliding Window Attention Mamba (SWAMamba), a novel framework that simultaneously leverages both proximal and distal codon effects on the A site. We conduct comprehensive evaluations on ribosome data across four species and find that SWAMamba significantly outperformed current state-of-the-art methods in predicting translation elongation rates.

Introduction

Translation elongation is a critical phase of protein synthesis that significantly affects cellular function and protein homeostasis (Brar 2016; Ingolia 2016; Shao et al. 2024). This process involves ribosomes moving along mRNAs at variable rates, recruiting tRNAs to the ribosomal A site for accurate codon-anticodon pairing, and sequentially adding amino acids to the growing polypeptide chain (Figure 1) (Tian et al. 2021). Designing coding sequences based on translation elongation rates can lead to substantial differences in protein output (Tunney et al. 2018; Zhao, Yu, and Liu 2017). Therefore, accurately predicting these rates is crucial for understanding and modulating protein synthesis.

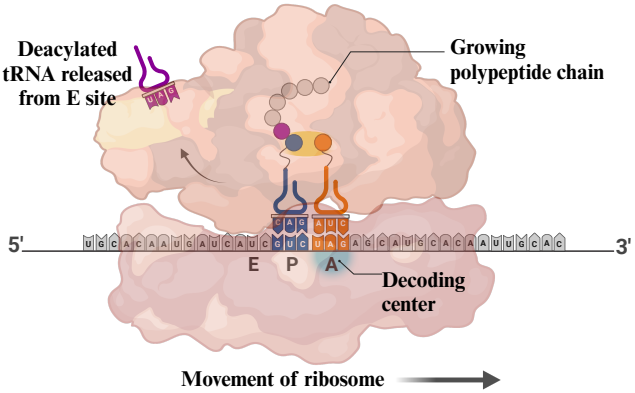

Figure 1: Translation elongation phase

The emergence of ribosome profiling has significantly advanced our comprehension of mRNA translation elongation (Andreev et al. 2017; Ingolia, Hussmann, and Weissman 2019). Ribosome profiling involves sequencing the reads of ribosome-protected fragments (RPFs) (Ingolia et al. 2009; Ingolia 2014). After normalization, these RPFs counts directly reflect the ribosome density distribution along the CDS. Moreover, Codons with higher ribosome density have slower translation rates, while those with lower ribosome density translate faster (Tunney et al. 2018).

A variety of computational approaches have been developed from the accumulating public ribosome profiling data. Tunney et al. (2018) demonstrated the critical role of codons near the A site in accurately predicting ribosome density, emphasizing their importance in computational models. In contrast, Tian et al. (2021) found that codons far from the A site also affect ribosome density predictions, highlighting the need to model ribosome density across the entire CDS region of mRNAs to capture long-range codon influences. However, models modeling the CDS region of mRNA overlook crucial codons near the A site due to insufficient effective mechanisms for focusing on these codons. Meanwhile, architectural limitations with long mRNA sequences com-

*Corresponding author.

pel methods to simplify inputs by conditioning solely on proximal codons around the A site, which fails to capture the influence of long-range codons. These conflicting issues arise from the complexity of ribosome elongation and highlight the need to integrate both the A site proximal and distal codons effect into computational models.

Predicting translation elongation rates by integrating both A site proximal and distal codon effects into computational models presents two key challenges: efficiently processing long mRNA sequences, and considering both proximal and distal codon effects simultaneously. For the first challenge, Tian et al. (2021) used a bidirectional gated recurrent unit (Bi-GRU) network to extract information from the CDS region. However, recurrent neural networks (RNNs) struggle with long-distance dependencies due to the vanishing gradient problem (Al-Selwi et al. 2023). While Transformer-based models are widely used for sequence modeling (Vaswani et al. 2017), these models face significant issues when processing long mRNA sequences due to the quadratic computational complexity of self-attention (Hatamizadeh and Kautz 2024). For the second challenge, a straightforward solution is to design a mechanism that effectively focuses on proximal codons near the A site while also capturing distal codon effects. Traditional methods addressing proximal codons at the A site typically use data segmentation rather than incorporating this focus into the model, leading to models that cannot effectively emphasize proximal codons or capture distant codon influences.

Recently, the Mamba model has emerged as a promising alternative to Transformer-based architectures, demonstrating significant advantages in modeling long-range dependencies without incurring quadratic computational costs (Gu and Dao 2023; Dao and Gu 2024). Mamba has outperformed Transformer-based models in DNA sequencing problems (Schiff et al. 2024), showcasing its potential for mRNA modeling. However, despite its strengths in handling long sequences, Mamba struggles with effectively focusing on important proximal codons from long mRNA sequences. Consequently, applying Mamba to translation elongation rate prediction without modifications is far from sufficient. Therefore, a mechanism to further enhance Mamba's focus on codons near the A site is necessary.

Inspired by the natural progression of ribosomes along mRNA sequences, which protect approximately 28 to 29 nucleotides, and recognizing the significant influence of these protected regions on ribosome density predictions at the A site (Tunney et al. 2018), we propose a sliding window mechanism to model codons near the A site during translation elongation. This natural alignment between biological processes and computational techniques offers an effective approach for capturing localized effects on translation dynamics. The window constraint allows us to incorporate local inductive bias (Pan et al. 2023) while preserving Mamba's efficiency in processing long-range dependencies without quadratic complexity. As ribosomes move, the composition of codons within the window continuously evolves, presenting a challenge in accurately capturing the changing relevance of each codon. To alleviate this dilemma, we further augment the sliding window with an attention-based mechanism, which dynamically learns to focus on the most relevant codons within each window, allowing the model to adaptively weigh the importance of proximal codons as the ribosome progresses, thereby capturing the changing critical codons for accurate translation elongation prediction.

In this work, we introduce SWAMamba, a novel framework designed to emphasize codons near the A site while capturing the long-range influence of mRNA codons to predict translation elongation rates accurately. Our approach innovatively combines the Mamba model's capability to model long-range codon effects with a sliding window attention mechanism that focuses on proximal codons at the A site, enabling a comprehensive consideration of both proximal and distal codon impacts. Additionally, we implement a dual-stage attention mechanism. The first stage focuses on the fundamental properties of codons near the A site, while the second stage refines the representation by re-emphasizing relevant codon features proximal to the A site in light of the global context. To the best of our knowledge, this work presents the first integration of the Mamba model with sliding window attention for biological sequence modeling, making it particularly suitable for predicting mRNA translation elongation rates. The major contributions of our work are as follows:

- We introduce SWAMamba, a novel framework integrating Mamba's long-range dependency modeling with a biologically-inspired sliding window attention mechanism, enhancing translation elongation rate prediction by capturing both distal and proximal codon effects.
- We propose a sliding window attention mechanism to effectively balance proximal and distal codon influences on ribosomal A site, addressing the trade-off between long-range and local effects in translation dynamics.
- The comprehensive experiments across four species show SWAMamba significantly outperforms state-of-the-art methods, offering a powerful tool for investigating mRNA translation dynamics and protein synthesis.

Related Work

Translation Elongation Rates Problem Prediction. O'Connor, Andreev, and Baranov (2016) introduces a robust normalization method RUST for Ribo-seq data, which prediction ribosome footprint density and reveals sequence features that affect ribosome footprint density. Liu and Song (2016) presents a kernel smoothing method to predict ribosome footprint profiles from transcript sequences in yeast datasets and suggests diverse applications. Zhang et al. (2017) introduces ROSE, a deep learning framework for analyzing ribosome profiling data to predict ribosome stalling events, to further decipher the complex regulatory mechanisms behind the translation elongation dynamics encoded in mRNA sequences. Tunney et al. (2018) presents a neural network model predicting ribosome density based on codon sequence, revealing factors influencing translation elongation and demonstrating its application in protein expression control. Hu et al. (2021) introduces Riboexp, a deep reinforcement learning framework for predicting ribosome density, and demonstrates practical applications

in codon optimization for increased protein production. Tian et al. (2021) presents RiboMIMO, a deep learning approach for modeling full-length CDS ribosome density distributions, outperforming existing methods and revealing long-range codon impacts on translation elongation. Shao et al. (2024) introduces Riboformer, a transformer-based deep learning model for predicting ribosome densities, and correcting experimental artifacts. Despite these advancements, existing methods can't effectively utilize both proximal and distal codon effects at the ribosomal A site.

Structured State-Space Sequence Model. Gu, Goel, and Ré (2021) proposes a Structured State-Space Sequence (S4) model, which is used to handle long-range dependency and is an alternative to transformer (Gu et al. 2022; Zhu et al. 2024). Gu and Dao (2023) introduces Mamba, a new neural network architecture using selective state space models, outperforming transformers in efficiency and performance across various sequence modeling tasks. Schiff et al. (2024) introduces Caduceus, a reverse complement equivariant bi-directional long-range DNA language model based on MambaDNA, outperforming larger models in long-range genomic sequence tasks. Despite these advancements and the enhanced capabilities of existing models for the S4 model, they still face challenges in effectively addressing localized issues that require concentrated attention.

Problem Formulation and Method

In this section, we first define the problem of predicting translation elongation rates. Then, we describe in detail the key components of the SWAMamba framework.

Problem Formulation

The mRNA translation elongation rates prediction problem involves using an mRNA sequence $s = (x_1, x_2, ..., x_T)$ of length T to predict ribosome density at each codon, denoted as $y = (v_1, v_2, ..., v_T)$. Here, x_i represents a codon in the mRNA sequence, and v_i represents the ribosome density at codon x_i. We denote the set of mRNA sequences as $\mathcal{S} = \{s_1, s_2, ..., s_n\}$ and the corresponding ribosome densities as $\mathcal{Y} = \{y_1, y_2, ..., y_n\}$. Here, n represents the number of mRNA sequences. Together, these sets $\mathcal{S}$ and $\mathcal{Y}$ form our dataset $\mathcal{D} = \{(s_1, y_1), (s_2, y_2), ..., (s_n, y_n)\}$.

Overview of SWAMamba

We present the SWAMamba framework, a novel approach for predicting translation elongation rates in mRNA sequences. The main workflow of our algorithm is illustrated in Figure 2A. The model takes an mRNA sequence as input, where codons, nucleotides, and amino acids are identified and encoded. These three encoding forms constitute the fused encoding, which is processed by our SWAMamba module to capture both proximal and distal codon effects. Finally, a fully connected layer predicts ribosome density for each codon in the mRNA sequence. The ribosome density value reflects the rate of translation at that codon. Figure 2B shows the architecture of the SWAMamba module. The process begins by capturing the properties of codons near the A site using sliding window attention. These features are then processed by the Bidirectional Mamba (BiMamba) (Schiff et al. 2024) module to capture the distal codons effect. Sliding window attention is applied again to the BiMamba output to re-emphasize relevant proximal codons features. This iterative approach allows the model to refine its understanding of the mRNA sequence, moving from proximal to distal codon effects and then back to proximal information. In the following sections, we provide a detailed explanation of each component of our framework.

mRNA Sequence Encoding. To effectively model ribosome density, a comprehensive encoding of mRNA sequences is essential. mRNA sequences are composed of codons, each consisting of three nucleotides selected from A, U, C, and G. This arrangement results in 64 possible codon combinations. To represent each codon x_i, We represent each codon x_i using one-hot encoding, denoted as $e_i^{codon} \in \{0,1\}^{64}$. However, this codon-level encoding alone is insufficient to capture nucleotide-level differences and the relationships between encoded amino acids. Following previous work (Tian et al. 2021), we incorporate additional encodings for both nucleotides and amino acids to enable more fine-grained representation. For nucleotide encoding, we use one-hot encoding for each of the four nucleotides and concatenate the three nucleotide encodings within a codon, resulting in $e_i^{nt} \in \{0,1\}^{(4\times3)}$. To capture amino acid properties, we encode the 20 standard amino acids and a stop codon, represented as $e_i^{aa} \in \{0,1\}^{21}$. The fused encoding for each codon x_i is obtained by concatenating these three representations:

$$e_i = \text{concat}([e_i^{codon}, e_i^{nt}, e_i^{aa}]) \in \mathbb{R}^{97}. \quad (1)$$

The complete encoded mRNA sequence is thus represented as $s_e = (e_1, e_2, ..., e_T)$, This comprehensive encoding strategy captures the full spectrum of sequence information from nucleotide to codon to amino acid level.

Focusing on Proximal Codons' Effects with Sliding Window Attention. Drawing inspiration from the well-established importance of codons proximal to the ribosomal A site in translation elongation (Hu et al. 2021; Tunney et al. 2018), we incorporate this crucial biological insight through an innovative sliding window attention mechanism, emulating the process of ribosome translation elongation. This approach is specifically designed to capture and emphasize the effects of codons near the A site. Specifically, for each codon e_i in the sequence, we define a fixed-size attention window of width w, enabling the codon to attend to $\frac{1}{2}w$ tokens on either side. We then apply three distinct linear transformations to obtain query, key, and value vectors: $Q = s_e W^Q$, $K = s_e W^K$, $V = s_e W^V$, reflecting the standard attention mechanism paradigm but tailored to our sliding window context. The attention scores for codon e_i for its neighboring codons within the window are calculated as:

$$a_{ij} = \frac{(Q_i K_j^T)}{\sqrt{d_k}}, \quad (2)$$

where $j \in [i - w/2, i + w/2]$ and d_k is the key dimension. The weighted sum of attention within the sliding window at

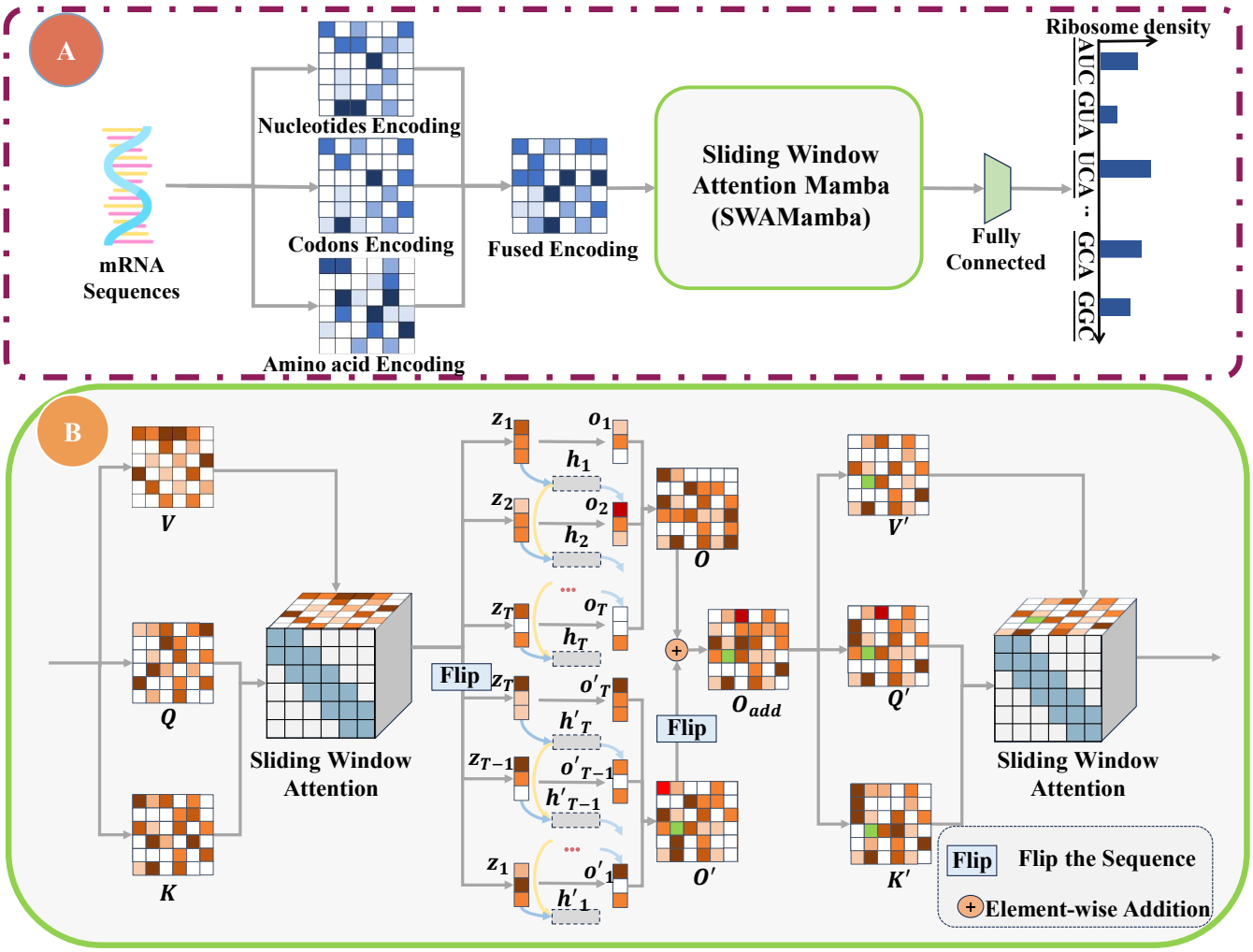

Figure 2: The main workflow of SWAMamba. (A) This figure shows how SWAMamba works. It takes an mRNA sequence as input and outputs the ribosome density for each codon, which reflects the translation elongation rates. (B) The diagram illustrates the Sliding Window Attention Mamba process, focusing on both proximal and distal codons relative to the A site.

position i is then computed as:

$$z_i = \sum_{j \in [i-w/2, i+w/2]} softmax(a_{ij}) V_j. \quad (3)$$

This formulation allows the model to dynamically weigh the importance of each codon proximal to the A site. This operation aggregates information from proximal codons, resulting in a new more representative codon feature representation $Z = \{z_1, z_2, ..., z_T\}$. By incorporating this sliding window attention mechanism, our model gains the ability to dynamically focus on the most relevant proximal codons, enhancing its capacity to capture translation dynamics.

Capturing Distal Codon Effects with BiMamba Modeling. While proximal codons play a crucial role in translation elongation, the influence of distal codons on the A site cannot be overlooked. To comprehensively model these long-range dependencies, we introduce a Mamba module, extending our model's capacity to capture complex, distant interactions within the mRNA sequence. The Mamba module, based on the State Space Model (SSM), maps input $z_t \in \mathbb{R}^D$ to output $o_t \in \mathbb{R}^D$ through an intermediate representation $h_t \in \mathbb{R}$, where D is the dimension of z_t and o_t. For discrete data inputs, the Mamba module can be expressed as:

$$\boldsymbol{h}_{t+1} = \overline{\boldsymbol{A}} h_t + \overline{\boldsymbol{B}} z_t, \quad o_{t+1} = \boldsymbol{C} h_t + \boldsymbol{D} z_t, \quad (4)$$

where $\boldsymbol{C}$ and $\boldsymbol{D}$ are system parameters, and $\overline{\boldsymbol{A}}$ and $\overline{\boldsymbol{B}}$ are discretized versions of continuous parameters $\boldsymbol{A}$ and $\boldsymbol{B}$:

$$\overline{\boldsymbol{A}} = \exp(\Delta \boldsymbol{A}), \quad \overline{\boldsymbol{B}} = \boldsymbol{A}^{-1}(\exp(\Delta \boldsymbol{A}) - \boldsymbol{I})\boldsymbol{B}, \quad (5)$$

where Δ is an additional time scale parameter. A selective SSM (Gu and Dao 2023) is incorporated into the models by allowing their parameters to be input-dependent, enabling $\boldsymbol{B}$, $\boldsymbol{C}$, and Δ to depend on the input z_t:

$$\begin{aligned} \boldsymbol{B}_t &= \text{Linear}_{\boldsymbol{B}}(z_t) \quad \boldsymbol{C}_t = \text{Linear}_{\boldsymbol{C}}(z_t) \\ \Delta_t &= \text{softplus}(\text{Linear}_{\Delta}(z_t)), \end{aligned} \quad (6)$$

where $\mathrm{Linear}(\cdot)$ represents a linear projection and $\mathrm{softplus}(\cdot) = \log(1 + \exp(\cdot))$.

We input the sequence representation $Z = \{z_1, z_2, ..., z_T\}$ into the selective SSM to obtain outputs $O = \{o_1, o_2, ..., o_T\}$ that aggregate upstream codon information. However, this approach only captures information from upstream codons in a unidirectional manner. Given that downstream codons also influence the translation elongation rate at the A site, we introduce a BiMamba to capture bidirectional mRNA sequence, crucial for understanding both upstream and downstream influences on translation. We subsequently construct $Flip(Z) = \{z_T, z_{T-1}, ..., z_1\}$. After processing by the Mamba model, we obtain outputs $O' = \{o'_T, o'_{T-1}, ..., o'_1\}$. We then fuse this bidirectional information using element-wise addition:

$$O_{add} = O + Flip(O') = \{o_1 + o'_1, o_2 + o'_2, ..., o_T + o'_T\}. \tag{7}$$

By incorporating this BiMamba module, our model gains the ability to capture and integrate long-range dependencies in the mRNA sequence, thereby providing a more comprehensive and detailed understanding of the factors affecting translation elongation at the A site.

Context-Informed Local Refinement with Sliding Window Attention. The final stage of our SWAMamba framework represents a crucial synthesis of global and local information, allowing for a more nuanced interpretation of proximal codon effects in light of the broader sequence context. Having captured long-range dependencies through the BiMamba module, we now return our focus to the local environment of each codon. This second application of sliding window attention operates on globally informed representations, enabling a context-rich analysis of local features. This approach is motivated by the fact that while distant codons can influence translation, the immediate context of the A site remains critically important. We apply three new linear transformations to O_{add} to obtain query, key, and value vectors: $Q' = O_{add}W'^{Q}$, $K' = O_{add}W'^{K}$, $V' = O_{add}W'^{V}$, similar to the first stage of sliding window attention, we focus on a window of size w around each position. The weighted sum of attention within this window at position i is computed as:

$$m_i = \sum_{j \in [i-w/2, i+w/2]} \mathrm{softmax}\left(\frac{Q'_i {K'_j}^T}{\sqrt{d_k}}\right) V'_j \tag{8}$$

This process yields a refined local codon feature representation $M = \{m_1, m_2, ..., m_T\}$, which encapsulates both the proximal codon effects and the relevant global context for each codon. The final step in our prediction pipeline involves processing each refined feature vector m_i through a fully connected neural network. This network maps the multi-scale features captured by our model to a scalar value representing the predicted ribosome density at each codon position. The output of this network is a sequence of predicted ribosome densities $y' = (v'_1, v'_2, ..., v'_T)$, where each v'_i corresponds to the predicted density at the i-th codon. The time and space complexity of sliding window attention is $O(n \cdot w)$ (Beltagy, Peters, and Cohan 2020), avoiding the $O(n^2)$ complexity of the self-attention mechanism in transformers. By incorporating sliding window attention, the Mamba model still maintains its efficient time and space complexity. The SWAMamba framework uses Mean Squared Error as the loss function to minimize the difference between predicted and observed ribosome densities (Appendix B.3 for details).

Experiments

In this section, we demonstrate SWAMamba's effectiveness in translation elongation rate prediction across four species.

Baseline

- **IXnos** (Jan et al. 2018) models translation elongation using a feed-forward neural network to predict ribosome density for each codon based on codons near the A site.
- **Riboexp** (Hu et al. 2021) uses policy networks in reinforcement learning to perform context-sensitive feature selection near the A site for ribosome density modeling.
- **RiboMIMO** (Tian et al. 2021) utilizes two layers of the Bi-GRU network to model full-length mRNA CDS regions for predicting ribosome density.
- **Riboformer** (Shao et al. 2024) uses near the A site codons with a transformer model to predict ribosome density.

Performance Comparison

We evaluated SWAMamba on datasets from E. coli (Mohammad, Green, and Buskirk 2019), S. cerevisiae (Stein et al. 2022), C. elegans (Stein et al. 2022), and Humans (Iwasaki, Floor, and Ingolia 2016) using five-fold cross-validation (data details in Appendix A[1]). Table 1 shows the performance measured using the average pearson correlation coefficients between predicted and experimental ribosome densities. SWAMamba consistently outperformed baseline methods across four species, due to its powerful ability to capture distal codons while maintaining a focus on proximal codons. It is noteworthy that Riboformer's performance declined significantly compared to its original report, primarily due to the limitations of the data-splitting method employed in the original Riboformer study, which oversimplified the task. To keep a more fair setting consistent with Inxnos and Riboexp (Tunney et al. 2018; Hu et al. 2021), we applied a more appropriate and widely accepted data splitting method for Riboformer. We introduce this method in Appendix B.1 and outline the limitations of the original Riboformer data-splitting approach in Appendix B.2. To enable a more in-depth comparison, we modified Riboformer to model the entire mRNA CDS, as the original version only considers codons near the A site. This modification allowed us to evaluate the performance of a Transformer-based model on full-length mRNA CDS. Although this adjustment improved Riboformer's performance, it still fell short of SWAMamba, which better captures long-range sequence dependencies due to its Mamba architecture. Full results and im-

[1]https://github.com/reset001/SWAMambaappendix.

	C. elegans	S. cerevisiae	Humans	E. coli
IXnos	0.504	0.502	0.507	0.549
Riboexp	0.531	0.587	0.572	0.566
RiboMIMO	0.576	0.612	0.607	0.572
Riboformer	0.495	0.482	0.476	0.552
SWAMamba	**0.596**	**0.630**	**0.641**	**0.640**

Table 1: Comparison of average pearson correlation coefficients between predicted and experimental ribosome densities across methods using four datasets.

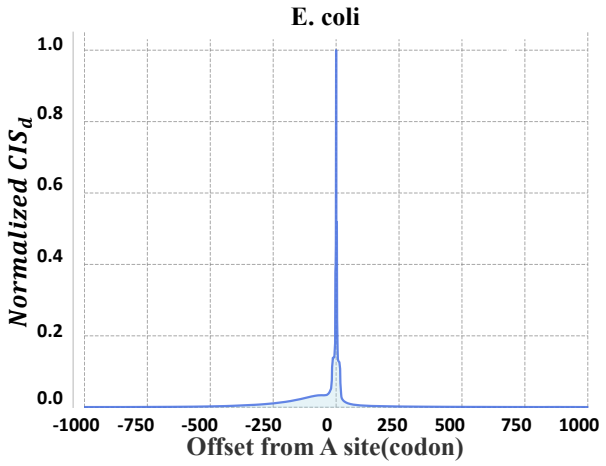

Figure 3: Relationship between codon contributions and distance on predicted ribosomal density at the A site in E. coli.

plementation details are provided in Appendix C.2 and B.5, respectively. Details on SWAMamba are in Appendix B.4, and the experimental setup is in Appendix B.6.

SWAMamba Enhances Codon Attention Near the A Site

To evaluate the impact of codons on prediction, we used the Codon Impact Score (CIS) (Tian et al. 2021), which quantifies the effect of a codon feature at position i on the predicted ribosome density at position j in the same gene, defined as:

$$CIS_{ij} = \frac{\partial y_j}{\partial e_i}, \tag{9}$$

where e_i represents the input codon feature at position i, and y_j denotes the ribosome density at position j predicted by SWAMamba. To evaluate how the distance between codons at positions i and j, with a relative distance $d = i - j$, affects the predicted ribosome density at position j, we utilized CIS_d (Tian et al. 2021). Specifically, CIS_d quantifies the contribution of the distance d between codons to the ribosome density prediction at position j, and is defined as:

$$CIS_d = \frac{\sum_{i,j} |CIS_{i,j}| \cdot \mathbb{I}(i - j = d)}{\sum_{i,j} \mathbb{I}(i - j = d)}, \tag{10}$$

where $\mathbb{I}(i - j = d)$ equals 1 if $i - j = d$ and 0 otherwise.

To conduct a more comprehensive evaluation, Figure 3 analyzes the relationship between codon contributions and distance on predicted ribosomal density at the A site in E. coli, with similar trends observed for other species (details in Appendix C.1). The horizontal axis represents the offset from the input codon to the A site codon, while the vertical axis denotes the normalized CIS_d value, which represents the extent of influence on the predicted ribosome density at the A site. These findings align with previous research, indicating that codons surrounding the A site typically dominate ribosome density prediction (Tunney et al. 2018; Tian et al. 2021). We further examined codons near the A site, as depicted in Figure 4, using normalized CIS_d for comparison. This analysis highlights the critical importance of these nearby codons in predicting ribosome density at the A site and demonstrates the effect of our sliding window attention mechanism. Graphs a-d in Figure 4 present zoomed-in views of codon contributions to predicted ribosome density around the A site across four datasets. These magnified views underscore the significance of nearby codons in predicting ribosome density at the A site. Graphs e-h compare codon contributions near the A site with and without sliding window attention. As evident in these graphs, the model with sliding window attention (blue line) focuses more on codons near the A site compared to the model without this mechanism (green line). This visualization effectively illustrates how the sliding window attention mechanism enhances the model's ability to capture the importance of codons near the A site, thereby improving the prediction of ribosome density.

SWAmamba Captures Distal Codon Influences

Although the codons' contribution diminishes with distance in both directions, codons located far from the A site can still influence the predicted ribosome density. To quantify this effect, we used CIS_d to assess the contribution at a distance d from the A site. We randomly selected a CDS sequence from E. coli and randomly designated position 97 as the A site. Figure 5 illustrates the CIS_d values at different positions relative to the predicted ribosome density at position $j = 97$. The influence of positions around 80 codons away from the A site appears as strong as codons near it. The results demonstrate that codons distal from the A site can indeed influence the predicted ribosome density over extended ranges, corroborating previous observations (Tian et al. 2021). This finding shows that our model effectively captures distal effects between codons.

Ablation Experiment

To evaluate the effectiveness of our sliding window attention mechanism and the inclusion of amino acid and nucleotide features in improving the model's prediction accuracy, we conducted an ablation study. Table 2 presents the average pearson correlation coefficients between the predicted and experimental ribosome densities. The results show that the dual-stage sliding window attention mechanism improves performance compared to the model without it. We further ablated the models by removing the first sliding window attention (1^{st} atten) and second sliding window attention (2^{nd} atten), demonstrating that dual-stage attention yields better results. Additionally, we ablated the models by excluding

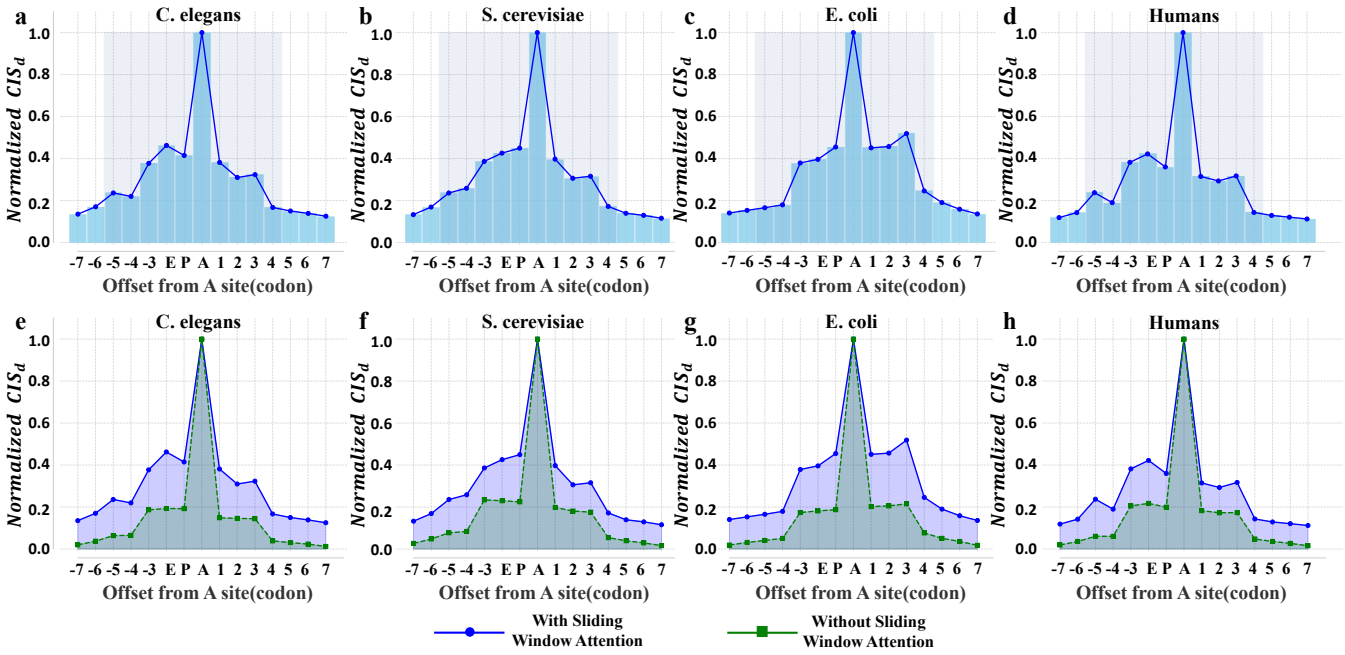

Figure 4: Graphs a-d show codon contributions to ribosome density prediction at the A site. The shading indicates the ribosome-protected region, which is considered to have a significant influence on ribosome density predictions at the A site. Graphs e-h compare these contributions with (blue line) and without (green line) sliding window attention.

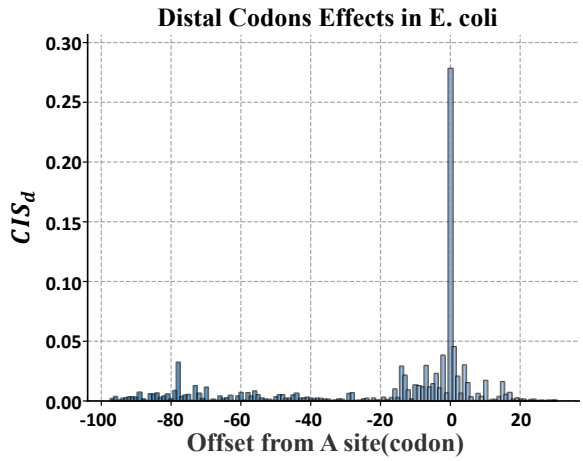

Figure 5: Codons impact on ribosome density prediction in E. coli. The graph illustrates CIS_d values for codons relative to A site (position 97) in a randomly chosen CDS sequence.

amino acid and nucleotide features, finding that these features also enhance prediction accuracy.

Conclusion and Future Work

In this paper, we introduce SWAMamba, a novel method for predicting translation elongation rates. SWAMamba builds on Mamba's strength in capturing long-range dependencies by effectively accounting for the influence of distant codons on the A site. Additionally, it integrates a sliding window attention mechanism to emphasize the role of nearby codons at the A site. As a result, SWAMamba effectively addresses both proximal and distant codon effects on the A site. To evaluate SWAMamba's effectiveness, we compared it with state-of-the-art methods in ribosome density prediction tasks across four species. Our analysis confirmed that SWAMamba can capture both proximal and distal codon effects at the A site simultaneously. Furthermore, ablation experiments showed that the sliding window attention mechanism significantly improved SWAMamba's performance. Overall, SWAMamba captures complex relationships between mRNA sequence features and ribosome densities, ultimately providing a powerful tool for investigating the intricate dynamics of mRNA translation to understand and modulate protein synthesis. In future work, we plan to refine SWAMamba further and apply it to a broader range of species and experimental conditions to enhance its prediction accuracy and generalization ability.

	C. elegans	S. cerevisiae	Humans	E. coli
SWAMamba	**0.596**	**0.630**	**0.641**	**0.640**
w/o dual atten	0.593	0.621	0.638	0.626
w/o 1^{st} atten	0.589	0.612	0.638	0.629
w/o 2^{nd} atten	0.591	0.627	**0.641**	0.639
w/o nucleotide	0.590	0.622	0.630	0.632
w/o amino acid	0.591	0.626	0.636	**0.640**

Table 2: Comparison of pearson correlation coefficients between predicted and experimental ribosome density values with and without different modules.

Acknowledgments

The work is supported by the National Natural Science Foundation of China (62072376), Guangdong Basic and Applied Basic Research Foundation (2022A1515010144), Innovation Capability Support Program of Shaanxi (2022KJXX-75). Figure 1 was created in BioRender. Zeng, X. (2025) https://BioRender.com/z40s368.

References

Al-Selwi, S. M.; Hassan, M. F.; Abdulkadir, S. J.; and Muneer, A. 2023. LSTM inefficiency in long-term dependencies regression problems. *Journal of Advanced Research in Applied Sciences and Engineering Technology*, 30(3): 16–31.

Andreev, D. E.; O'Connor, P. B.; Loughran, G.; Dmitriev, S. E.; Baranov, P. V.; and Shatsky, I. N. 2017. Insights into the mechanisms of eukaryotic translation gained with ribosome profiling. *Nucleic acids research*, 45(2): 513–526.

Beltagy, I.; Peters, M. E.; and Cohan, A. 2020. Longformer: The long-document transformer. *arXiv preprint arXiv:2004.05150*.

Brar, G. A. 2016. Beyond the triplet code: context cues transform translation. *Cell*, 167(7): 1681–1692.

Dao, T.; and Gu, A. 2024. Transformers are SSMs: Generalized models and efficient algorithms through structured state space duality. *arXiv preprint arXiv:2405.21060*.

Gu, A.; and Dao, T. 2023. Mamba: Linear-time sequence modeling with selective state spaces. *arXiv preprint arXiv:2312.00752*.

Gu, A.; Goel, K.; Gupta, A.; and Ré, C. 2022. On the parameterization and initialization of diagonal state space models. *Advances in Neural Information Processing Systems*, 35: 35971–35983.

Gu, A.; Goel, K.; and Ré, C. 2021. Efficiently modeling long sequences with structured state spaces. *arXiv preprint arXiv:2111.00396*.

Hatamizadeh, A.; and Kautz, J. 2024. MambaVision: A Hybrid Mamba-Transformer Vision Backbone. *arXiv preprint arXiv:2407.08083*.

Hu, H.; Liu, X.; Xiao, A.; Li, Y.; Zhang, C.; Jiang, T.; Zhao, D.; Song, S.; and Zeng, J. 2021. Riboexp: an interpretable reinforcement learning framework for ribosome density modeling. *Briefings in Bioinformatics*, 22(5): bbaa412.

Ingolia, N. T. 2014. Ribosome profiling: new views of translation, from single codons to genome scale. *Nature reviews genetics*, 15(3): 205–213.

Ingolia, N. T. 2016. Ribosome footprint profiling of translation throughout the genome. *Cell*, 165(1): 22–33.

Ingolia, N. T.; Ghaemmaghami, S.; Newman, J. R.; and Weissman, J. S. 2009. Genome-wide analysis in vivo of translation with nucleotide resolution using ribosome profiling. *science*, 324(5924): 218–223.

Ingolia, N. T.; Hussmann, J. A.; and Weissman, J. S. 2019. Ribosome profiling: global views of translation. *Cold Spring Harbor perspectives in biology*, 11(5): a032698.

Iwasaki, S.; Floor, S. N.; and Ingolia, N. T. 2016. Rocaglates convert DEAD-box protein eIF4A into a sequence-selective translational repressor. *Nature*, 534(7608): 558–561.

Jan, A.; Jansonius, B.; Delaidelli, A.; Bhanshali, F.; An, Y. A.; Ferreira, N.; Smits, L. M.; Negri, G. L.; Schwamborn, J. C.; Jensen, P. H.; et al. 2018. Activity of translation regulator eukaryotic elongation factor-2 kinase is increased in Parkinson disease brain and its inhibition reduces alpha synuclein toxicity. *Acta neuropathologica communications*, 6: 1–17.

Liu, T.-Y.; and Song, Y. S. 2016. Prediction of ribosome footprint profile shapes from transcript sequences. *Bioinformatics*, 32(12): i183–i191.

Mohammad, F.; Green, R.; and Buskirk, A. R. 2019. A systematically-revised ribosome profiling method for bacteria reveals pauses at single-codon resolution. *Elife*, 8: e42591.

O'Connor, P. B.; Andreev, D. E.; and Baranov, P. V. 2016. Comparative survey of the relative impact of mRNA features on local ribosome profiling read density. *Nature communications*, 7(1): 12915.

Pan, X.; Ye, T.; Xia, Z.; Song, S.; and Huang, G. 2023. Slide-transformer: Hierarchical vision transformer with local self-attention. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2082–2091.

Schiff, Y.; Kao, C. H.; Gokaslan, A.; Dao, T.; Gu, A.; and Kuleshov, V. 2024. Caduceus: Bi-Directional Equivariant Long-Range DNA Sequence Modeling. In *International Conference on Machine Learning*, 43632–43648. PMLR.

Shao, B.; Yan, J.; Zhang, J.; Liu, L.; Chen, Y.; and Buskirk, A. R. 2024. Riboformer: a deep learning framework for predicting context-dependent translation dynamics. *Nature communications*, 15(1): 2011.

Stein, K. C.; Morales-Polanco, F.; van der Lienden, J.; Rainbolt, T. K.; and Frydman, J. 2022. Ageing exacerbates ribosome pausing to disrupt cotranslational proteostasis. *Nature*, 601(7894): 637–642.

Tian, T.; Li, S.; Lang, P.; Zhao, D.; and Zeng, J. 2021. Full-length ribosome density prediction by a multi-input and multi-output model. *PLOS Computational Biology*, 17(3): e1008842.

Tunney, R.; McGlincy, N. J.; Graham, M. E.; Naddaf, N.; Pachter, L.; and Lareau, L. F. 2018. Accurate design of translational output by a neural network model of ribosome distribution. *Nature structural & molecular biology*, 25(7): 577–582.

Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A. N.; Kaiser, Ł.; and Polosukhin, I. 2017. Attention is all you need. *Advances in neural information processing systems*, 30.

Zhang, S.; Hu, H.; Zhou, J.; He, X.; Jiang, T.; and Zeng, J. 2017. Analysis of ribosome stalling and translation elongation dynamics by deep learning. *Cell systems*, 5(3): 212–220.

Zhao, F.; Yu, C.-h.; and Liu, Y. 2017. Codon usage regulates protein structure and function by affecting translation elongation speed in Drosophila cells. *Nucleic acids research*, 45(14): 8484–8492.

Zhu, L.; Liao, B.; Zhang, Q.; Wang, X.; Liu, W.; and Wang, X. 2024. Vision mamba: Efficient visual representation learning with bidirectional state space model. *arXiv preprint arXiv:2401.09417*.