A Dataset Details

A.1 Data Availability

To evaluate SWAMamba's performance in predicting ribosome density, we utilized datasets from Shao et al. (2024), including those from E. coli (Mohammad, Green, and Buskirk 2019), S. cerevisiae (Stein et al. 2022), and C. elegans (Stein et al. 2022). These datasets were processed using a lysis buffer containing high magnesium concentrations, which addresses issues encountered with traditional methods. Specifically, traditional approaches that use chloramphenicol (Cm) buffer to halt elongation can alter translation and cause pauses at Ser and Gly codons (Marks et al. 2016). The high-magnesium method mitigates these problems (Mohammad, Green, and Buskirk 2019), providing a clearer view of the in vivo translational landscape (Shao et al. 2024). Additionally, we incorporated a human dataset (Iwasaki, Floor, and Ingolia 2016) for a more comprehensive evaluation of SWAMamba. Ribosome profiling datasets for four species were obtained from publicly available sequencing data from the Gene Expression Omnibus. The dataset for E. coli is under the accession number GSE77617. The datasets for S. cerevisiae and C. elegans are under the accession number GSE152850. These datasets were used for model training by Shao et al. (2024). The human datasets are under the accession numbers SRR2075925 and SRR2075926 and were used for model training by Tunney et al. (2018) and Hu et al. (2021). All data processing adhered to established protocols from previous research.

A.2 Data Analysis

We further analyzed the composition of our dataset. Table 3 shows the number of mRNA sequences in the dataset and the range of mRNA sequence lengths. Figure 6 shows the mRNA sequence length distribution across the four datasets. The x-axis represents the sequence length, and the y-axis indicates the count of mRNA sequences.

B Models Details and Experiment Settings

B.1 Data Division Considering Codons Near the A Site in Traditional Models

In predicting ribosome density through the entire coding CDS region of the mRNA sequence, we consider a model that inputs an mRNA sequence $s = (x_1, x_2, ..., x_T)$ of length T to predict ribosome density at each codon, denoted as $y = (v_1, v_2, ..., v_T)$. Traditional models, such as those used in Riboexp (Hu et al. 2021) and iXnos (Tunney et al. 2018), focus on codons near the A site by setting a window size w, with the input defined as $l_t^i = (x_{t-w/2}^i, x_{t-w/2+1}^i, ..., x_t^i, x_{t+1}^i, ..., x_{t+w/2}^i)$ and the output as v_t^i, where i refers to the i_{th} mRNA and t represents the t_{th} codon on the i_{th} mRNA. The data for an mRNA sequence is expressed as $mRNA_i = \{(l_1^i, v_1^i), (l_2^i, v_2^i), ..., (l_T^i, v_T^i)\}$, and a dataset as:

$$\mathcal{D} = \{mRNA_1, mRNA_2, ..., mRNA_n\}, \tag{12}$$

where n is the number of mRNA sequences. Most methods considering codons near the A site divide training and test sets by entire mRNA sequences. Specifically, a part of the mRNA sequences is divided into a training set, and another part of the mRNA sequences is divided into a test set.

B.2 Limitations of Data Division in Riboformer

Compared to the traditional data division method described in Appendix B.1, Riboformer uses:

$$\mathcal{D}_{Riboformer} = \{(l_1^1, v_1^1), (l_2^1, v_2^1), ..., (l_T^1, v_T^1), ..., (l_t^i, v_t^i), (l_{t+1}^i, v_{t+1}^i), ..., (l_T^n, v_T^n)\}, \tag{13}$$

which divides the dataset into separate fragments rather than using entire mRNA sequences. This approach oversimplifies the task. The feature l_t^i is composed of w codons within the window, and the features of adjacent codons, l_t^i and l_{t+1}^i, differ by only one codon, resulting in highly similar features. As the codons are adjacent, v_t^i and v_{t+1}^i are relatively close, as codons with similar characteristics and ribosome densities may appear in both the training and test sets. While the model's prediction performance improves when it has already encountered the ribosome densities of adjacent codons within the same mRNA sequence and uses this information to train the model, it diminishes the model's ability to accurately predict unseen mRNA sequences. We argue that dividing datasets by entire mRNA sequences is more reasonable and universally applicable. To maintain a fair comparison consistent with IXnos (Tunney et al. 2018) and Riboexp (Hu et al. 2021), we divided the training and test sets based on the entire mRNA sequences. As a result, the Riboformer's performance dropped significantly, demonstrating that the Riboformer's data partitioning method indeed oversimplifies the ribosome density prediction problem. Therefore, we adopted the same data partitioning approach as Riboexp and IXnos to more fairly evaluate the model's performance in predicting ribosome density.

	C. elegans	S. cerevisiae	Humans	E. coli
Number	3076	2265	2813	1145
Length Interval	30∼4576	33∼2241	103∼5144	68∼1521

Table 3: The number and length range of the mRNA sequences in the dataset.

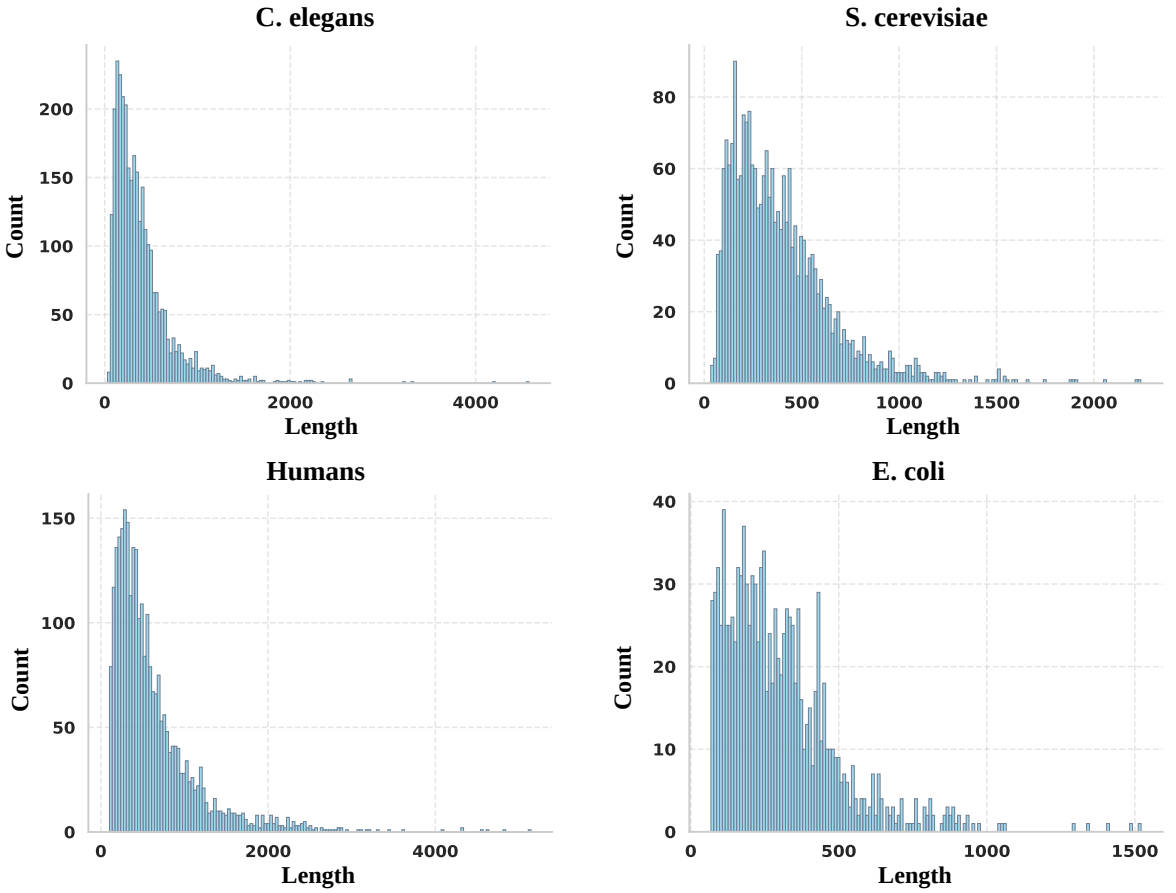

Figure 6: Distribution of sequence lengths across four datasets.

B.3 Optimization and Loss Function

For training the SWAMamba framework, we adopt Mean Squared Error (MSE) as a loss function to minimize the difference between predicted and observed ribosome densities, defined as:

$$\mathcal{L}_{MSE}(y, y') = \frac{1}{n}\sum_{i=1}^{n}(y_i - y_i')^2, \tag{14}$$

where $y = (v_1, v_2, ..., v_n)$ represents the observed ribosome density labels, and $y' = (v_1', v_2', \ldots, v_n')$ denotes the model's predicted output. Each v_i and v_i' correspond to the observed and predicted ribosome densities at the i-th codon position along the mRNA sequence, respectively.

B.4 SWAMamba Details

The model parameters were consistently applied across all datasets. The learning rate was fixed at 0.001, and a batch size of 4 was employed. The sliding window attention mechanism utilized a window size of 32. The model architecture incorporated 11 Mamba layers, and the maximum sequence length was set to 1024 codons. Given that each codon consists of three nucleotides, this corresponds to a maximum mRNA sequence length of 3072 nucleotides. As evident from Figure 6, this maximum length encompassed the vast majority of sequences in the dataset, with only a few exceptionally long mRNA sequences exceeding this threshold. In cases where the input sequence was shorter than the maximum length, padding was applied; conversely, sequences exceeding this length were truncated.

B.5 Riboformer Details

Riboformer. The initial Riboformer framework necessitated the input of ribosome profiling data processed using chloramphenicol buffer, along with mRNA sequence, to predict ribosome profiling data processed by the high-magnesium method.

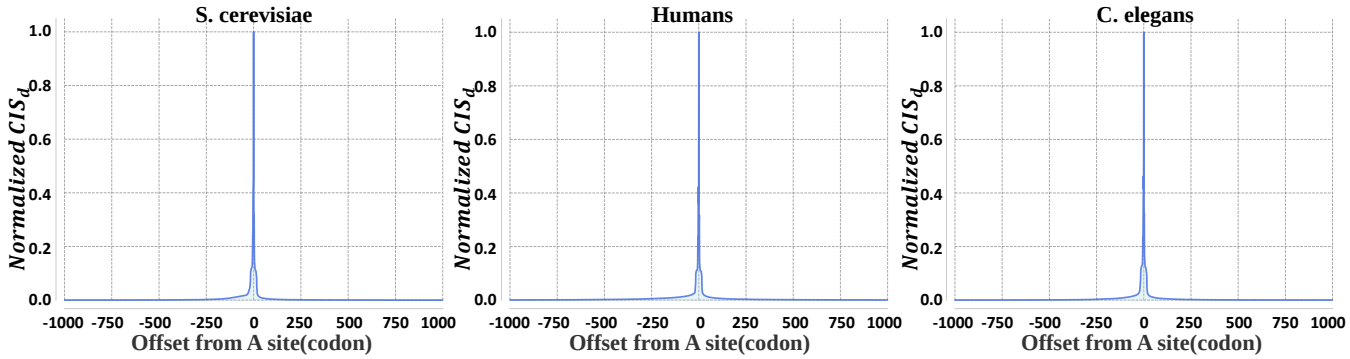

Figure 7: The relationship between codon contributions and distance on predicted ribosomal density at the A site is examined in S. cerevisiae, humans, and C. elegans. The figure shows the average contribution of codons at a distance d from the A site in all mRNA sequences in the dataset.

However, Obtaining both chloramphenicol buffer-processed data and high-magnesium method data is inconsistent across species, limiting the model's broad applicability. To ensure a fair comparison of all models, we adhered to the original parameters provided in the Riboformer code, modifying the input to accept only mRNA sequences. Additionally, we abandoned Riboformer's data division methodology, which is detailed in Appendix B.2. Instead, we implemented a reasonable and universally applicable data division approach, aligning with established practices in the field described in Appendix B.1.

Riboformer (Entire CDS). To facilitate a direct comparison between the transformer model and SWAMamba in using the entire CDS, we modified the Riboformer to utilize the entire CDS. The Riboformer (Entire CDS) model was configured with settings identical to those of SWAMamba. Input features comprised codon, amino acid, and nucleotide information, with the maximum codon sequence length constrained to 1024. Sequences shorter than this threshold were padded, while those exceeding it were truncated. This adaptation enabled a more equitable assessment of the two models' performance when processing complete CDS data.

B.6 Experiment Settings

This study employed a rigorous five-fold cross-validation methodology to assess and compare various algorithms. The dataset was initially partitioned into five equal subsets. In each fold, four subsets (80% of the data) were allocated for the training process, while the fifth subset (20% of the data) served as the test set. From the four training subsets, 10% was randomly selected to act as a validation set, with the remaining 90% used for actual training. This procedure was iterated five times, ensuring each subset functioned once as the test set. For each iteration, the model demonstrating superior performance on the validation set was selected and subsequently evaluated on the corresponding test set. The final reported results represent the average performance metrics of these five models on their respective test sets, thereby providing a more comprehensive and robust comparison of the algorithms' efficacy.

C Additional Results

C.1 Relationship Between Codon Contribution and A Site Distance

Figure 7 analyzes the relationship between codon contributions and distance on predicted ribosomal density at the A site in S. cerevisiae, humans, and C. elegans, respectively. Consistent with our initial findings, these datasets demonstrate that codons near the A site significantly influence ribosome density prediction. Moreover, this influence diminishes progressively as the distance from the A site increases.

C.2 Comparison of Riboformer (Entire CDS) and SWAMamba

To facilitate a more thorough comparison between SWAMamba and Transformer models, we adapted Riboformer to incorporate the entire CDS, referred to as Riboformer (Entire CDS). The comparative analysis, detailed in Table 4, reveals that integrating full-length mRNA CDS features enhanced Riboformer's performance across most datasets. However, despite these improvements, Riboformer's performance remained inferior to that of SWAMamba. These findings further substantiate the superiority of the SWAMamba approach. Even after removing the dual-stage sliding window attention, SWAMamba still outperforms Riboformer (Entire CDS), demonstrating that Mamba is more effective than Transformer in modeling long mRNA sequences.

C.3 Comparison of Average Pearson Correlation Coefficients with Variance

Due to page constraints in the main text, the variance was not displayed in Table 1. To provide a more comprehensive evaluation, the variance is included here. Table 5 presents the comparison of average Pearson correlation coefficients between predicted

	C. elegans	S. cerevisiae	Humans	E. coli
Riboformer (Entire CDS)	0.516±0.005	0.551±0.008	0.579±0.017	0.542±0.010
SWAMamba (w/o dual atten)	0.593±0.004	0.621±0.003	0.638±0.006	0.626±0.010
SWAMamba	**0.596±0.004**	**0.630±0.003**	**0.641±0.007**	**0.640±0.008**

Table 4: Mean Pearson correlation coefficients and standard deviations between predicted and experimental ribosome density values.

	C. elegans		S. cerevisiae		Humans		E. coli	
Method	r	R^2	r	R^2	r	R^2	r	R^2
IXnos	0.504±0.008	0.253±0.008	0.502±0.003	0.252±0.002	0.507±0.007	0.250±0.008	0.549±0.015	0.300±0.017
Riboexp	0.531±0.004	0.273±0.006	0.587±0.008	0.338±0.015	0.572±0.026	0.307±0.037	0.566±0.008	0.312±0.016
RiboMIMO	0.576±0.005	0.330±0.006	0.612±0.007	0.373±0.008	0.607±0.020	0.361±0.030	0.572±0.023	0.324±0.026
Riboformer	0.495±0.004	0.231±0.006	0.482±0.004	0.224±0.006	0.476±0.019	0.224±0.018	0.552±0.011	0.282±0.014
SWAMamba	**0.596±0.004**	**0.345±0.004**	**0.630±0.003**	**0.385±0.007**	**0.641±0.007**	**0.389±0.008**	**0.640±0.008**	**0.401±0.013**

Table 5: Comparison of Pearson correlation coefficient and R^2 between predicted and experimental ribosome density values for different methods.

and experimental ribosome densities across methods, using four datasets. Additionally, the R^2 values are reported to further assess the prediction accuracy, with both metrics contributing to a more detailed evaluation of the methods' performance.