(A) Cross-species Foundational Representation

$\mathcal{L}_{MLM}$

C A A A G T T A G C A T ... G C A C T G

Tokenization

Mask

[CLS] Token 1 C A A A G T Token 2 T A G C A T ... Token N G C A C T G [SEP]

[CLS] Token 1 C A A A G T Token 2 [MASK] ... Token N G C A C T G [SEP]

Embed Layer

MambaSeq Block

Norm

Self-attention

Norm

MLP

Attention Block

[CLS] Token 1 C A A A G T Token 2 T A G C A T ... Token N G C A C T G [SEP]

Input Sequence

Token

Masked Token

Copy Weights

Copy Weights

Copy Weights

Predicted Token

(B) Functional Representation Enhancement

Initialization

Token Embedding

Masked Token Embedding

MambaSeq Block

Norm

Self-attention

Norm

Router

MLP 1

MLP 2

MLP 5

Attention Block

$\mathcal{L}_{reg}$

$\mathcal{L}_{MLM}$

$\mathcal{L}_{aux}$

Upcycling

Predicted Track

Target Track

$\mathcal{L}_{MLM}$ MLM Loss

$\mathcal{L}_{reg}$ Sequence to Signal Loss

$\mathcal{L}_{aux}$ Load Balancing Loss

MLP Copies