

Run Summary

riboHMM Validation Report

Run Timestamp: 2025-07-28T14:18:20.711388

Test Data File: /home/kai/Documents/00_RNA/01_riboHMM_simple_prod/02_Processed_Data/test_top1000_raw.npz

Trained Model File: /home/kai/Documents/00_RNA/01_riboHMM_simple_prod/04_Models/model_top1000_raw.joblib

Model Config File: /home/kai/Documents/00_RNA/01_riboHMM_simple_prod/model_config.json

ORF Annotations File: /home/kai/Documents/00_RNA/01_riboHMM_simple_prod/01_Reference_Files/MANE_CDS_coordinates.csv

Number of Transcripts in Test Set: 200

The heatmap displays a 16x16 matrix with a symmetric pattern. The diagonal elements are dark purple and labeled '0.010'. The off-diagonal elements are yellow and labeled '0.990' or '1.000'. The matrix is symmetric, with the values 0.990 and 0.010 appearing in the same positions as the diagonal elements.

Pre-INIT-1

Pre-INIT-2

Pre-INIT-3

INIT

Post-INIT-1

Post-INIT-2

T1

T2

T3

Pre-TFRM-1

Pre-TERM-2

Pre-TERM-3

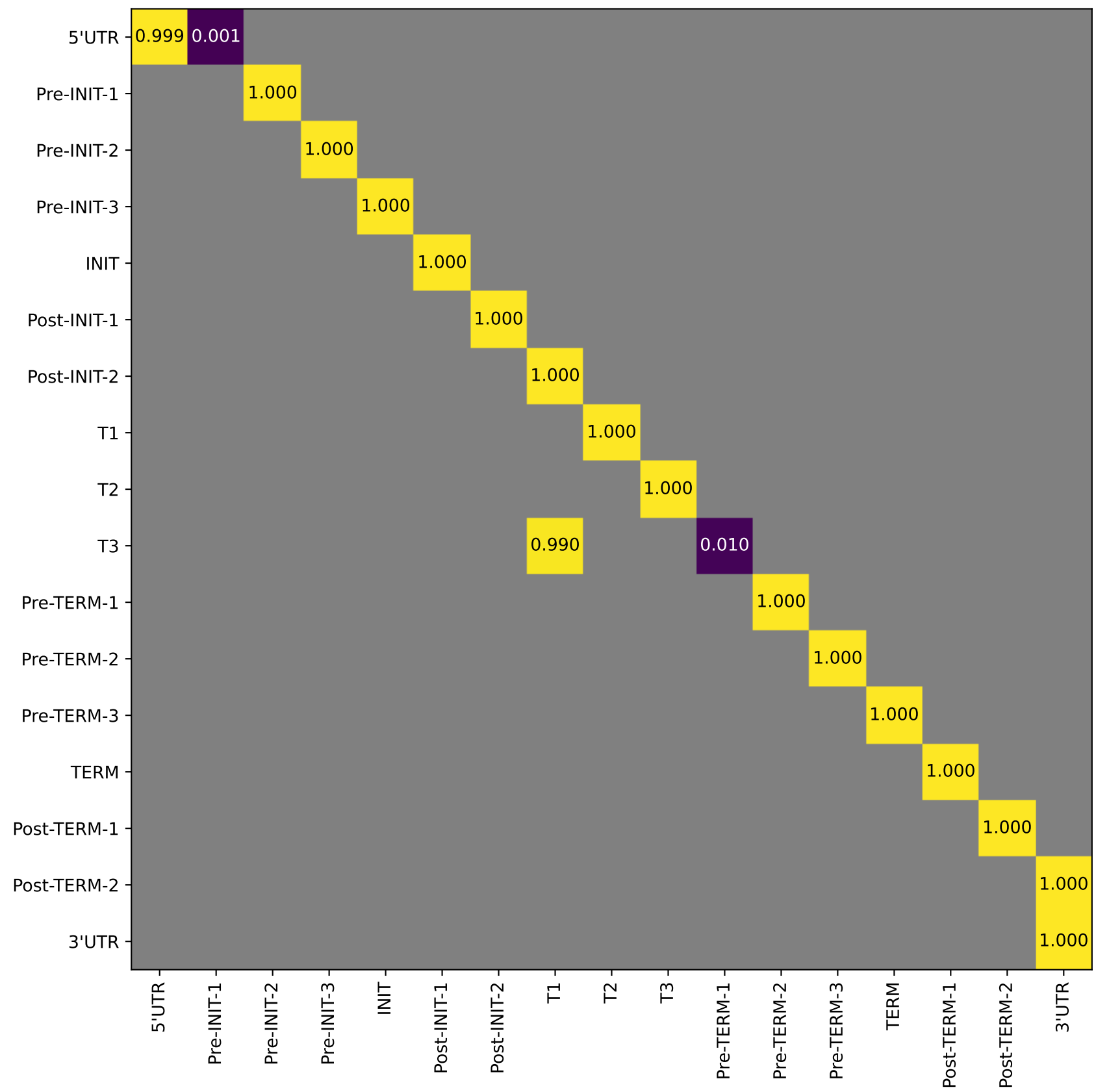
EDM

Project TERM 1

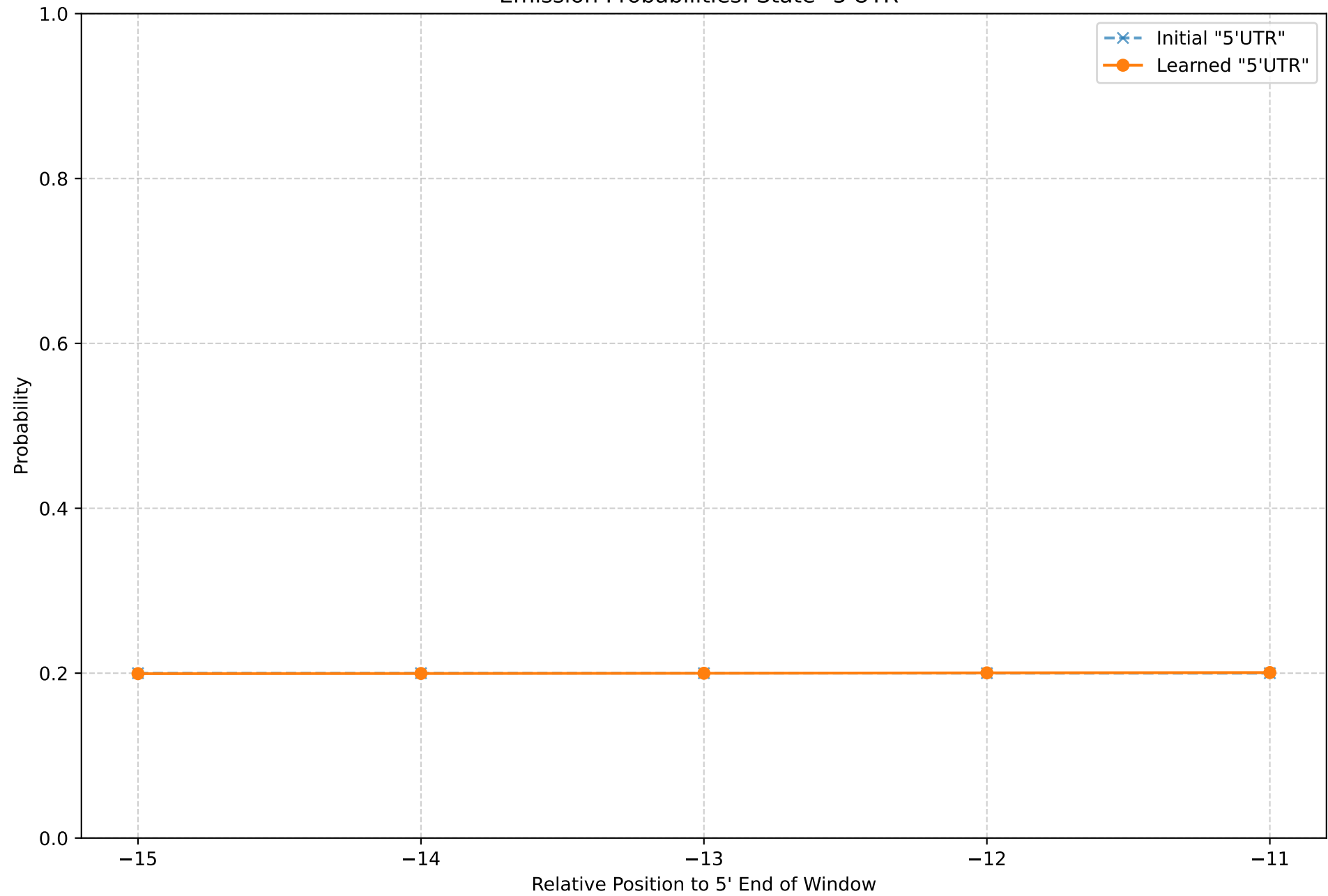
CONFIDENTIAL

DE

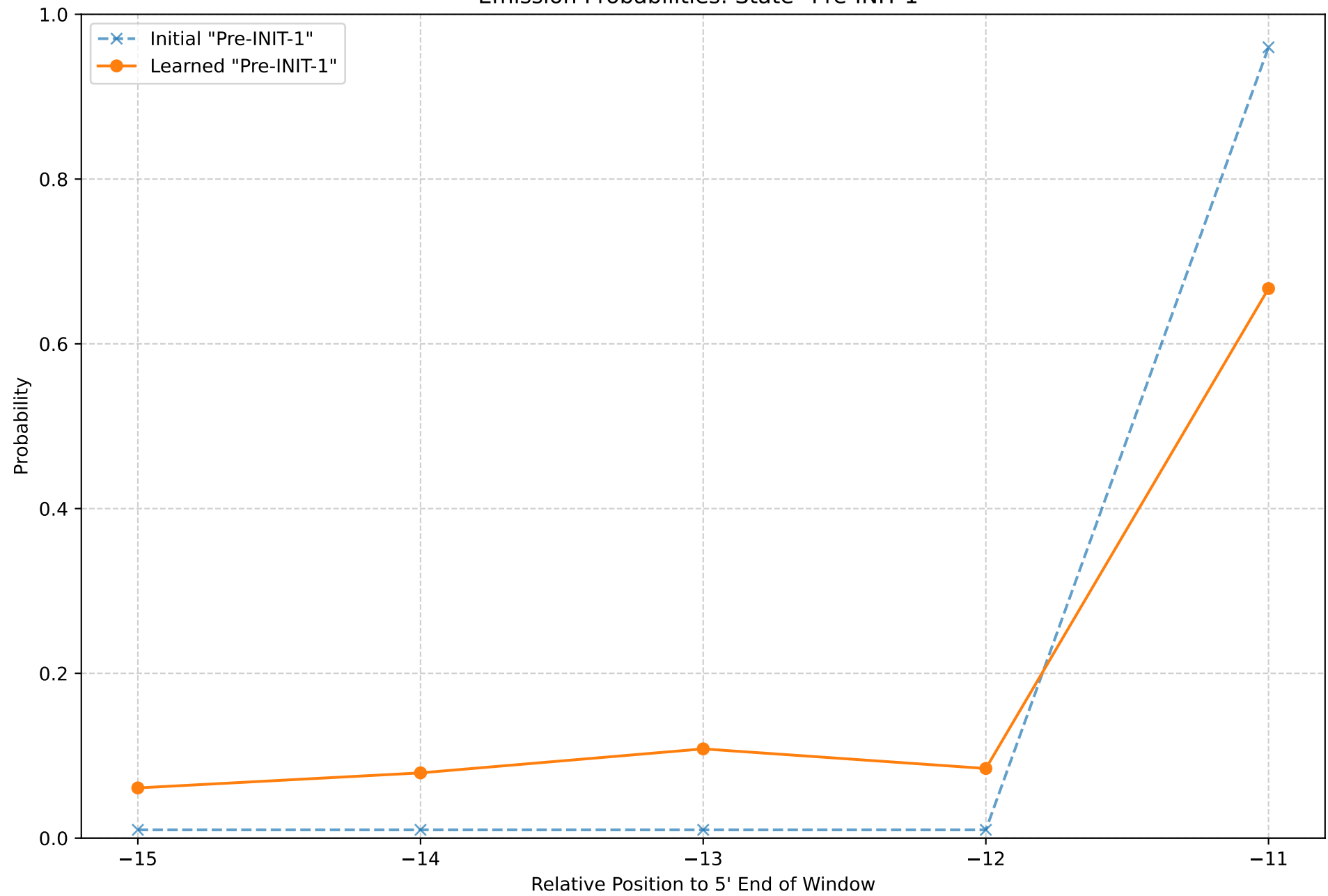
Learned Transition Matrix



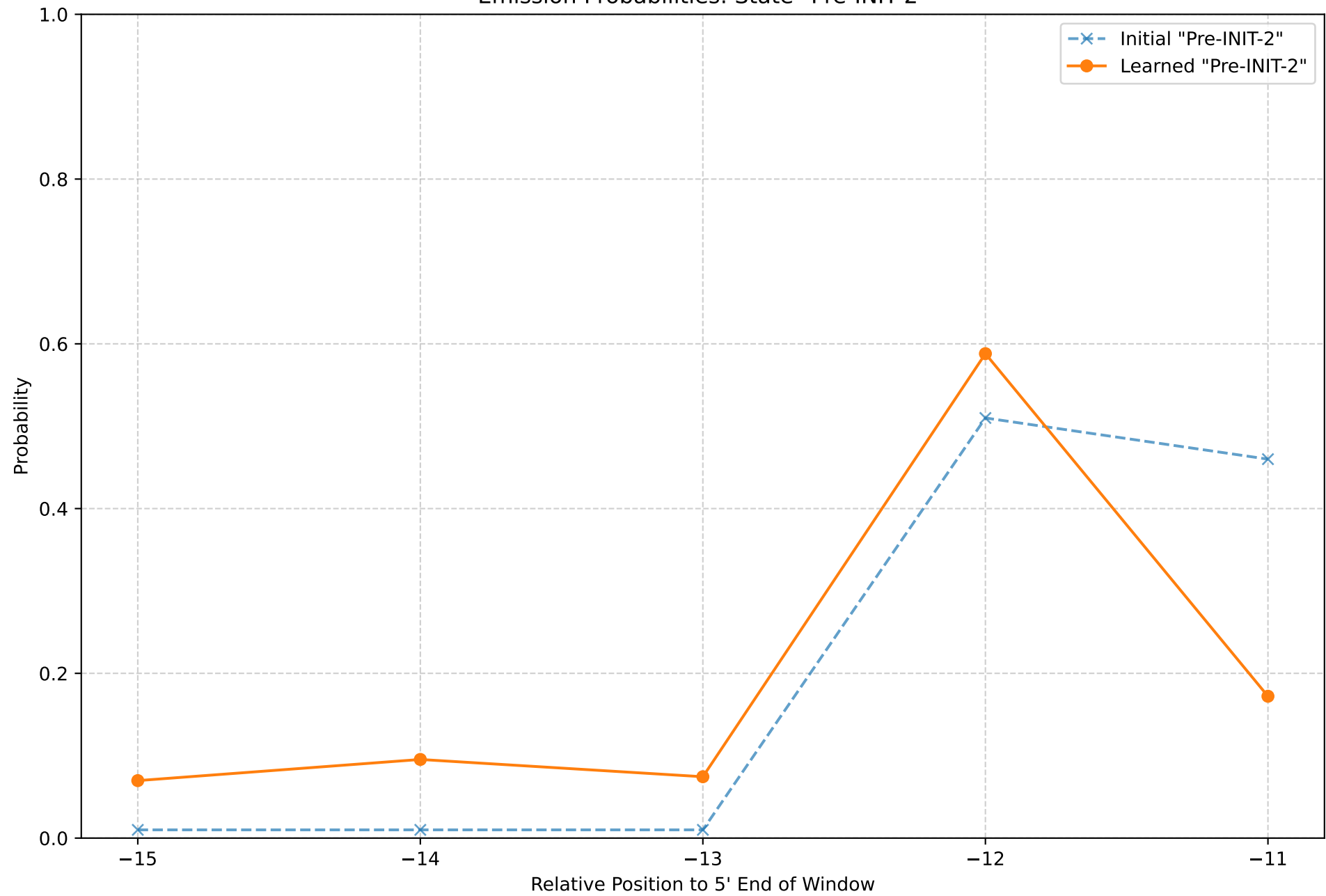
Emission Probabilities: State "5'UTR"



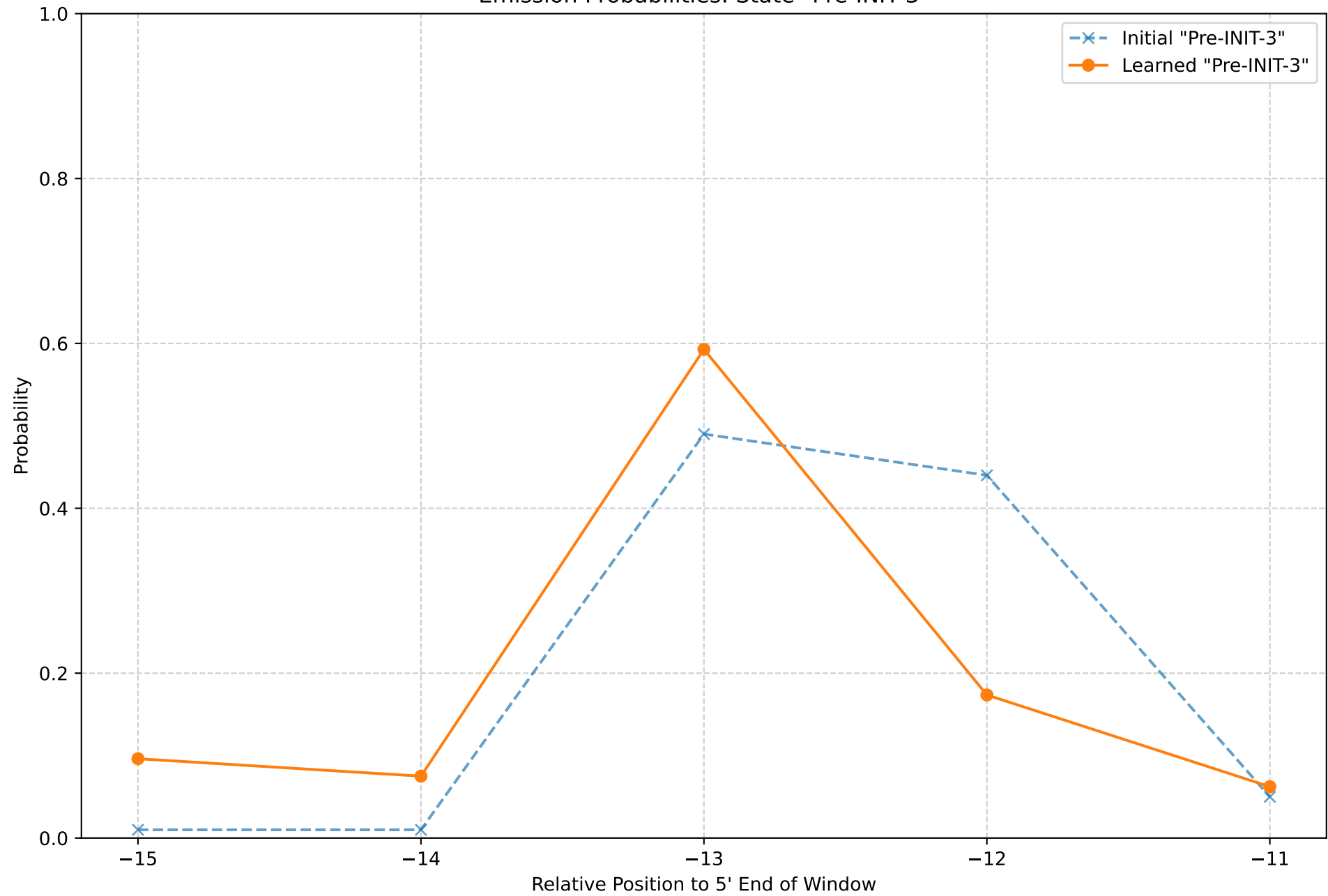
Emission Probabilities: State "Pre-INIT-1"



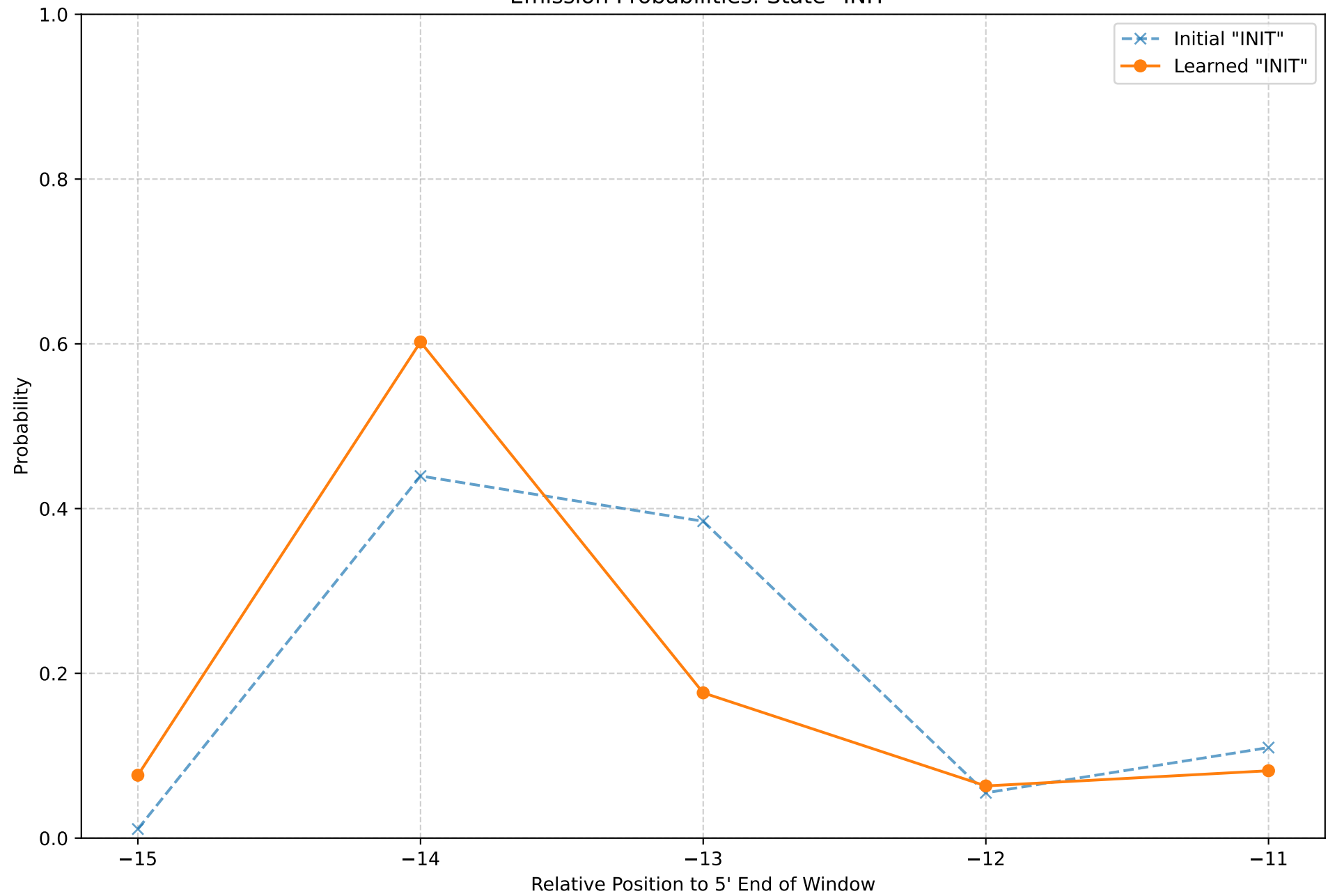
Emission Probabilities: State "Pre-INIT-2"



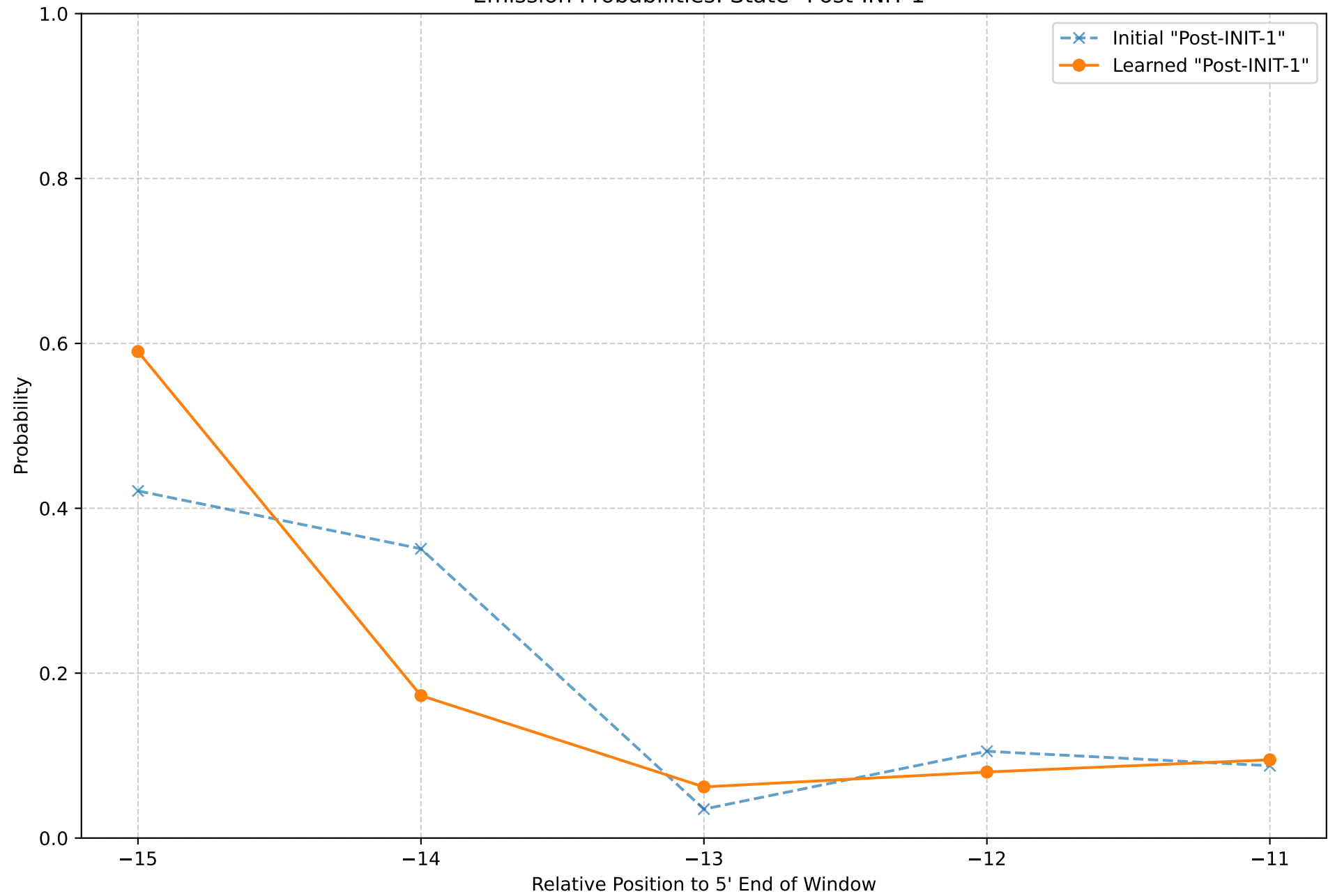
Emission Probabilities: State "Pre-INIT-3"



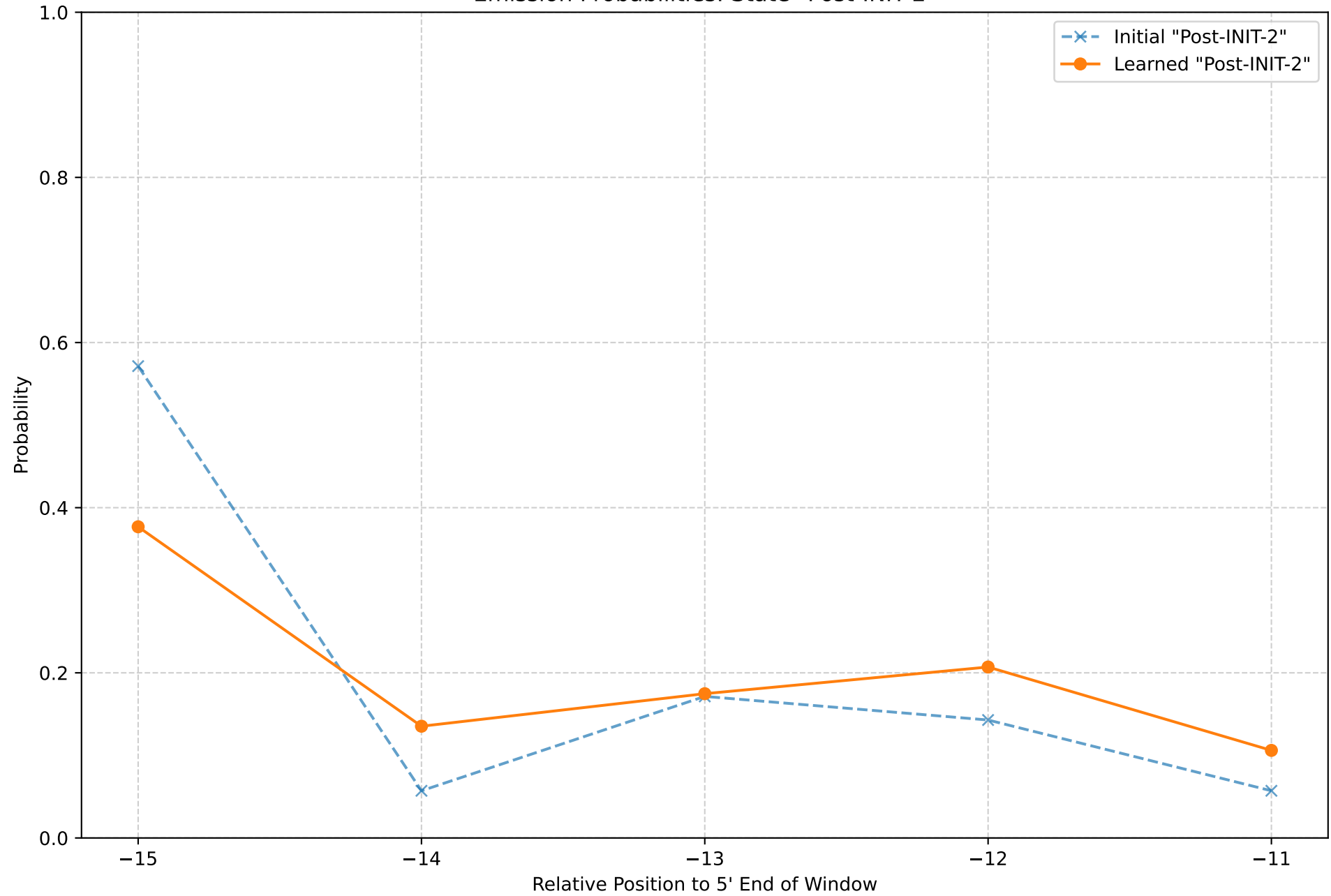
Emission Probabilities: State "INIT"



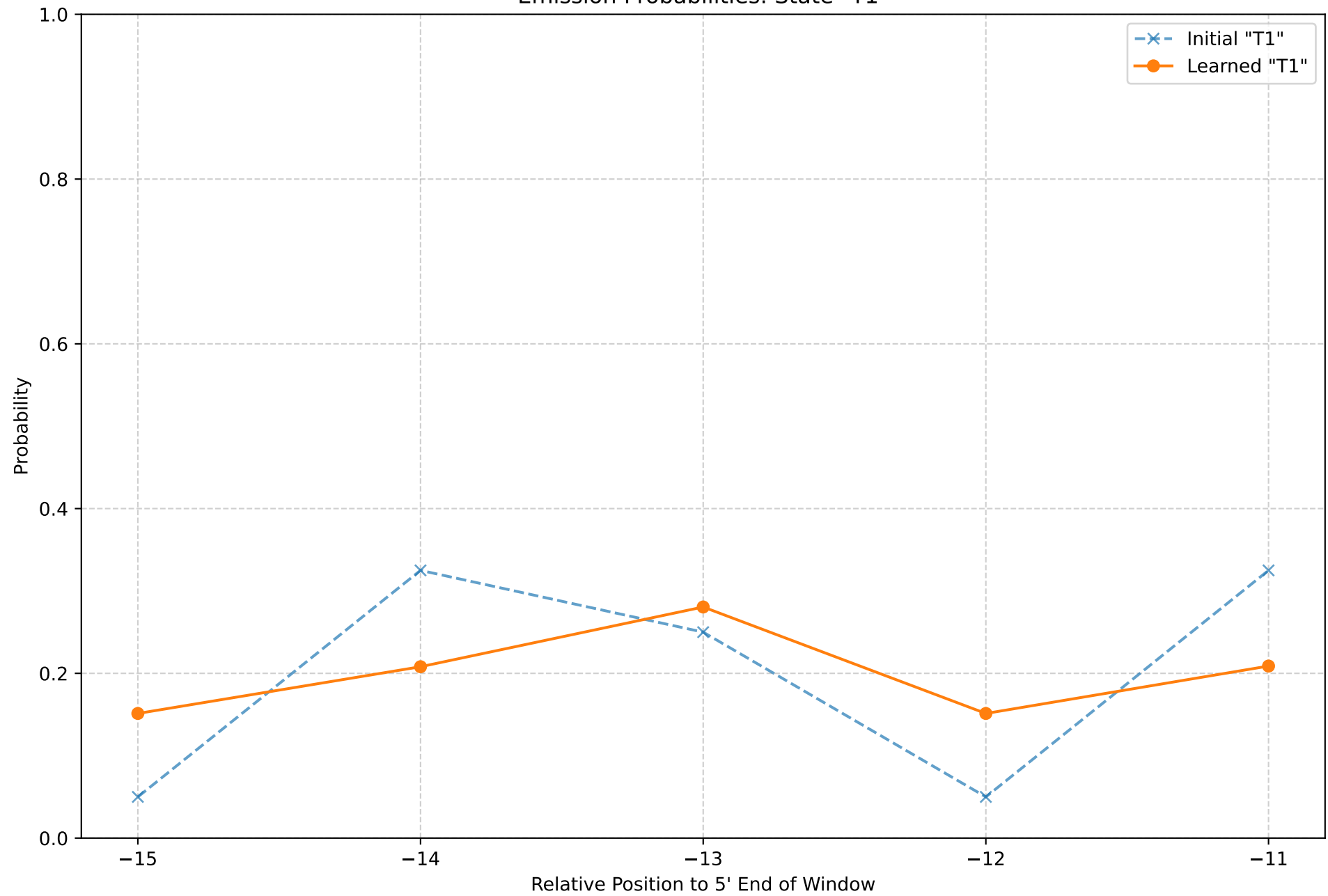
Emission Probabilities: State "Post-INIT-1"



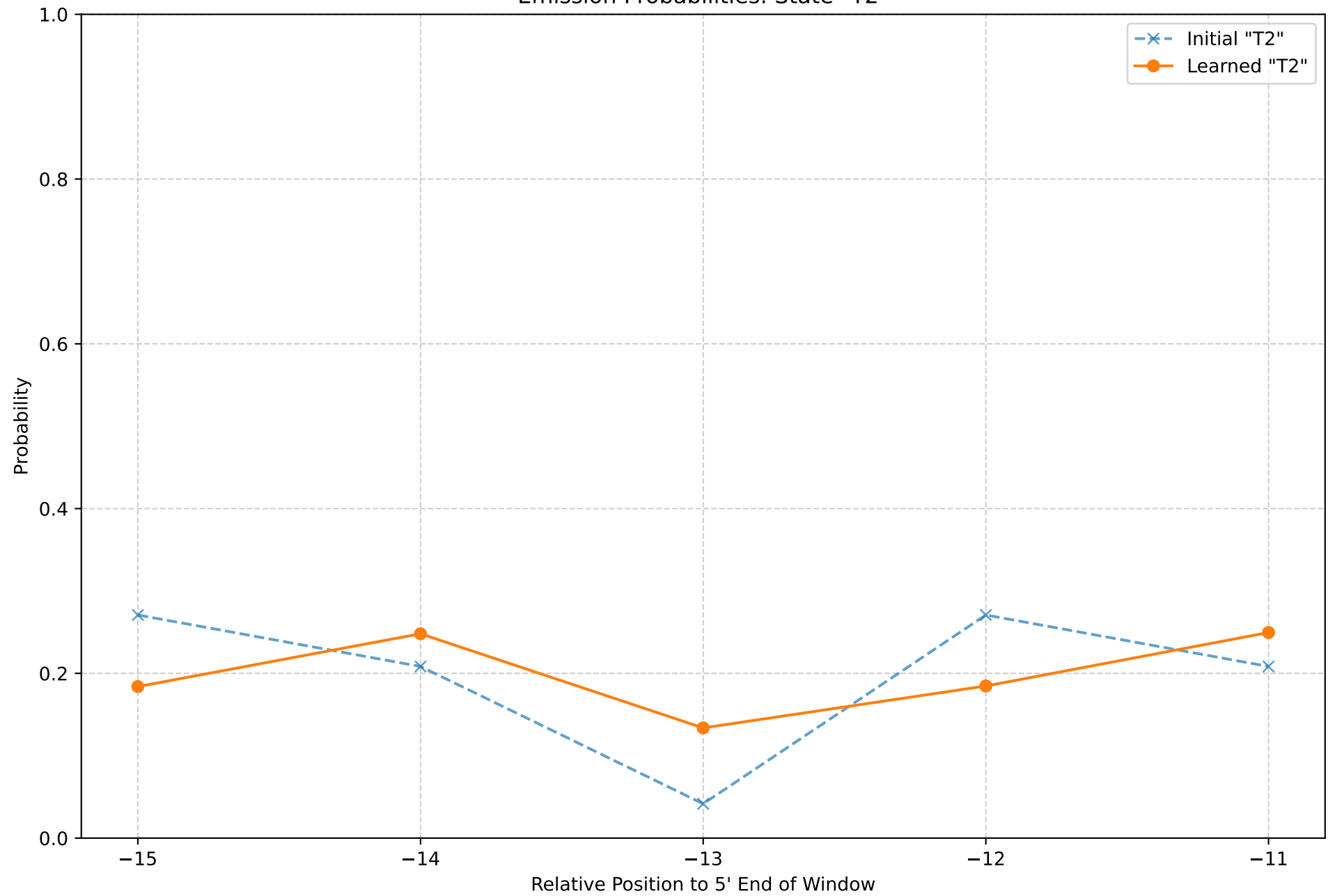
Emission Probabilities: State "Post-INIT-2"



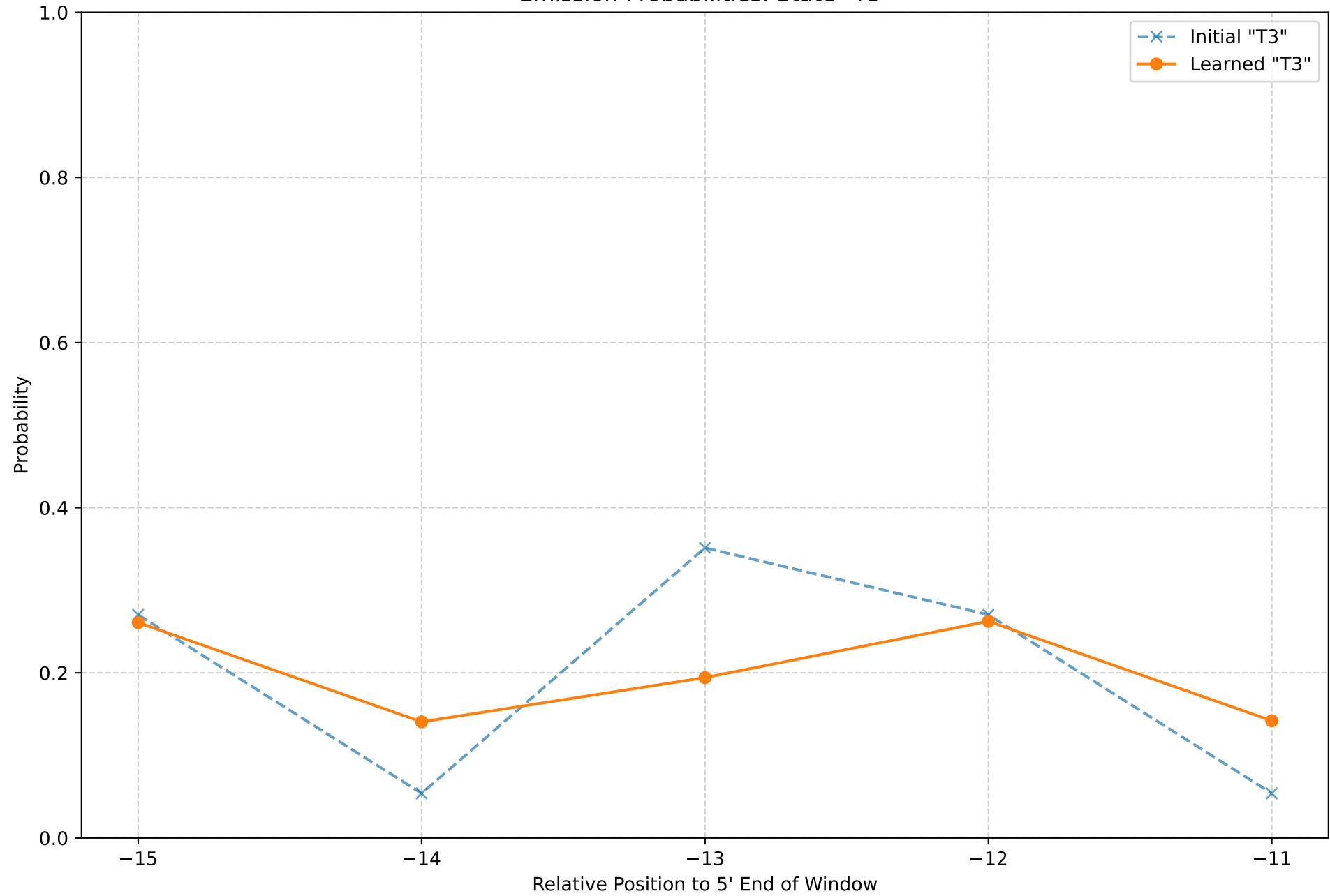
Emission Probabilities: State "T1"



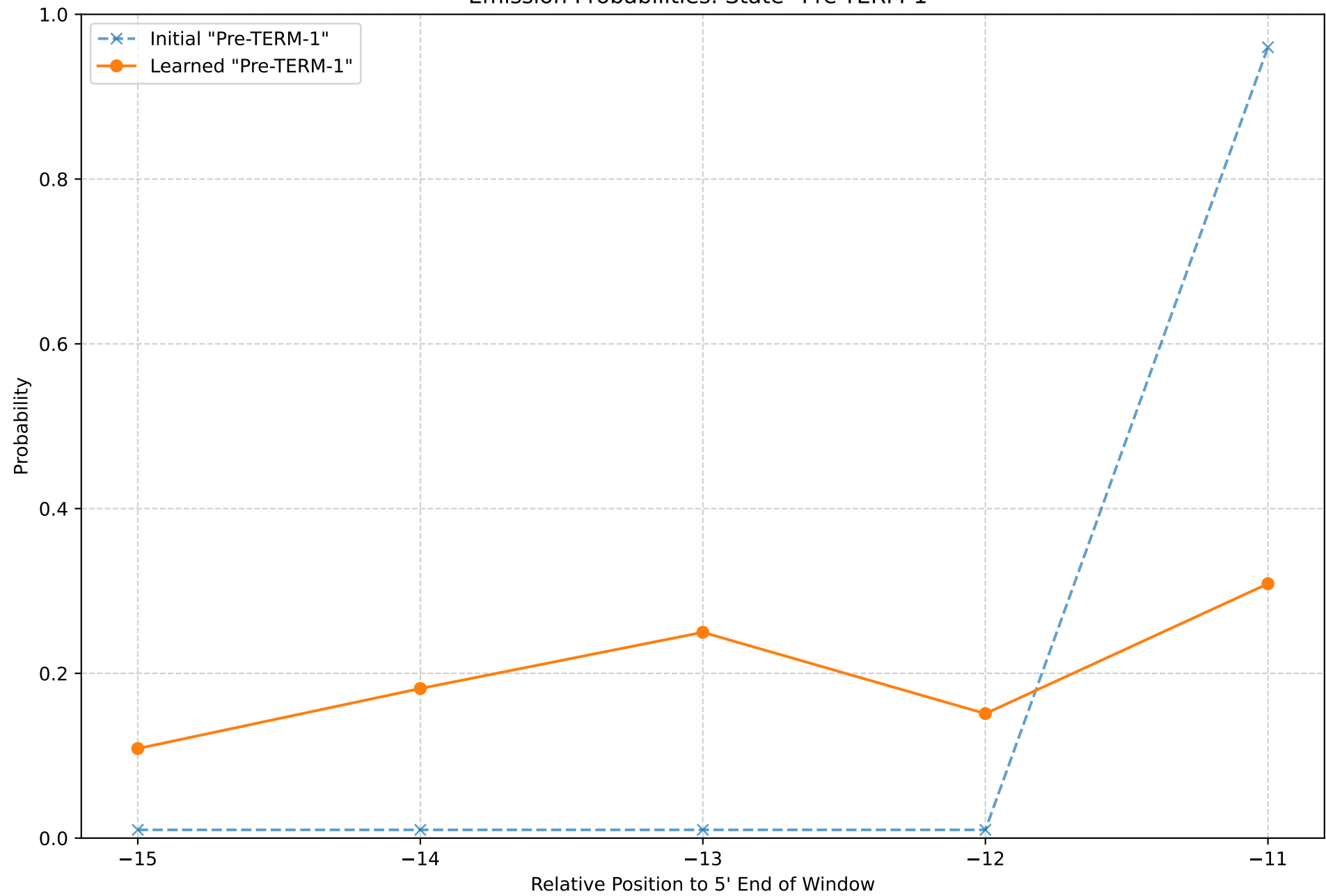
Emission Probabilities: State "T2"



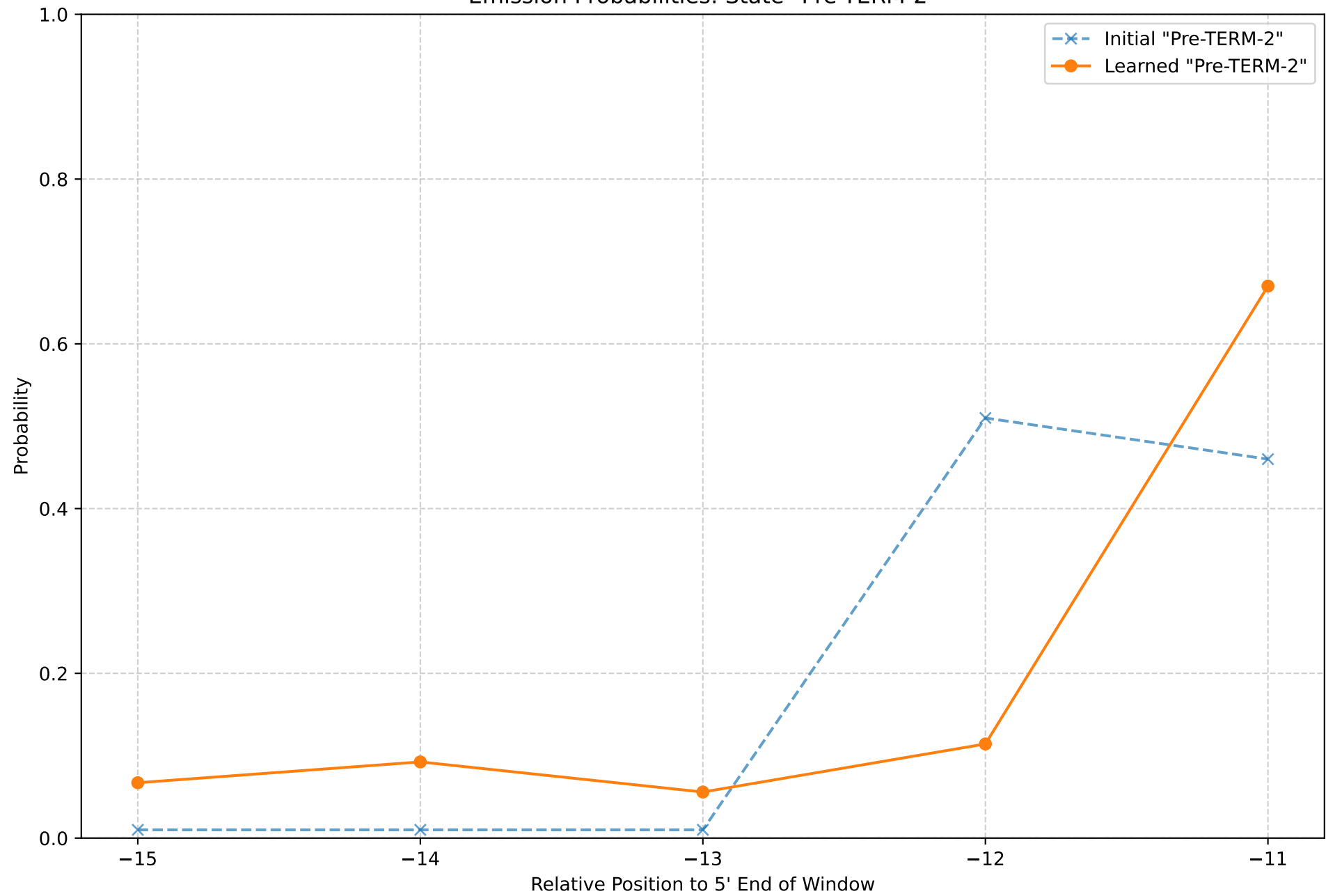
Emission Probabilities: State "T3"



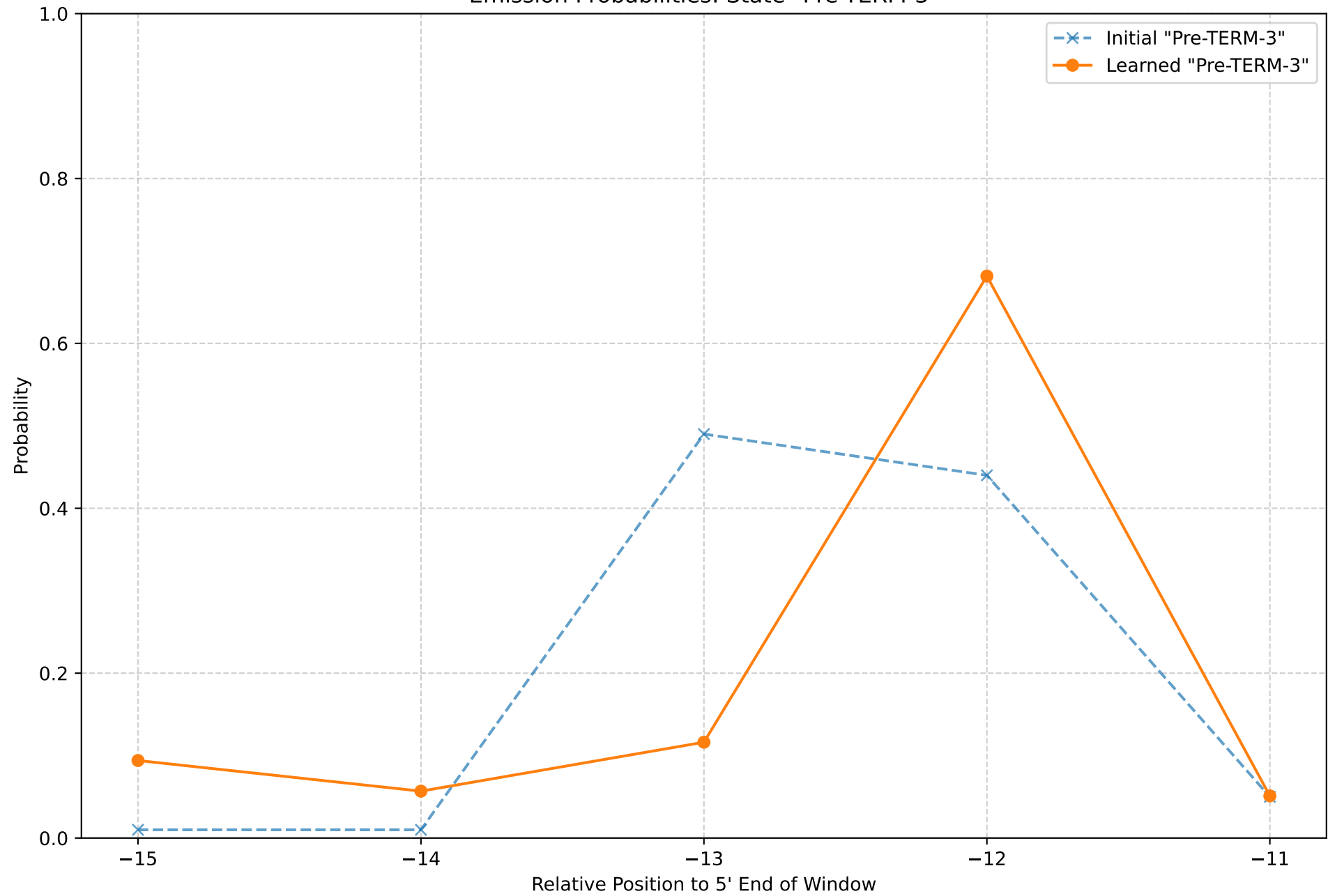
Emission Probabilities: State "Pre-TERM-1"



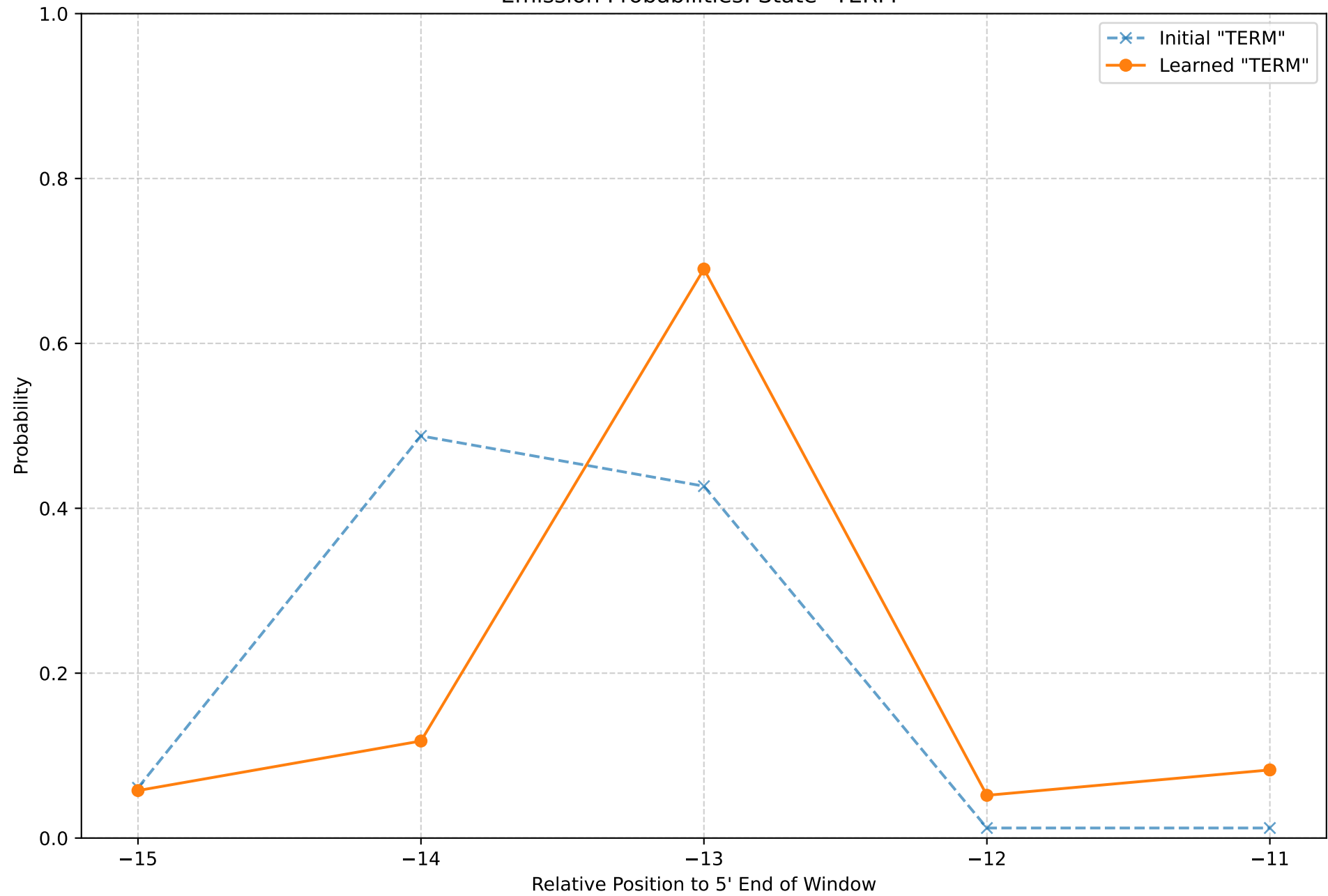
Emission Probabilities: State "Pre-TERM-2"



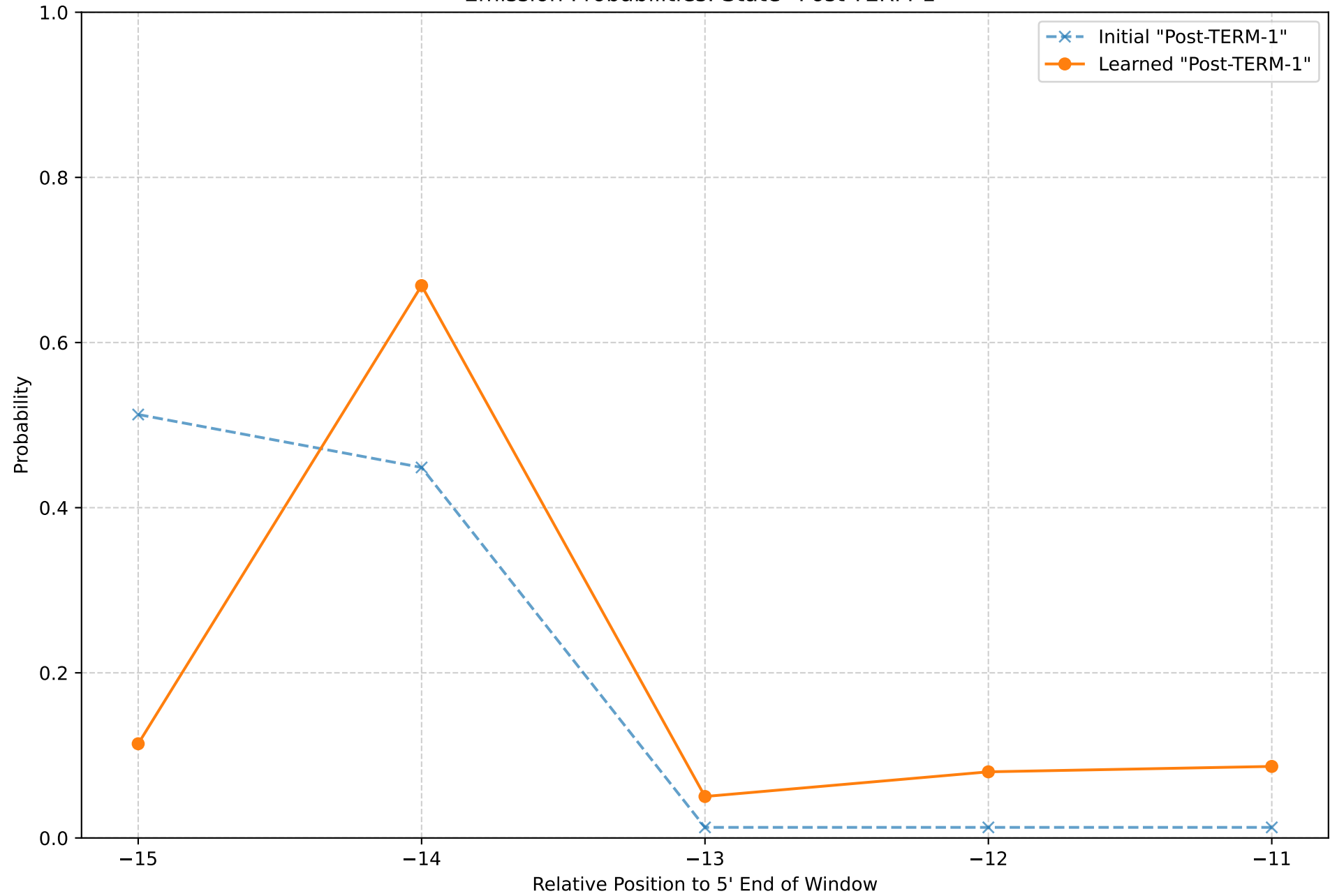
Emission Probabilities: State "Pre-TERM-3"



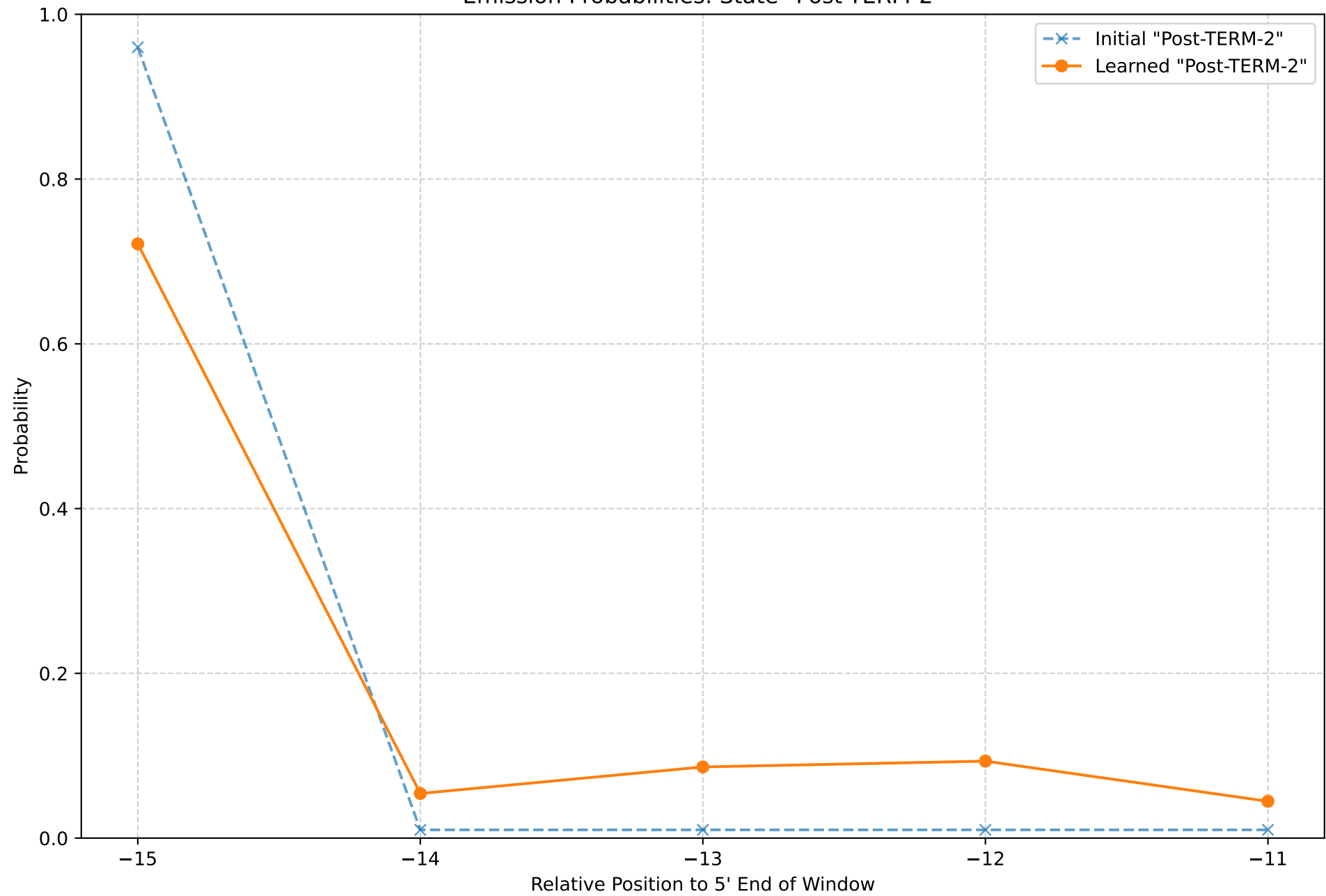
Emission Probabilities: State "TERM"



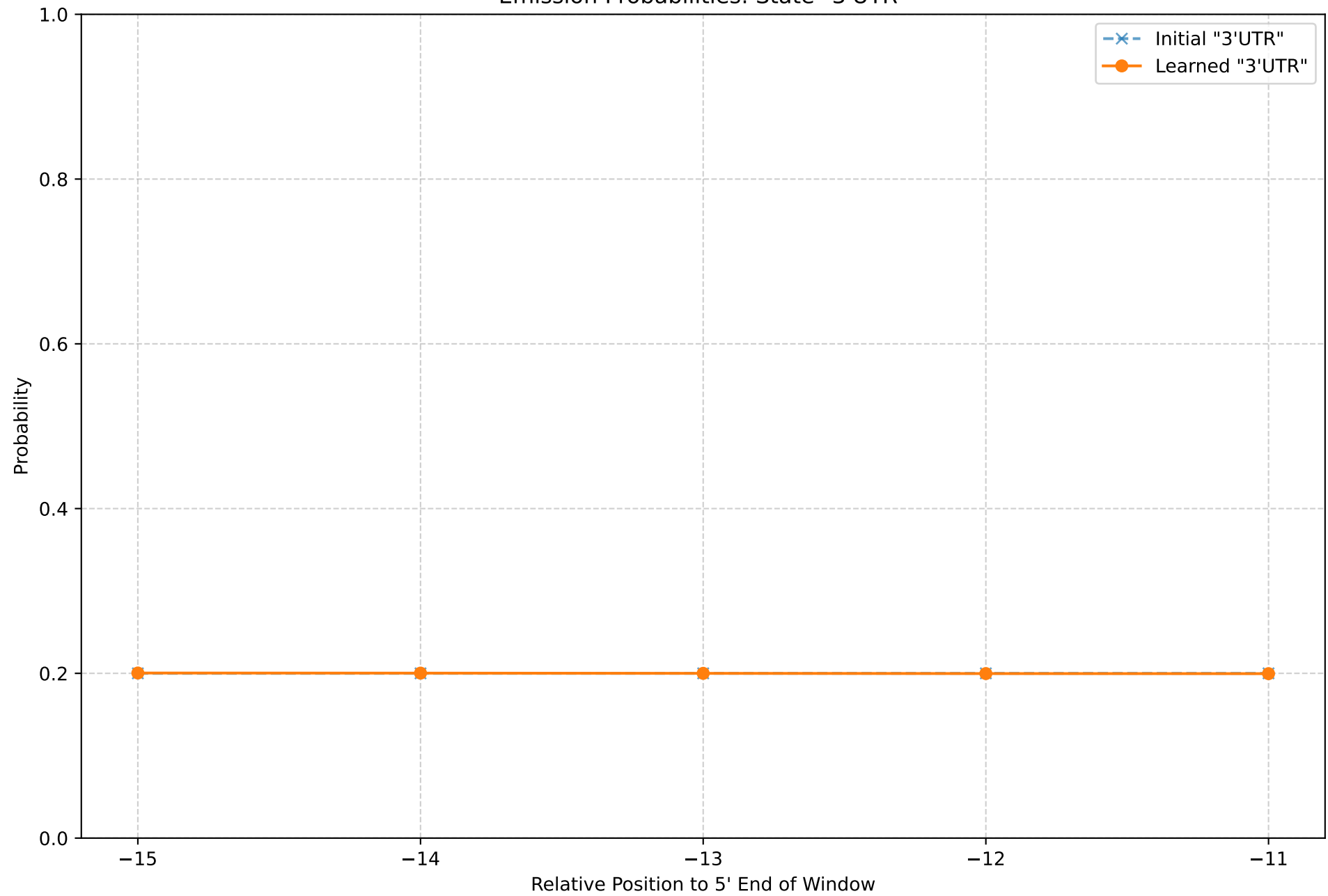
Emission Probabilities: State "Post-TERM-1"



Emission Probabilities: State "Post-TERM-2"

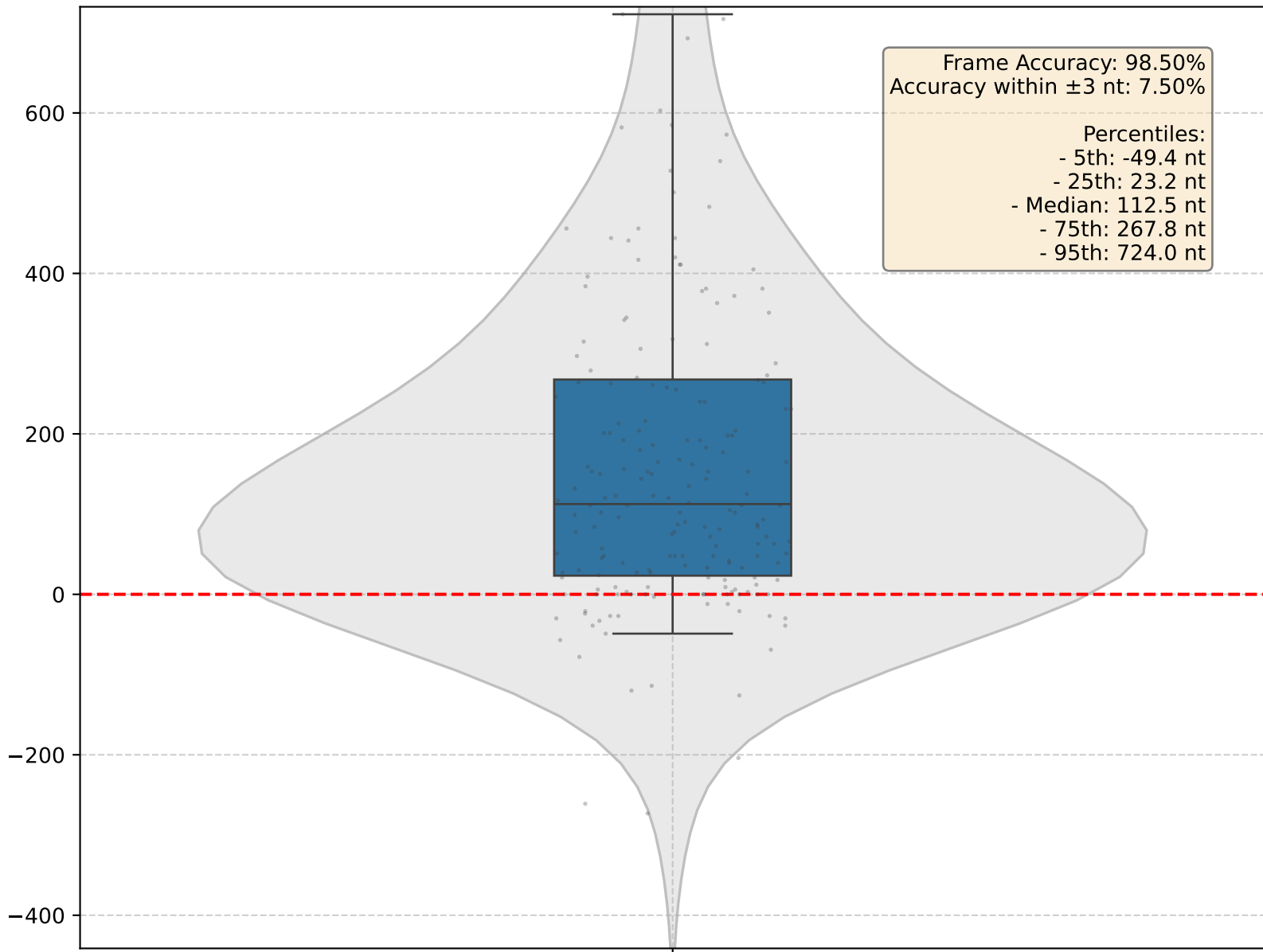


Emission Probabilities: State "3'UTR"

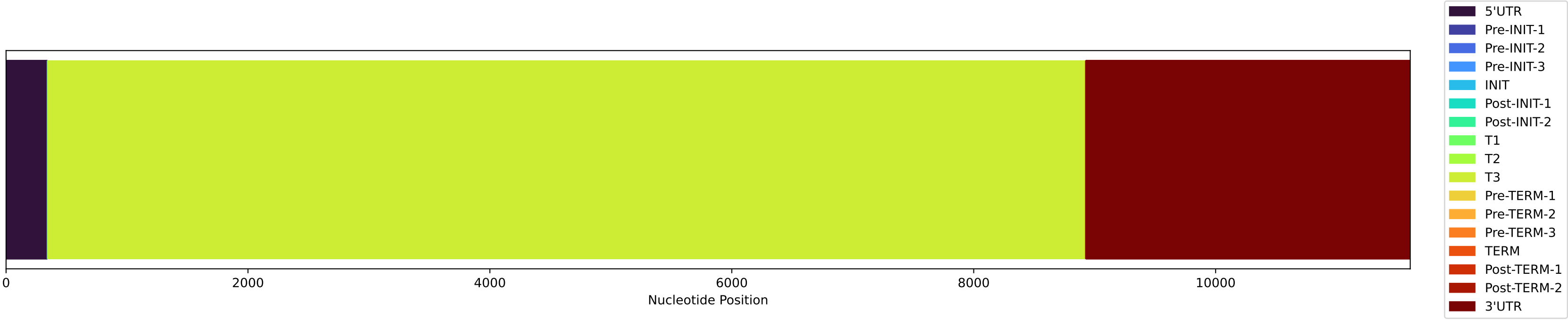
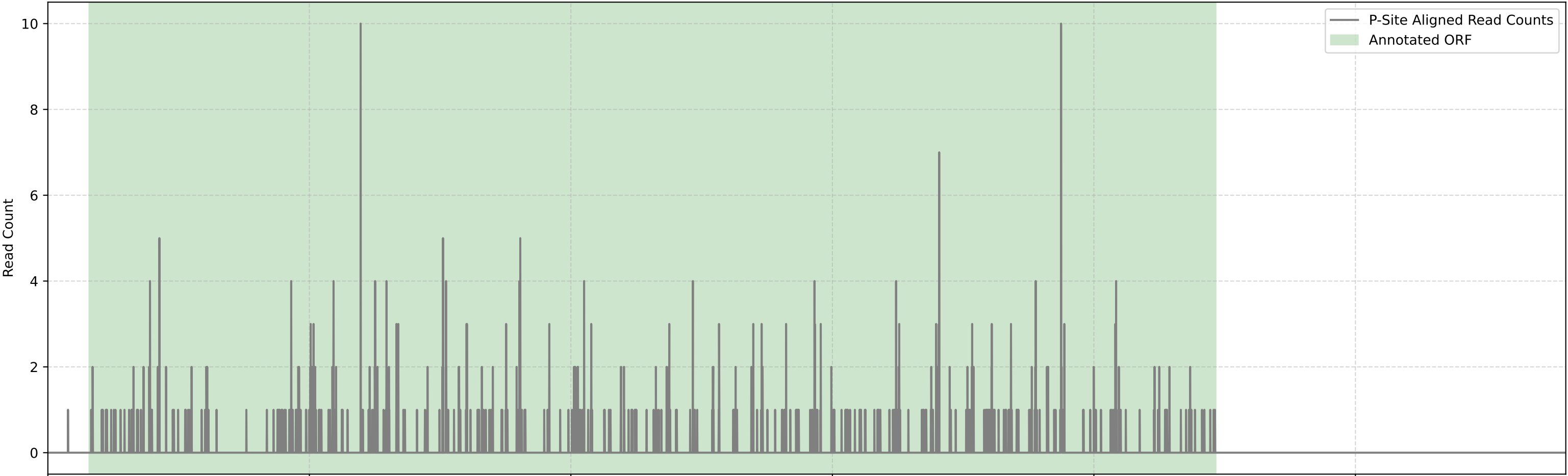


Distance: Predicted INIT State vs. True Start Codon

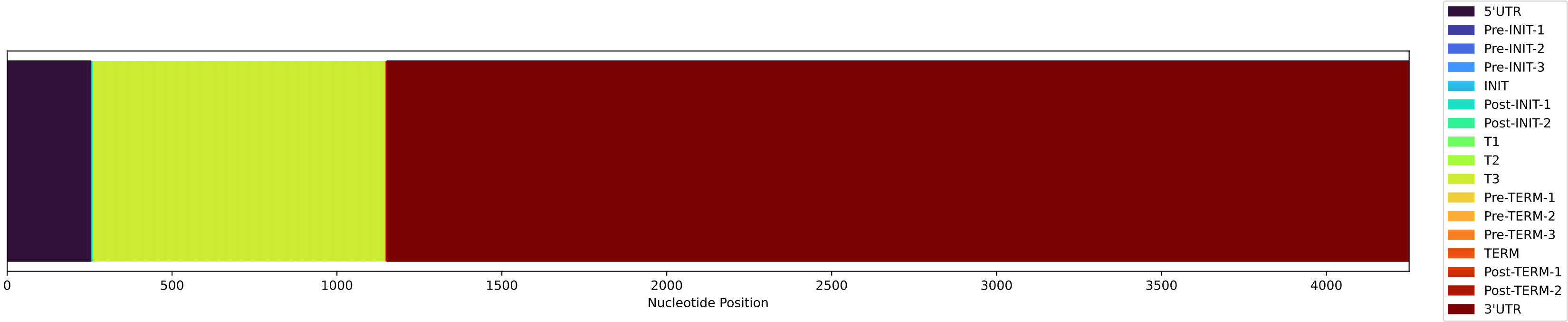
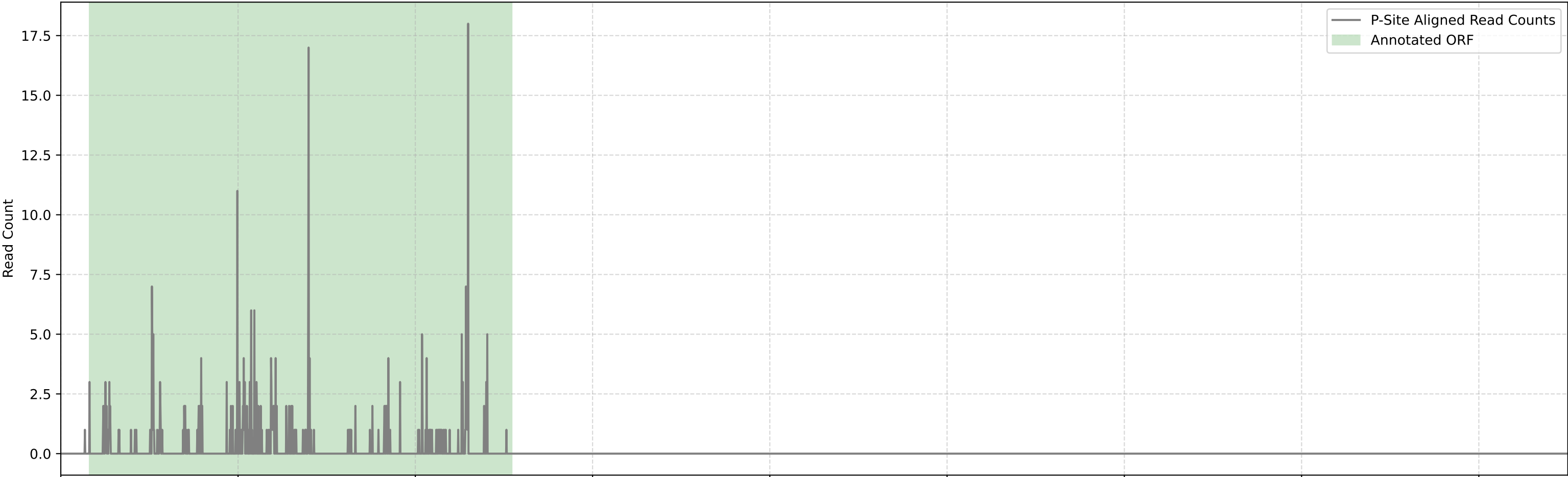
Distance (nt) [Predicted - True]



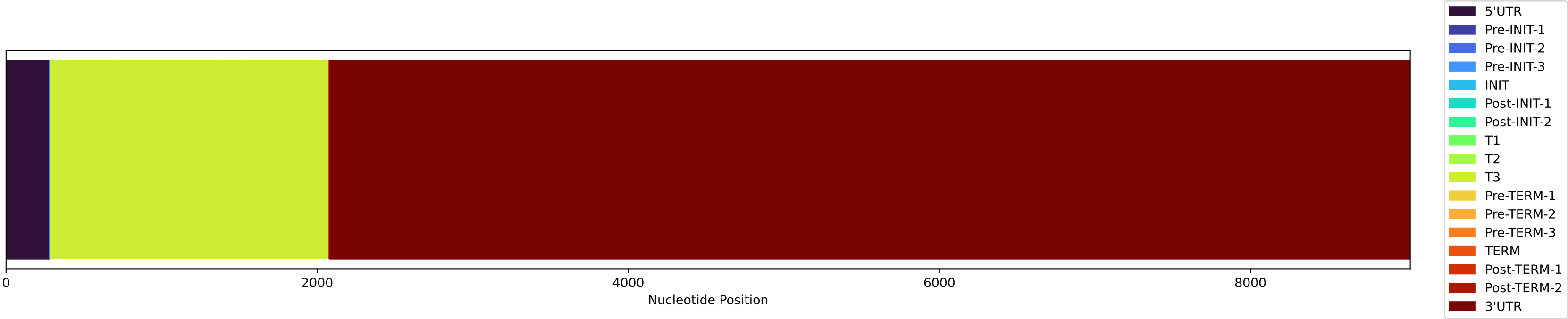
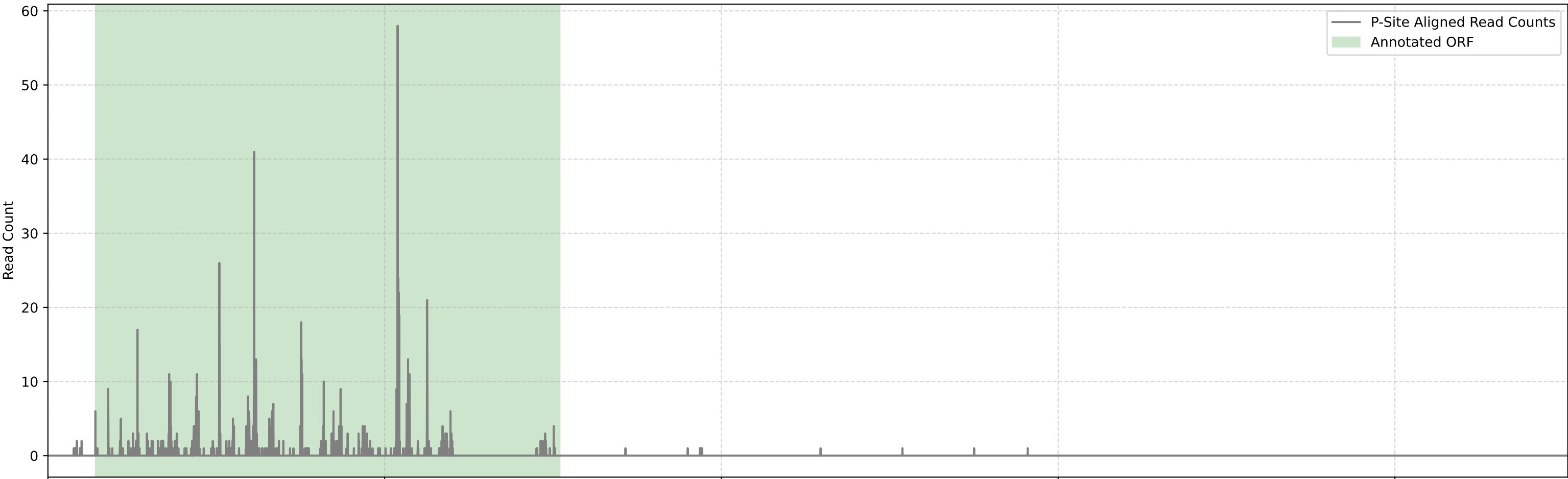
Example: 10th Percentile (by TPM) Transcript (ENST00000316623.10)
Total Reads: 740 | TPM: 788.07



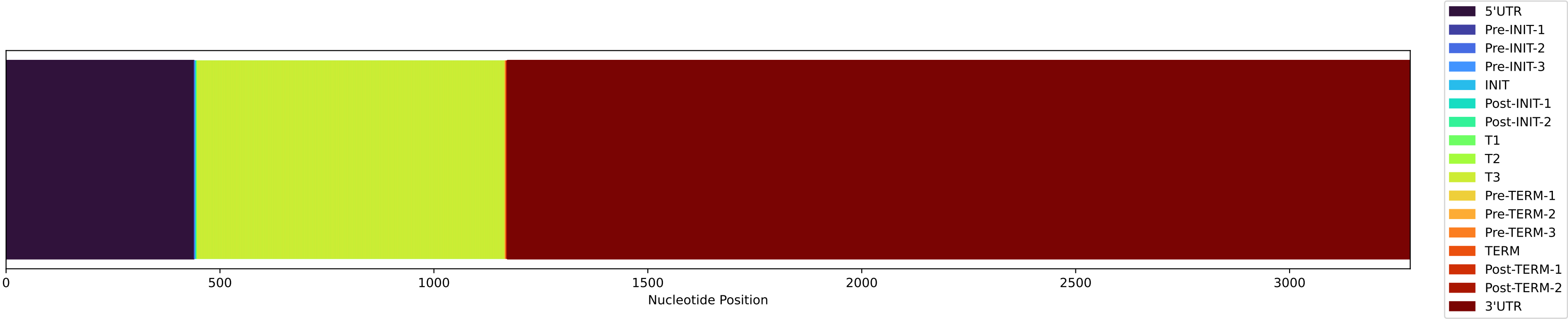
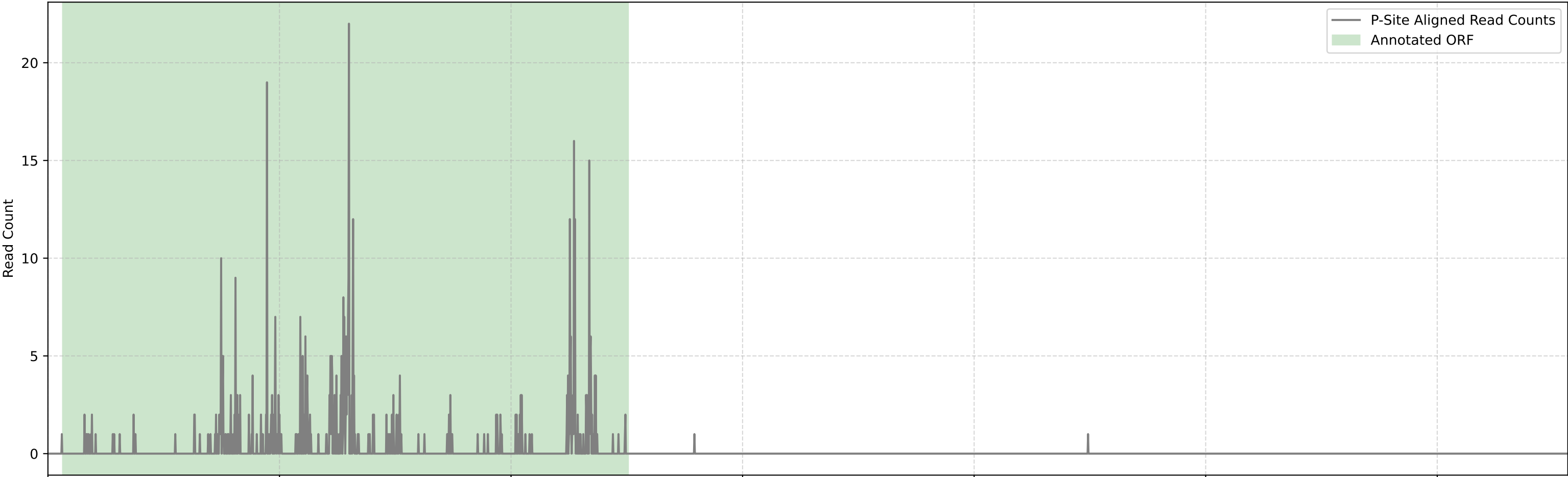
Example: 20th Percentile (by TPM) Transcript (ENST00000284719.8)
Total Reads: 377 | TPM: 1096.43



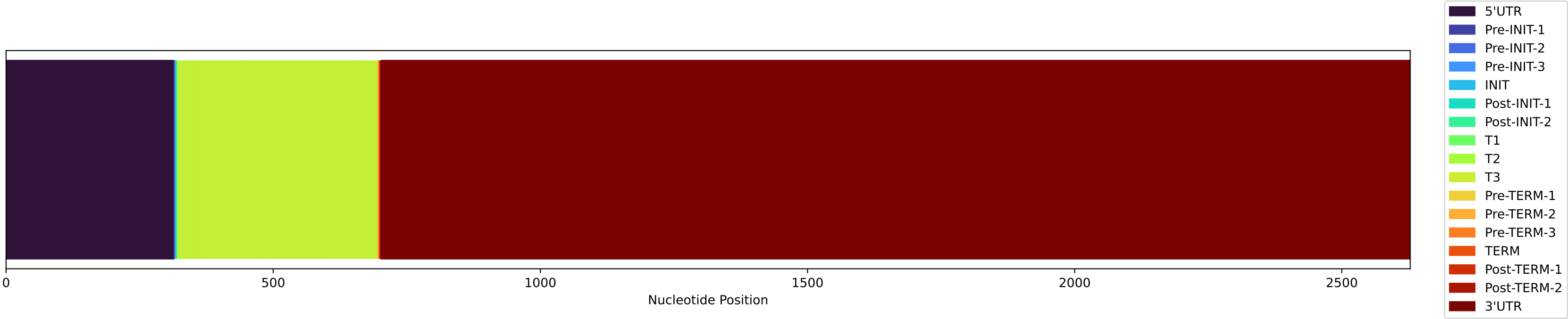
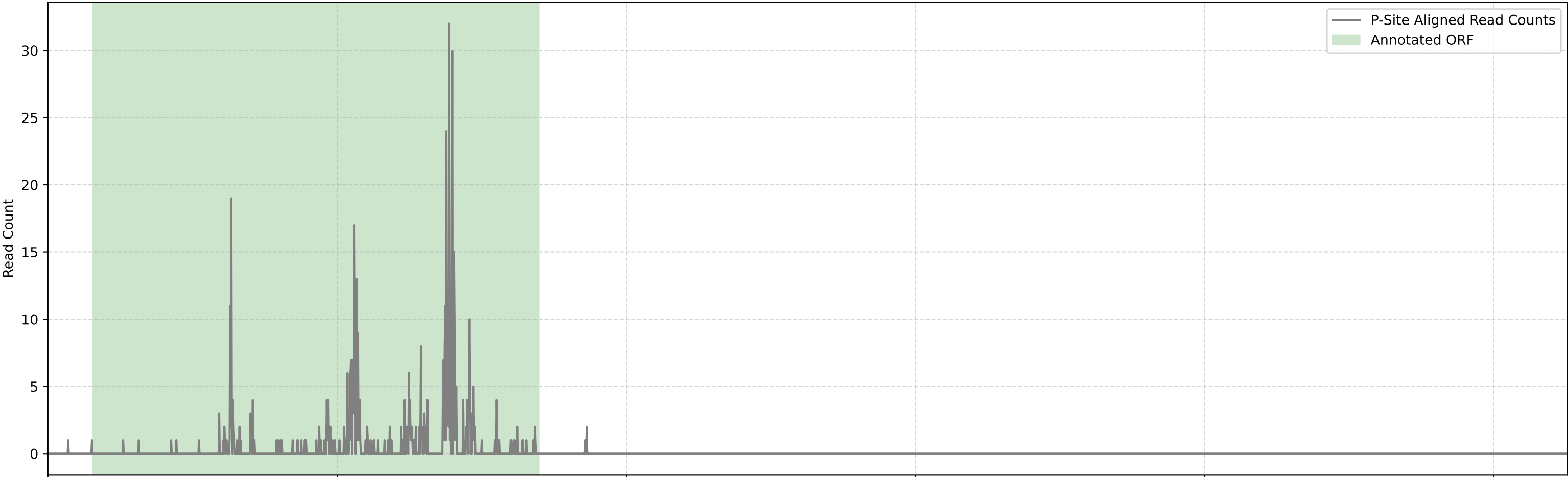
Example: 30th Percentile (by TPM) Transcript (ENST00000381298.7)
Total Reads: 1,156 | TPM: 1583.23



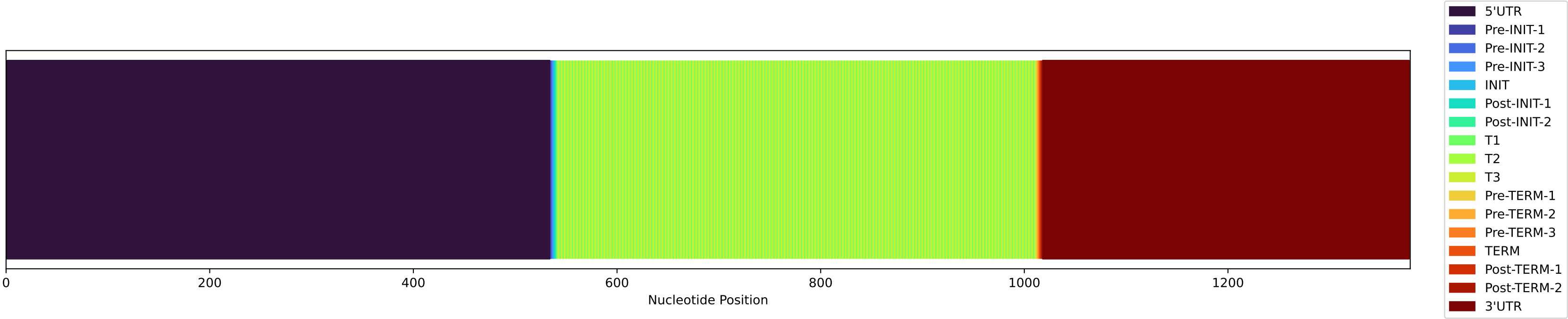
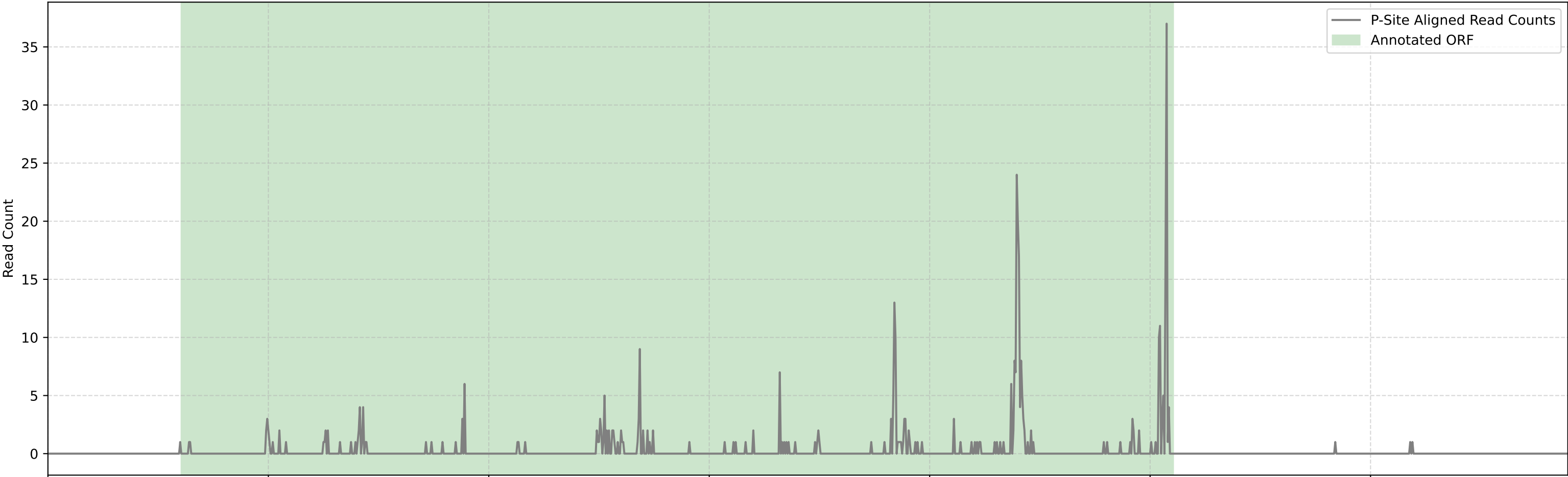
Example: 40th Percentile (by TPM) Transcript (ENST00000362031.10)
Total Reads: 569 | TPM: 2143.40



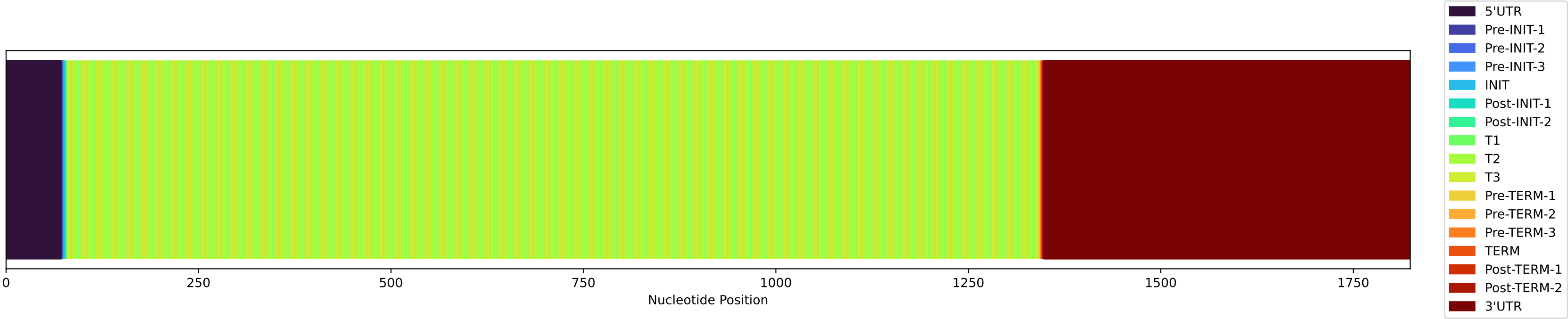
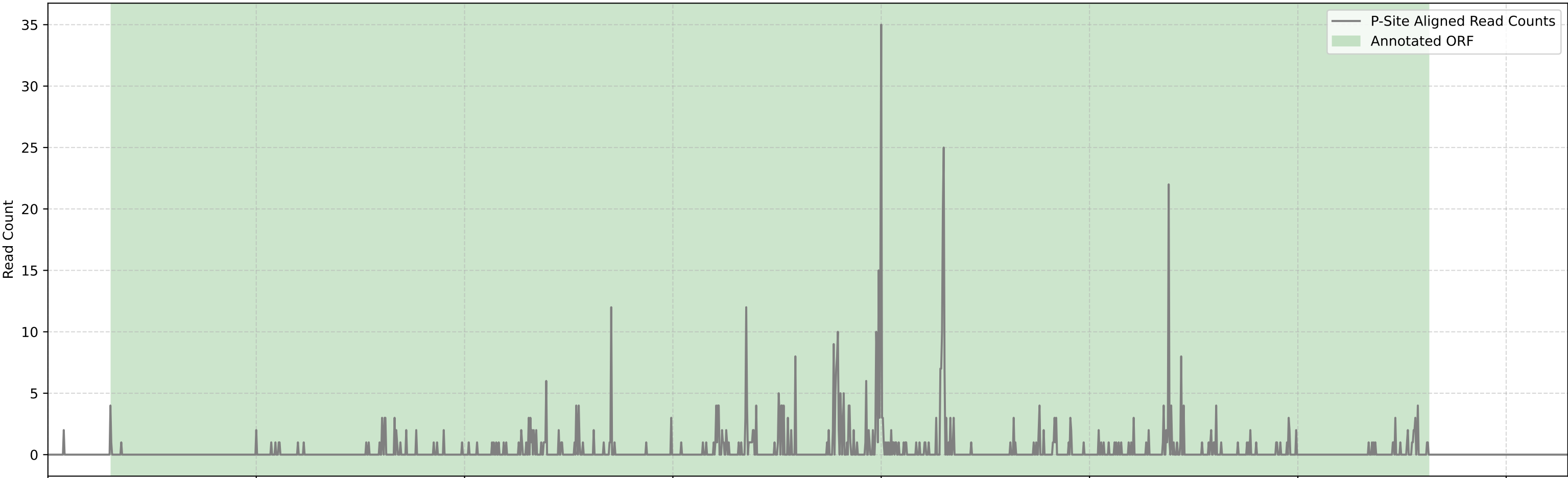
Example: 50th Percentile (by TPM) Transcript (ENST00000369722.8)
Total Reads: 547 | TPM: 2573.31



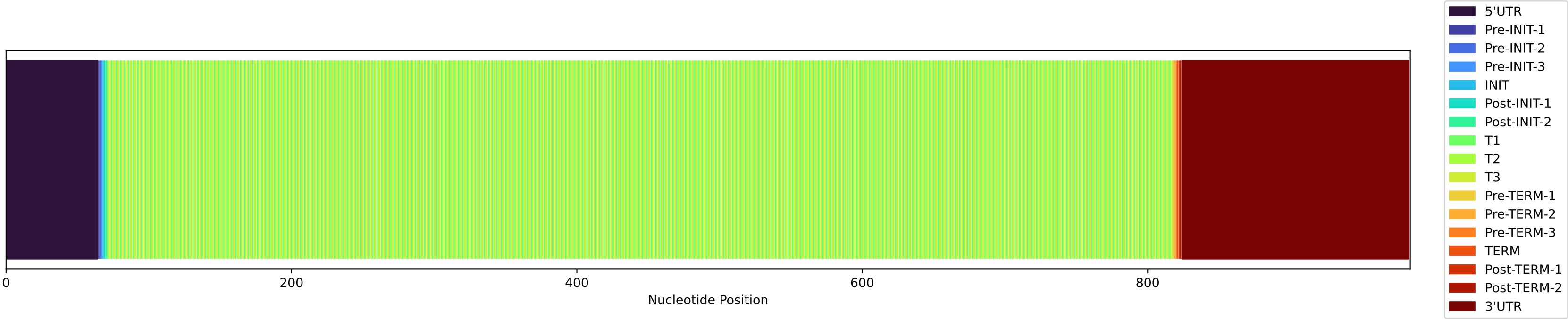
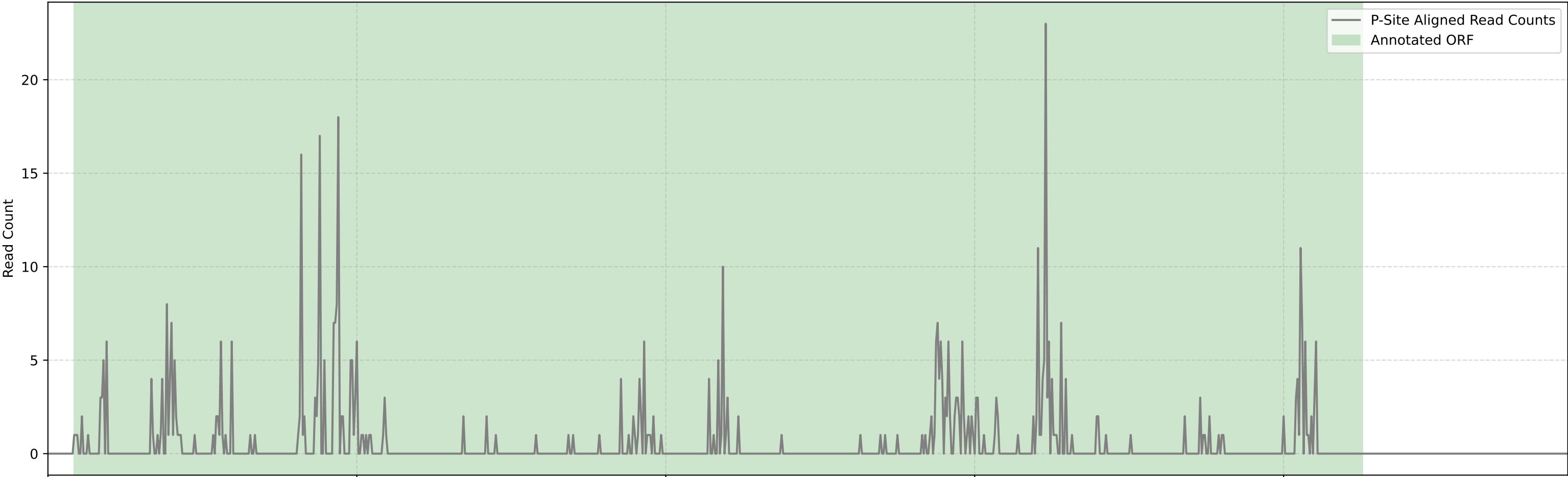
Example: 60th Percentile (by TPM) Transcript (ENST00000535923.6)
Total Reads: 404 | TPM: 3621.98



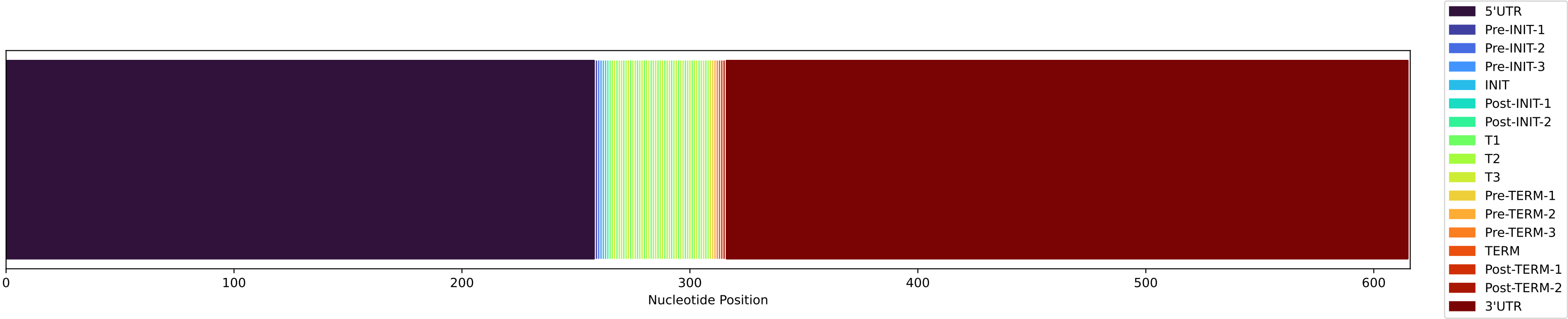
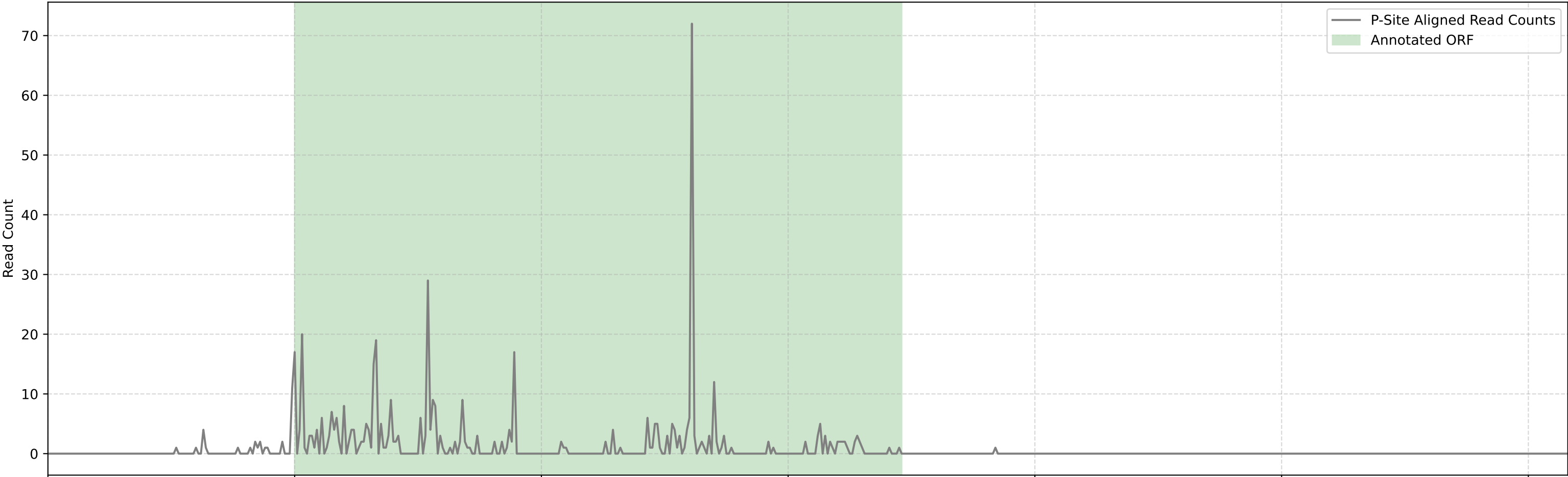
Example: 70th Percentile (by TPM) Transcript (ENST00000254108.12)
Total Reads: 665 | TPM: 4507.40



Example: 80th Percentile (by TPM) Transcript (ENST00000259457.8)
Total Reads: 510 | TPM: 6407.73



Example: 90th Percentile (by TPM) Transcript (ENST00000250495.10)
Total Reads: 513 | TPM: 10295.94



Example: Max Expression Transcript (ENST00000621411.3)
Total Reads: 2,480 | TPM: 60001.22

