

CellSpliceNet for Interpretable Multimodal Modeling of Alternative Splicing Across Neurons in *C. elegans*

Arman Afrasiyabi^{1,2,♡}, Jake Kovalic^{4,♡}, Chen Liu^{1,♡}, Egbert Castro^{3,♡},
Alexis Weinreb^{2,♡}, Erdem Varol^{6,♣}, David M. Miller III^{5,♠},
Marc Hammarlund^{2,♡}, Smita Krishnaswamy^{1,2,7,♡}

Department of {¹Computer Science, ²Genetics, ³Computational Biology and Biomedical Informatics, ⁴Applied Physics,
⁵Cell & Developmental Biology, ⁶Computer Science & Engineering, ⁷Applied Math, ⁸Biological Sciences}

♣New York University ♠Vanderbilt University ♡Yale University

Supplementary Materials

Attention in whole RNA sequences

Here we deepen the evidence for the conclusions drawn from Figure 3 of the main manuscript. Across a larger panel of events, they show that (i) **CellSpliceNet** consistently concentrates its attention on intron–exon boundaries, (ii) the precise distribution of those weights is flexibly re-tuned in a neuron-specific, event-dependent manner, and (iii) these attendant differences are captured by low-dimensional PHATE embeddings that reliably cluster neuron classes with similar regulatory signatures. Together, Supplementary Figures S1, S2, S3 demonstrate that the balance between universal splice-site recognition and context-aware modulation is a pervasive property of **CellSpliceNet** rather than an artefact of the examples chosen for the main text.

Attention in the structural modality

Supplementary Figures S4 and S5 extend the structural-modality analysis of Figure 4 in the main text. The additional heat-map panels chart nucleotide-level attention across twenty further neurons, reaffirming that **CellSpliceNet** funnels salience toward the exonic midpoint and tapers sharply into the flanking introns (Fig. S4). Overlaying the attention scores on the secondary structures for twenty extra cassette-exon events (Fig. S5) reveals a consistent topological bias: single-stranded loop nucleotides inside the exon draw the highest weights, whereas paired stem regions remain comparatively muted. Together, these expanded examples demonstrate the robustness of **CellSpliceNet**’s preference for accessible, centrally located motifs when inferring PSI.

Attention in the ROI modality

We also extend the ROI-level interpretability analysis presented in Figure 5 of the main manuscript. Fig. S6 aggregates nucleotide-wise attention maps for 20 additional cassette-exon events, confirming that **CellSpliceNet** consistently concentrates weight over the exon core and upstream splice junction while rapidly decaying into the flanking introns. Fig. S7 lifts this view to the motif scale, displaying heat maps in which each subplot corresponds to a neuron type and each row to one of the top ROI motifs.

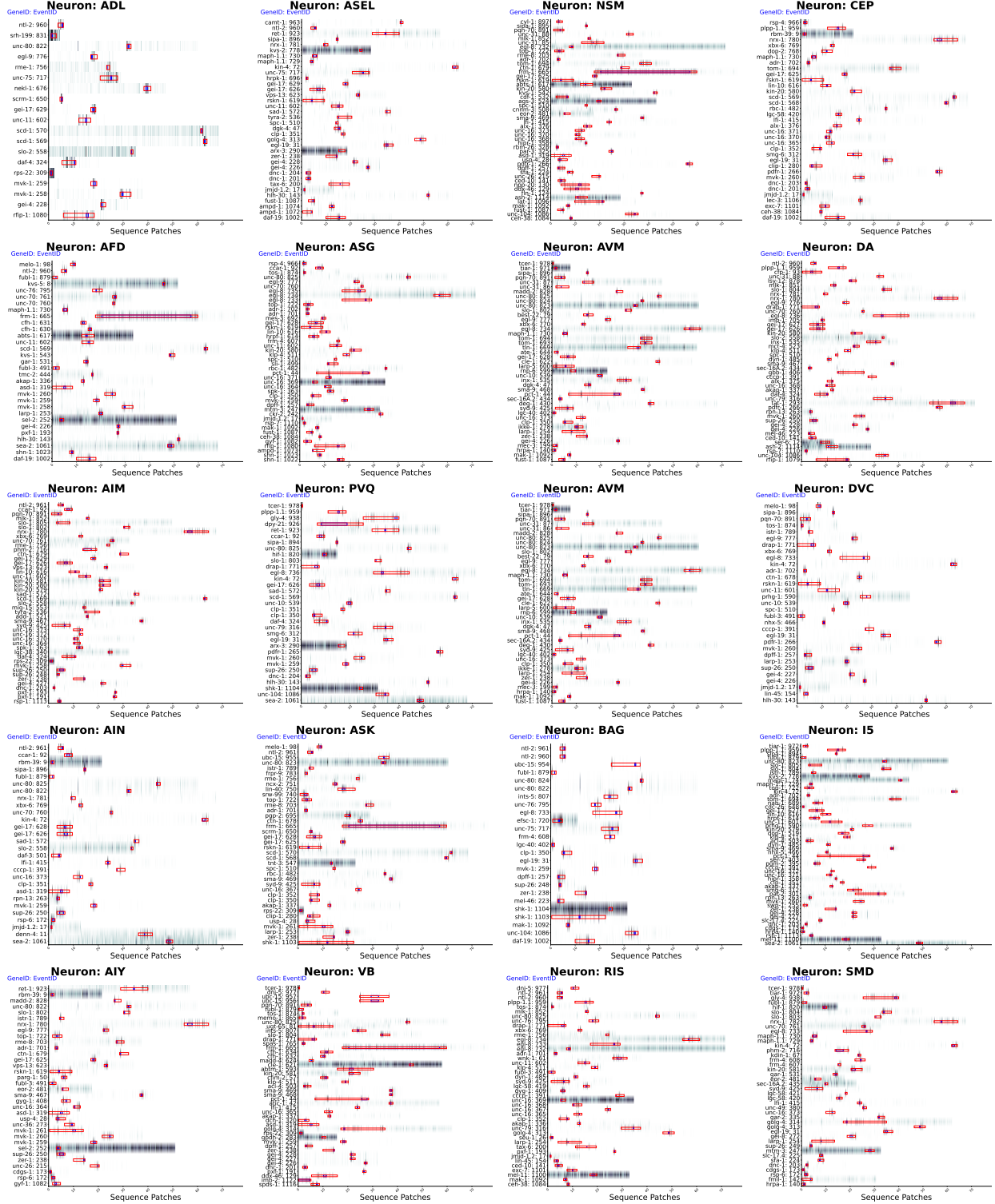


Fig. S1: Expanded RNA patch-level attention heat-maps. Heat maps for twenty additional neurons. Color encodes the mean attention weight assigned by CellSpliceNet. The blue boxes show the exon regions, and the red boxes show the introns.

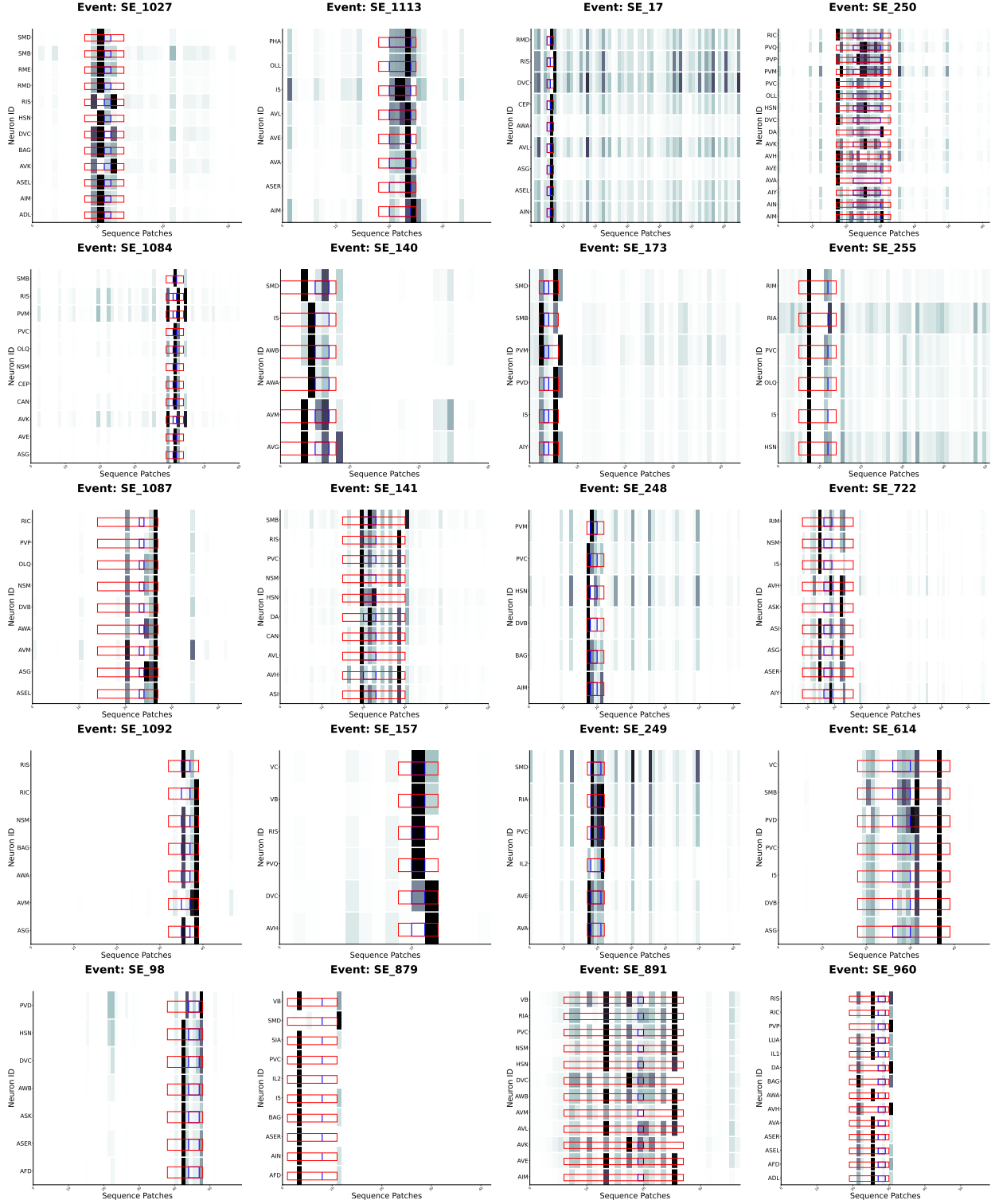


Fig. S2: **Neuron-specific redistribution of attention across additional events.** Stacked bar plots compare per-patch attention profiles for 20 unique events across different neuron types (in each subplot) in each subplot's y axis.

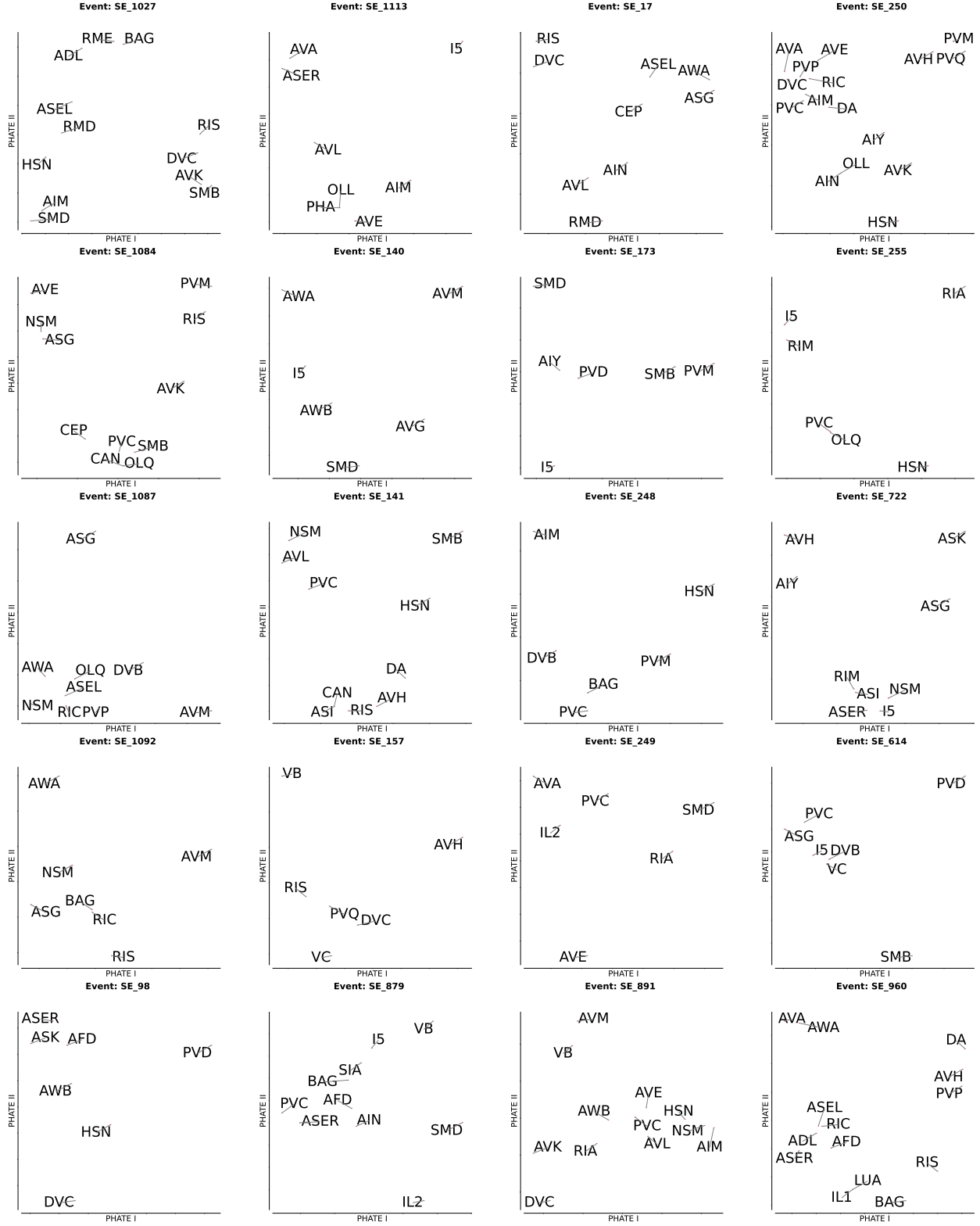


Fig. S3: **PHATE embedding different events across the neuron types.** Each point in a subplot represents a neuron class; axes are the first two PHATE dimensions computed from its attention vector for the corresponding splicing event. Neuron classes with similar attention distributions cluster tightly.

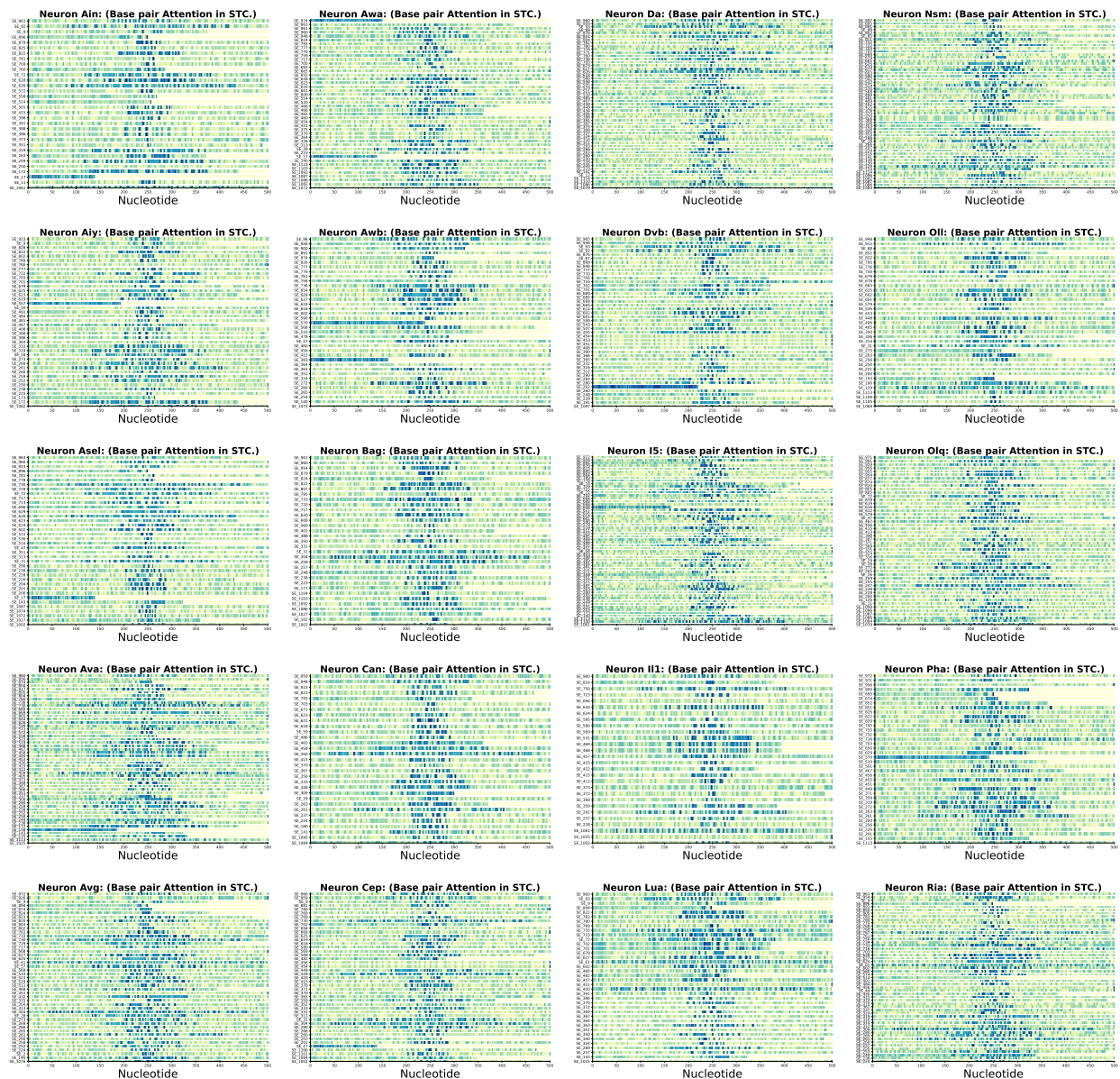


Fig. S4: **Structure-modality attention concentrates on exonic loop nucleotides** Heat maps depict nucleotide-level attention weights assigned by the structural modality of CellSpliceNet. Warmer colors indicate higher salience. Across all cases, attention peaks sharply within the exon center and decays monotonically toward splice junctions and intronic regions.

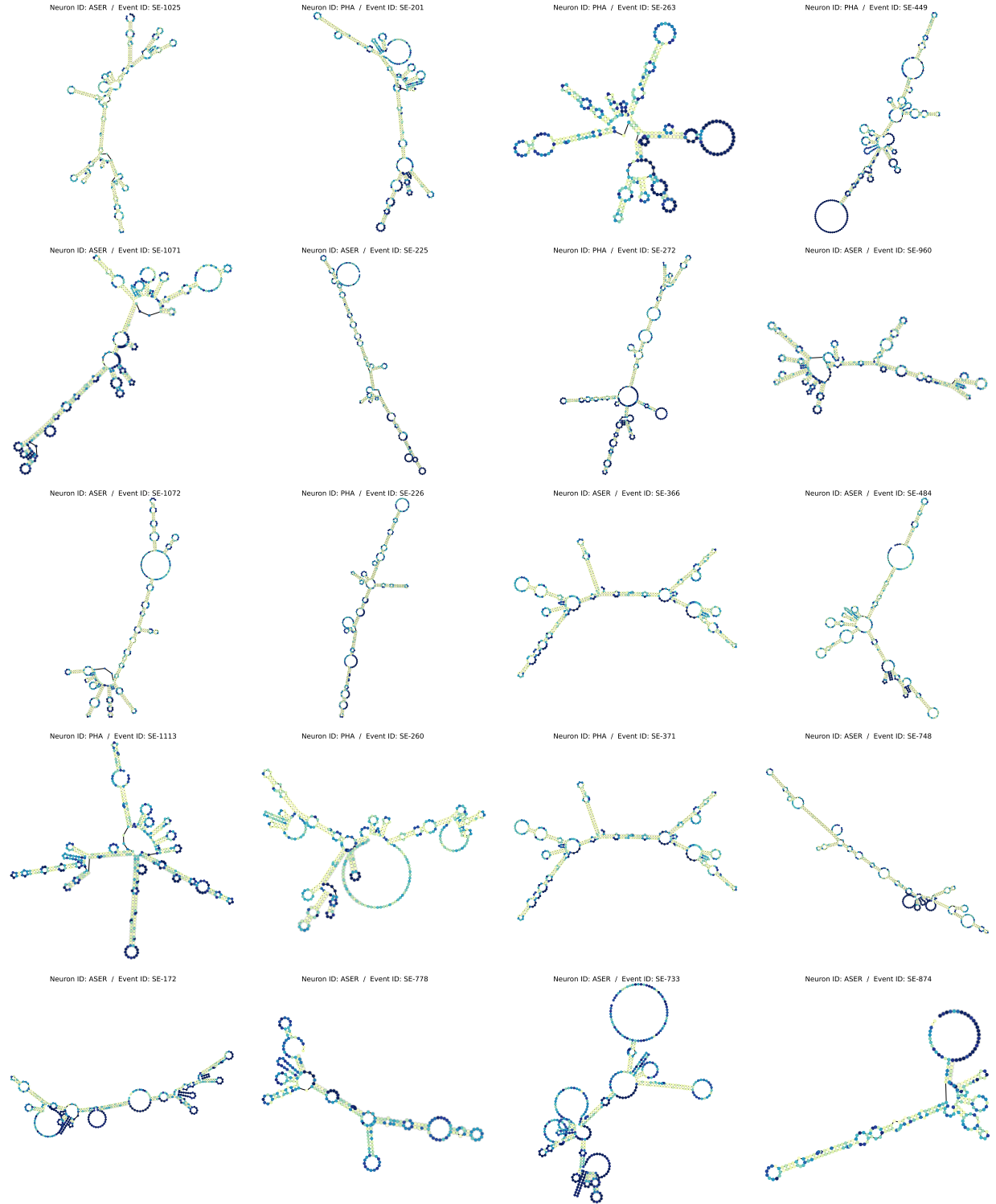


Fig. S5: **Structure-modality attention concentrates on secondary structures** Attention scores are overlaid on the secondary structures for twenty additional cassette-exon events; dark blue denotes high attention, whereas light yellow indicates low attention as inferred by CellSpliceNet.



Fig. S6: **Heightened attention at ROI.** Nucleotide-resolution attention maps for representative events of 20 different neuron classes reveal a pronounced focus on the upstream splice boundary.

