SUPPLEMENTAL MATERIALS

Supplemental Figures.

Supplemental Figure 1. Distance to 3'ss for composite exon poly(A) sites (A) and skipped exon poly(A) sites (B).

Supplemental Figure 2. Exon size for composite terminal exons (A) and skipped terminal exons (B). Internal exons and 3'-most exons are also plotted.

Supplemental Figure 3. C-terminal protein sequences of two cyclin C isoforms. Hs.cyclinC.a and Mm.cyclinC.a are protein isoforms encoded by human and mouse transcripts containing exon 12', a skipped terminal exon. Hs.cyclincC.b and Mm.cyclinC.b are isoforms encoded by human and mouse transcripts containing exon 12, a 3'-most exon. See Supplemental Figure 4 for details. Protein sequences encoded by exon 11 and exon 12/12' are separated by a dotted line. The similarity between two protein isoforms is indicated by symbols above their alignment, with "*" representing identical aa; ":" representing highly similar aa; and "." representing similar aa. The PEST motif was found by the pestfind program from EMBOSS.

Supplemental Figure 4. Primers used for confirmation of intronic polyadenylation for 9 genes. The exons and introns relevant to the intronic polyadenylation are shown for each gene. Poly(A) sites are indicated by vertical arrows and pA. QPCR primers are indicated by horizontal arrows, and their sequences are shown.

Supplemental Figure 5. Potential base-pairing of 5'ss sequence with U1 snRNA for introns with composite exon poly(A) sites (A) and introns with skipped exon poly(A) sites (B). For (A) and (B), the left panel is ΔG and the right panel is number of potential base pairs.

Supplemental Figure 6. (A) Distributions of intron size for three groups of introns. (B) mKS test for introns without poly(A) sites vs. introns with composite exon poly(A) sites. For both groups, only introns >512 nt were used for the comparison.

Supplemental Figure 7. mKS tests for skipped exon poly(A) sites. (A) Upstream 3'ss scores of introns containing skipped exon poly(A) sites vs. 3'ss scores of introns without poly(A) sites. (B) Downstream 3'ss scores of introns containing skipped exon poly(A) sites vs. 3'ss scores of introns without poly(A) sites. (C) Downstream 3'ss scores vs. upstream 3'ss scores.

Supplemental Figure 8. Upstream (A) and downstream (B) poly(A) region scores. Ten groups of poly(A) sites are shown, as indicated. P-values from the Wilcoxon tests comparing each group with group 1 and group 2 are also shown.

Supplemental Figure 9. Conservation of PAS type between human and rat orthologous poly(A) site pairs. See the figure legend for Figure 5B for details.

Supplemental Tables.

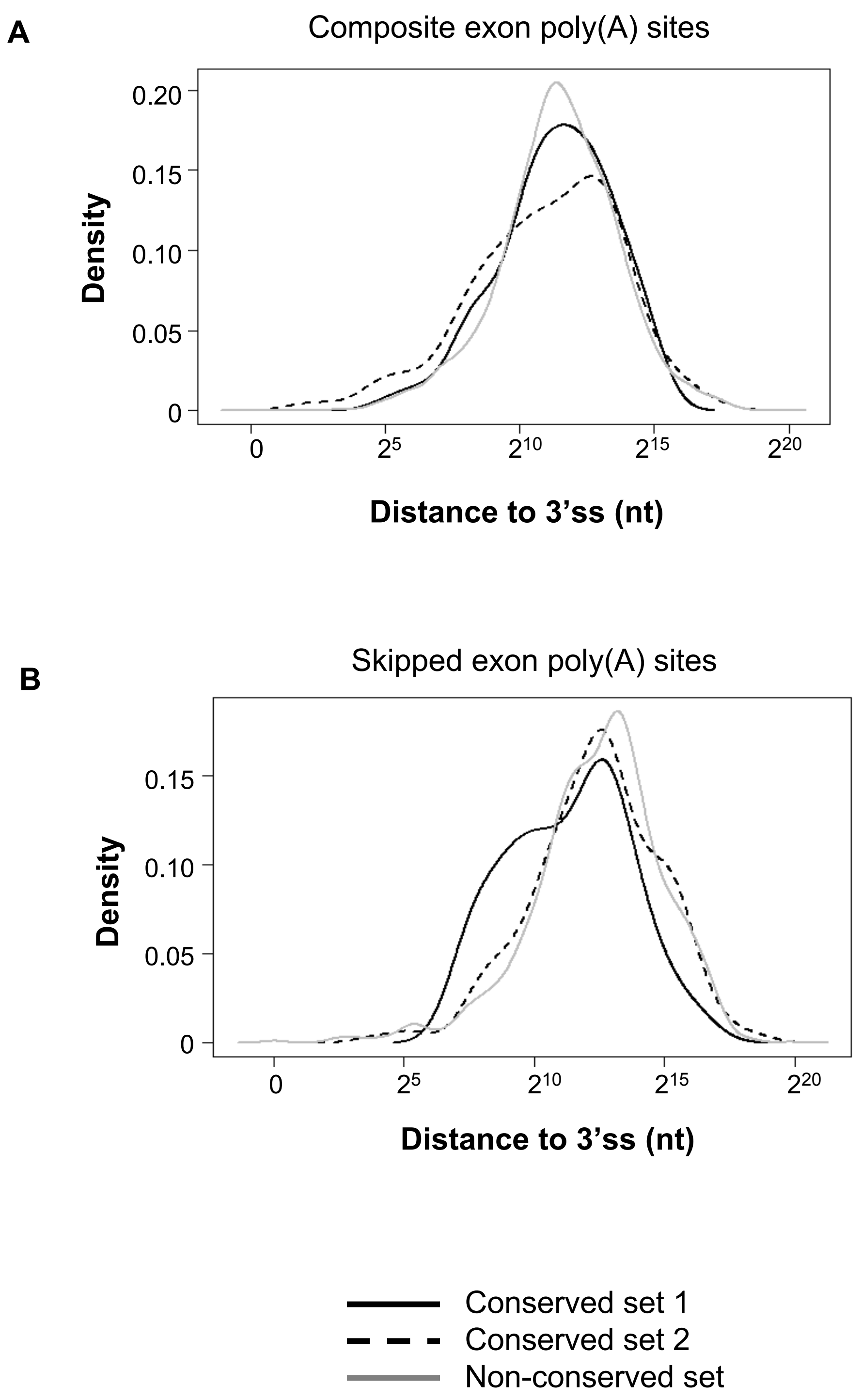
1. Supplemental Table 1. Intronic polyadenylation events for poly(A) sites in the conserved set 1. Site ID, poly(A) site ID in the PolyA_DB 2 database (http://polya.umdnj.edu/PolyA_DB2/); Intron, genomic location of the intron containing poly(A) site, shown as chromosome number: 5' splice site position: 3' splice site position; Poly(A) site, genomic location of poly(A) site; Type, type of terminal exon resulting from intronic polyadenylation (b, both; c, composite terminal exon; s, skipped terminal exon); NM/KG, RefSeq or KnownGene Accession; Inum, intron number; Region, mRNA region affected by intronic polyadenylation (5UTR, 5' UTR; 3UTR, 3' UTR; CDS, Coding sequence); Gene

ID, gene identification number in the NCBI Gene database; Gene Symbol, Gene symbol in the NCBI Gene database; Gene Annotation, gene annotation obtained from the NCBI Gene database.

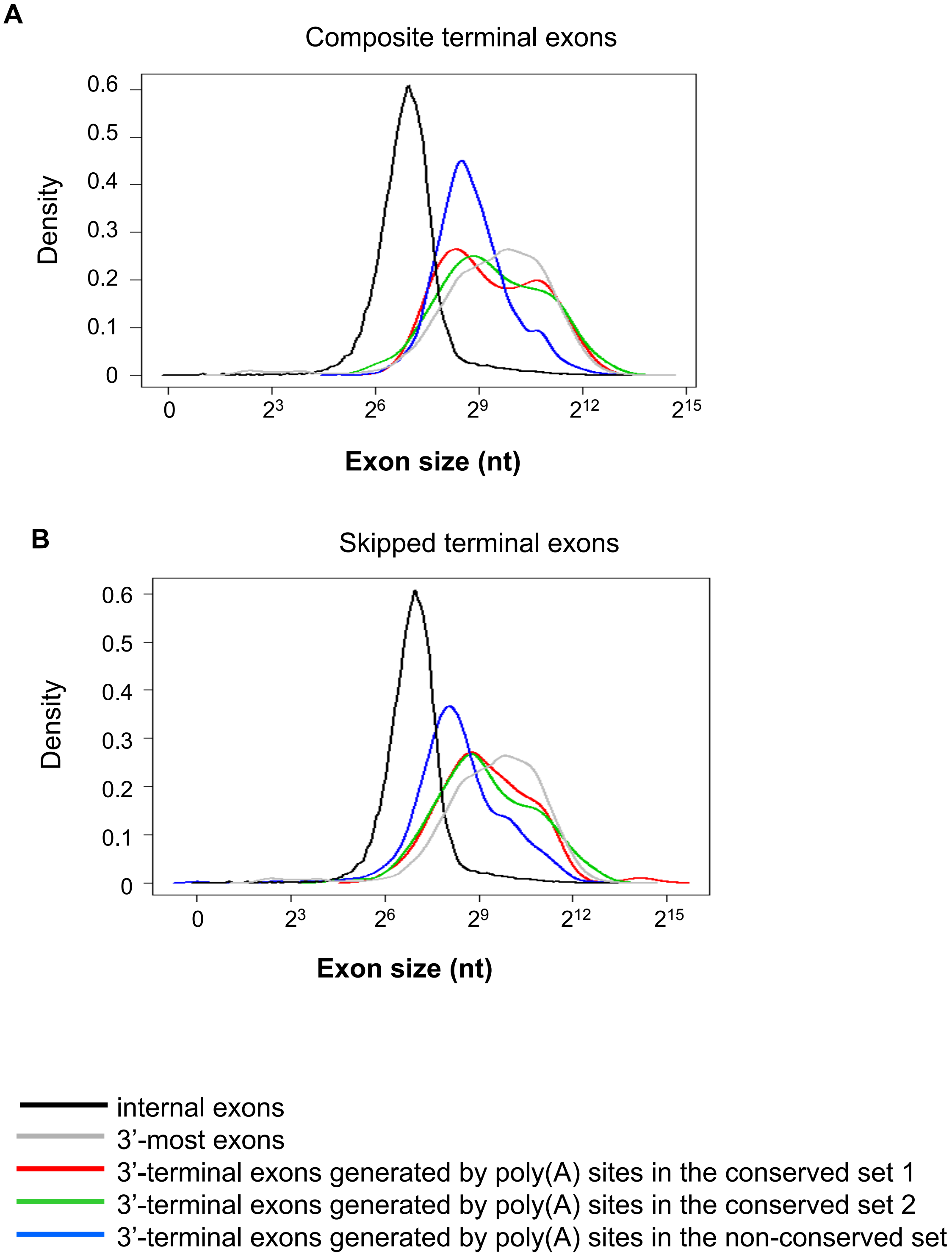
2. Supplemental Table 2. Intronic polyadenylation events for poly(A) sites in the conserved set 2. See Supplemental Table 1 for annotations.

3. Supplemental Table 3. Intronic polyadenylation events for poly(A) sites in the non-conserved set. See Supplemental Table 1 for annotations.

Supplemental Figure 1.



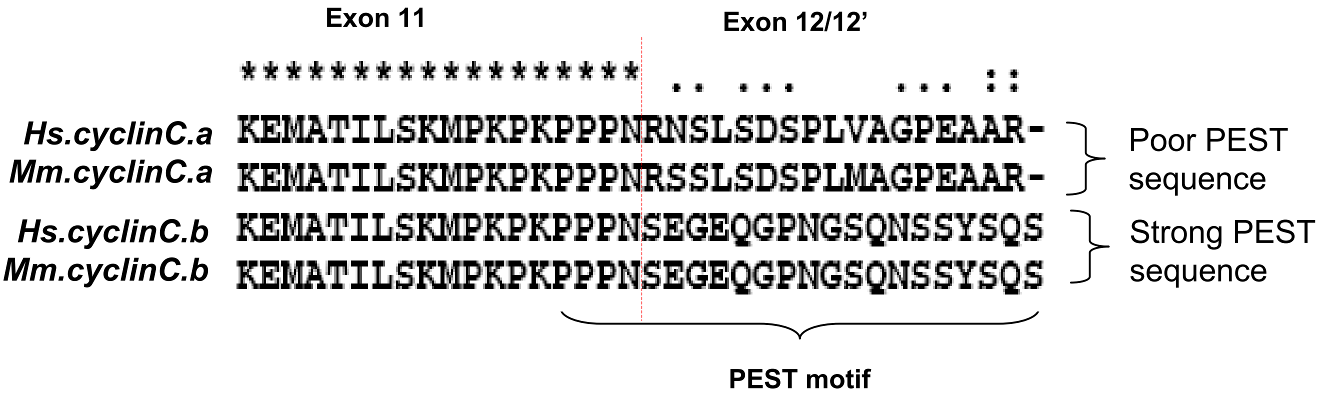
Supplemental Figure 2.



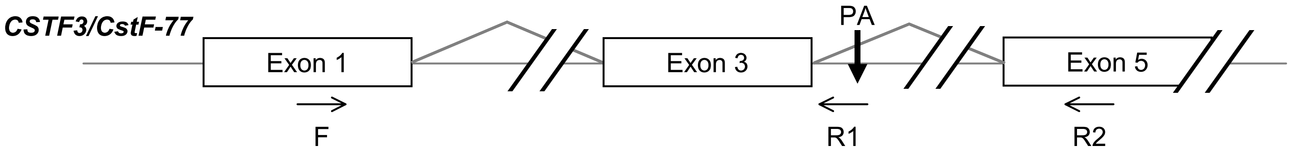
A Composite terminal exons

B Skipped terminal exons

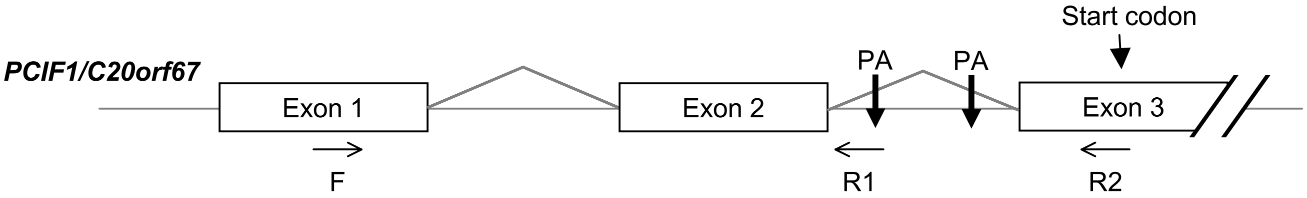
Supplemental Figure 3.



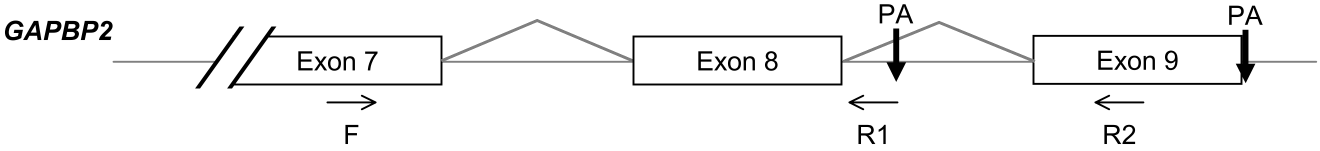
Supplemental Figure 4.



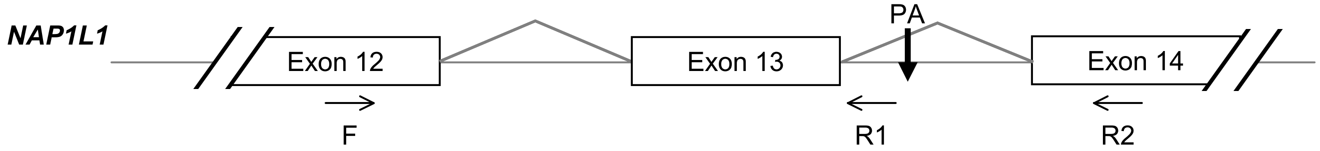
F: 5’-GAGGCCATGTCAGGAGAC
R1: 5’-GCAACTCCAAAATGCAACAA
R2: 5’-CATAAATCAATGTGCAAAACC



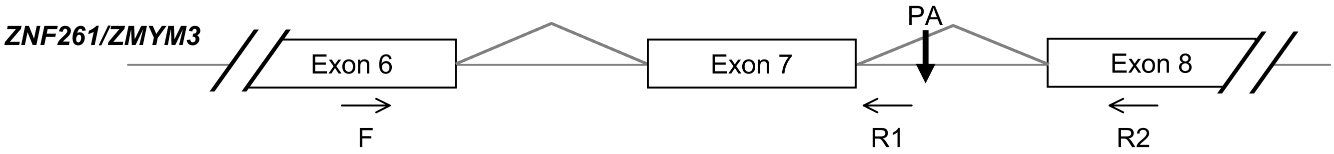
F: 5’-CGGTCGAGCAGAACGTGT
R1: 5’-TGAAGTCCCCCTTTCTCTCA
R2: 5’-CTGCCGTGATTCTCATTGG



F: 5’-GGTCAGCCCATCATTGTGAC
R1: 5’-CAATATTTATTTATTGAGGGCTTGC
R2: 5’-GCTTCATCCAGCTGTTTCTG



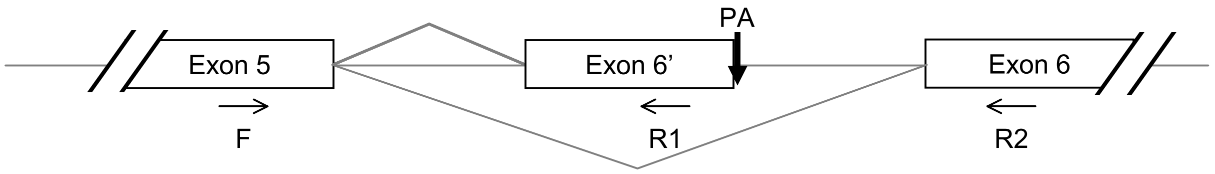
F: 5’-GAAATTGGTCACTTTTTACGTGAG
R1: 5’-TTATTTGCTCATTTGGTAAACATT
R2: 5’-CATCTCCTTCTTCTTCCCCTTC



F: 5’-TGCAGCATATGCCAGAAGAC
R1: 5’-AGCCCTCTAATGCACTACTCC
R2: 5’-TTACACAGGGTCTTGCACCA

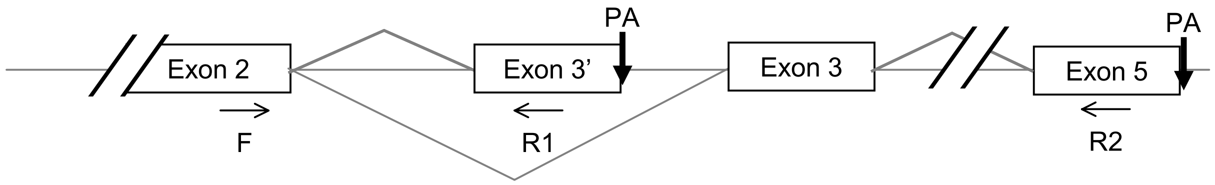
Supplemental Figure 4.

CDC42



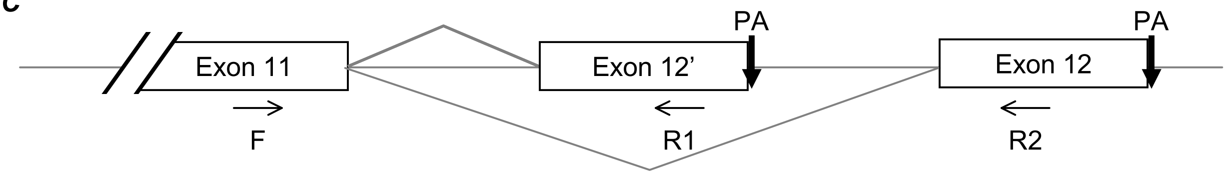
F: 5’-CTGAAGGCTGTCAAGTATGTGG
R1: 5’-TCTTTCTACAGTAGTGGGACAGGA
R2: 5’-CTGTGCAGAAAGGGCTCTG

TAF9



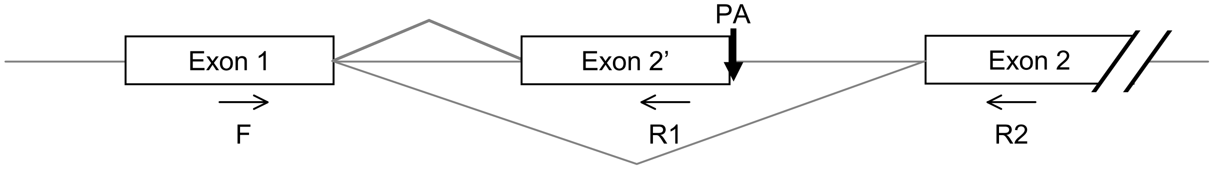
F: 5’-TACACCAGGGGTTGGAAAAA
R1: 5’-TGCCATCATCTGTGCATCTT
R2: 5’-AAAGAACTTGAAAAATCTCACACTGA

Cyclin C



F: 5’-GATGAGAGAAAAGAGATGGCAAC
R1: 5’-CAGGTATCACTCAACATAGAGATTACA
R2: 5’-CCAATATGCTTGACAGAAACAAG

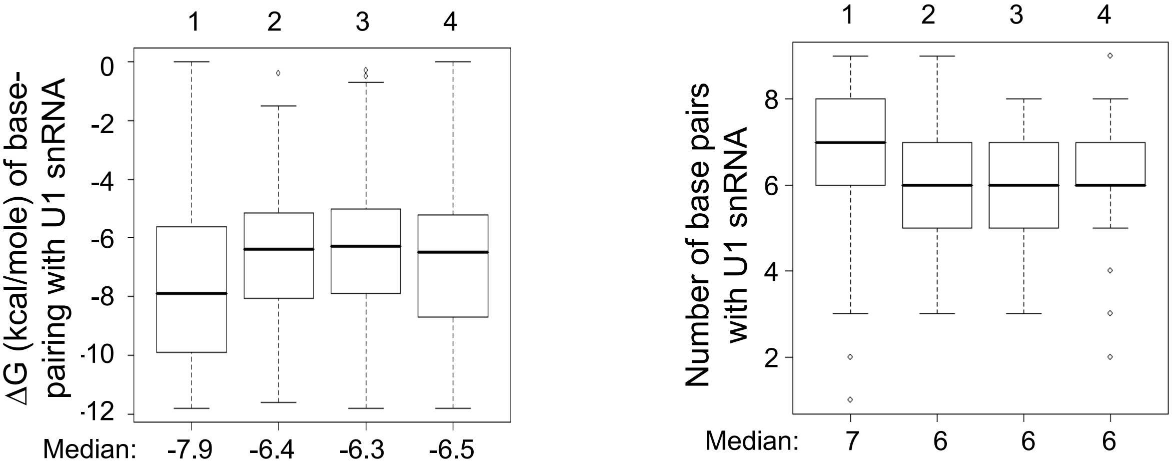
EPC1



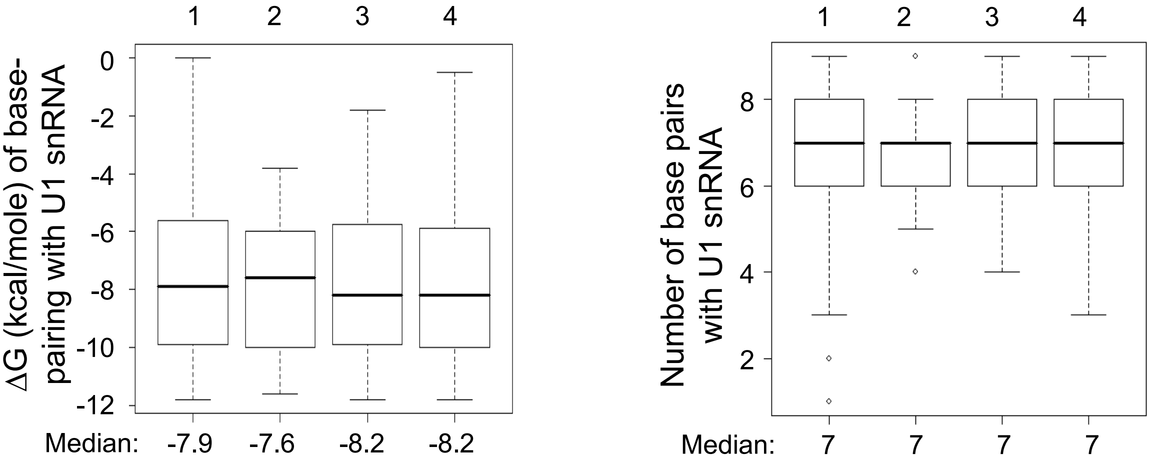
F: 5’-GCACGAATACGCCTCGATAA
R1: 5’-CAAAATTGTAGCACACAAATCCA
R2: 5’-CATACACCTGCTGTGCTGAAA

Supplemental Figure 5.

A. Introns containing composite exon poly(A) sites.



B. Introns containing skipped exon poly(A) sites.



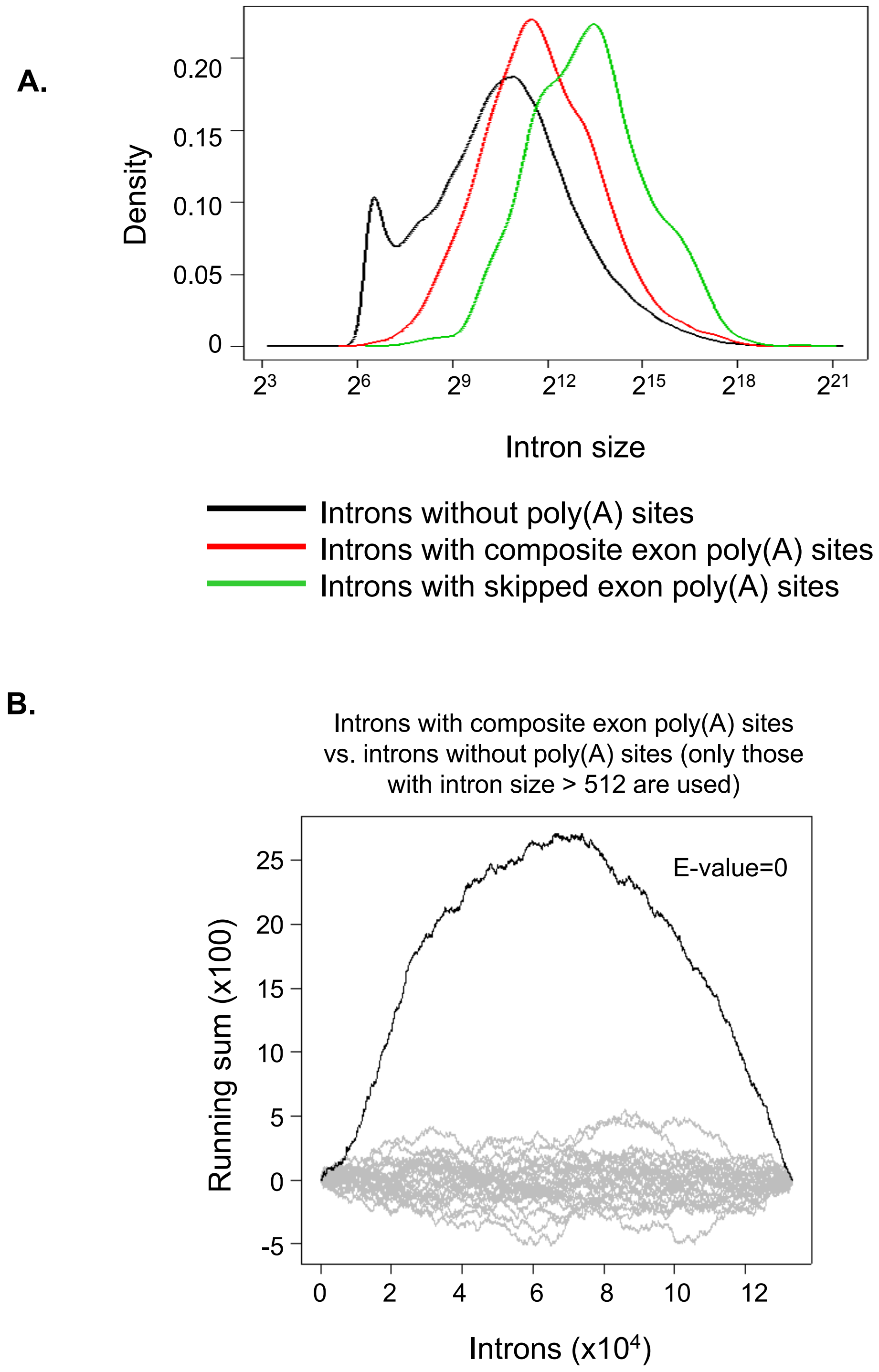
Group 1: Introns without poly(A) sites

Group 2: Conserved set 1

Group 3: Conserved set 2

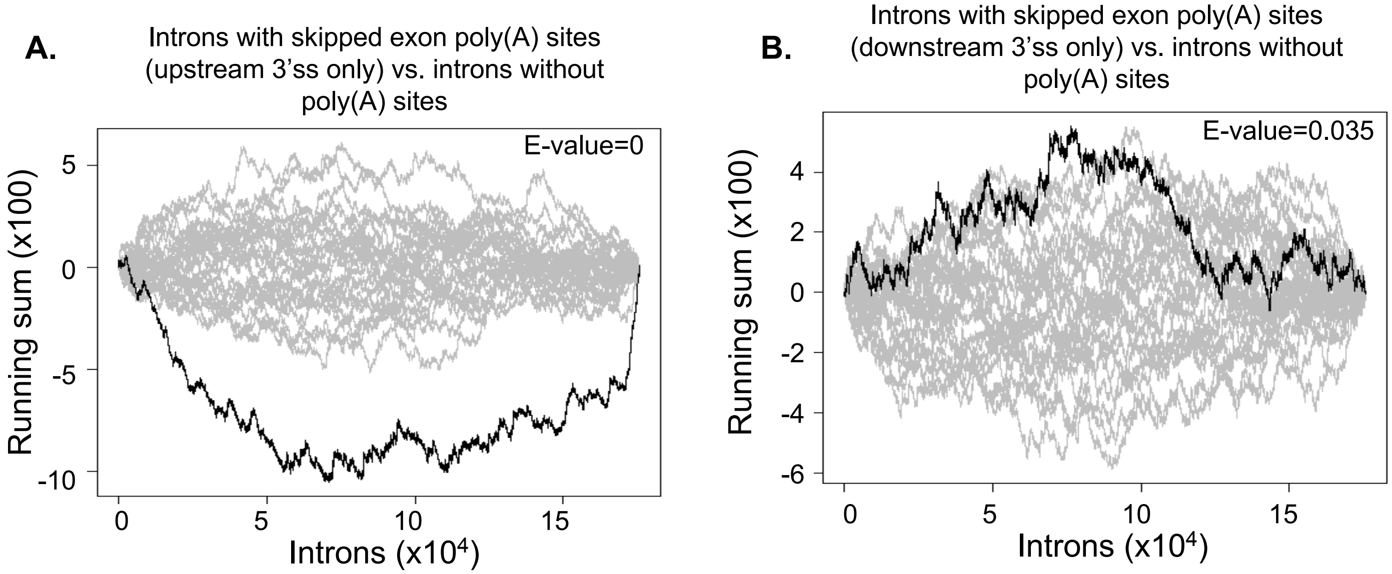
Group 4: Non-conserved set

Supplemental Figure 6.



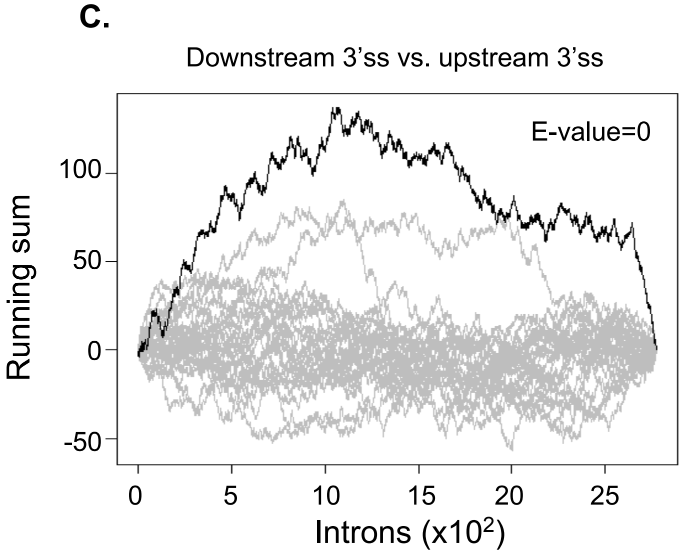
Introns with composite exon poly(A) sites vs. introns without poly(A) sites (only those with intron size > 512 are used)

Supplemental Figure 7.



A. Introns with skipped exon poly(A) sites (upstream 3’ss only) vs. introns without poly(A) sites

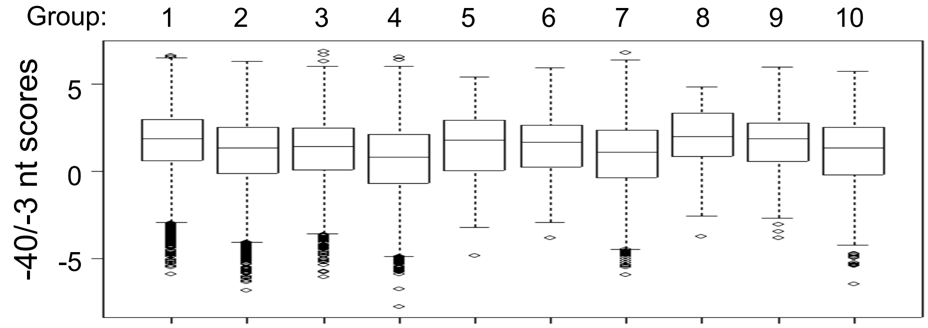
B. Introns with skipped exon poly(A) sites (downstream 3’ss only) vs. introns without poly(A) sites



C. Downstream 3’ss vs. upstream 3’ss

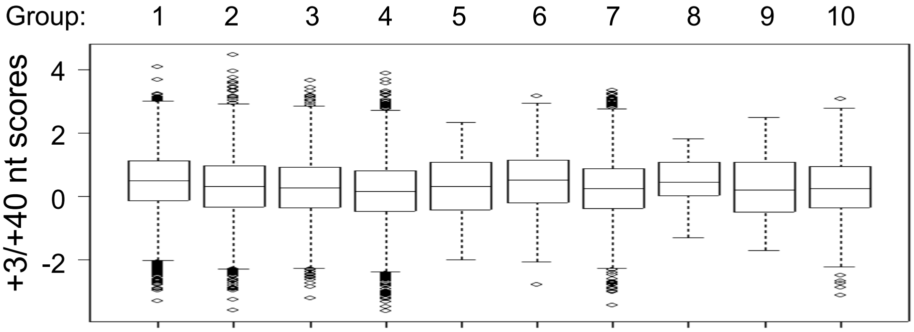
Supplemental Figure 8.

A. Upstream (-40 to -3 nt) scores



Group:	1	2	3	4	5	6	7	8	9	10
Median:	1.87	1.31	1.40	0.80	1.76	1.66	1.07	1.97	1.84	1.34
P-value (vs. group 1):	-	0	0	0	0.52	0.06	0	0.58	0.25	0
P-value (vs. group 2):	0	-	0.03	0	0.07	0.05	$6x10^{-7}$	$9x10^{-3}$	$9x10^{-3}$	0.61

B. Downstream (+3 to +40 nt) scores



Group:	1	2	3	4	5	6	7	8	9	10
Median:	0.49	0.32	0.27	0.16	0.32	0.52	0.25	0.45	0.21	0.26
P-value (vs. group 1):	-	0	0	0	0.23	0.97	0	0.86	$6x10^{-3}$	$1x10^{-10}$
P-value (vs. group 2):	0	-	$5x10^{-3}$	0	0.78	0.04	$1x10^{-3}$	0.13	0.43	0.42

Group 1: Conserved 3'-most poly(A) sites
Group 2: Non-conserved 3'-most poly(A) sites
Group 3: Conserved other poly(A) sites in 3'-most exons
Group 4: Non-conserved other poly(A) sites in 3'-most exons
Group 5: Composite exon poly(A) sites in the conserved set 1
Group 6: Composite exon poly(A) sites in the conserved set 2
Group 7: Composite exon poly(A) sites in the non-conserved set
Group 8: Skipped exon poly(A) sites in the conserved set 1
Group 9: Skipped exon poly(A) sites in the conserved set 2
Group 10: Skipped exon poly(A) sites in the non-conserved set

Supplemental Figure 9.