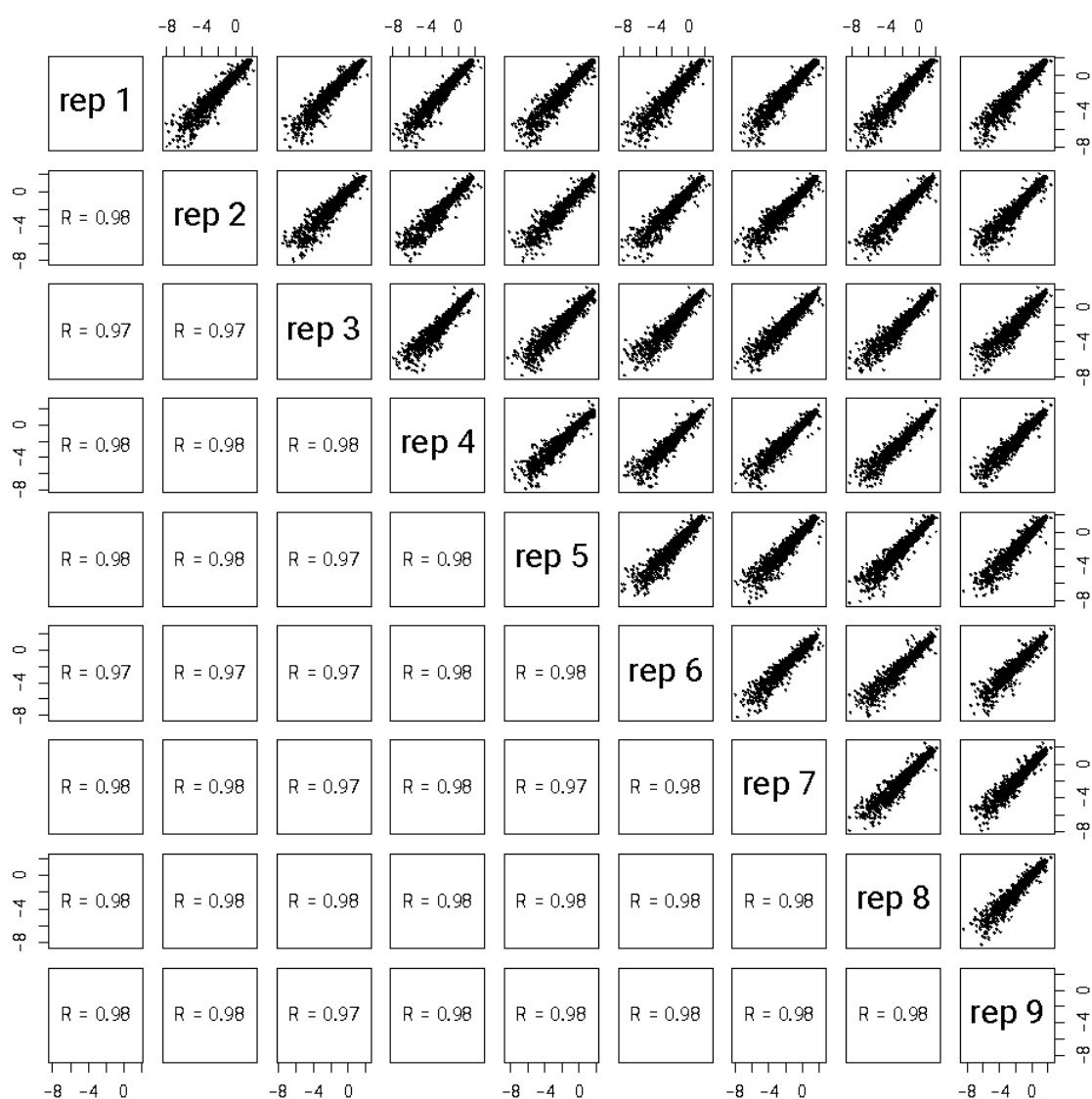
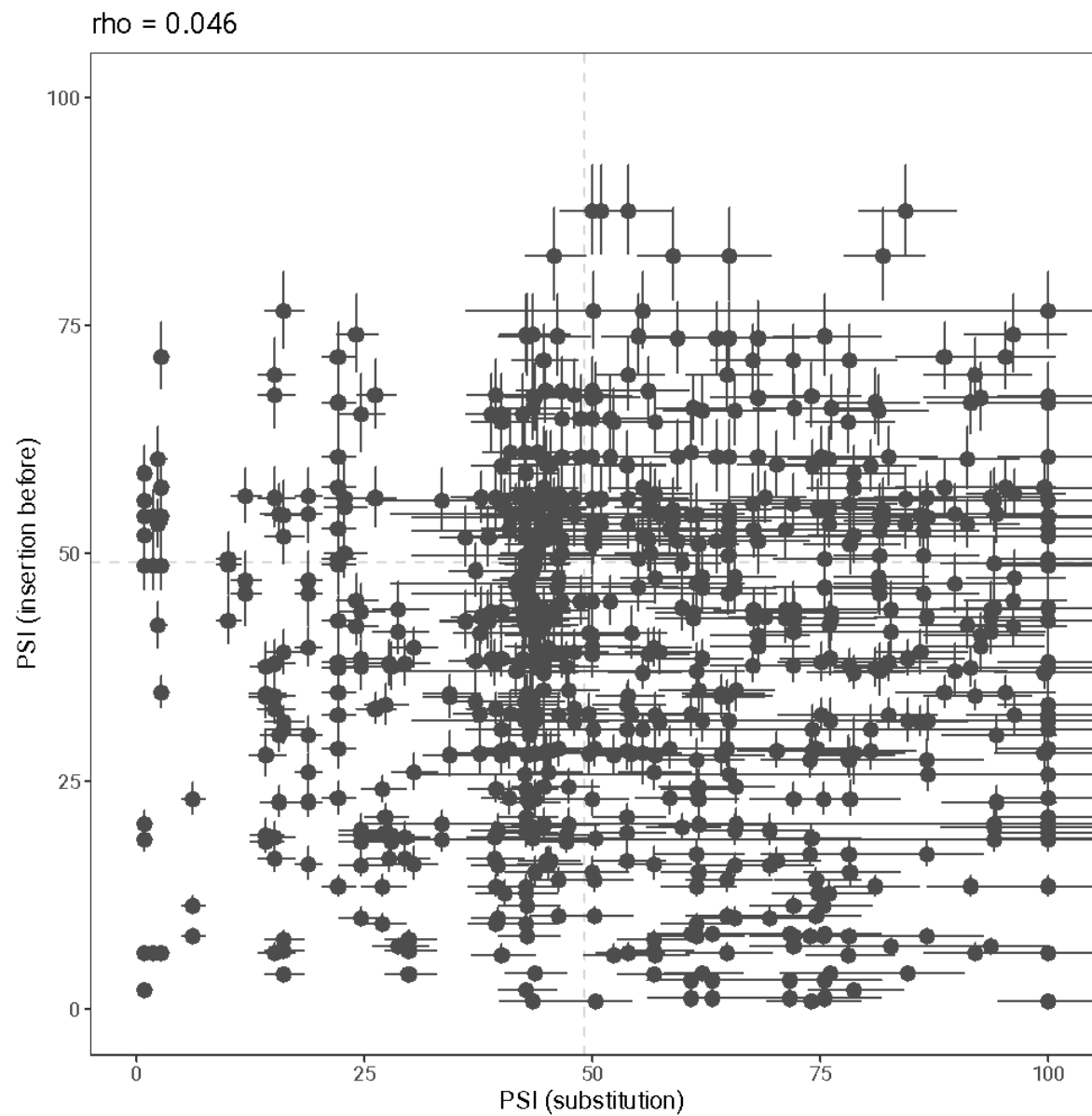


Supplementary Figure 1



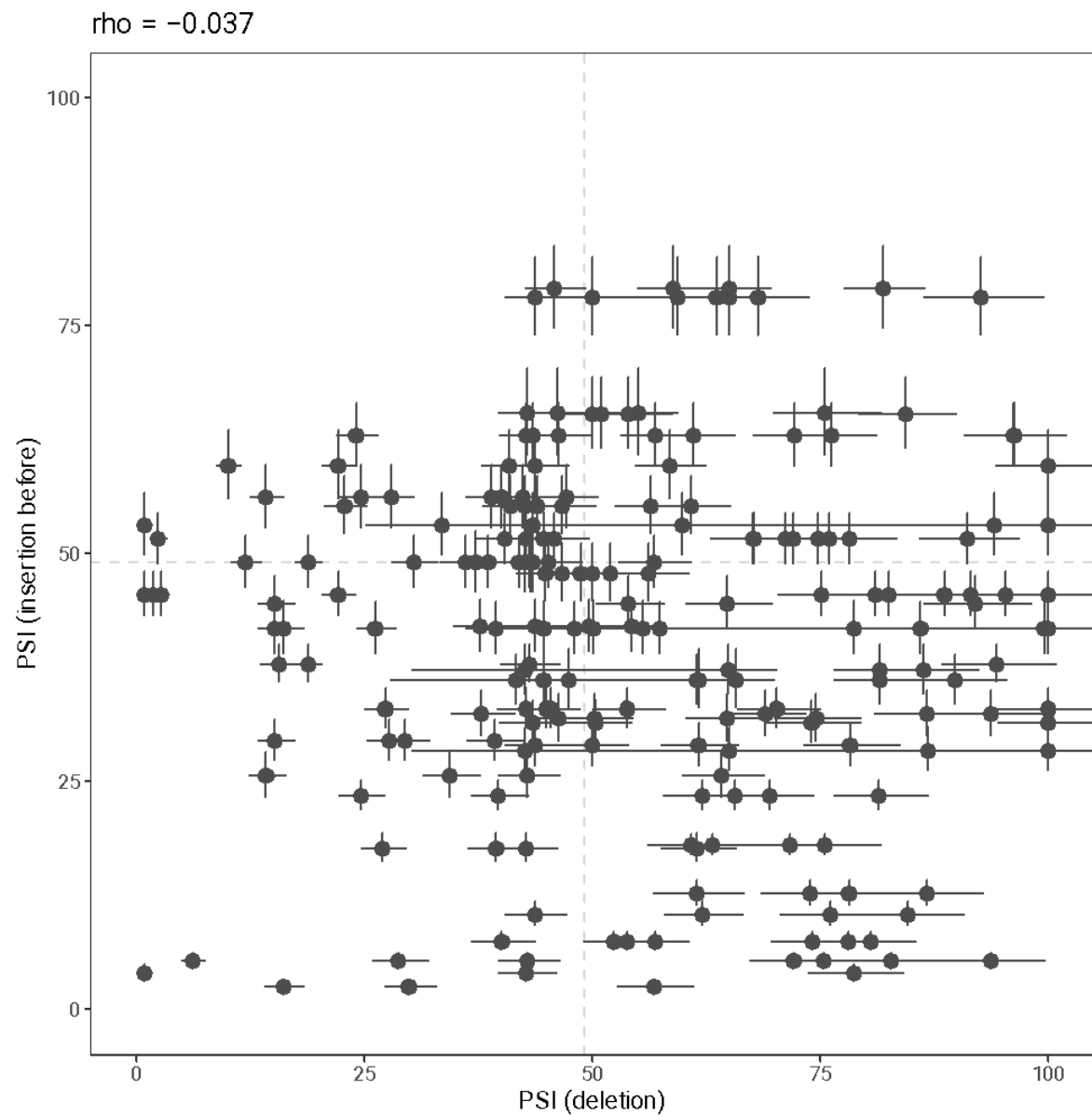
Pairwise correlations of enrichment scores for all exon variants in our library across nine experimental replicates.

Supplementary Figure 2



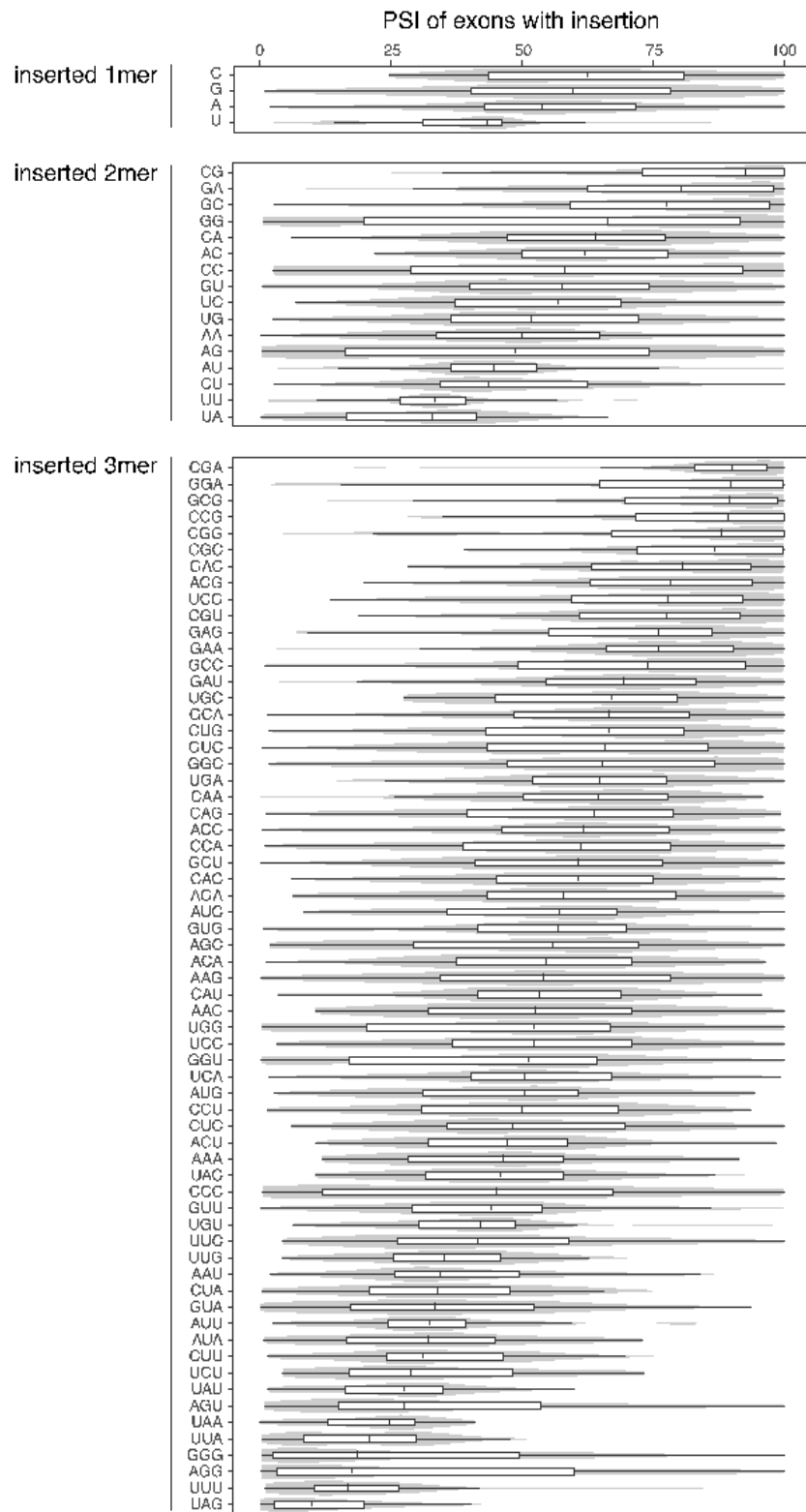
Correlation between PSI values for 1-nt insertions before a given position and substitutions at that position.

Supplementary Figure 3



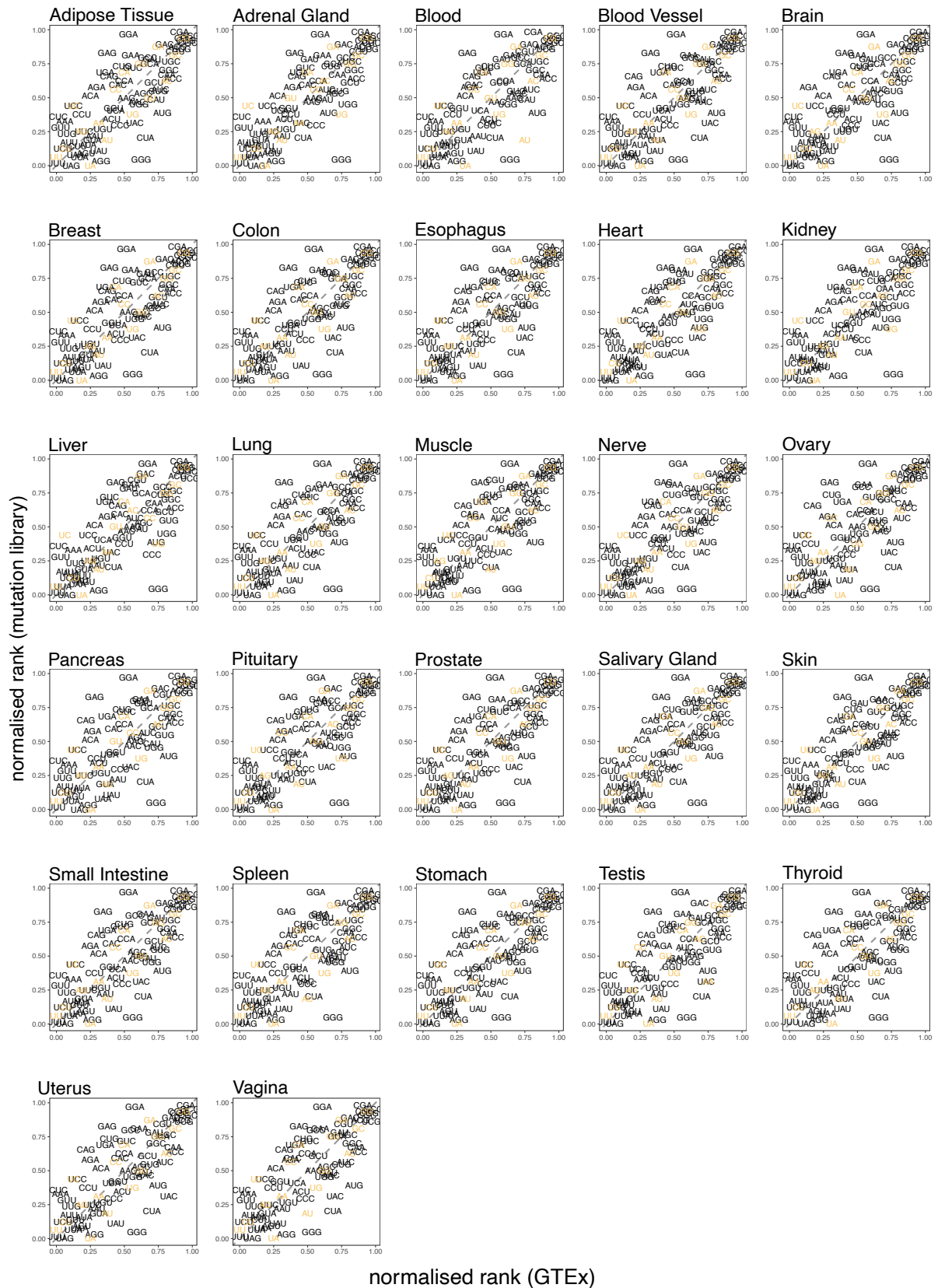
Correlation between PSI values for 1-nt insertions before a given position and deletions at that position.

Supplementary Figure 4



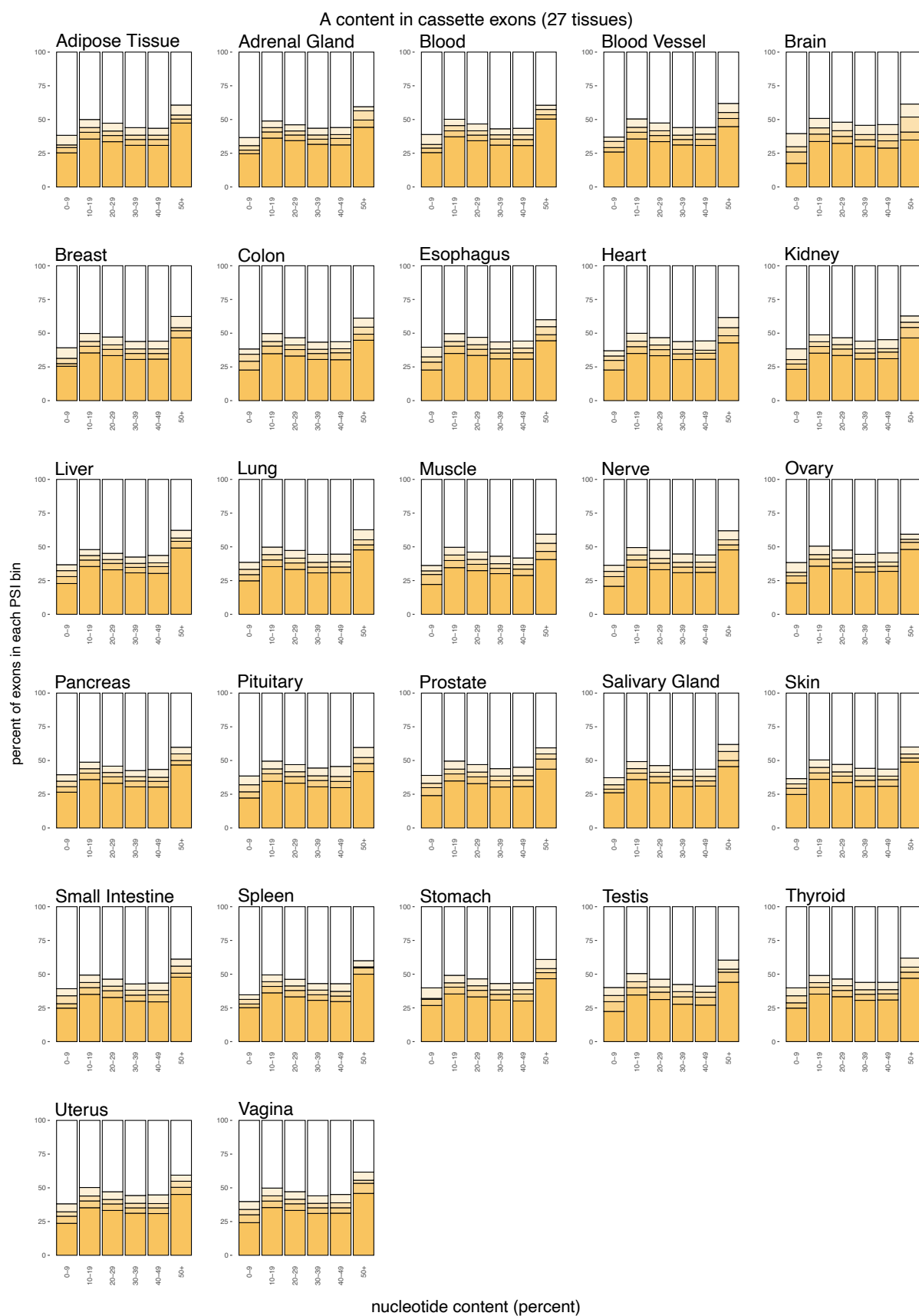
Distribution of inclusion values for exons with each possible 1-, 2-, and 3-nt insertion.

Supplementary Figure 5



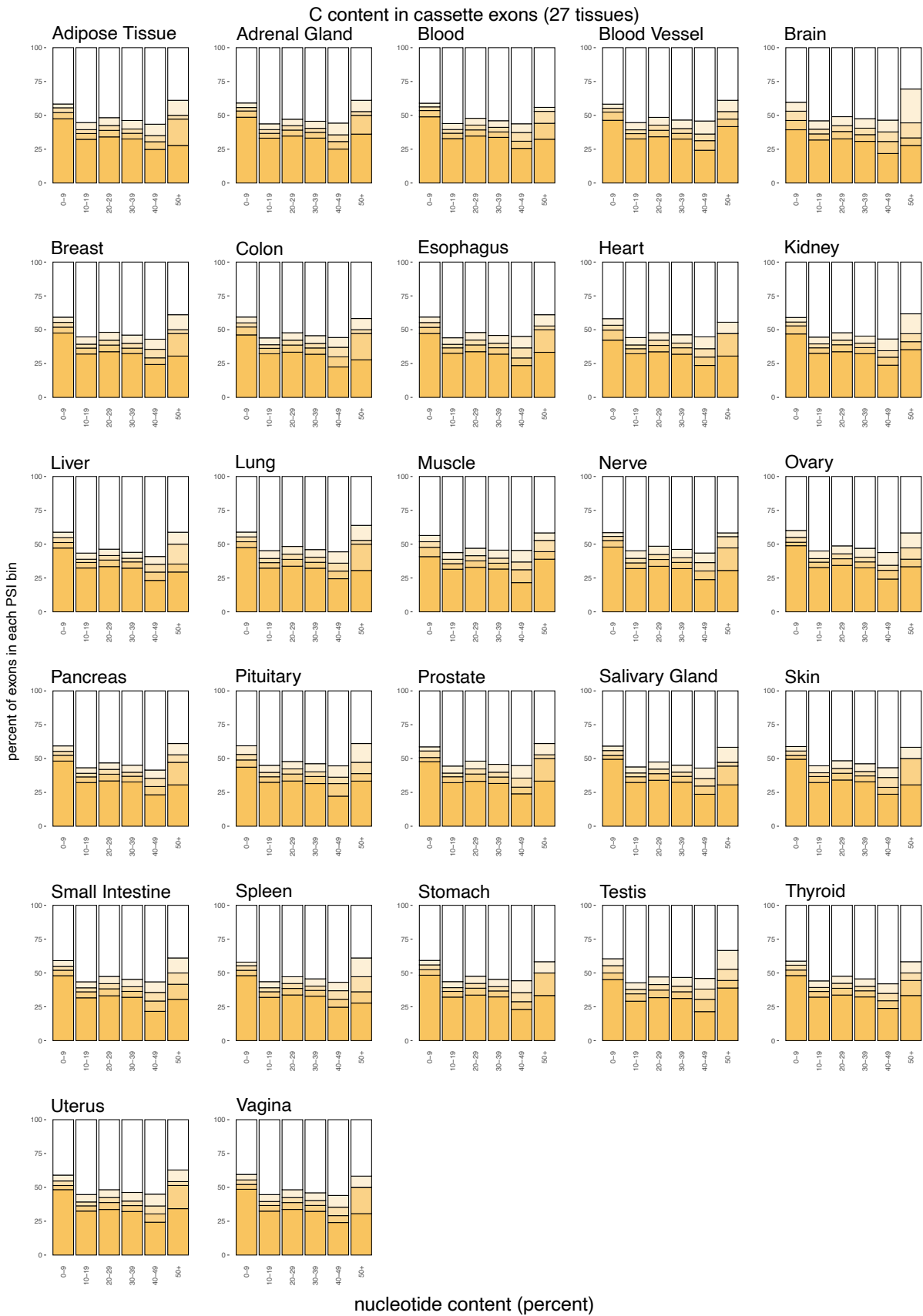
Correlation between 2 (or 3) nt sequences ranked by the median PSI of exon variants in our library, with this sequence inserted, and their rank based on median PSI of exons containing 30 (or 20) such 2mers (or 3mers) in their sequence (all GTEx tissues).

Supplementary Figure 6



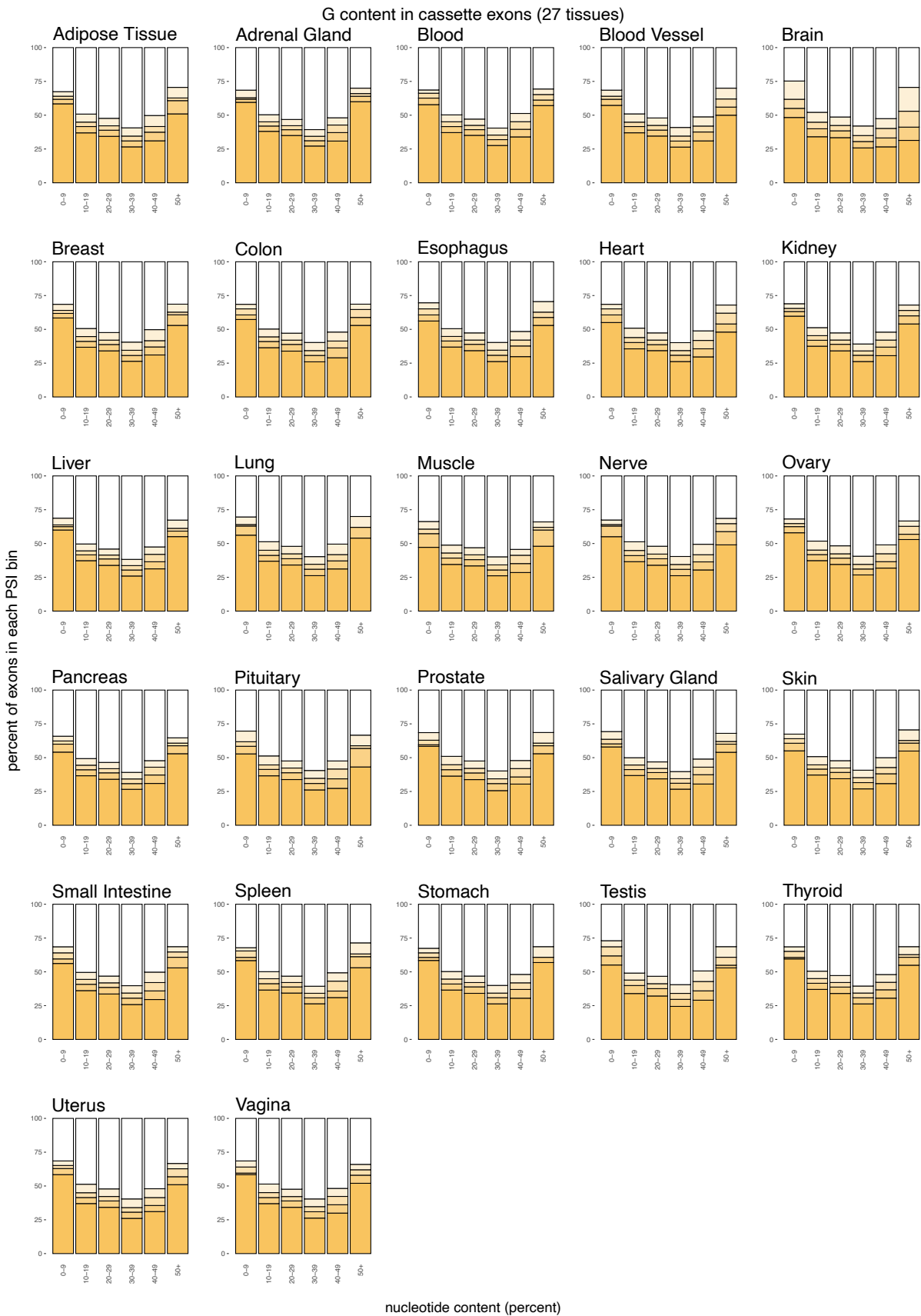
Distribution of PSI values in cassette exons (all GTEx tissues) relative to the percentage of each adenines present in the exon sequence.

Supplementary Figure 7



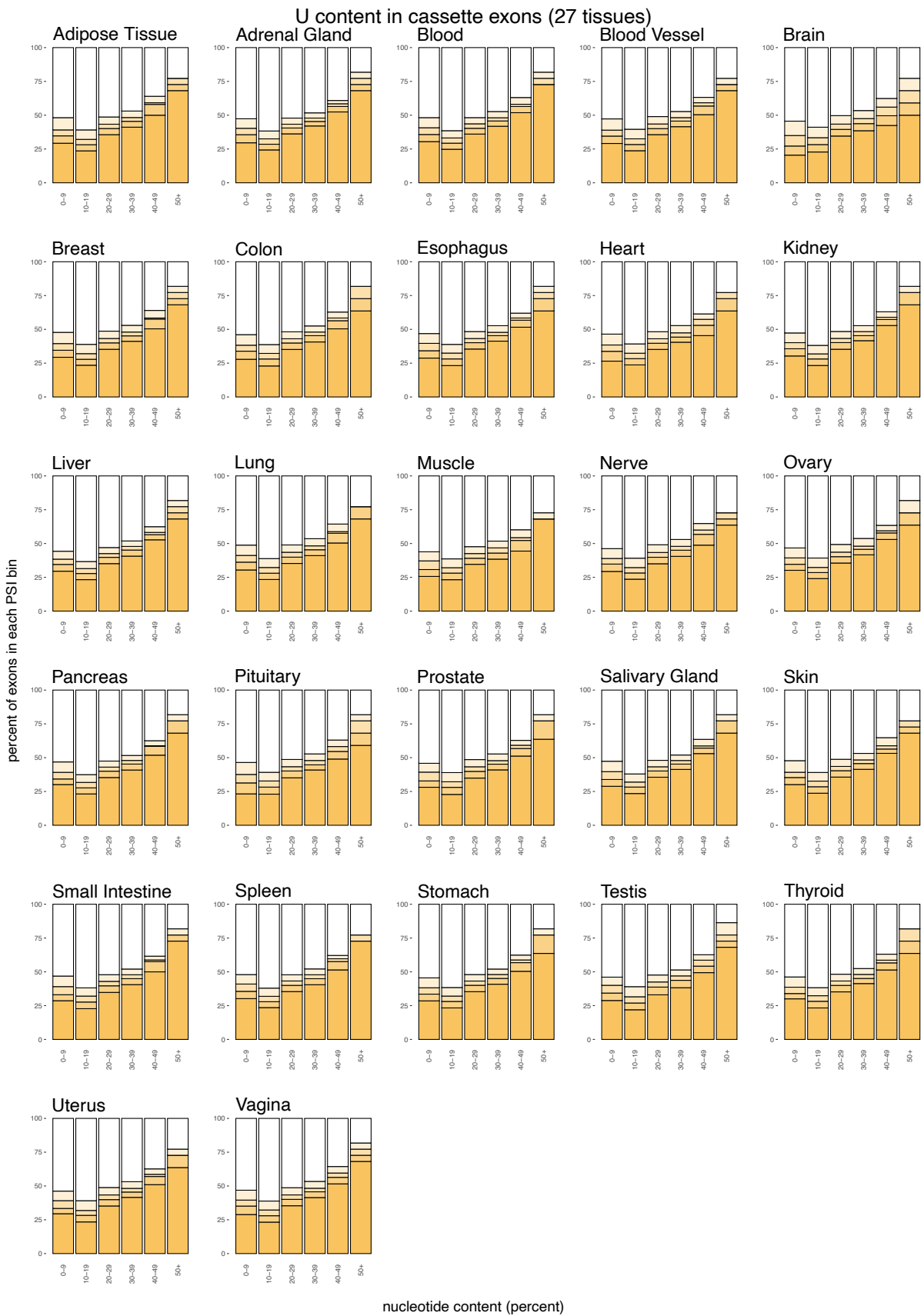
Distribution of PSI values in cassette exons (all GTEx tissues) relative to the percentage of each cytosines present in the exon sequence.

Supplementary Figure 8



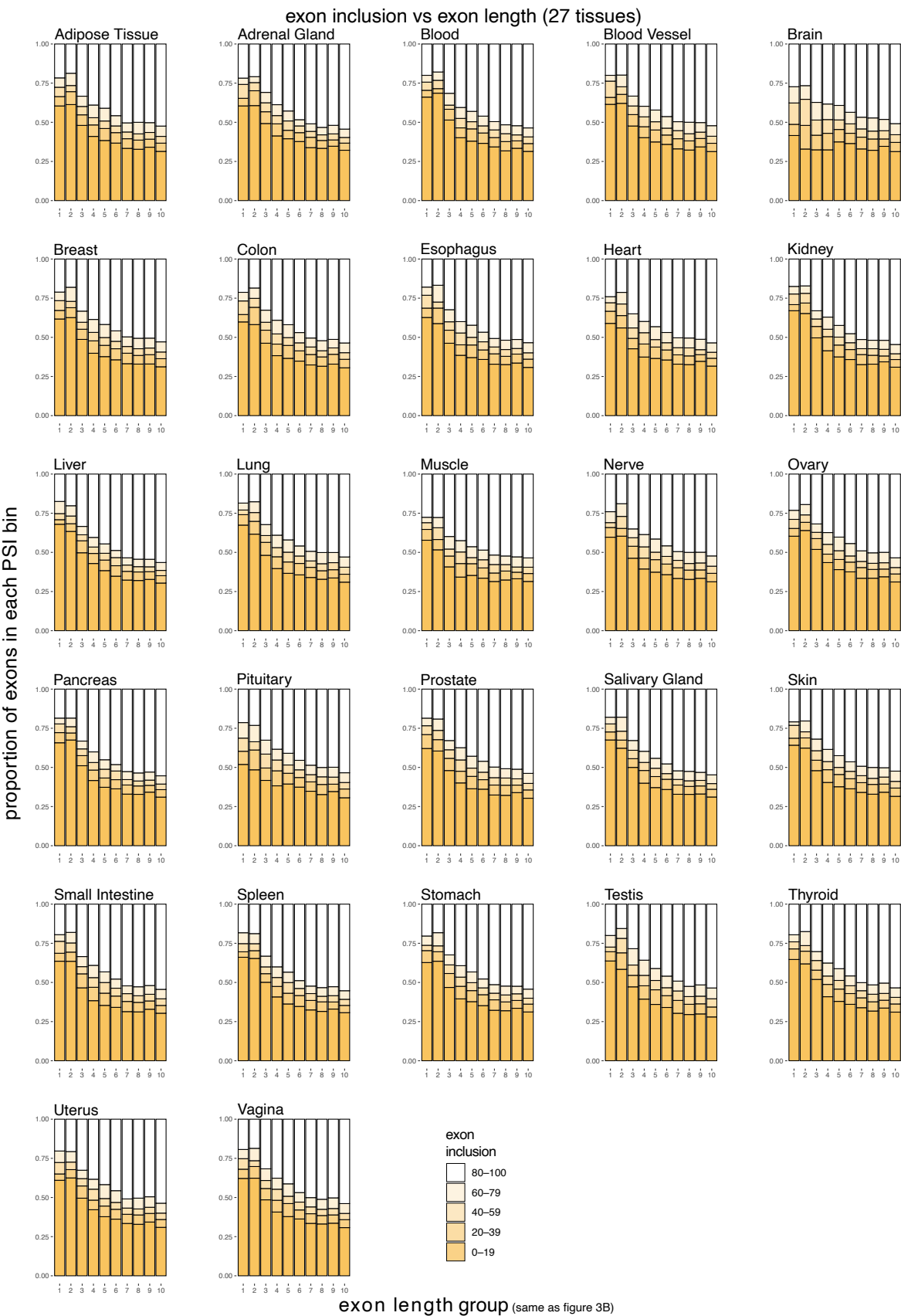
Distribution of PSI values in cassette exons (all GTEx tissues) relative to the percentage of each guanines present in the exon sequence.

Supplementary Figure 9



Distribution of PSI values in cassette exons (all GTEx tissues) relative to the percentage of each uracils present in the exon sequence.

Supplementary Figure 10



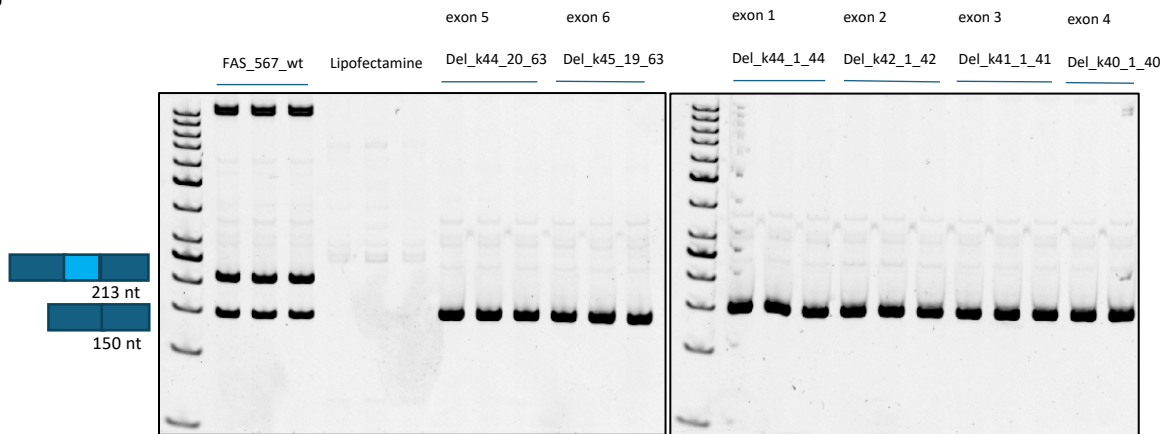
Distribution of cassette exon inclusion values versus exon length (all GTEx tissues). Exon length bins are the same as in figure 3B.

Supplementary Figure 11

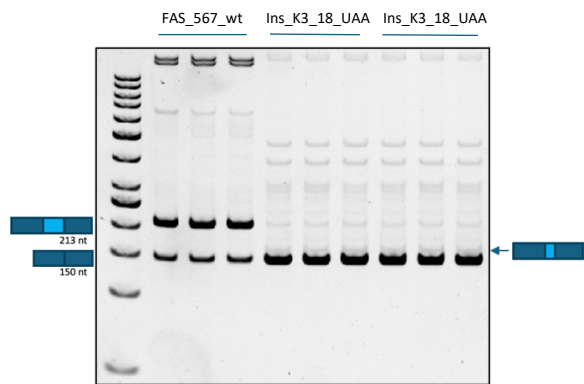
A

Full length FAS exon 6
GAUCCAGAUCAACUUGGGUGGCUUUGUCUUCUUCUUUUUGCCAAUCCACUAAUUGUUUGGG
Detected microexons
Exon 1 del (1-44) -----AUUCCACUAAUUGUUUGGG
Exon 2 del (1-42) -----CAAUCCACUAAUUGUUUGGG
Exon 3 del (1-41) -----CCAAUCCACUAAUUGUUUGGG
Exon 4 del (1-40) -----GCCAAUCCACUAAUUGUUUGGG
Exon 5 del (20-63) GAUCCAGAUCAACUUGGG-----
Exon 6 del (19-63) GAUCCAGAUCAACUUGG-----

B



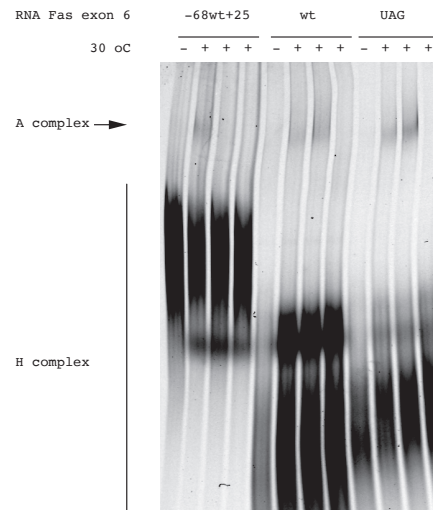
C



Ins_k3_18_UAA GAUCCAGAUCAACUUGG**UAA**GGUGGCUUUGUCUUCUUCUUUUUGCCAAUCCACUAAUUGUUUGGG

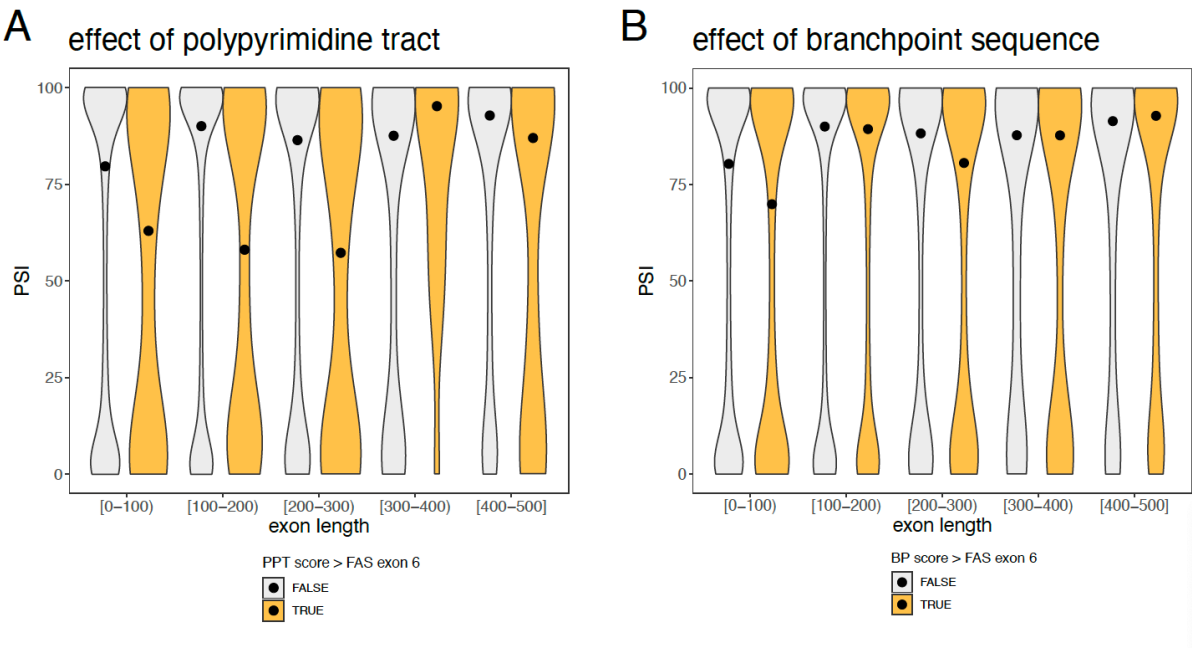
A. Sequences of microexons in our library with detectable levels of inclusion.
B. RT-PCR analysis of the inclusion of exons with a sequence corresponding to the detectable microexons.
C. RT-PCR analysis of the inclusion of FAS exon 6 with a UAA insertion after exon position 18. Results from two independently obtained clones are shown. Arrow indicates the position of microexon-inclusion product.

Supplementary Figure 12



Validation of spliceosome assembly assays and enhanced A (U2 snRNP-containing) complex formation on Fas exon 6 upon inclusion of a 3' splice site (UAG). Fluorescently-labeled RNAs corresponding to Fas exon 6 (wt), Fas exon 6 with a mutation that creates a 3' splice site after the internal polypyrimidine tract (UAG) (see Figure 3) or Fas exon 6 flanked by intronic sequences (-68wt+25, which includes the 3' 68 nucleotides of intron 5 and the 5' 25 nucleotides of intron 6) were incubated with HeLa nuclear extracts under in vitro splicing conditions at 30 °C (+) or on ice as a control (-) and the ribonucleoprotein complexes formed were fractionated by electrophoresis on a composite agarose-polyacrylamide gel. The electrophoretic positions of U2 snRNP-containing complexes (A complex) and hnRNP-containing complexes (H complex) are indicated. Controls of U1 / U2 snRNP inactivation by RNase H-mediated digestion of the 5' end of U1 snRNA or of the branch point recognition sequence of U2 snRNA are included. Inactivation of U2 snRNP (but not of U1 snRNP) reduces A complex formation on WT and UAG RNAs, while complexes formed on the -68wt+25 RNA are sensitive to inactivation of either U1 or U2 snRNPs because of exon definition-mediated effects (Izquierdo et al Mol Cell 19: 475-484, 2005). Note the significant increase in A/H complex ratio upon inclusion of a 3' splice site (UAG) in Fas exon 6 compared to the wild type sequence (WT).

Supplementary Figure 13



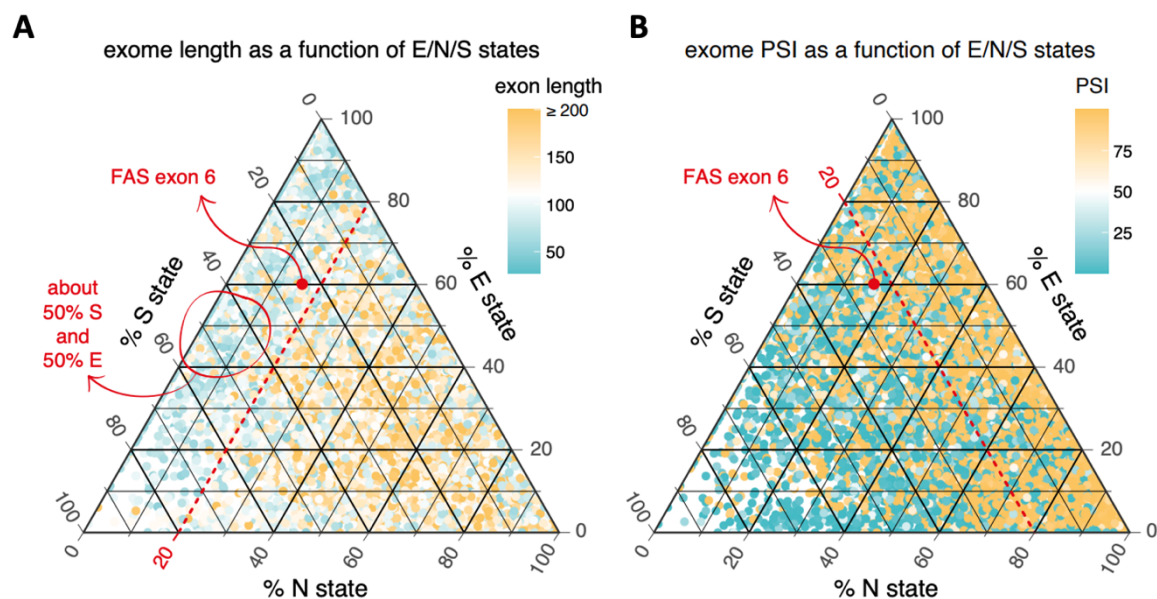
Inclusion of exons categorized into different length groups, divided by whether they have an exonic sequence more or less similar to a 3' splice site than FAS exon 6 (as in Figure 4C), split in the effect of the polypyrimidine tract (A) or branchpoint sequence (B) according to Ref. 48.

Supplementary Figure 14



Fraction of nucleotides in E/N/S states at exonic 5' or 3' terminal 20 4mers for internal exons of 100-150 nucleotides (A) or over 150 nucleotides (B). Note the higher proportion of E states at the 5' and 3' ends of the exon, likely due to the contribution of exonic nucleotides to splice site signals.

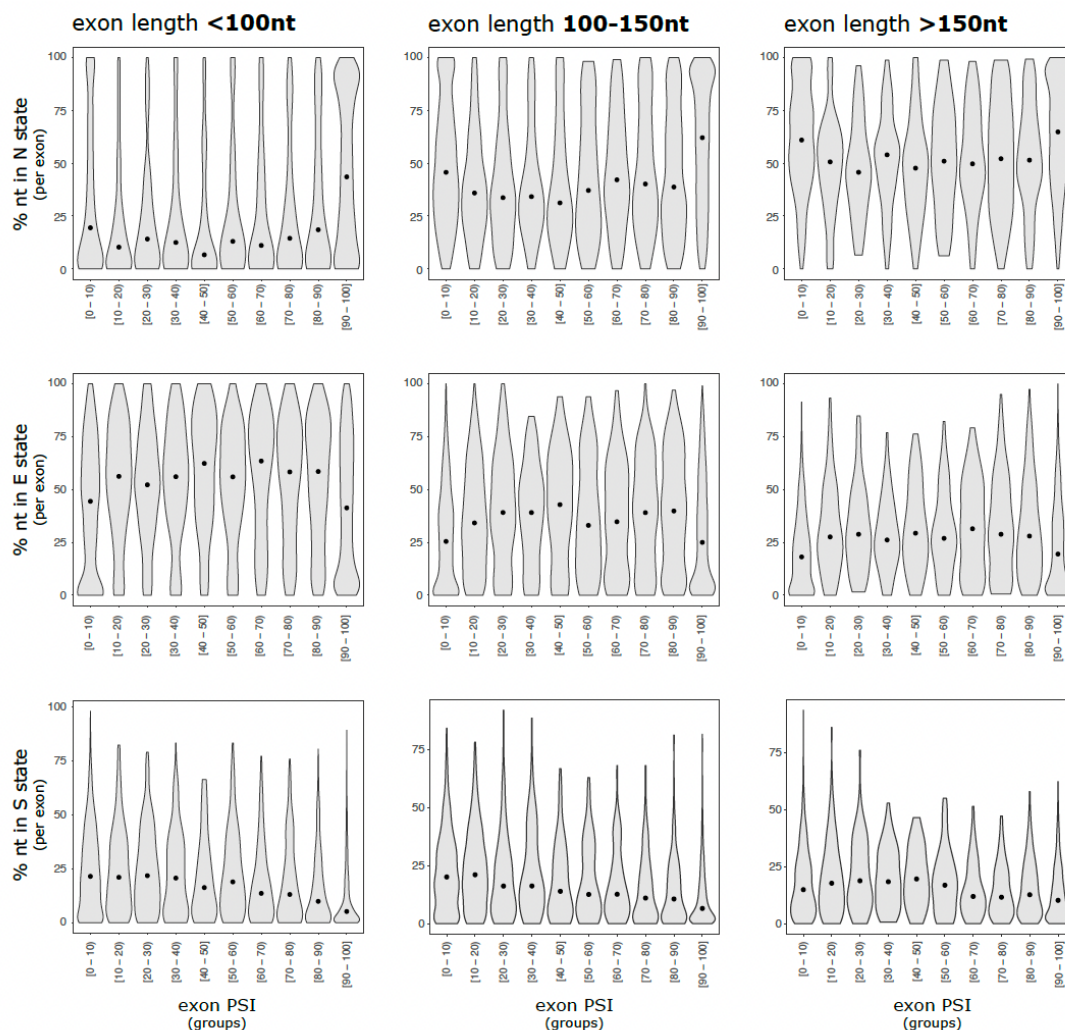
Supplementary Figure 15



A. Ternary plot illustrating the relative composition of E/N/S states along the sequences of all 18,551 exons in our dataset. The colour of each point corresponds to the exon's length. A local regression model (loess) fit over the raw data is shown in figure 6F to help to visualize trends in the data.

B. Ternary plot illustrating the relative composition of E/N/S states along the sequences of all 18,551 exons in our dataset. The colour of each point corresponds to the exon's PSI value. A local regression model (loess) fit over the raw data is shown in figure 6G to help to visualize trends in the data.

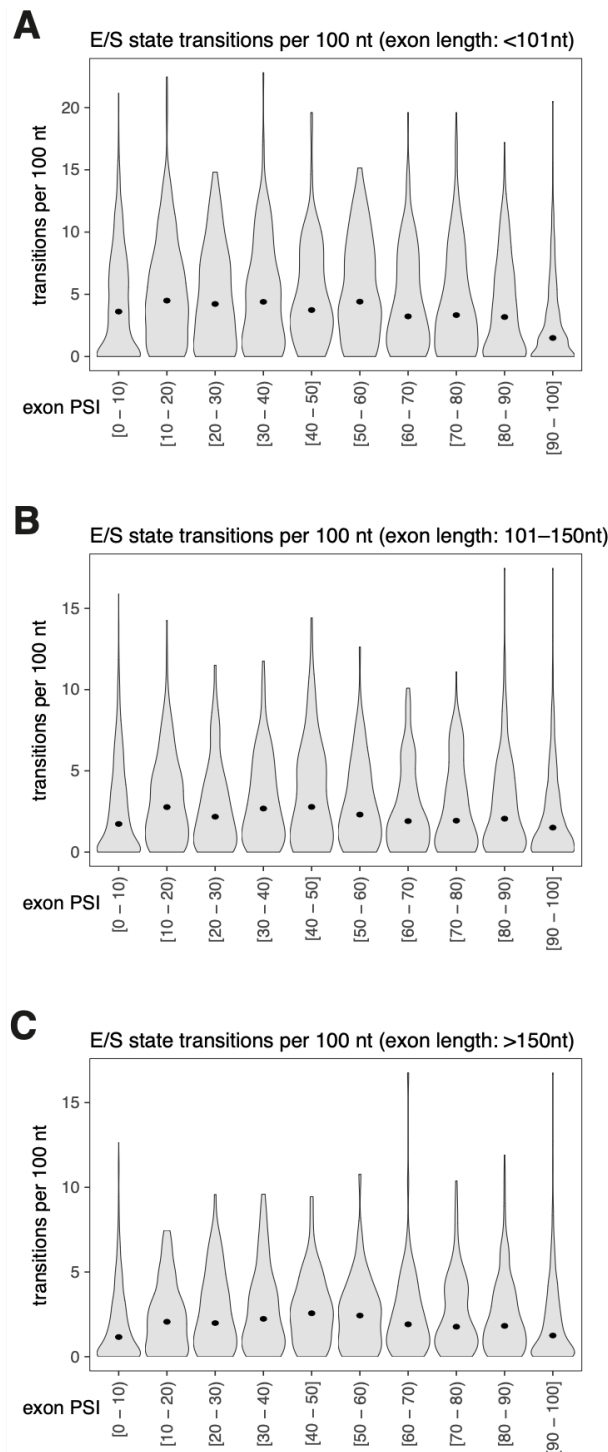
Supplementary Figure 16



Percentage of nucleotides in N, E or S states for exons of different lengths and inclusion levels. The distribution of the fraction of nucleotides in Neutral (N), Enhancer (E) or Silencer (S) states is shown for the indicated exon lengths and PSI values. Data source as for Figures 6 and S15. Main conclusions are:

- Highly included exons display lower density of regulatory elements, in all likelihood related to stronger flanking splice sites characteristic of highly included/constitutive exons. This is in line with multiple previous reports, including PMID: 33112234.
- The density of nucleotides within splicing enhancer elements generally decreases with exon length, suggesting that splicing enhancers are required to sustain the inclusion of alternative exons flanked by weaker splice sites, but the density of enhancer sequences does not determine the extent of exon inclusion.
- In contrast, the density of splicing silencers generally correlates inversely with the levels of exon inclusion, suggesting that splicing silencers, while generally less frequent than enhancers, are key determinants of exon inclusion.

Supplementary Figure 17

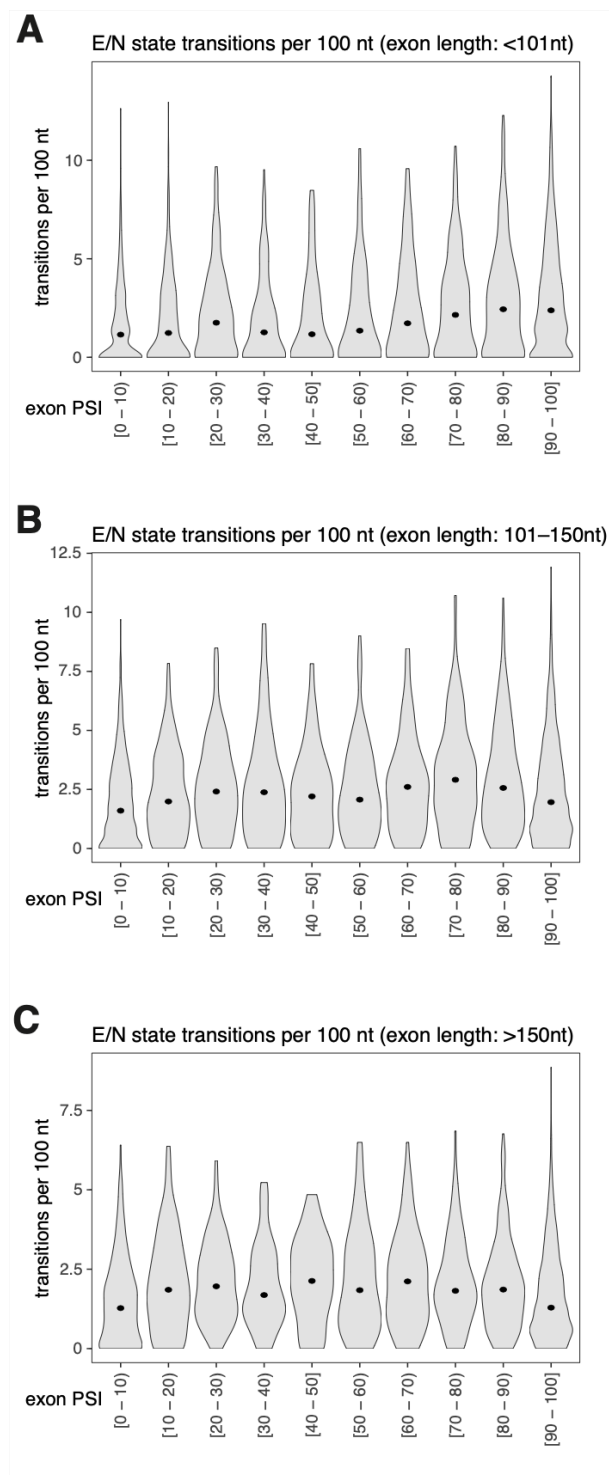


A. Distribution of the number of enhancer-silencer ("E"/"S") state transitions per 100 nucleotides in all exons shorter than 101 nt, divided into groups based on exon inclusion levels.

B. Distribution of the number of enhancer-silencer ("E"/"S") state transitions per 100 nucleotides in all exons with a length between 101 and 150 nt, divided into groups based on exon inclusion levels.

C. Distribution of the number of enhancer-silencer ("E"/"S") state transitions per 100 nucleotides in all exons longer than 150 nt, divided into groups based on exon inclusion levels.

Supplementary Figure 18

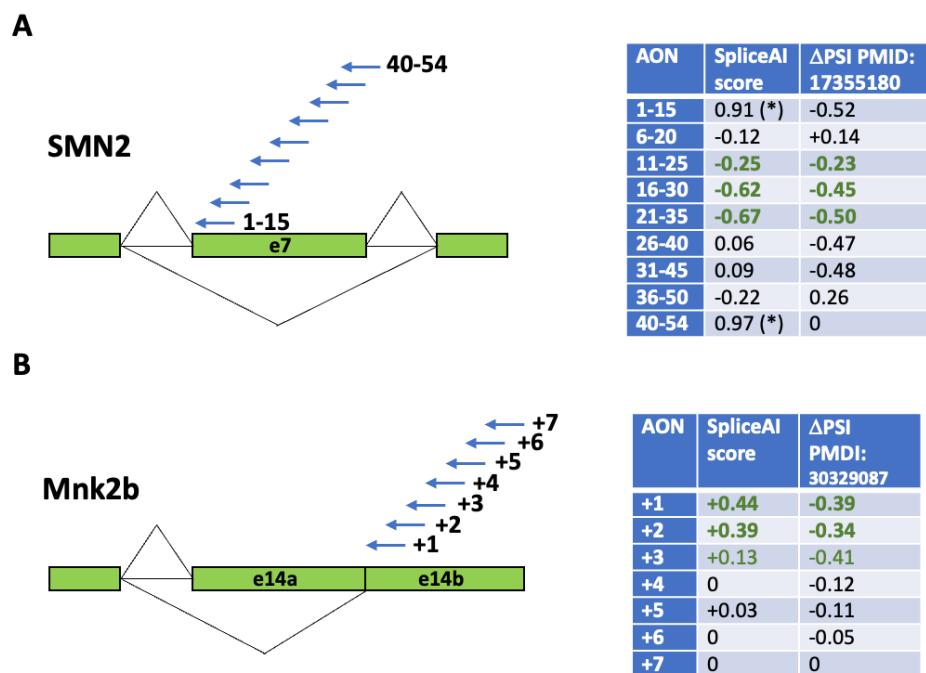


A. Distribution of the number of enhancer-neutral ("E"/"N") state transitions per 100 nucleotides in all exons shorter than 101 nt, divided into groups based on exon inclusion levels.

B. Distribution of the number of enhancer-neutral ("E"/"N") state transitions per 100 nucleotides in all exons with a length between 101 and 150 nt, divided into groups based on exon inclusion levels.

C. Distribution of the number of enhancer-neutral ("E"/"N") state transitions per 100 nucleotides in all exons longer than 150 nt, divided into groups based on exon inclusion levels.

Supplementary Figure 19



Benchmarking of previously published effects on alternative splicing regulation of overlapping antisense oligonucleotides (AON walks) and SpliceAI-based predictions of the effects of deletion of exonic sequences complementary to the corresponding AONs. For each series of overlapping AONs, schematically represented above the indicated genomic regions (left schemes), SpliceAI scores for the expected effects of deletion of the complementary exonic region are tabulated along with the experimentally determined values of PSI changes upon treatment with the corresponding AON. Green indicates correlation between the two classes of values. **A. SMN2 exon 7** (Hua et al PLoS Biol. 5, e73, 2007 PMID: 17355180). (*) indicates high values of SpliceAI scores likely related to effects of deletions on the sequences of the splice sites themselves -occurring upon deletions involving the first and last nucleotides of the exon-. SpliceAI detects the presence of a splicing enhancer(s) between nucleotides 11 and 35 but fails to detect other regulatory sequences located in the 3' half of the exon. **B. Mnk2b exon e14b** (Mogilevsky et al Nucleic Acids Res. 46, 11396-11404, 2018 PMID: 30329087). SpliceAI predicts the presence of regulatory sequences in only the 5' 1/3 of the exon, although it predicts opposite effects for the deletion of these sequences and the effects of AONs complementary to them.

Supplementary Table 1

Indel library amplification primers

FAS_i5_GC_F 5'-tgtccaatgttccaacctacag-3'

FAS_i6_GC_R 5'-ctacttcccaagttatttcaatctg-3'

pCMV FAS 567 Vector Backbone amplification primers

PT1 5'-GTCGACGACACTTGCTCAAC-3'

PT2 5'-AAGCTTGCATCGAATCAGTAG-3'

Primers used for amplifying technical replicates (Input library)

Forward primer	Barcode	Forward primer (FAS_i5_TR_F)	Reverse primer	Barcode	Reverse Primer (PT2)	Amplicon Size (nt)
FAS_TR1	GCCGAATT	aaaatgtccaatgttccaacc	PT2_TR1	GCCGAATT	aagcttgcacgaatcagtag	231+16
FAS_TR2	CGGCAATT	aaaatgtccaatgttccaacc	PT2_TR2	CGGCAATT	aagcttgcacgaatcagtag	231+16
FAS_TR3	GAACGTTC	aaaatgtccaatgttccaacc	PT2_TR3	GAACGTTC	aagcttgcacgaatcagtag	231+16

Primers used for amplifying biological replicates (Output library)

Forward primer	Barcode	sequence	Reverse primer	Barcode	Reverse Primer (PT2)	Inclusion amplicon Size (nt)
FAS_BR1	CAACCATG	CAGCAACACCAAGTGCAAAG	PT2_BR1	CAACCATG	aagcttgcacgaatcagtag	195 +16
FAS_BR2	CGTACCTT	CAGCAACACCAAGTGCAAAG	PT2_BR2	CGTACCTT	aagcttgcacgaatcagtag	195 +16
FAS_BR3	ACACTGTG	CAGCAACACCAAGTGCAAAG	PT2_BR3	ACACTGTG	aagcttgcacgaatcagtag	195 +16
FAS_BR4	GACACACT	CAGCAACACCAAGTGCAAAG	PT2_BR4	GACACACT	aagcttgcacgaatcagtag	195 +16
FAS_BR5	ACGACTTG	CAGCAACACCAAGTGCAAAG	PT2_BR5	ACGACTTG	aagcttgcacgaatcagtag	195 +16
FAS_BR6	ATCGATCG	CAGCAACACCAAGTGCAAAG	PT2_BR6	ATCGATCG	aagcttgcacgaatcagtag	195 +16
FAS_BR7	AACCAACC	CAGCAACACCAAGTGCAAAG	PT2_BR7	AACCAACC	aagcttgcacgaatcagtag	195 +16
FAS_BR8	CCATCGTT	CAGCAACACCAAGTGCAAAG	PT2_BR8	CCATCGTT	aagcttgcacgaatcagtag	195 +16
FAS_BR9	GCGCATAT	CAGCAACACCAAGTGCAAAG	PT2_BR9	GCGCATAT	aagcttgcacgaatcagtag	195 +16

Supplementary Table 2

Experimental validation of individual clones and PSI determination by Hek293 transfection

SampleID	sequenceID	sequence	PSI	SD
Fas wt		GATCCAGATCTAACTGGGGTGGCTTTGTCTTCTTTTGCCAATCCACTAATTGTTGGG	49.9	1.3
Indel1	EXK513	ATCCAGATCTAACTGGGGTGGCTTTGTCTTCTTTTGCCAATCCACTAATTGTTGGG	53.6	0.5
Indel3	EXK515	GATCCAGATCTAACTGGGGTGGCTTTGTCTTCTTTTGCCAATCCACTAATTGAGATTGGG	81.2	1.2
Indel4	EXK516	GATCCAGATCTAACTGGGGCCTGGCTTTGTCTTCTTTTGCCAATCCACTAATTGTTGGG	97.4	1.3
Indel5	EXK517	GATCCAGATCTAACTGGGGTGGCTTTGTCTTCTTTTGCCAATCCCTGACTAATTGTTGGG	56.2	1.4
Indel6	EXK518	TTTGCCAATCCACTAATTGTTGGG	11.7	4.5
Indel8	EXK520	GATCCAGATCTAACTGGGGTGGCTGTTGGG	45.1	0.4
Indel9	EXK521	GATCCAGATCTAACTGGGGTGGCTTTGTCTTCTTTTGCCAATCCAATTTCTAATTGTTGGG	13.2	1.5
Indel10	EXK522	GATCCAGATCTAACTGGGGTGGCTTTGTCTTCTTTTGCCAATCCACTAATTGTTGTACGG	82.7	1.3
Indel13	EXK525	GATCCAGATCTAACTGGGGTGGCTTTGTCTTCTATCCTTTTGCCAATCCACTAATTGTTGGG	59.1	0.8
Indel16	EXK528	GATCCAGATCTAACTGGGGTGGCTTTGTCTTCTTTTGCCAATCCACTGTTGGG	63.8	1.9
Indel18	EXK530	GATCCAGATCTAACTGGGGTGGCCACTAATTGTTGGG	34.7	2.1
Indel19	EXK531	GATCCAGATCTAACAATCCACTAATTGTTGGG	13.2	2.3
Indel20	EXK532	GATCCAGATCTAACTGGGGTGGCTTTGTCTTCTCTCTTTTGCCAATCCACTAATTGTTGGG	57.4	3.3
Indel22	EXK534	GATCCAGTTTGCCAATCCACTAATTGTTGGG	30.1	0.4
Indel23	EXK535	GATCCAGATCTAACGGATTGGGGTGGCTTTGTCTTCTTTTGCCAATCCACTAATTGTTGGG	77.8	0.5
Indel24	EXK536	GATCCAGATCTAACTGGGGTGGCTTTGTCTTCTTTTGCCAATCCACTAATTGCGTTTGGG	56.6	0.6

Supplementary Table 2. Percent Spliced In (% of exon inclusion) measurements of individual minigene transfections, determined by RT-PCR of RNA isolated from Hek293 cells transfected with the indicated exon sequence variants. SD indicates standard deviation of at least 3 independent measurements. These data correspond to the graph shown in Fig. 1C and were used to convert enrichment scores (es) into PSI values for all the variants in the library.

Supplementary Table 3

Sequences of antisense oligonucleotides used in the AON Walk

RNA oligo 2' O-Me PS	Sequence DNA FAS exon 6	Size	Position	Sequence RNA 2' O-MePS
FAON_2	GATCCAGATCTAACTGGGGT	21	0_21	ACCCAAGUUAGAUCUGGAUC
FAON_3	AGATCTAACTGGGGTGGCTT	21	6_26	AAGCCACCCAAGUUAGAUCU
FAON_4	TAACTGGGGTGGCTTTGTCT	21	11_31	AGACAAAGCCACCCAAGUUA
FAON_5	TGGGGTGGCTTTGTCTTCTC	21	16_36	GAAGAAGACAAAGCCACCCA
FAON_6	TGGCTTGTCTTCTCTTTTG	21	21_41	CAAAAGAAGAAGACAAAGCCA
FAON_7	TTGTCTTCTTTTGCCAAT	21	26_46	AUUGGCAAAAGAAGAAGACAA
FAON_8	TTCTTCTTTTGCCAATTCCAC	21	31_51	GUGGAUUUGGCAAAAGAAGAA
FAON_9	CTTTTGCCAATTCCACTAATT	21	36_56	AAUUAGUGGAAUUGGCAAAAG
FAON_10	GCCAATCCACTAATTGTTTG	21	41_61	CAACAAUUAGUGGAAUUGGC

Supplementary Table 3. Antisense oligonucleotides (21 nucleotides in length, harbouring 2'-O Me phosphorothioate modifications) complementary to 21 nucleotide regions of Fas exon 6 from nucleotide 1 to 61, thus walking the exon with a resolution of 5 nucleotides. The 5'-3' sequence of the exon covered, the position and the 5'-3' sequence of the antisense oligonucleotide are indicated.

