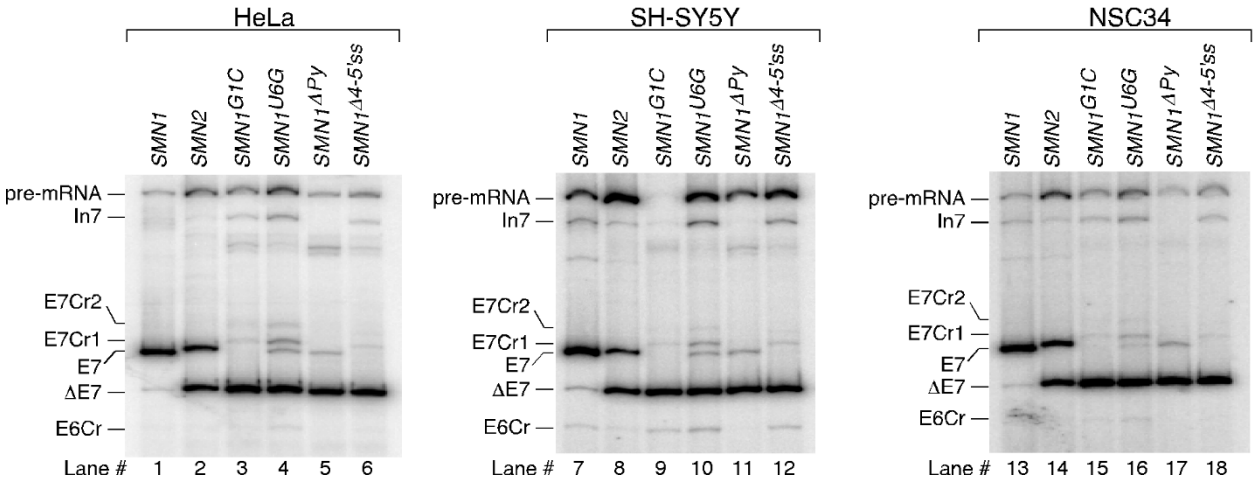
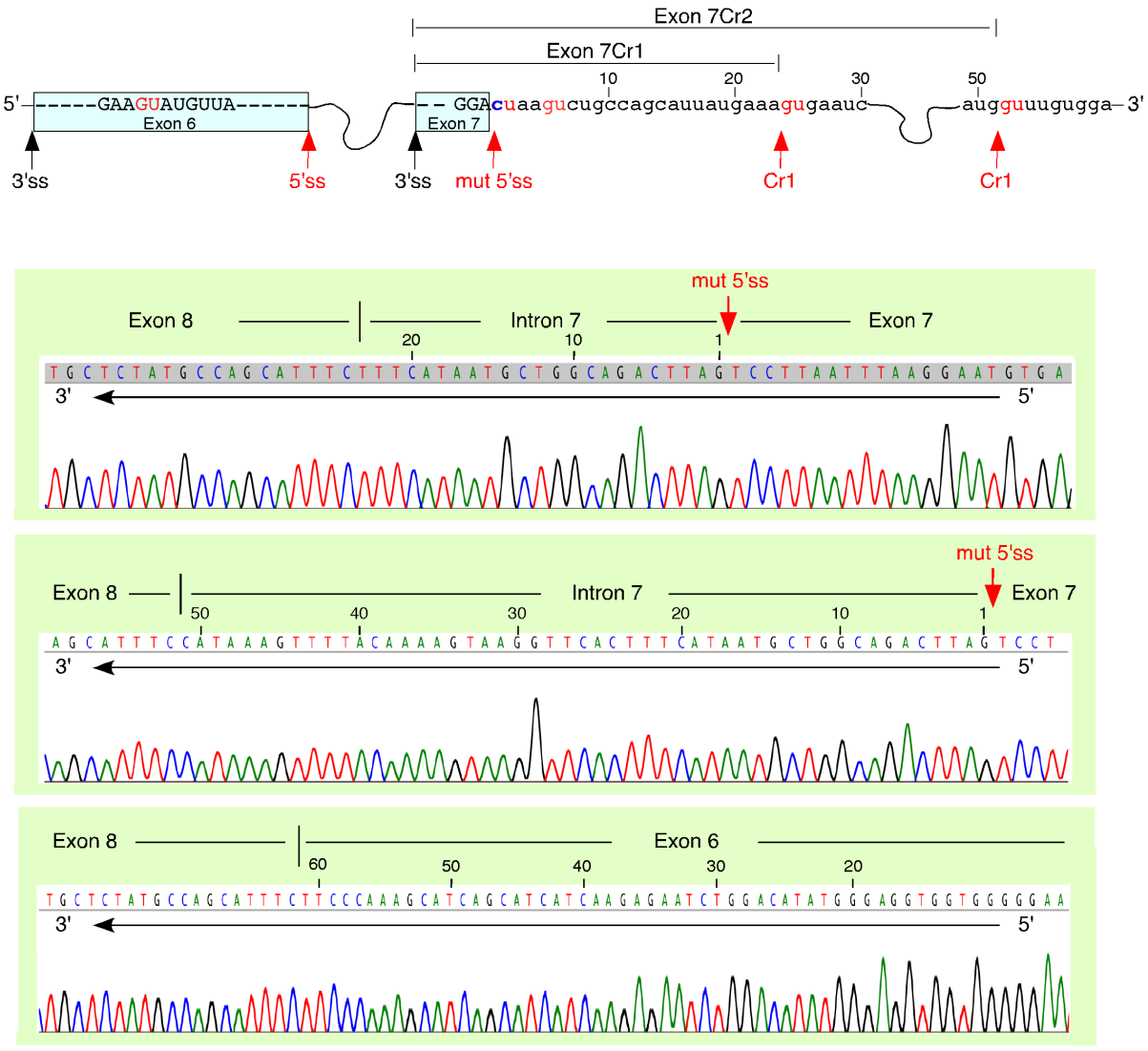
Supplementary Figures

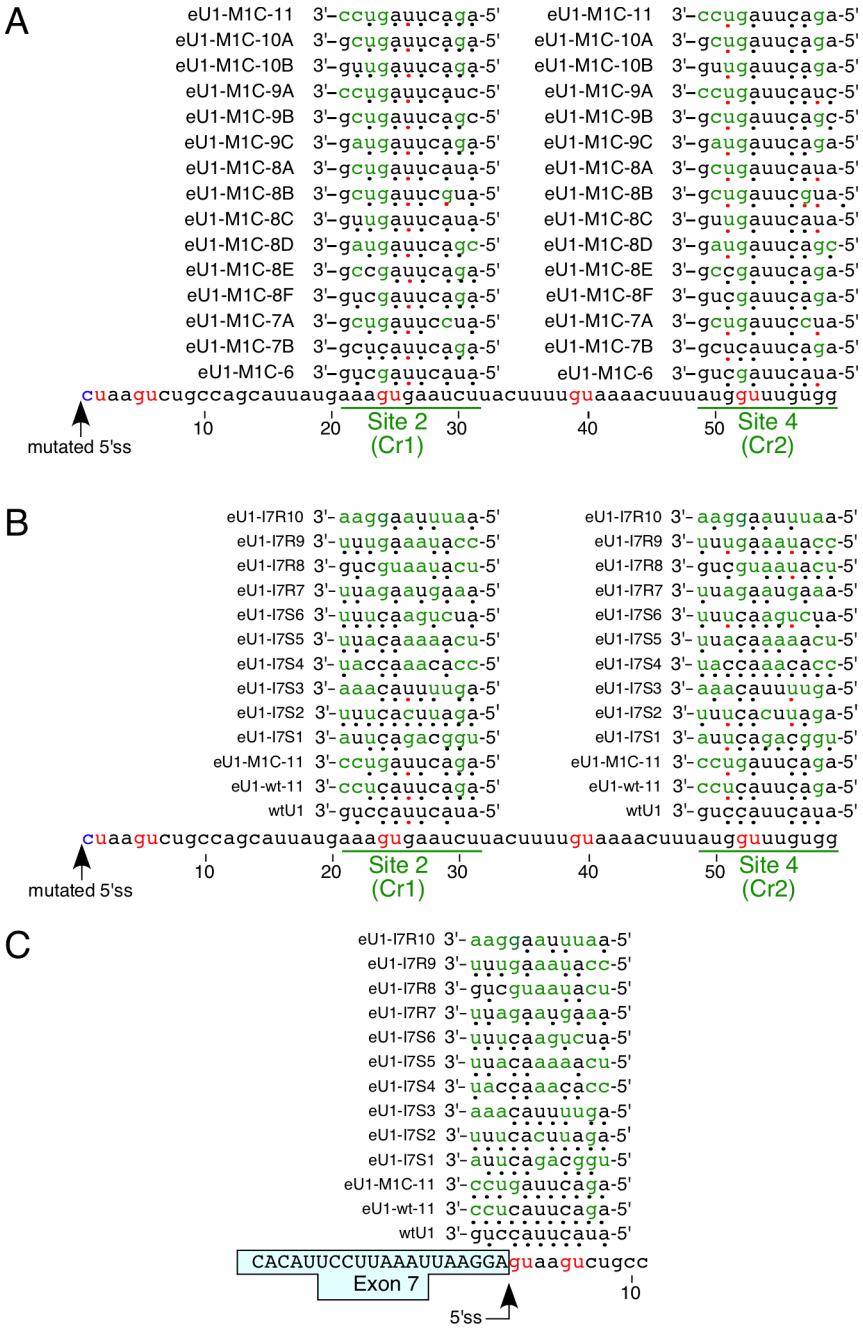
Supplementary Figure S1



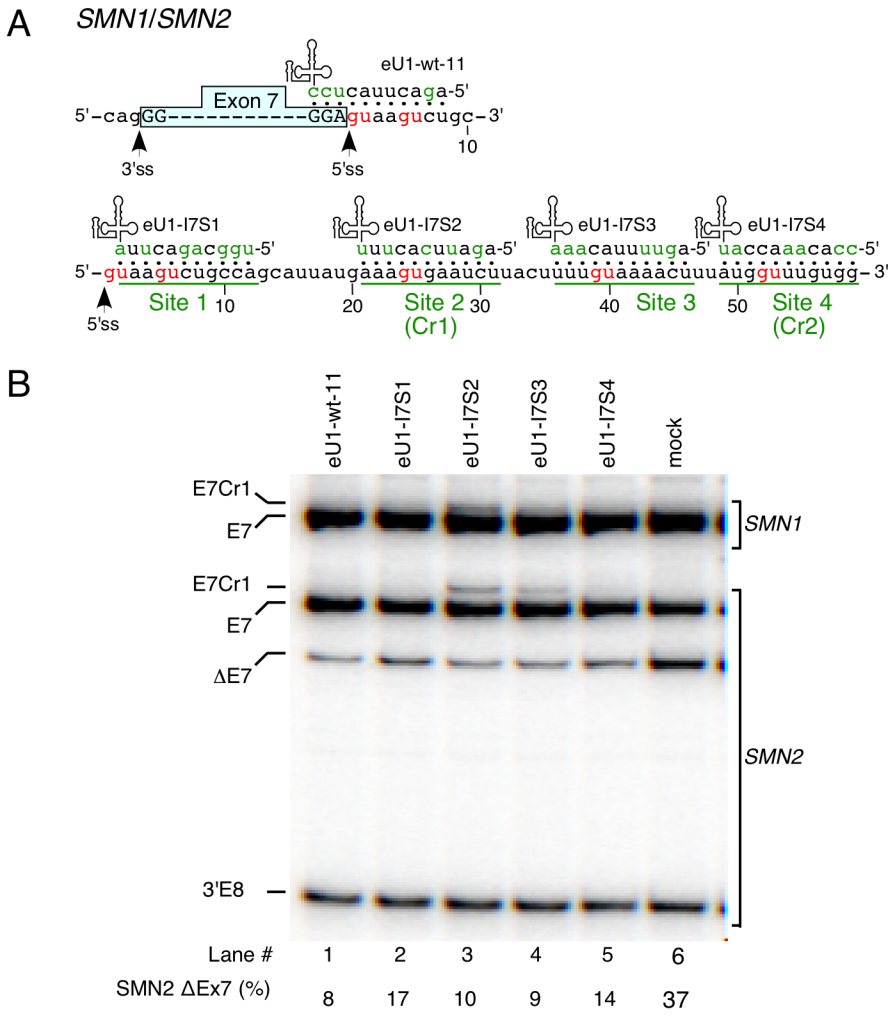
Supplementary Figure S2



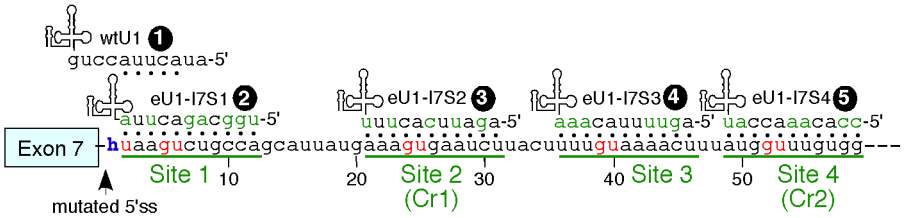
Supplementary Figure S3

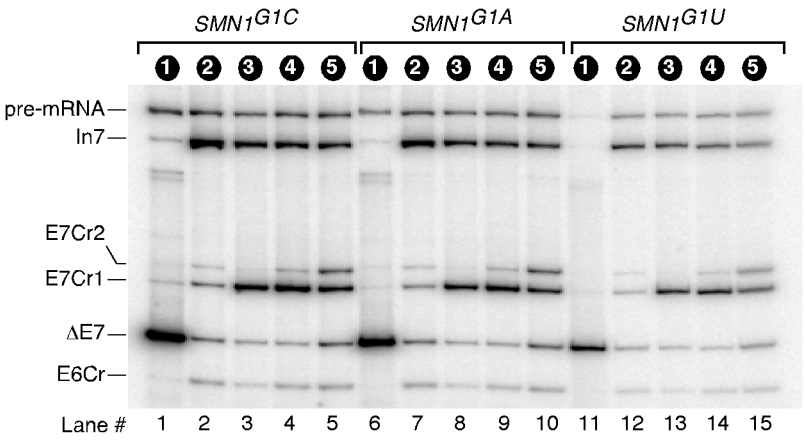


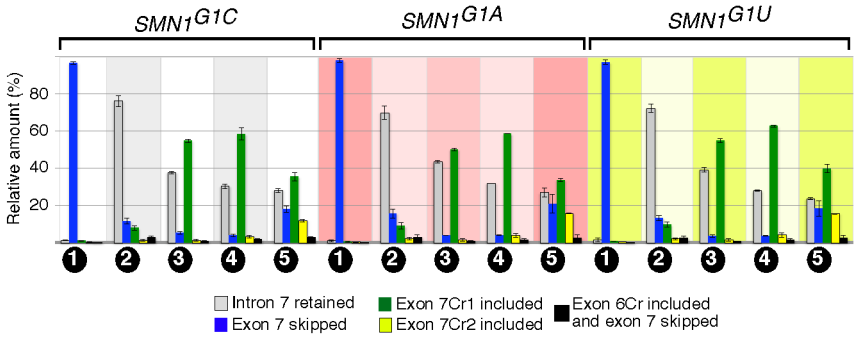
Supplementary Figure S4



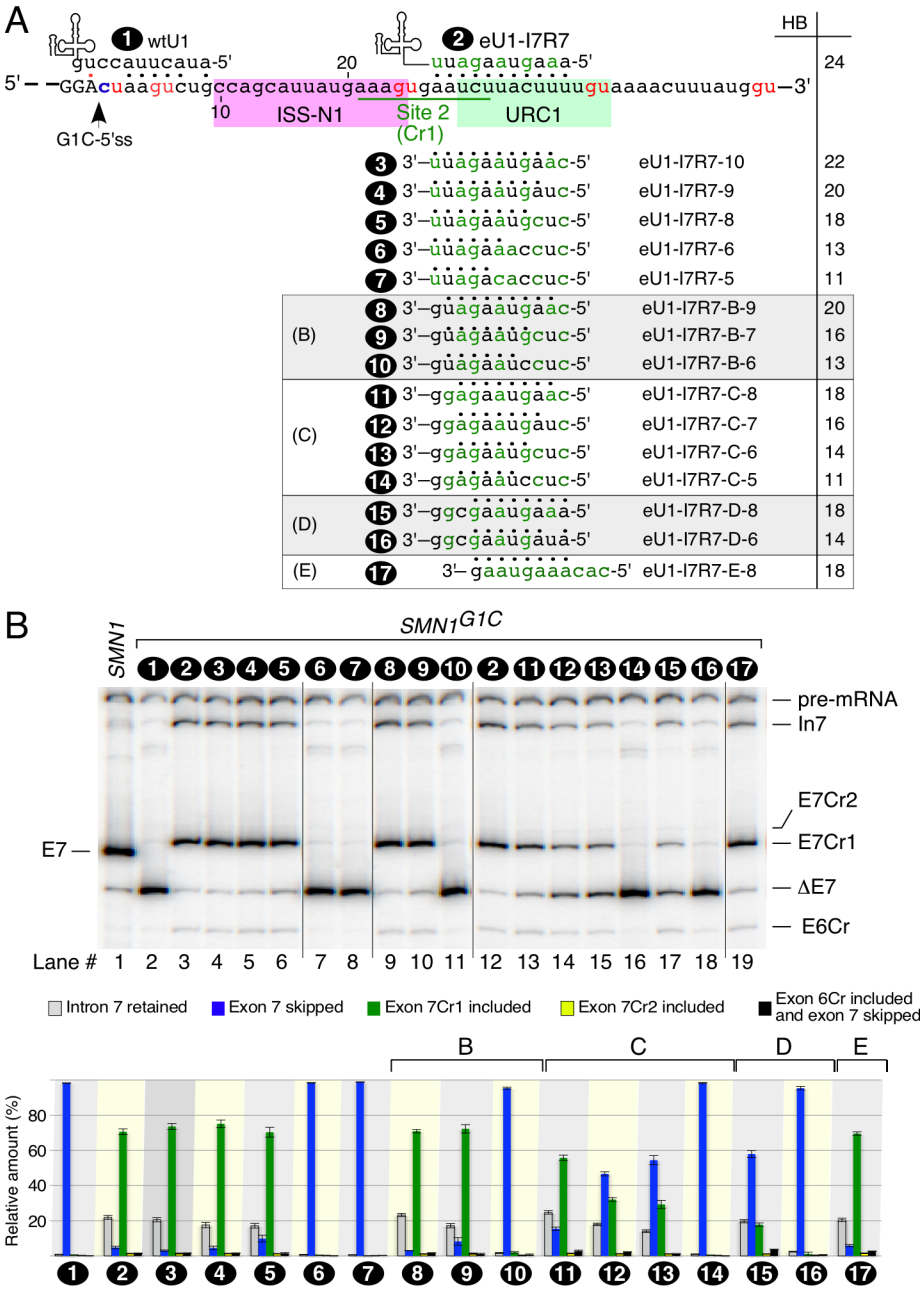
Supplementary Figure S5



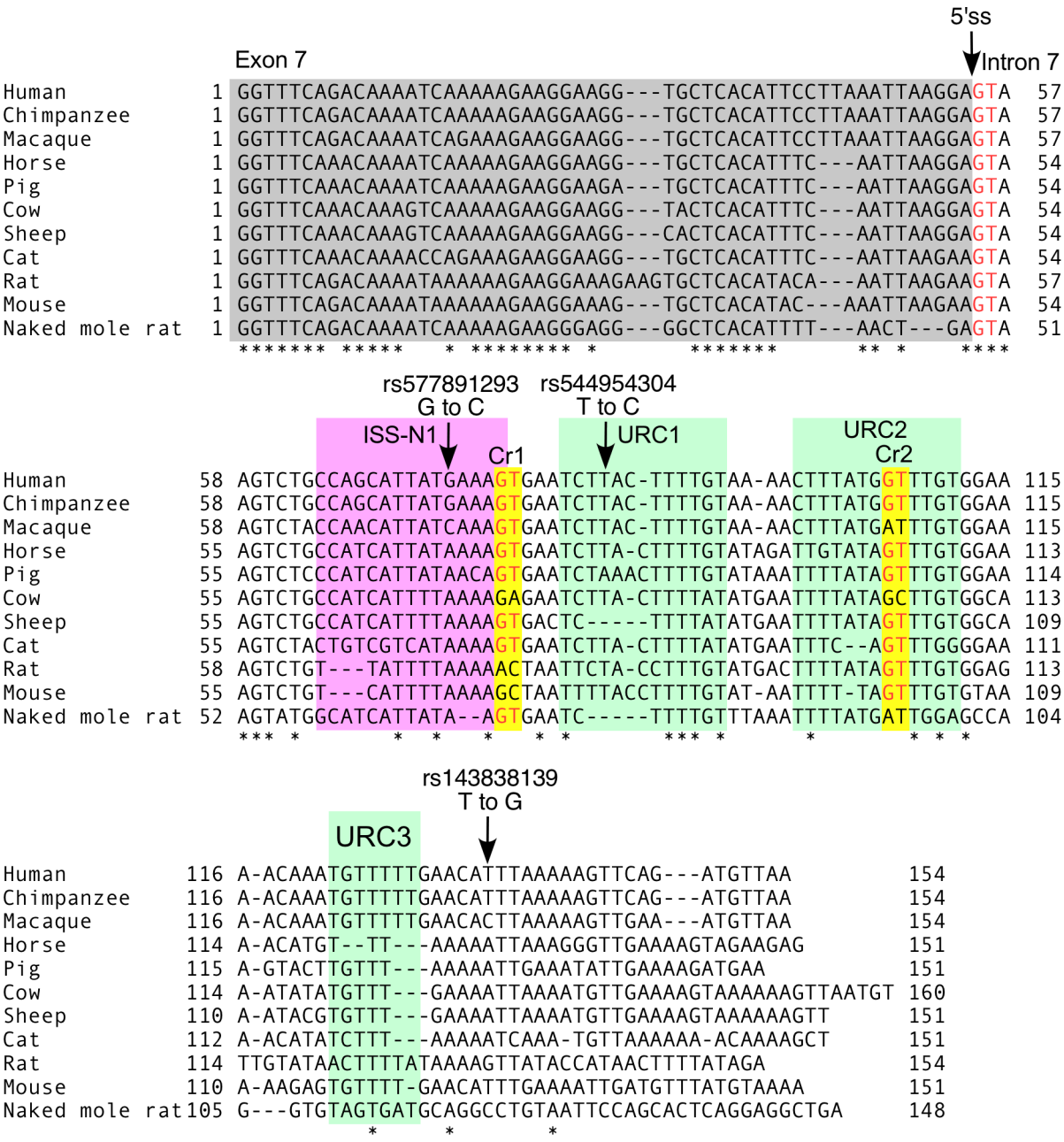




Supplementary Figure S6



Supplementary Figure S7

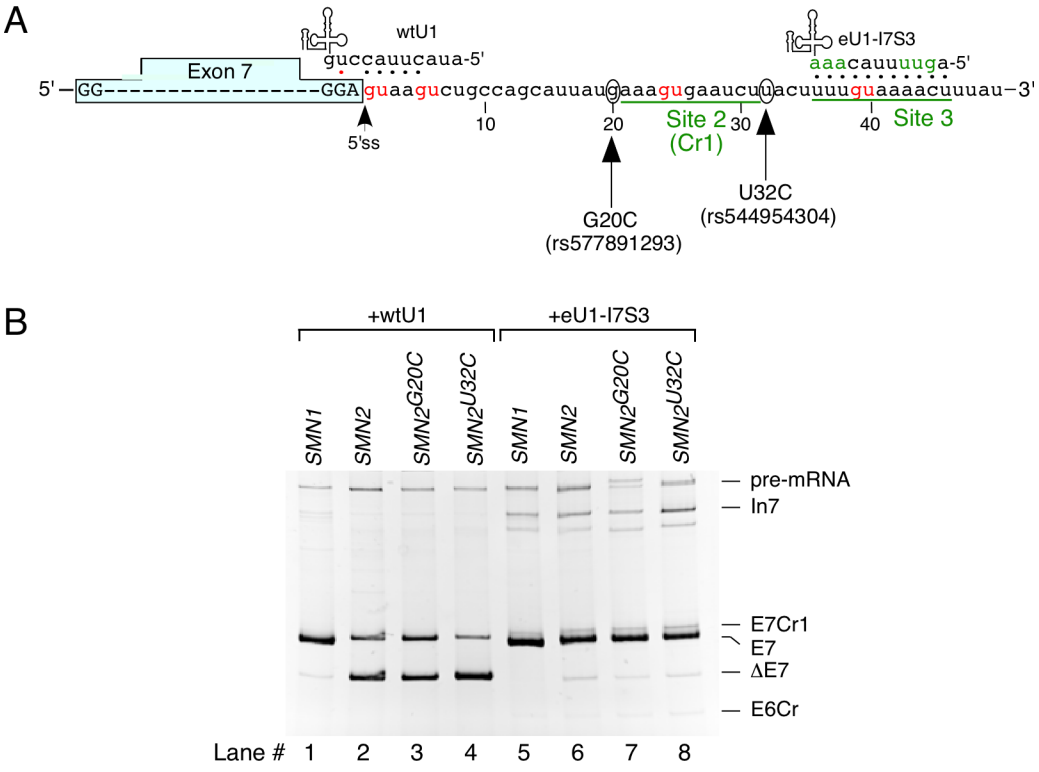



```
Human                1 GGTTTCAGACAAAATCAAAAAGAAGGAAGG---TGCTCACATTCCTTAAATTAAGGAGTA  57
Chimpanzee           1 GGTTTCAGACAAAATCAAAAAGAAGGAAGG---TGCTCACATTCCTTAAATTAAGGAGTA  57
Macaque              1 GGTTTCAGACAAAATCAGAAAGAAGGAAGG---TGCTCACATTCCTTAAATTAAGGAGTA  57
Horse                1 GGTTTCAAACAAAATCAAAAAGAAGGAAGG---TGCTCACATTTC---AATTAAGGAGTA  54
Pig                  1 GGTTTCAAACAAAATCAAAAAGAAGGAAGA---TGCTCACATTTC---AATTAAGGAGTA  54
Cow                  1 GGTTTCAAACAAAGTCAAAAAGAAGGAAGG---TACTCACATTTC---AATTAAGGAGTA  54
Sheep                1 GGTTTCAAACAAAGTCAAAAAGAAGGAAGG---CACTCACATTTC---AATTAAGGAGTA  54
Cat                  1 GGTTTCAAACAAAACCAGAAAGAAGGAAGG---TGCTCACATTTC---AATTAAGAAGTA  54
Rat                  1 GGTTTCAGACAAAATAAAAAAGAAGGAAAGAAGTGCTCACATACA---AATTAAGAAGTA  57
Mouse                1 GGTTTCAGACAAAATAAAAAAGAAGGAAAG---TGCTCACATAC---AAATTAAGAAGTA  54
Naked mole rat       1 GGTTTCAGACAAAATCAAAAAGAAGGGAGG---GGCTCACATTTT---AACT---GAGTA  51
                         ******* *****   * ******** *       *******      ** *    ****

Human               58 AGTCTGCCAGCATTATGAAAGTGAATCTTAC-TTTTGTAA-AACTTTATGGTTTGTGGAA 115
Chimpanzee          58 AGTCTGCCAGCATTATGAAAGTGAATCTTAC-TTTTGTAA-AACTTTATGGTTTGTGGAA 115
Macaque             58 AGTCTACCAACATTATCAAAGTGAATCTTAC-TTTTGTAA-AACTTTATGATTTGTGGAA 115
Horse               55 AGTCTGCCATCATTATAAAAGTGAATCTTA-CTTTTGTATAGATTGTATAGTTTGTGGAA 113
Pig                 55 AGTCTCCCATCATTATAACAGTGAATCTAAACTTTTGTATAAATTTTATAGTTTGTGGAA 114
Cow                 55 AGTCTGCCATCATTTTAAAAGAGAATCTTA-CTTTTATATGAATTTTATAGCTTGTGGCA 113
Sheep               55 AGTCTGCCATCATTTTAAAAGTGACTC-----TTTTATATGAATTTTATAGTTTGTGGCA 109
Cat                 55 AGTCTACTGTCGTCATAAAAGTGAATCTTA-CTTTTATATGAATTTC--AGTTTGGGGAA 111
Rat                 58 AGTCTGT---TATTTTAAAAACTAATTCTA-CCTTTGTATGACTTTTATAGTTTGTGGAG 113
Mouse               55 AGTCTGT---CATTTTAAAAGCTAATTTTACCTTTTGTAT-AATTTT-TAGTTTGTGTAA 109
Naked mole rat      52 AGTATGGCATCATTATA--AGTGAATC-----TTTTGTTTAAATTTTATGATTGGAGCCA 104
                         *** *       *  *   *   * *       *** *      *       * * *

Human              116 A-ACAAATGTTTTTGAACATTTAAAAAGTTCAG---ATGTTAA          154
Chimpanzee         116 A-ACAAATGTTTTTGAACATTTAAAAAGTTCAG---ATGTTAA          154
Macaque            116 A-ACAAATGTTTTTGAACACTTAAAAAGTTGAA---ATGTTAA          154
Horse              114 A-ACATGT--TT---AAAAATTAAAGGGTTGAAAAGTAGAAGAG         151
Pig                115 A-GTACTTGTTT---AAAAATTGAAATATTGAAAAGATGAA            151
Cow                114 A-ATATATGTTT---GAAAATTAAAATGTTGAAAAGTAAAAAAGTTAATGT  160
Sheep              110 A-ATACGTGTTT---GAAAATTAAAATGTTGAAAAGTAAAAAAGTT       151
Cat                112 A-ACATATCTTT---AAAAATCAAA-TGTTAAAAAA-ACAAAAGCT       151
Rat                114 TTGTATAACTTTTATAAAAGTTATACCATAACTTTTATAGA             154
Mouse              110 A-AAGAGTGTTTT-GAACATTTGAAAATTGATGTTTATGTAAAA          151
Naked mole rat     105 G---GTGTAGTGATGCAGGCCTGTAATTCCAGCACTCAGGAGGCTGA      148
                           *     *       *
```

Supplementary Figure S8



Supplementary Figure S9

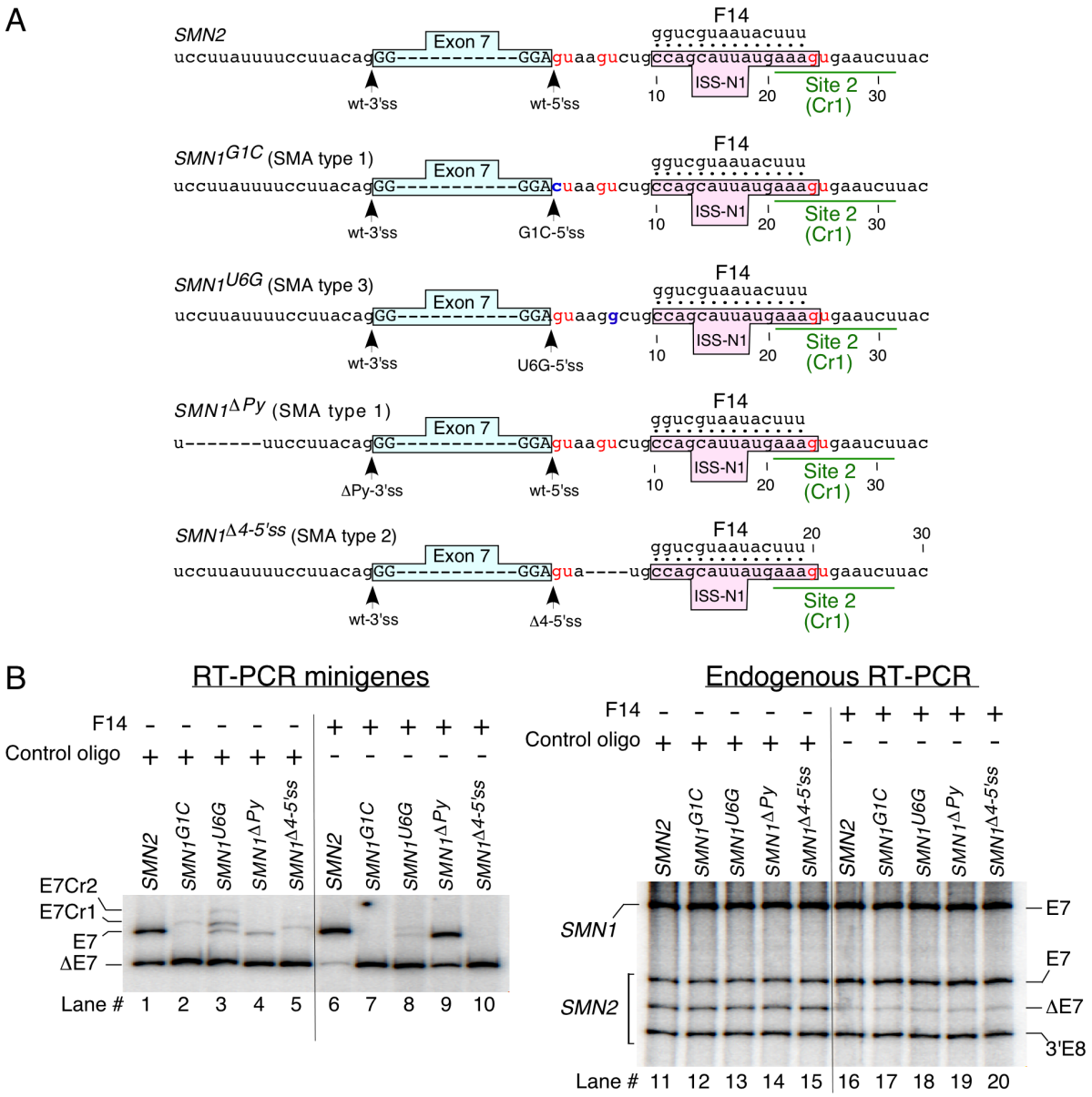


Figure Legends

Supplementary Figure S1. In vivo splicing pattern of the *SMN* minigenes in different cell lines. HeLa, SH-SY5Y and NSC-34 cells were transfected with 0.05 μg of a given minigene. Samples were collected 24 hour later and the splicing products were amplified by RT-PCR as in (52). The identity of the minigenes used for transfections is indicated at the top of the gels. Splicing products are indicated at the left of the gels. Abbreviations are described in Supplementary Table S1.

Supplementary Figure S2. Portions of cloned DNA sequences confirming usage of the cryptic splice site 1, cryptic site 2 and cryptic site within exon 6 during exon 7 splicing of the $SMN1^{G1C}$ minigene. Sequence context of these cryptic sites is indicated in the top panel, where exonic sequences are shown in capital letters, and intronic in lower case letters. In addition, exons are represented by colored boxes. Numbering starts from the first position of intron 7. The GU dinucleotides are indicated in red. The 5′ ss of exon 6 and exon 7 (wild-type, mutated and cryptic) are indicated by red arrows. The 3′ ss of exon 6 and exon 7 are indicated by black arrows. Mutated nucleotide at the first position of intron 7 is shown in blue. Cryptic site 1 and cryptic site 2 usage increased the length of exon 7 by 23 and 51 nucleotides, respectively. In sequencing chromatograms numbering of nucleotides starts either with the first position of intron 7 or the first position of exon 6. Abbreviations are described in Supplementary Table S1.

Supplementary Figure S3. Diagrammatic representation of potential base pairing formed between different *SMN* targets and eU1s. (A) Potential complementarity between the cryptic 5′ splice sites of exon 7, Cr1 and Cr2, and eU1s designed against the G1C-5′ss. Numbering starts from the first position of intron 7. The GU dinucleotides are indicated in red. Cr1 (Site 2) and Cr2 (Site 4) are underlined. The mutated 5′ ss of exon 7 is indicated by an arrow. Mutated nucleotide at the fist position of intron 7 is shown in blue. The identity of each eU1 is indicated on the left of each sequence. The mutated nucleotides at the 5′ end of eU1s are shown in green. Base pairing is marked by black circles; red circles represent wobble base pairing. Abbreviations are described in Supplementary Table S1. (B) Potential complementarity between the cryptic 5′ splice sites of exon 7, Cr1 and Cr2, and eU1s designed against different targets within intron 7

and the 5′ ss of exon 7, native and mutated. Wild-type U1 (wtU1) was included as well. Labeling is the same as (A). (C) Potential complementarity between the wild-type 5′ss of exon 7 and the same eU1s as in (B). Labeling is the same as in (A). Abbreviations are described in Supplementary Table S1.

Supplementary Figure S4. Effect of eU1s on splicing of endogenous *SMN* exon 7. (A) Diagrammatic representation of eU1 constructs and their targets within intron 7. Labeling is the same as in Figures 1B and 2A. (B) Splicing of endogenous *SMN* exon 7 in the presence of the overexpressed eU1s shown in (A). The identity of each eU1 is indicated at the top of the gel. Splice products amplified by RT-PCR were digested with DdeI to distinguish between the transcripts originated from *SMN1* and *SMN2* (29). The identity of products is indicated on the left of the gel. Whether splice products originate from *SMN1* or *SMN2* is marked on the right of the gel. The percentage of *SMN2* exon skipping was calculated as in (48). Abbreviations are described in Supplementary Table S1.

Supplementary Figure S5. Effect of eU1s on splicing of exon 7 with mutated 5′ss. (A) Diagrammatic representation of eU1 constructs and their target sites within intron 7. Labeling is the same as in Figures 1A and 2B. Of note, the *SMN1* minigene mutants used in this study carry a G-to-C, G-to-A or G-to-U mutation at the first position of intron 7. H shown in blue represents C or A or U. (B) In vivo splicing pattern of the *SMN1* minigenes that carry an indicated mutation at the first position of intron 7 in the presence of the overexpressed eU1s shown in (A). The identity of the minigenes and co-transfected eU1 constructs is marked at the top of the gel. The identity of splice products is indicated on the left of the gel. Quantification of the relative amount of the indicated splice intermediate/products is given in the bottom panel as a bar diagram. Error bars represent standard error. Abbreviations are described in Supplementary Table S1.

Supplementary Figure S6. Effect of decreased base pairing on the ability of the eU1-I7R7 variants to promote inclusion of exon 7 in *SMN1*^{G1C} transcripts. (A) Diagrammatic representation of eU1-I7R7 variants that form different numbers of continuous base pairs with eU1-I7R7 target site. Labeling is the same as in Figures 1B and 2A. The eU1-I7R7 variants were grouped in five sets according to what nucleotides at their 5′ ends were involved in base pairing

with the target. The names of the eU1 variants as well as the number of hydrogen bonds they form are given on the right of the panel. Abbreviations: HB, hydrogen bond. (B) In vivo splicing pattern of the *SMN1*^{G1C} minigene in the presence of the overexpressed eU1s shown in (A). The identity of U1 constructs is marked at the top of the gel. The identity of splice products is indicated on the left and the right of the gel. Quantification of the relative amount of the indicated splice intermediate/products is given in the bottom panel as a bar diagram. Error bars represent standard error. Abbreviations are described in Supplementary Table S1.

Supplementary Figure S7. Alignment of mammalian *SMN* sequences. All sequences were obtained from Genbank. The entire exon 7 and the first 100 bases of intron 7 from each species were aligned using the ClustalW algorithm. The species from which each sequence was derived is indicated on the left. Exon 7 is indicated with a gray box. The 5′ss is indicated by an arrow. GT dinucleotide of the 5′ss is highlighted in red letters. ISS-N1 and URCs are indicated with pink and green colored boxes, respectively. Cr1 and Cr2 sites are highlighted in yellow, and conserved nucleotides are indicated with red letters. Location and identities of 3 known SNPs are indicated.

Supplementary Figure S8. Effect of SNPs near Cr1 site on exon 7 splicing. (A) Diagrammatic representation of SNPs and the annealing sites of co-transfected U1 snRNAs within intron 7. Labeling is the same as in Figures 1B and 2A. The location and identity of each mutation designed to mimic SNPs are indicated. Accession numbers are given. (B) In vivo splicing pattern of minigenes co-transfected with either wtU1 or eU1-I7S3 expression constructs. Minigenes and U1 expression constructs used for transfection are indicated at the top of the gel. Labeling is the same as in Figure 1C. Abbreviations are described in Supplementary Table S1.

Supplementary Figure S9. Effect of ISS-N1-targeting ASO on usage of Cr1 in transcripts derived the *SMN1* minigenes carrying various pathogenic mutations. (A) Diagrammatic representation of ISS-N1 target in the context of SMA-associated splicing mutations. Name of the minigenes as well as type of SMA caused by a given mutation are given. Numbering of nucleotides starts from the first position of intron 7. Exonic sequences are shown in capital letters, intronic in lower case letters. The GU dinucleotides are marked in red. The 3′ss and the

5′ss of exon 7, native and mutated, are indicated by arrows. Mutated nucleotide within the 5′ss of exon 7 are shown in blue. Cr1 (Site 2) is underlined. Base pairing between F14 ASO and ISS-N1 is marked by black circles. (B) Effect of F14 on in vivo splicing pattern of exon 7 in the SMN minigenes (left panel) and in endogenous *SMN1*/*SMN2* (right panel). The identity of the minigenes and the presence of F14/control oligo is indicated at the top of each gel. Labeling is the same as in Figure 11. Abbreviations are described in Supplementary Table S1.

Supplementary Tables

Table S1. Abbreviations used in this study

Abbreviation	Full Name	Relevant First Figure
3′ss	3′ splice site	
5′ss	5′ splice site	
Δ4-5′ss	5′ss carrying a 4-nt deletion from the 4th to the 7th positions of *SMN1* intron 7	5
ΔE7	Transcript with skipped *SMN* exon 7	1
ΔPy	Deletions in the polypyrimidine tract at the 3′ss of *SMN1* exon 7	5
ASO	Antisense oligonucleotide	
bp	Base pair	
C6U	A C-to-U substitution at the 6th position of *SMN2* exon 7	
Cr1	Cryptic 5′ss between the 23rd and the 24th positions of *SMN* intron 7	1
Cr2	Cryptic 5′ss between the 51st and the 52nd positions of *SMN* intron 7	1
E6Cr	Cryptic 5′ss between the 61st and the 62nd positions of *SMN* exon 6; Shortened transcript due to the activation of the E6Cr and skipping of exon 7	1
E7-3′Cr	Cryptic 3′ss between the 8th and the 9th positions of *SMN* exon 7	4
E7Cr1	Extended exon 7 due to the activation of Cr1	1
E7Cr2	Extended exon 7 due to the activation of Cr2	1
E7Mv-Cr1	Splice product generated by the usage of the Mv-Cr1 site	9
eU1	Engineered U1 snRNA	
G1C-5′ss	5′ss carrying a G-to-C mutation at the 1st position of intron 7	
hnRNP	Hetero-nuclear ribonucleoprotein	1
ISS-N1	Intronic splicing silencer N1	
Mv-Cr1	Cr1 site moved away from ISS-N1 due to insertion of Sequence 1 or 2	9
nt	Nucleotide	
R-series eU1s	eU1s targeting random sequences located within exon 7 and intron 7 of *SMN*	2
SMN (Italics)	Survival motor neuron gene or transcript	6
SMN	Survival motor neuron protein	
S-series eU1s	eU1s targeting GU dinucleotide-containing sites (sites 1 through 6) located within the first 90 nucleotides of *SMN* intron 7	2
TSL1	Terminal stem-loop 1 located within *SMN* exon 7	
U1 snRNA	U1 small nuclear RNA	
U1 snRNP	U1 small nuclear ribonucleoprotein	
U6G-5′ss	5′ss carrying an U-to-G mutation at the 6th position of intron 7	
URC1	U-rich cluster 1	1
URC2	U-rich cluster 2	1
wt	Wild-type	
wt-3′ss	Wild-type 3′ splice site (Native 3′ splice site)	1
wt-5′ss	Wild-type 5′ splice site (Native 5′ splice site)	1
wtU1	Wild-type U1 snRNA	1

Supplementary Table S2. Strength of the splice sites as determined by MaxEntScan scoring algorithm (59)*

5′ splice site					
Name	Sequence	MaxNet	MDD	MM	WMM
Native (wt) site of exon 7	ggaGTAAGT	8.57	12.28	6.36	6.42
G1C-5′ss	gga**C**TAAGT	-3.94	1.20	-3.61	-2.87
U6G-5′ss	ggaGTAAG**G**	4.33	9.48	4.67	5.49
Δ4-5′ss	ggaGTA----TGC	1.87	4.38	2.64	2.25
Intron 7 site 1	taaGTCTGC	-7.82	-1.62	-6.51	-0.82
Intron 7 site 2 (Cr1)	aaaGTGAAT	0.76	4.18	3.31	4.71
Intron 7 site 3	tttGTAAAA	-6.62	-5.32	0.2	0.08
Intron 7 site 4 (Cr2)	atgGTTTGT	4.88	5.38	4.03	2.77
Intron 7 site 5	aatGTTTTT	-13.14	-7.72	-10.55	-3.21
Intron 7 site 6	aaaGTTCAG	-9.44	0.48	-4.40	-2.78
Mutation 23G (Cr1)	aa**g**GTGAAT	6.38	10.28	6.84	7.87
Mutation 26A (Cr1)	aaaGT**A**AAT	4.06	7.28	4.04	6.09
Mutation 28G (Cr1)	aaaGTGA**G**T	8.40	12.68	7.78	8.16
Mutation 23/28G (Cr1)	aa**g**GTGA**G**T	10.47	15.28	11.31	11.33
Mutation 38G	tt**g**GTAAAA	3.23	6.58	3.9	4.2
Mutation 50A (Cr2)	a**a**gGTTTGT	7.81	11.58	6.11	4.81
Mutation 54A (Cr2)	atgGT**A**TGT	8.35	11.98	7.95	7.72
Mutation 55A (Cr2)	atgGTT**A**GT	7.07	8.88	5.13	5.72
Native 5′ss of exon 6	atgGTAAGT	11.01	15.48	10.01	10.67
Cryptic in exon 6 (E6Cr)	gaaGTATGT	6.97	8.68	5.94	5.76
Cryptic in exon 7	aagGTGCTC	3.68	8.18	3.62	3.21
Mutation -1G (5′ss)	gg**g**GTAAGT	9.65	13.78	8.77	9.59
3′ splice site					
Name	Sequence		MaxNet	MM	WMM
Native (wt) site of exon 7	ttcctttattttccttacAGggt		10.92	13.08	15.51
Cryptic site of exon 7 in *SMN1*-3′ss mutant	ttttccttaca**t**ggtttcAGaca		6.98	6.66	6.30
Cryptic site of exon 7 in *SMN2*-3′ss mutant	ttttccttaca**t**ggttttAGaca		6.82	6.86	6.01
ΔPy-3′ss	ttttaacttttttccttacAGggt		10.97	11.89	13.9
Native (wt) site of exon 8	ttctaatttctcatttgcAGgaa		10.77	10.86	10.52

*Exonic sequences are highlighted in red, mutations are shown in bold and are underlined, deletion of four nucleotides at the 5′ss of exon 7 is indicated by a dashed line. Abbreviations: MaxNet, Maximum Entropy Model; MDD, Maximum Dependence Decomposition Model; MM, First-order Markov Model; WMM, Weight Matrix Model. The above-mentioned models were used to score the strength of the splice sites. The input sequences for the 5′ ss scoring were 9-nucleotide-long: three "exonic" (upstream of a GU dinucleotide, shown in low case letters) and six "intronic" bases (shown in capital letters). The input sequences for the 3′ ss scoring were 23 nucleotides long, including twenty "intronic" and three "exonic" (three nucleotides downstream of the cleavage site) bases. A hallmark AG dinucleotide of the 3′ is shown in capital letters.

Supplementary Table S3. Strength of the 5′ splice sites as determined by HBond scoring algorithm*

Name	Sequence	H-Bond score
Native (wt) 5′ss of exon 7	ggaGTAAGTcT	14.50
U6G-5′ss	ggaGTAAG**g**cT	12.00
Δ4-5′ss	ggaGTA----tGcca	8.70
Intron 7 site 1	tAaGTctGcca	3.00
Intron 7 site 2 (Cr1)	aAaGTgAaTcT	5.70
Intron 7 site 3	tttGTAAaacT	9.00
Intron 7 site 4 (Cr2)	atGGTttGTgg	9.40
Intron 7 site 5	aAtGTtttTga	2.20
Intron 7 site 6	aAaGTtcagAT	3.00
Mutation 23G (Cr1)	aA**G**GTgAaTcT	10.90
Mutation 26A (Cr1)	aAaGT**A**AaTcT	9.60
Mutation 28G (Cr1)	aAaGTGA**G**TcT	13.00
Mutation 23/28G (Cr1)	aA**G**GTGA**G**TcT	18.20
Mutation 30A (Cr1)	aAaGTgAaTc**a**	5.50
Mutation 50A (Cr2)	a**A**GGTttGTgg	11.50
Mutation 54A (Cr2)	atGGT**A**tGTgg	13.70
Mutation 55A (Cr2)	atGGTt**A**GTgg	13.10
Cryptic in exon 6 (Cr)	ggaGTAtGTta	10.70
Cryptic in exon 7	aAGGTgctcAc	10.10
Mutation -1G (5′ss)	gg**G**GTAAGTcT	17.50

*Exonic sequences are highlighted in red, mutations are shown in bold and are underlined, deletion of four nucleotides at the 5′ss of exon 7 is indicated by a dashed line. Each sequence also displays H-Bond pattern. The input sequences were 11-nucleotide-long: three “exonic” (upstream of a GU dinucleotide) and nine “intronic” bases.