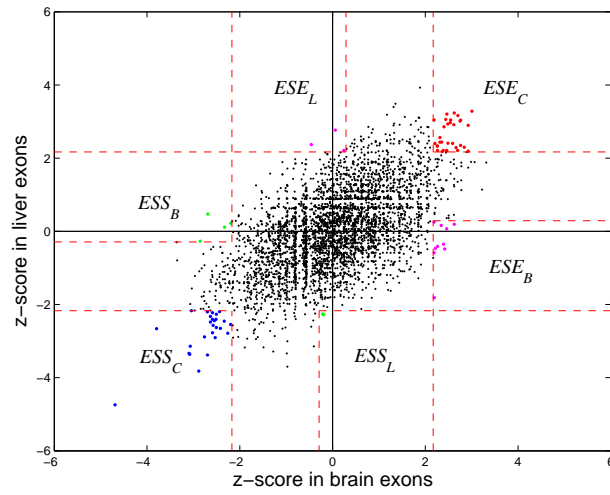
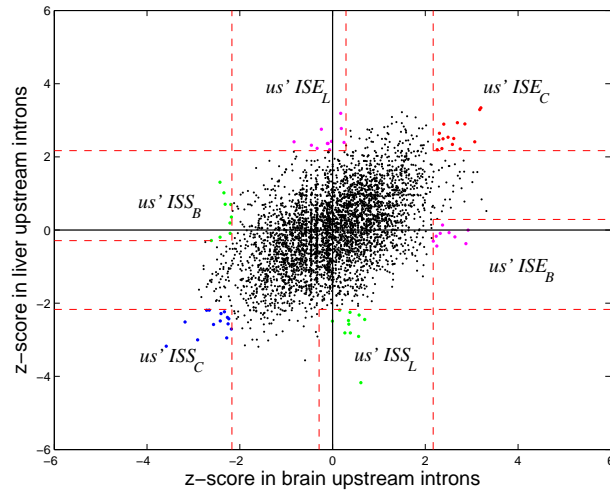


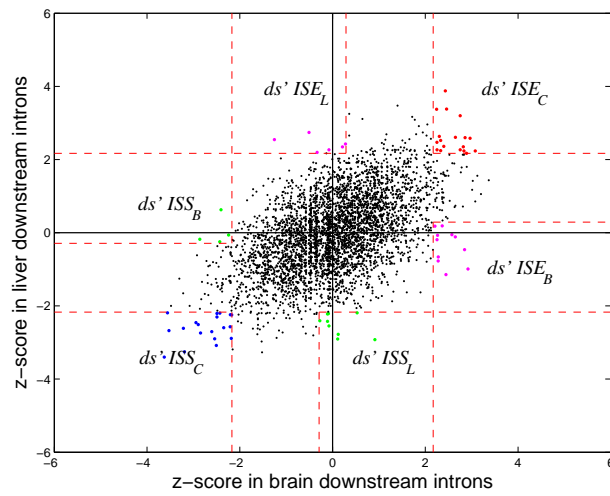
(a)



(b)

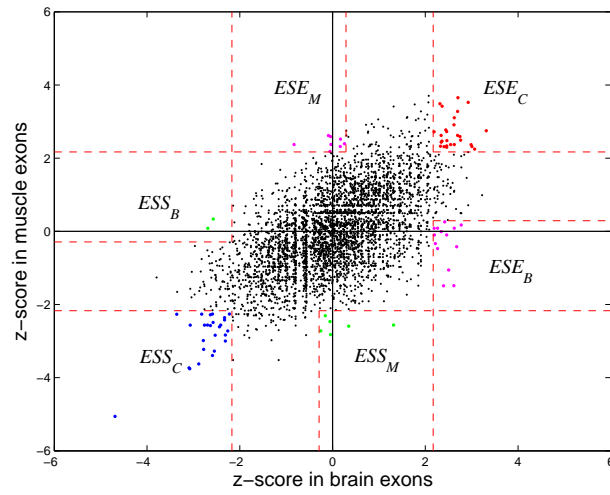


(c)

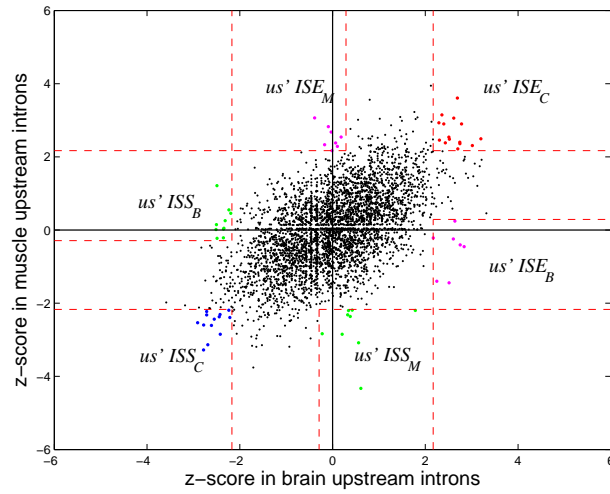


Supplementary Figure 1. z-scores for all hexamers in brain and liver. (a) z-scores in exons. ESE_C , ESE_B and ESE_L stand for common ESE, brain-specific ESE and liver-specific ESE, respectively. (b) z-scores in 400 nt intronic sequences upstream of the exons. (c) z-scores in 400 nt intronic sequences downstream of the exons.

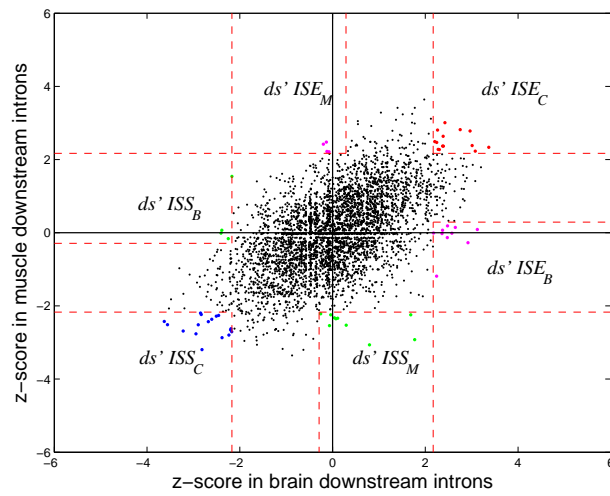
(a)



(b)



(c)



Supplementary Figure 2. z-scores for all hexamers in brain and muscle. (a) z-scores in exons. ESE_C , ESE_B and ESE_M stand for common ESE, brain-specific ESE and muscle-specific ESE, respectively. (b) z-scores in 400 nt intronic sequences upstream of the exons. (c) z-scores in 400 nt intronic sequences downstream of the exons.

Supplementary Table 1: List of 456 common and tissue-specific SREs and comparison with existing results. Columns 2 to 6 specify the type of the SREs. Column 7 indicates whether an SRE is also a RESCUE-ESE. Last column contains the information of an SRE in SpliceAid database, which includes the binding factors, PMID reference and the number of the sequences (in parentheses) that our SRE can match to. If more than two records can be found in SpliceAid, only first two are kept..

Motif	ESE	ESS	5' ISE	5' ISS	3' ISE	3' ISS	RescueESE	Selected SpliceAid
AACUGC	BL?	-	
AAGAAG	BLM	AAGAAG	HTra2alpha,9546399(2);HTra2beta1,9546399(2);
AAGCAG	BLM	AAGCAG	SC35,10094314(2);
AUCUAU	BL?	.	.	.	BL-	.	-	
ACUUCG	BL?	-	9G8,10094314(1);SRp20,10094314(1);
ACGGCA	BLM	-	
AGAAGC	BLM	AGAAGC	SC35,7543047(1);
AGCAGC	BL?	AGCAGC	FMRP,15805463(1);
AGCUGC	BLM	-	
AGGAAC	BLM	AGGAAC	SF2/ASF,7543047(2);FMRP,15805463(1);
UAUGAC	BLM	-	
UCGACU	BL?	-	SRp20,10094314(4);
UGUUAG	BLM	-	
UGGAGC	BLM	-	
UGGUGU	BL?	-	
UGGGCA	BL?	-	
CACGGC	BLM	-	
CAGCAA	BL?	-	
CUCAUA	BL?	-	
CUGGUG	BL?	-	FMRP,15805463(1);
CGACUG	BL?	.	.	.	-L-	.	-	FMRP,15805463(4);
CGGCAC	BLM	-	SF2/ASF,16825284(1);
CGGCCA	BL?	-	
GACUAU	BLM	-	
GUGAUA	BLM	-	
GUGGCU	BLM	-	FMRP,15805463(1);
GCAGAA	BL?	GCAGAA	
GCCAAC	BL?	-	
GGAUUU	BL?	-	
GGAUGA	BL?	GGATGA	SF2/ASF,7543047(1);
GGAGCA	BLM	-	SRp40,9037021(4);
GGUCAG	BL?	-	SC35,10629063(1);
AUGAGC	B--	-	
ACGCGC	B--	-	SF2/ASF,7543047(4);
AGCCUG	B--	-	
UCUUGC	B-?	.	.	.	UCUUGC	.	-	
CUGAAA	B--	CTGAAA	
CGCUGC	B--	-	SC35,10629063(1);
GCAUUC	B--	-	
GCGCGC	B--	--M	-	SF2/ASF,7543047(1);
GGGGAC	B-?	-	
GGGGCC	B--	.	.	B?M	.	.	-	
AAAUCA	-L?	-	
AUGUAC	-L?	-	
UCAGGC	-L?	-	
AGGGUG	B?M	-	
UUUUGG	B?M	-	
UUCGAG	B?M	-	SC35,7543047(3);SC35,10094314(3);et al.
CAAUGA	B?M	-	
CAACCA	B?M	-	
CUUCGA	B?M	-	9G8,10094314(1);SRp20,10094314(2);
CCUGCG	B?M	-	SC35,10629063(1);
CGGAGA	B?M	CGGAGA	SC35,7543047(2);
GCUAAU	B?M	-	
GGAGCG	B?M	-	SC35,7543047(1);SF2/ASF,16825284(1);
GGUACC	B?M	-	
GGUGGA	B?M	-	
AUCCCG	B?-	-	SC35,10094314(1);
UACCUC	B?-	?L-	-	
CAGCGC	B?-	-	
CCCUUA	B?-	-	

CCGACG	B?-	-	9G8,10094314(1);SF2/ASF,16825284(1);
GAUGAG	B?-	GATGAG	
AUGGCU	-?M	-	
AGAGCG	-?M	-	
UGGAGA	-?M	TGGAGA	SC35,10094314(1);
CAGGCA	-?M	-	
CGGGCU	-?M	-	
GCAACA	-?M	GCAACA	
GGAAGG	--M	-	
GGCGGG	--M	-	
AACGAG	?LM	-	SC35,10094314(1);
AAGGAG	?LM	AAGGAG	SRp40,9037021(1);FMRP,15805463(3);
AUAUGC	?LM	-	
AGAAGA	?LM	AGAAGA	SF2/ASF,7543047(2);SRp40,9037021(1);et al.
AGCAGA	?LM	AGCAGA	SC35,7543047(4);
AGCGAG	?LM	-	
UCGAUA	?LM	-	SRp20,10094314(1);
UCGGGC	?LM	-	SF2/ASF,7543047(1);
UGUGGC	?LM	-	
UGCAAC	?LM	-	
CAACGA	?LM	-	9G8,10094314(2);SRp20,10094314(3);
CAGACU	?LM	-	
CUCGGU	?LM	-	
CCAAUG	?LM	-	SF2/ASF,16825284(1);
CGAUAU	?LM	-	
CGAGUA	?LM	-	SC35,7543047(3);SC35,10094314(2);
GAUGGC	?LM	-	
GAGCUG	?LM	-	SRp40,9037021(1);SC35,10094314(1);
GAGCGA	?LM	-	FMRP,15805463(3);
GCAGAU	?LM	GCAGAT	
GGAGAU	?LM	GGAGAT	SC35,7543047(1);SC35,10094314(2);
AAACUG	?L-	AAACTG	
ACCGCC	?L-	-	
ACGUUC	?L-	ACGTTC	
UAUAGG	?L-	-	
UUGGGA	?L-	-	
CUCAAU	?L-	-	
CCGCAA	?L-	-	
CGCUCC	?L-	-	
GUGUCU	?L-	-	
GGCGUA	?L-	-	
CACGUG	?-M	-	
GAUGCC	?-M	-	
GAGCGG	?-M	-	FMRP,15805463(1);SF2/ASF,16825284(1);
GAGGAG	?-M	GAGGAG	Added,10022858(1);
GCUGCG	?-M	-	
GGGCUG	?-M	-	FMRP,15805463(3);
AAAUAA	.	BLM	-	
AAUUAA	.	BLM	-	
ACAAUA	.	BL?	-	
ACUCAC	.	BL?	-	
ACCGUC	.	BLM	-	SRp55,9649504(1);
AGAUUA	.	BL?	BL?	.	.	.	-	
AGGGGA	.	BL?	-	
UAAGAU	.	BLM	-	
UAGGCU	.	BL?	-	hnRNP A1,7510636(1);SC35,10094314(1);
UUAAAG	.	BLM	-	
UUUAAA	.	BLM	-	
UUUUAA	.	BL?	-	
UUUGAA	.	BL?	-	
UUCUUU	.	BL?	-	
UUGAAA	.	BL?	-	
UCUUAG	.	BLM	-	
UGUAGC	.	BL?	-	
UGCUUA	.	BL?	-	
UGCCUG	.	BLM	-	
CUGCCU	.	BLM	-	
CCAUGU	.	BL?	-	
CCUGCC	.	BLM	-	
GAACUC	.	BL?	-	
GUAUAA	.	BL?	-	

GUGUUU	.	BL?	-	
GCCUGC	.	BLM	-	
AAUGUU	.	B-?	-	
AUAAUA	.	B-M	-	
AUGAAU	.	B--	-	SC35,10094314(1);
AGAGUG	.	B-?	-	SC35,7543047(1);
ACAUCC	.	-L?	-	
ACCUUU	.	-L?	-	
AAUAAU	.	B?M	-	
AACACU	.	B?M	-	
AUAAGA	.	B?M	-	
AUAUUG	.	B?M	-	
UAAUAA	.	B?M	?-M	-	ZRANB2,19304800(1);
UAUUUA	.	B?M	-	
UUGUUU	.	B?M	-	
CAUUUG	.	B?M	-	
CUUAAG	.	B?M	-	
CCCAUU	.	B?M	-	
CGCUGA	.	B?M	-	
GAUAAA	.	B?M	-	
GACUCA	.	B?M	-	SC35,10629063(1);
GGGUCU	.	B?M	.	?-M	.	.	.	-	
GUUCUG	.	B?-	-	SC35,10629063(1);
AAUUC	.	--M	-	
AAGCCU	.	-?M	-	
UCCUUG	.	--M	-	
CCCUGC	.	-?M	-	
CCGCC	.	-?M	-	
CCGCGC	.	-?M	-	SF2/ASF,7543047(1);
AUACAA	.	?LM	-	
ACCUCU	.	?LM	-	
UAACUU	.	?LM	-	
UUAUAA	.	?LM	-	
CACUCU	.	?LM	-	
CAGACC	.	?LM	-	
CUCUGU	.	?LM	.	.	.	?LM	.	-	
CCUCCC	.	?LM	-	
GUAAAG	.	?LM	-	ZRANB2,19304800(2);
GGUCUU	.	?LM	-	
AUUCUU	.	?L-	-	
CCCCAG	.	?L-	-	
CGGCAG	.	?L-	-	SF2/ASF,16825284(1);
CCGCUC	.	?-M	-	
GGGUUG	.	?-M	-	
UAACGU	.	.	BLM	-	
UAUGCA	.	.	BL?	-	
UUGUAC	.	.	BLM	-	
UGAUUG	.	.	BL?	-	
UGUGCG	.	.	BL?	-	
CACAU	.	.	BL-	-	
CUAAAG	.	.	BLM	-	
CUUGGG	.	.	BLM	-	
CCUAAA	.	.	BLM	-	
CGACAG	.	.	BL?	-	SF2/ASF,16825284(1);
GAGAUU	.	.	BL?	-	GAGATT SRp20,10094314(1);
GUGUGC	.	.	BL?	-	
GCAGCA	.	.	BLM	-	
GGAGUC	.	.	BL?	-	
AACUCA	.	.	B--	-	
AUAUGU	.	.	B--	-	
UAGUCG	.	.	B-?	-	
UUGGUG	.	.	B-?	-	
UCCUGA	.	.	B-?	-	
CAUGCA	.	.	B--	-	Nova-1,9154818(1);
CUGGGC	.	.	B-?	-	FMRP,15805463(2);
CCUACU	.	.	B--	-	
GGUGCC	.	.	B-?	-	
ACACUG	.	.	-L?	-	
ACUUGA	.	.	-L?	-	
AGGAGA	.	.	-L?	-	AGGAGA SC35,7543047(4);SF2/ASF,7543047(1);et al.
UACACU	.	.	-L-	-	

UUGUAG	.	.	-L?	.	B?M	.	-	
CUAUAA	.	.	-L?	.	.	.	-	
CCGAAC	.	.	-L?	.	.	.	-	
CGCAUG	.	.	-L?	.	.	.	-	
GAUGGU	.	.	-L?	.	.	.	-	
GACACA	.	.	-L?	.	.	.	-	
GCACGC	.	.	-L?	.	.	.	-	
AAGUGA	.	.	B?M	.	.	.	-	
ACUUAG	.	.	B?M	.	.	.	-	
AGCUUA	.	.	B?M	.	.	.	-	
UUCAUG	.	.	B?M	.	.	.	-	
UCCUGC	.	.	B?M	.	.	.	-	SC35,10094314(1);
CACUUA	.	.	B?M	.	.	.	-	
CAGGGG	.	.	B?M	.	.	.	-	
GUGAUU	.	.	B?M	.	.	.	-	
GCUCCC	.	.	B?M	.	.	.	-	SC35,10629063(1);
AUGUGU	.	.	B?-	.	.	.	-	ETR-3,15657417(1);
UGAGAU	.	.	B?-	.	.	.	-	
AUAAAU	.	.	-?M	.	.	.	-	
ACCGGA	.	.	-?M	.	.	.	-	
AGCAAG	.	.	-?M	.	.	.	-	AGCAAG
AGCUGA	.	.	-?M	.	.	.	-	AGCTGA SRp40,9037021(1);
AGGCCG	.	.	-?M	.	.	.	-	
UAAUUA	.	.	-?M	.	.	.	-	
UCCCCU	.	.	--M	.	.	.	-	SC35,10629063(1);
GCUGAC	.	.	--M	.	.	.	-	
AGACAC	.	.	?LM	.	.	.	-	
UAUAAA	.	.	?LM	.	.	.	-	
UCAGCU	.	.	?LM	.	.	.	-	
UGAUUU	.	.	?LM	.	.	.	-	
UGUCAC	.	.	?LM	.	.	.	-	
UGCGUG	.	.	?LM	.	.	.	-	SRp55,9649504(1);
CACUUG	.	.	?LM	.	.	.	-	
CUUGAA	.	.	?LM	.	.	.	-	
CCAGAG	.	.	?LM	.	.	.	-	
GUGCGU	.	.	?LM	.	.	.	-	
AAGUUU	.	.	?L-	.	.	BL?	-	
CAA AUG	.	.	?L-	.	.	.	-	
CAGCCA	.	.	?-M	.	.	.	-	
GCAGCC	.	.	?-M	.	.	.	-	
GGCAGC	.	.	?-M	.	?LM	.	-	
AAGAUU	.	.	.	BLM	.	.	-	AAGATT
UUAUCA	.	.	.	BL?	.	.	-	
UUACAA	.	.	.	BL?	.	.	-	
UCCAAU	.	.	.	BL?	.	.	-	
CAUUUU	.	.	.	BL?	.	.	-	Nova-1,9154818(1);
CUAGAA	.	.	.	BLM	.	.	-	SC35,10629063(1);
CUUUCU	.	.	.	BL?	.	.	-	
CUGCAA	.	.	.	BL?	.	.	-	CTGCAA
CUGGAA	.	.	.	BL?	.	.	-	CTGGAA
CCCACC	.	.	.	BL?	.	.	-	
CCCUUC	.	.	.	BL?	.	.	-	
CCCCGC	.	.	.	BL?	.	-LM	-	SF2/ASF,16825284(1);
GCCGAU	.	.	.	BLM	.	.	-	
GGGAAA	.	.	.	BL?	.	.	-	
AACCAG	.	.	.	B-?	.	.	-	SC35,10094314(2);
AACCGU	.	.	.	B-?	.	.	-	AACCGT
ACUAUU	.	.	.	B-?	.	--M	-	
CAAAAA	.	.	.	B--	.	.	-	
CAUCCG	.	.	.	B-M	.	.	-	
CAGUCA	.	.	.	B-?	.	.	-	
CUUACA	.	.	.	B-?	.	.	-	
GUAUCG	.	.	.	B--	.	.	-	
GUCAGU	.	.	.	B-?	.	.	-	SC35,10629063(1);
AUUCAU	.	.	.	-L-	.	.	-	Nova-1,9154818(5);SRp20,10094314(1);
AUCCCU	.	.	.	-L?	.	.	-	
UAAGGC	.	.	.	-L?	.	.	-	
UCUCUC	.	.	.	-LM	.	.	-	
UGUAUG	.	.	.	-L?	.	.	-	
CUUCUU	.	.	.	-L?	.	BL?	-	hnRNP I (PTB),9214659(1);SRp20,10094314(1);
CUCUCU	.	.	.	-LM	.	.	-	

CCCCCA	.	.	.	-L?	.	.	-	
GAGAAG	.	.	.	-L?	.	.	GAGAAG	SRp40,9037021(1);
GCAUGC	.	.	.	-L?	.	.	-	
GGAUAC	.	.	.	-L?	.	.	-	
AUAUGA	.	.	.	B?M	.	.	-	
AUUUAG	.	.	.	B?M	.	.	-	
UCCUCA	.	.	.	B?M	.	.	-	SC35,10629063(1);
UGAGAA	.	.	.	B?M	.	.	TGAGAA	SRp40,9037021(1);
UGCCGA	.	.	.	B?M	.	.	-	
CAGUCC	.	.	.	B?M	.	.	-	SC35,10094314(1);
CAGGAC	.	.	.	B?M	.	.	-	
GGGGCG	.	.	.	B?M	.	.	-	
AUCGUU	.	.	.	B?-	.	.	-	SRp20,10094314(2);
ACAGGC	.	.	.	B?-	.	.	-	
CAAGGA	.	.	.	B?-	.	.	-	SF2/ASF,7543047(1);SRp40,9037021(1);
CAUCAC	.	.	.	B?-	.	.	-	Nova-1,9154818(1);SRp20,10094314(2);
CAGCGG	.	.	.	B?-	.	.	-	SF2/ASF,9649504(2);
CAGGCU	.	.	.	B?-	.	.	-	
GUUCCG	.	.	.	B?-	.	.	-	SC35,10094314(1);
GGUAGC	.	.	.	B?-	.	.	-	ZRANB2,19304800(1);
AAUUGC	.	.	.	-?M	.	.	-	
AUAUAC	.	.	.	--M	.	.	-	
AUGCAG	.	.	.	-?M	.	.	-	
ACACAC	.	.	.	--M	.	.	-	added,15889141(1);
UAUGUG	.	.	.	-?M	.	.	-	ETR-3,15657417(1);
CUAAUU	.	.	.	-?M	.	.	-	
GUAUAU	.	.	.	-?M	.	.	-	
UCUAAU	.	.	.	?LM	.	.	-	
UCUCUU	.	.	.	?LM	.	.	-	
UCUGCA	.	.	.	?LM	.	.	-	
UCUGCG	.	.	.	?LM	.	.	-	
CAAUAC	.	.	.	?LM	.	.	-	
CUAUAU	.	.	.	?LM	.	.	-	
CUUCCU	.	.	.	?LM	B-?	.	-	
CUCUUU	.	.	.	?LM	.	.	-	
CUCCCG	.	.	.	?LM	.	.	-	
CCUAUG	.	.	.	?LM	.	.	-	
GUUAUC	.	.	.	?LM	.	.	-	
GUGUAU	.	.	.	?LM	.	.	-	
GGAAAC	.	.	.	?LM	.	.	GGAAAC	
CUGCGG	.	.	.	?L-	.	.	-	
CCGACC	.	.	.	?L-	B-?	.	-	SF2/ASF,16825284(1);
CGACCC	.	.	.	?L-	.	.	-	
GGCUUU	.	.	.	?L-	.	.	-	
CAUCAA	.	.	.	?-M	.	.	CATCAA	SRp20,10094314(1);
AAUCAU	BL?	.	-	Nova-1,9154818(1);
AUUGAU	BL?	.	-	9G8,10094314(1);
AUGGAG	BLM	.	ATGGAG	SC35,7543047(1);SC35,10629063(1);et al.
UUACUG	BLM	.	-	
UUGAUU	BL?	.	-	
UUGCUU	BLM	.	-	
UGAGGG	BL?	.	-	
CAUUAC	BL?	.	-	
CUUGCC	BL?	.	-	
CUGCGU	BL?	.	-	
CCGAAG	BL?	.	-	
GAACCG	BLM	.	-	
GAGGGA	BL?	.	-	
GUCGUC	BL?	.	-	
GCCUUU	BL?	.	-	
GGAUGC	BL?	.	-	
GGAGGG	BL?	.	-	SRp40,9037021(2);
ACGCC	B--	.	-	
UUGAGG	B-?	.	-	
UGCGCC	B-?	.	-	
CACCCU	B-?	.	-	
CACCGU	B-?	.	-	
GAGUUG	B-M	.	-	
GCGCCU	B--	.	-	
GGAAUG	B-?	.	-	SF2/ASF,7543047(1);SC35,10094314(2);
GGGAUG	B--	.	-	

AUUACA	-L-	.	-	
AUGAUC	-L?	.	-	
UGAACA	-L?	.	-	SF2/ASF,16825284(1);
GCUUAC	-L?	.	-	
GCGGAU	-L?	.	-	
AUUACU	B?M	.	-	
CCAAAC	B?M	.	-	
CGCUAA	B?M	.	-	
GUUAAA	B?M	.	-	
GCGUGC	B?M	.	-	
GGCAGG	B?M	.	-	
GGCGGC	B?M	.	-	
GGGGGA	B?M	.	-	Added,10022858(1);
AGGGGC	B?-	.	-	
UUCUAU	B?-	.	-	
CAGUUA	B?-	.	-	
CUAGUA	B?-	.	-	hnRNP DL,12406575(1);
CGACAU	B?-	.	-	
GGAACU	B?-	.	-	
UUUAGU	-?M	.	-	
UGCAUG	--M	-L-	-	Fox-1,16537540(1);Fox-2,16537540(1);et al.
UGGUUU	-?M	.	-	
GUCUCG	--M	.	-	
ACGUUA	?LM	.	-	
UUAAAA	?LM	.	-	
UGUGGA	?LM	.	-	
CGUUUAU	?LM	.	-	
GUGGAG	?LM	.	-	
AACUGG	?L-	.	-	AACTGG
UAUCUA	?L-	.	-	
UUGCGA	?L-	.	-	
UCAUCU	?L-	.	-	Nova-1,9154818(2);SRp20,10094314(1);
UCUAUC	?L-	--M	-	
CGCGCU	?L-	.	-	SF2/ASF,7543047(1);
CGGUGC	?L-	.	-	
GAUCAG	?L-	.	-	
AAUGGU	?-M	.	-	
AGGGCU	?-M	.	-	
UGACAU	?-M	.	-	
CUGACA	?-M	.	-	
GCGAAC	?-M	.	-	
AAACUU	BL?	-	
AUUUUG	BL?	-	
ACUCGG	BL?	-	
AGUUUC	BL?	-	
UACGUA	BLM	-	
UUCUUC	BLM	-	SRp20,10094314(2);
UCAUCA	BLM	TCATCA	Nova-1,9154818(2);SRp20,10094314(3);
UCUAAA	BL?	-	
CACAUG	BL?	-	
CUAAAA	BLM	-	
CUUCUC	BL?	-	
CUGGAG	BLM	-	
CCUUCU	BL?	-	
CGAGAA	BLM	CGAGAA	SC35,7543047(1);SC35,10094314(1);
GAAUAU	BL?	-	
GUUUGG	BL?	-	
GCAUGG	BL?	-	
CACACA	B-?	-	added,15889141(1);
GAAGAG	B-?	GAAGAG	SF2/ASF,7543047(1);
GCACAG	B-?	-	
GGAAGA	B-?	GGAAGA	
ACGGCU	-L?	-	
UGAAGG	-L-	TGAAGG	FMRP,15805463(2);
UGGGGA	-L?	-	
CGCCCC	-L-	-	
GCAUGU	-L-	-	
GCUCGG	-L?	-	
GCCCCG	-L?	-	
AAAAGC	B?M	AAAAGC	
AAGCCA	B?M	AAGCCA	

AGAUAA	B?M	-	SC35,7543047(1);
AGCCAU	B?M	-	
UAAAAG	B?M	-	ZRANB2,19304800(1);
UAGAAC	B?M	-	
UCGUUC	B?M	-	SRp20,10094314(2);
CUACCC	B?M	-	
CUUCAAA	B?M	-	SRp20,10094314(3);
CCAUCU	B?M	-	
GCAUAA	B?M	-	
ACGAUG	B?-	-	9G8,10094314(1);
UCCUUA	B?-	-	
GUUUCU	B?-	-	
GGAUCC	B?-	-	SC35,10629063(1);
AUCAUC	-?M	ATCATC	SRp20,10094314(2);
AGCAUU	-?M	-	
CUAUCU	--M	-	
CUAGCC	-?M	-	hnRNP DL,12406575(2);
CGCGCC	-?M	-	SF2/ASF,7543047(1);SF2/ASF,16825284(1);
GAGGCU	--M	-	SRp40,9037021(1);
AUAAGU	?LM	-	Nova-1,9154818(1);
ACUCAA	?LM	ACTCAA	SC35,10629063(1);
UAGAAG	?LM	-	
UCGGGU	?LM	-	
CAUGGG	?LM	-	
CCGGAU	?LM	-	SC35,10094314(1);
GGCCCU	?LM	-	
UAUGGC	?L-	-	
UUCAUU	?L-	-	Nova-1,9154818(4);
UGAACG	?L-	-	
CUCUUC	?L-	-	SRp20,10094314(1);
CCCCGA	?L-	-	SF2/ASF,16825284(1);
GUGCAU	?L-	-	
GCUAUG	?L-	-	FMRP,15805463(1);
ACUCUA	?-M	-	
AGCUAG	?-M	-	
AGGUAG	?-M	-	ZRANB2,19304800(2);
UAUAUU	?-M	-	
UUUUUU	?-M	-	
UUCCCC	?-M	-	
GCCAUG	?-M	-	Nova-1,9154818(1);

Supplementary Table 2: List of 71 SREs with p -value < 0.01 in the position bias test. Column 2 and 3 represent the annotation of the SRE. Column 4 gives the total number of occurrence of the SRE in the data analyzed. Column 5 gives the total number of intronic or exonic sequences used in analyses. The last three columns list p -value, chi-square statistic and degree of freedom, respectively.

Elements	Anno.	Anno.	No.Occu.	No. Seq.	p-value	statistic	DF
CTCTCT	5ISS	-LM	301	721	1.33E-16	158.1739	38
TCTCTC	5ISS	-LM	279	721	2.59E-14	144.3971	38
ATGGAG	3ISE	BLM	142	994	1.83E-13	139.1915	38
AAGATT	5ISS	BLM	168	1030	4.69E-12	130.4024	38
TTGTAC	5ISE	BLM	93	964	6.85E-11	122.9677	38
TCTCTT	5ISS	?LM	263	721	1.58E-10	120.6236	38
TTCTTC	3ISS	BLM	316	1044	1.61E-09	113.9873	38
AGATAA	3ISS	B?M	101	658	4.69E-08	104.0396	38
CTGCCT	ESS	BLM	43	454	5.26E-08	51.65116	9
CCTAAA	5ISE	BLM	116	964	1.20E-07	101.1897	38
TGATTT	5ISE	?LM	196	639	1.51E-07	100.4796	38
TTACTG	3ISE	BLM	177	994	2.83E-07	98.55172	38
CTTGGG	5ISE	BLM	202	964	5.63E-07	96.4	38
CTAAAA	3ISS	BLM	189	1044	1.41E-06	93.49206	38
TTAAAG	ESS	BLM	33	454	1.63E-06	43.66667	9
GTGATT	5ISE	B?M	113	663	2.21E-06	92.05556	38
TTCATG	5ISE	B?M	136	663	2.41E-06	91.77778	38
AGGAAC	ESE	BLM	38	483	2.63E-06	42.52632	9
CTCTTT	5ISS	?LM	226	721	3.50E-06	90.57399	38
AGTTTC	3ISS	BL?	136	702	4.21E-06	89.97059	38
TATGCA	5ISE	BL?	90	626	6.83E-06	88.38636	38
GGGAAA	5ISS	BL?	151	686	1.08E-05	86.87417	38
TGTCAC	5ISE	?LM	88	639	1.58E-05	85.58621	38
AGATAT	5ISE	BL?	89	626	1.97E-05	84.84091	38
GAGGGA	3ISE	BL?	144	643	2.81E-05	83.63636	38
GTTTGG	3ISS	BL?	134	702	3.27E-05	83.1194	38
TTGATT	3ISE	BL?	120	643	4.17E-05	82.28814	38
CTCTGT	3ISE	?LM	239	663	4.97E-05	81.68103	38
TAAAAG	3ISS	B?M	135	658	5.42E-05	81.37778	38
TCATCA	3ISS	BLM	142	1044	5.43E-05	81.36765	38
GAATAT	3ISS	BL?	107	702	8.62E-05	79.75	38
TCCTGC	5ISE	B?M	134	663	1.06E-04	79.00752	38
CTGGAG	3ISS	BLM	245	1044	1.11E-04	78.86885	38
AGACAC	5ISE	?LM	90	639	1.51E-04	77.75	38
GGGGGA	3ISE	B?M	117	682	1.78E-04	77.16814	38
GCAGCA	5ISE	BLM	163	964	1.87E-04	76.9816	38
CACTTA	5ISE	B?M	86	663	2.23E-04	76.34884	38
GGCCCT	3ISS	?LM	99	728	2.25E-04	76.30612	38
TAGAAC	3ISS	B?M	96	658	2.29E-04	76.25	38
TTGTAG	3ISE	B?M	118	682	2.94E-04	75.33333	38
TCCTCA	5ISS	B?M	135	653	2.95E-04	75.31298	38
GGAGGG	3ISE	BL?	189	643	3.28E-04	74.92064	38
GCAGAA	ESE	BL?	25	320	3.47E-04	30.6	9
CCTTCT	3ISS	BL?	163	702	4.21E-04	74	38
TGATTG	5ISE	BL?	89	626	5.04E-04	73.31818	38
CACTTG	5ISE	?LM	109	639	5.32E-04	73.11927	38
CTAGAA	5ISS	BLM	177	1030	5.37E-04	73.08475	38
CTTTCT	5ISS	BL?	223	686	6.37E-04	72.43243	38
ATTACT	3ISE	B?M	100	682	7.93E-04	71.6	38
TATAAA	5ISE	?LM	150	639	8.70E-04	71.24324	38
TGAGAA	5ISS	B?M	151	653	1.03E-03	70.60403	38
CATGGG	3ISS	?LM	161	728	1.10E-03	70.3354	38
GAGATT	5ISE	BL?	89	626	1.51E-03	69.09302	38
TAGAAG	3ISS	?LM	100	728	1.78E-03	68.42424	38
AAAAGC	3ISS	B?M	108	658	1.88E-03	68.22222	38
GATGGC	ESE	?LM	21	306	1.94E-03	26.14286	9
CCCACC	5ISS	BL?	123	686	2.21E-03	67.56098	38
GTAAAA	3ISE	B?M	109	682	2.30E-03	67.3945	38
ACTTAG	5ISE	B?M	95	663	2.99E-03	66.33684	38
AGCCAT	3ISS	B?M	102	658	3.13E-03	66.14	38
GAAAC	5ISS	?LM	100	721	3.13E-03	66.14	38
ACTCAA	3ISS	?LM	85	728	3.28E-03	65.95294	38

ATTTTG	3ISS	BL?	207	702	3.47E-03	65.71707	38
GCATGG	3ISS	BL?	136	702	3.52E-03	65.65672	38
CTTCCT	5ISS	?LM	290	721	4.32E-03	64.80282	38
GACTCA	ESS	B?M	26	294	5.70E-03	23.23077	9
CATTTT	5ISS	BL?	256	686	6.19E-03	63.27059	38
TCAGCT	5ISE	?LM	102	639	6.55E-03	63.02	38
AAGAAG	ESE	BLM	45	478	6.72E-03	22.77778	9
CTAAAG	5ISE	BLM	131	964	6.88E-03	62.80916	38
CTTCTC	3ISS	BL?	169	702	8.49E-03	61.89157	38

Supplementary Table 3: List of clustering result for 6 types of SREs (ESE, ESS, upstream ISE, upstream ISS, downstream ISE and downstream ISS). Three columns give cluster ID, cluster annotation, and the SREs in a cluster and their original annotations, respectively.

ESE clusters

1	?LM	AGCGAG ?LM GAGCGA ?LM
2	?LM	CAGACT ?LM
3	BLM	AGGAAC BLM
4	BLM	AGCTGC BLM GAGCTG ?LM
5	BLM	CAATGA B?M CCAATG ?LM
6	B?-	CCCTTA B?-
7	--M	AGAGCG -?M GAGCGG ?-M
8	?LM	CTCGGT ?LM
9	BL-	CAGCAA BL? CGGCCA BL? CCGCAA ?L-
10	BL?	GGATTT BL? GGATGA BL?
11	B-M	CCTGCG B?M GCTGCG ?-M
12	?-M	CACGTG ?-M
13	?L-	GTGTCT ?L-
14	--M	GGAAGG --M
15	?L-	TATAGG ?L-
16	-L?	ATGTAC -L?
17	B--	CTGAAA B--
18	BLM	GTGGCT BLM TGTGGC ?LM
19	-LM	ATGGCT -?M GATGGC ?LM
20	B?-	TACCTC B?-
21	B--	CGCTGC B--
22	BL?	ATCTAT BL?
23	BLM	TATGAC BLM ATATGC ?LM
24	B--	GCATTC B--
25	BLM	ACTTCG BL? TTCGAG B?M CTTCGA B?M

26	-L?	AAATCA -L?
27	B--	GGGGAC B-? GGGGCC B--
28	BLM	TGTTAG BLM GGTCAG BL?
29	-LM	GCAACA -?M TGCAAC ?LM
30	B-M	TCTTGC B-? TTTTGG B?M
31	BLM	CAACCA B?M AACGAG ?LM CAACGA ?LM
32	BLM	TCGACT BL? CGACTG BL? GACTAT BLM CGAGTA ?LM
33	BLM	TGGAGC BLM GGAGCA BLM CGGAGA B?M GGAGCG B?M
34	BLM	AAGAAG BLM AAGCAG BLM AGAAGC BLM AGCAGC BL? GCAGAA BL? AAGGAG ?LM AGAAGA ?LM AGCAGA ?LM
35	BLM	GCCAAC BL? GCTAAT B?M
36	BL-	CTCATA BL? CTCAAT ?L-
37	--M	CGGGCT -?M GGCGGG --M GGGCTG ?-M
38	BLM	ACGGCA BLM CACGGC BLM CGGCAC BLM

39	BL-	CCGACG B?- ACCGCC ?L-
40	BLM	GTGATA BLM TCGATA ?LM CGATAT ?LM
41	B-M	GGTACC B?M GATGCC ?-M
42	B-M	GGTGA B?M GAGGAG ?-M
43	?L-	GGCGTA ?L-
44	BL-	TGGCA BL? TTGGGA ?L-
45	B--	ATGAGC B-- GATGAG B?-
46	BLM	TGGTGT BL? CTGGTG BL? AGGGTG B?M
47	-LM	TCAGGC -L? CAGGCA -?M TCGGGC ?LM
48	?L-	ACGTTC ?L- CGCTCC ?L-
49	B--	ACGCGC B-- GCGCGC B-- CAGCGC B?-
50	BL-	AACTGC BL? AAACTG ?L-
51	B--	AGCCTG B-- ATCCCC B?-
52	-LM	TGGAGA -?M GCAGAT ?LM GGAGAT ?LM

ESS clusters

1	BLM	TTTAAA BLM TTTTAA BL? TTTGAA BL? TTGAAA BL?
2	BLM	AAATAA BLM AATTAT BLM GTATAA BL? AATAAT B?M TAATAA B?M TTATAA ?LM
3	-LM	ACCTTT -L? ACCTCT ?LM
4	BLM	ACTCAC BL? GACTCA B?M
5	B?M	CGCTGA B?M
6	B-?	AGAGTG B-?
7	BLM	GGTCT B?M GGTCTT ?LM
8	B?M	AACACT B?M
9	BL?	AGATAT BL?
10	?LM	CAGACC ?LM
11	-?M	AAGCCT -?M
12	?-M	GGTTG ?-M
13	-L?	ACATCC -L?
14	BLM	ACCGTC BLM
15	--M	CCGCCC -?M CCGCGC -?M CCGCTC ?-M
16	BL?	TGTAGC BL?
17	BL?	AGGGGA BL?
18	B-M	ATAATA B-M ATAAGA B?M
19	--M	AATTCC --M
20	BLM	CCATGT BL? CATTG B?M CCCATT B?M

21	BL-	TTCTTT BL? GTTCTG B?- ATCTTT ?L-
22	B--	ATGAAT B--
23	BLM	TTAAAG BLM CTTAAG B?M GTAAAG ?LM
24	B?M	TATTTA B?M GATTAA B?M
25	?LM	CACTCT ?LM CTCTGT ?LM
26	BLM	GAACTC BL? TAACTT ?LM
27	BLM	TCTTAG BLM TGCTTA BL?
28	B-M	AATGTT B-? ATATTG B?M
29	BLM	GTGTTT BL? TTGTTT B?M
30	BLM	ACAATA BL? ATACAA ?LM
31	BLM	TAAGAT BLM TAGGCT BL?
32	--M	TCCTTG --M CCCTGC -?M
33	BLM	TGCCTG BLM CTGCCT BLM CCTGCC BLM GCCTGC BLM CCTCCC ?LM
34	?L-	CCCCAG ?L- CGGCAG ?L-

5' ISE (upstream intronic enhancer) clusters

1	BLM	CTGGG	BLM
2	BLM	GCAGCA	BLM
3	BL?	TATGCA	BL?
4	-L?	CCGAAC	-L?
5	-L?	TTGTAG	-L?
6	BLM	TAACGT	BLM
7	?L-	AAGTTT	?L-
8	-?M	AGGCCG	-?M
9	B-?	TTGGTG	B-?
10	B?M	AAGTGA	B?M
11	BLM	TTGTAC	BLM
12	--M	AGCTGA GCTGAC	-?M --M
13	BL-	CACATA CAAATG	BL- ?L-
14	B--	AACTCA	B--
15	B-?	TAGTCG	B-?
16	BL?	GGAGTC	BL?
17	BL?	CGACAG	BL?
18	B--	CATGCA	B--
19	-LM	CTATAA ATAAAT TATAAA	-L? -?M ?LM
20	-L?	AGGAGA	-L?
21	B-M	GGTGCC GCTCCC	B-? B?M
22	-LM	GACACA AGACAC TGTAC	-L? ?LM ?LM
23	B-M	CTGGGC CAGGGG	B-? B?M
24	-L?	GATGGT	-L?
25	--M	TCCCCT	--M
26	--M	TAATAA TAATAT	?-M -?M

27	-L-	ACACTG TACACT	-L? -L-
28	BLM	ACTTAG CACTTA CACTTG	B?M B?M ?LM
29	BLM	AGCTTA TCAGCT	B?M ?LM
30	BLM	TGTGCG GTGTGC TGCGTG GTGCGT	BL? BL? ?LM ?LM
31	-LM	ACCGGA CCAGAG	-?M ?LM
32	-LM	ACTTGA CTTGAA	-L? ?LM
33	BLM	CTAAAG CCTAAA	BLM BLM
34	BL-	AGATAT GAGATT TGAGAT	BL? BL? B?-
35	?-M	CAGCCA GCAGCC GGCAGC	?-M ?-M ?-M
36	B-M	TCCTGA TTCATG TCCTGC	B-? B?M B?M
37	BLM	TGATTG GTGATT TGATTT	BL? B?M ?LM
38	B--	CCTACT	B--
39	B--	ATATGT ATGTGT	B-- B?-
40	-LM	CGCATG GCACGC AGCAAG	-L? -L? -?M

5' ISS (upstream intronic silencer) clusters

1	-LM	TGTATG -L? GTATAT -?M CTATAT ?LM GTGTAT ?LM
2	B-M	CATCCG B-M
3	-?M	AATGTC -?M
4	B?-	GTTCCG B?-
5	?LM	TCTGCA ?LM TCTGCG ?LM
6	B?M	ATATGA B?M
7	-LM	CTTCTT -L? CTTCCT ?LM
8	BL-	CCCACC BL? CCGACC ?L- CGACCC ?L-
9	B?M	CAGGAC B?M
10	BLM	TTACAA BL? TGAGAA B?M
11	-LM	CCCCCA -L? CTCCCG ?LM
12	B?-	ATCGTT B?-
13	-L?	TAAGGC -L?
14	BLM	GCCGAT BLM TGCCGA B?M
15	B-?	ACTATT B-?
16	-LM	GGATAC -L? GGAAAC ?LM
17	B-?	CTTACA B-?
18	BLM	CATTTT BL? ATTTAG B?M
19	BL?	GGGAAA BL?
20	B?-	CATCAC B?-
21	-L?	GAGAAG -L?
22	B?-	ACAGGC B?- CAGGCT B?-
23	BLM	AAGATT BLM
24	BL-	CAGCGG B?- CTGCGG ?L-

25	B-M	GGGGCC B?M GGGTCT ?-M GGGGCG B?M
26	B--	CAAAAA B-- CAAGGA B?-
27	-L-	ATTCAT -L- ATCCCT -L?
28	--M	ATATAC --M ACACAC --M
29	-LM	TATGTG -?M CCTATG ?LM
30	B-M	TCCTCA B?M CATCAA ?-M
31	-LM	CTAATT -?M TCTAAT ?LM
32	B--	GTATCG B-- GGTAGC B?-
33	BLM	CTAGAA BLM CTGCAA BL? CTGGAA BL?
34	BLM	TTATCA BL? GTTATC ?LM
35	-LM	TCTCTC -LM CTCTCT -LM TCTCTT ?LM CTCTTT ?LM
36	BL?	CCCTTC BL? CCCCGC BL?
37	B-M	CAGTCA B-? GTCAGT B-? CAGTCC B?M
38	BL-	CTTCT BL? GGCTTT ?L-
39	BLM	TCCAAT BL? CAATAC ?LM
40	B-?	AACCAG B-? AACCGT B-?
41	-LM	GCATGC -L? ATGCAG -?M

3' ISE (downstream intronic enhancer) clusters

1	BL?	GGATGC	BL?
2	BLM	GTAAAA TTAAAA	B?M ?LM
3	B?-	GGAAGT	B?-
4	B-?	CTTCCT	B-?
5	BL?	CTTGCC	BL?
6	BL?	AATCAT	BL?
7	BLM	ATTGAT TTGATT TTGCTT	BL? BL? BLM
8	BLM	TGAGGG GAGGGA GGAGGG GGGGGA	BL? BL? BL? B?M
9	B--	CCGACC CGACAT	B-? B?-
10	?-M	AGGCT	?-M
11	?L-	CGCGCT	?L-
12	BLM	GAACCG	BLM
13	?LM	CTCTGT	?LM
14	-?M	TTTAGT	-?M
15	B?-	CTAGTA	B?-
16	B-?	TTGAGG	B-?
17	B?-	CAGTTA	B?-
18	-L-	ATGATC TGAACA GATCAG	-L? -L? ?L-
19	B-M	GAGTTG	B-M
20	-L?	GCGGAT	-L?
21	--M	TGCATG	--M
22	B-M	GCGTGC GCGAAC	B?M ?-M
23	-L-	CGACTG AACTGG	-L- ?L-
24	BLM	GGCAGC GGCAGG GGCGGC	?LM B?M B?M

25	BLM	CGCTAA ACGTTA CGTTAT	B?M ?LM ?LM
26	B--	GGAATG GGGATG	B-? B--
27	B-?	CACCCT CACCGT	B-? B-?
28	BL-	CTGCGT TTGCGA	BL? ?L-
29	BL-	GTCGTC TCATCT	BL? ?L-
30	BLM	TTACTG CATTAC ATTACT	BLM BL? B?M
31	BLM	TTGTAG ATGGAG TGTGGA GTGGAG	B?M BLM ?LM ?LM
32	B--	ACGCCC TGCGCC GCGCCT	B-- B-? B--
33	--M	TGTTTT AATGGT	-?M ?-M
34	BL-	AGGGGC CGGTGC	B?- ?L-
35	BLM	CCGAAG CCAAAC	BL? B?M
36	BL?	GCCTTT	BL?
37	--M	GTCTCG	--M
38	BL-	ATCTAT TTCTAT TATCTA TCTATC	BL- B?- ?L- ?L-
39	?-M	TGACAT CTGACA	?-M ?-M
40	-L-	ATTACA GCTTAC	-L- -L?

3' ISS (downstream intronic silencer) clusters

1	BLM	TCTAAA BL?	CTAAAA BLM
2	BL?	ATTTTG BL?	
3	BLM	CTTCTT BL?	CTTCTC BL?
		CCTTCT BL?	CCATCT B?M
4	BLM	CGAGAA BLM	TAGAAC B?M
		TAGAAG ?LM	
5	--M	TCTATC --M	CTATCT --M
		CTAGCC -?M	
6	B?-	ACGATG B?-	
7	-L-	TGAAGG -L-	TGAACG ?L-
8	B?M	TCGTTC B?M	
9	-L?	GCTCGG -L?	
10	?-M	TATATT ?-M	TTTAT ?-M
11	BLM	GAATAT BL?	GCATAA B?M
12	B-M	AAGCCA B?M	AGCCAT B?M
		GCCATG ?-M	
13	BL?	AAACTT BL?	
14	-LM	CCCCGC -LM	
15	-L?	TGGGGA -L?	
16	?L-	TACCTC ?L-	
17	BL?	ACTCGG BL?	
18	BLM	CTGGAG BLM	
19	B-?	GAAGAG B-?	GGAAGA B-?
20	-?M	ATCATC -?M	AGCATT -?M
21	BL-	TCCTTA B?-	TTCATT ?L-
22	--M	GAGGCT --M	
23	BLM	TACGTA BLM	

24	B?-	GGATCC B?-	
25	BLM	CACATG BL?	GCATGG BL?
		CATGGG ?LM	
26	B-M	CTACCC B?M	TTCCCC ?-M
27	BLM	CTTCAA B?M	ACTCAA ?LM
28	BL-	AAGTTT BL?	AGTTTC BL?
		GTTTGG BL?	GTTTCT B?-
29	BLM	TTCTTC BLM	TCATCA BLM
30	BLM	AAAAGC B?M	TAAAAG B?M
		ATAAGT ?LM	
31	--M	ACTATT --M	ACTCTA ?-M
32	?L-	TATGGC ?L-	GCTATG ?L-
33	-L-	TGCATG -L-	GCATGT -L-
		GTGCAT ?L-	
34	-L-	CGCCCC -L-	GCCCCG -L?
		CCCCGA ?L-	
35	B-M	AGATAA B?M	AGCTAG ?-M
		AGGTAG ?-M	
36	?L-	CTCTTC ?L-	
37	?LM	GGCCCT ?LM	
38	-LM	ACGGCT -L?	TCGGGT ?LM
		CCGGAT ?LM	
39	B-?	CACACA B-?	GCACAG B-?
40	--M	GCGCGC --M	CYCGCC -?M