

Table S3: repeat families enriched in mouse cryptic exons

chr	start	end	geneID	exonID	strand	repeat family	repeat type
chr1	21229021	21229273	<i>Tmem14a</i>	E005i1	+	antisense B2 SINE	B3A
	24022406	24022770	<i>Fam135a</i>	E008i1	-	antisense B2 SINE	B2 Mm1t
	24022406	24022770	<i>Fam135a</i>	E008i1	-	antisense B2 SINE	B3
	24022406	24022770	<i>Fam135a</i>	E008i1	-	antisense B2 SINE	B3
	193368766	193369053	<i>Camk1g</i>	E017i2	-	simple repeat	(CA)n
	24022406	24022770	<i>Fam135a</i>	E008i1	-	simple repeat	(CA)n
chr2	58049385	58049662	<i>Ernm</i>	E001i1	-	antisense B2 SINE	B3
	24691859	24692466	<i>Cacna1b</i>	E050i1	-	low complexity	AT rich
	58049385	58049662	<i>Ernm</i>	E001i1	-	low complexity	AT rich
	165183138	165183472	<i>Cdh22</i>	E013i2	-	simple repeat	(TTCA)n
	24691859	24692466	<i>Cacna1b</i>	E050i1	-	simple repeat	(CA)n
	24691859	24692466	<i>Cacna1b</i>	E050i1	-	simple repeat	(TTTA)n
chr3	152292456	152292758	<i>Fam73a</i>	E015i2	-	antisense B2 SINE	B3
	152643769	152644016	<i>Ak5</i>	E014i6	-	simple repeat	(TGGGGG)n
chr5	52434150	52434525	<i>Ccdc149</i>	E009i3	-	low complexity	GA-rich
	122556832	122557092	<i>Ift81</i>	E003i1	-	low complexity	T-rich
	122556832	122557092	<i>Ift81</i>	E003i1	-	simple repeat	(CA)n
	52434150	52434525	<i>Ccdc149</i>	E009i3	-	simple repeat	(TG)n
chr6	119369759	119370019	<i>Adipor2</i>	E010i2	-	antisense B2 SINE	B3
	120204163	120204566	<i>B4galnt3</i>	E002i1	-	simple repeat	(CA)n
chr7	34110974	34111200	<i>Wtip</i>	E001i1	-	antisense Alu SINE	B1 Mus1
	34110974	34111200	<i>Wtip</i>	E001i1	-	antisense B2 SINE	B2 Mm2
	81490078	81490378	<i>Ap3b2</i>	E031i2	-	simple repeat	(CAA)n
	81490078	81490378	<i>Ap3b2</i>	E031i2	-	simple repeat	(CA)n
chr9	109837772	109838325	<i>Nme6</i>	E020i1	+	antisense B4 SINE	RSINE1
	78207096	78207339	<i>Gsta4</i>	E005i2	+	antisense B2 SINE	B3
	60637548	60637770	<i>Lrrc49</i>	E017i3	-	simple repeat	(CA)n
chr10	45639667	45639904	<i>Hace1</i>	E013i1	+	antisense B2 SINE	B3
	67013987	67014265	<i>Reep3</i>	E001i1	-	antisense ID SINE	ID4
	71263484	71263818	<i>Ube2d1</i>	E010i1	-	antisense B2 SINE	B2 Mm2
	71263484	71263818	<i>Ube2d1</i>	E010i1	-	antisense B2 SINE	B3A
	123150861	123151225	<i>Usp15</i>	E015i3	-	simple repeat	(CA)n
	67013987	67014265	<i>Reep3</i>	E001i1	-	simple repeat	(CCAA)n
	71263484	71263818	<i>Ube2d1</i>	E010i1	-	simple repeat	(AGGGGG)n
	71263484	71263818	<i>Ube2d1</i>	E010i1	-	simple repeat	(TAAA)n
chr11	29421254	29421502	<i>Ccdc88a</i>	E004i8	+	antisense Alu SINE	B1F1
	29421254	29421502	<i>Ccdc88a</i>	E004i8	+	antisense Alu SINE	B1 Mus1
	76468424	76468678	<i>Abr</i>	E026i1	-	simple repeat	(CA)n
	76468424	76468678	<i>Abr</i>	E026i1	-	simple repeat	(GGGA)n
chr12	55864889	55865192	<i>Brms1l</i>	E008i1	+	antisense B4 SINE	RSINE1
	81509727	81510152	<i>Synj2bp+Gm20498</i>	E036i2	-	antisense B2 SINE	B3A
	81509727	81510152	<i>Synj2bp+Gm20498</i>	E036i2	-	low complexity	AT rich
	81509727	81510152	<i>Synj2bp+Gm20498</i>	E036i2	-	simple repeat	(CA)n
chr13	103862427	103862688	<i>Erb2ip</i>	E026i1	-	antisense B2 SINE	B3
chr14	13957359	13957732	<i>Thoc7</i>	E007i1	-	antisense B2 SINE	B3A
	13957359	13957732	<i>Thoc7</i>	E007i1	-	antisense B4 SINE	B4A
chr15	12833121	12833490	<i>Drosha</i>	E009i1	+	antisense B4 SINE	B4
chr16	45389870	45390285	<i>Cd200</i>	E003i1	-	simple repeat	(CA)n
chr17	34739289	34739599	<i>C4b</i>	E032i1	-	antisense B4 SINE	B4A
	34739289	34739599	<i>C4b</i>	E032i1	-	simple repeat	(CA)n
chr18	6986038	6986336	<i>Mkx</i>	E007i7	-	simple repeat	(CA)n
	80138052	80138405	<i>Adnp2</i>	E002i1	-	simple repeat	(CA)n

Figure S5:  
The nucleotide sequence of B2 SINE in the antisense orientation

AAAGAUUUUAUUUAUUUAUUUAUUUAUUAUGUAAGUACACUGUAGC  
UGUCUUCAGACACUCCAGAAGAGGGAAUCAGAUCUCGUU  
ACGGAUGGUUGUGAGCCACCAUGUGGUUGCUGGGAUUU  
GAACUCCUGACCUUCGGAAGAGCAGUCGGGUGCUCUUA  
CCCACUGAGCCAUCUCACCAGCC

Table S4: repeat families enriched in human cryptic exons

chr	start	end	geneID	exonID	strand	repeat type	repeat class
chr1	19145498	19145779	<i>UBR4</i>	E084i1	-	simple repeat	(AC) <sub>n</sub>
chr2	111799429	111799800	<i>ANAPC1</i>	E020i1	-	simple repeat	(AATCCAA) <sub>n</sub>
	241668774	241669086	<i>ATG4B</i>	E047i1	+	simple repeat	(GT) <sub>n</sub>
	43942575	43942904	<i>LRPPRC</i>	E024i1	-	simple repeat	(ATTTT) <sub>n</sub>
chr3	121682816	121683064	<i>GOLGB1</i>	E013i1	-	simple repeat	(T) <sub>n</sub>
chr4	102324149	102325099	<i>SLC39A8</i>	E016i2	-	simple repeat	(ATAC) <sub>n</sub>
chr7	783125	783410	<i>DNAAF5</i>	E024i1	+	simple repeat	(CGTGAG) <sub>n</sub>
chr10	27098079	27098427	<i>ANKRD26</i>	E045i1	-	simple repeat	(TTTAATT) <sub>n</sub>
	3099456	3099920	<i>PFKP</i>	E010i1	+	simple repeat	(GT) <sub>n</sub>
chr11	58616892	58617155	<i>ZFP91+ZFP91-CNTF</i>	E012i1	+	simple repeat	(TG) <sub>n</sub>
chr12	116789933	116790188	<i>RNFT2</i>	E016i2	+	simple repeat	(TGGA) <sub>n</sub>
	88086609	88086890	<i>CEP290</i>	E024i1	-	simple repeat	(ATCC) <sub>n</sub>
chr14	53096016	53096319	<i>DDHD1</i>	E023i1	-	simple repeat	(AGGG) <sub>n</sub>
chr16	4367197	4367543	<i>CORO7-PAM16+CORO7</i>	E081i1	-	simple repeat	(GAGGG) <sub>n</sub>
chr19	4491914	4492253	<i>HDGFRP2</i>	E015i1	+	simple repeat	(GT) <sub>n</sub>
chr21	43792282	43792675	<i>RRP1</i>	E008i1	+	simple repeat	(GTGA) <sub>n</sub>