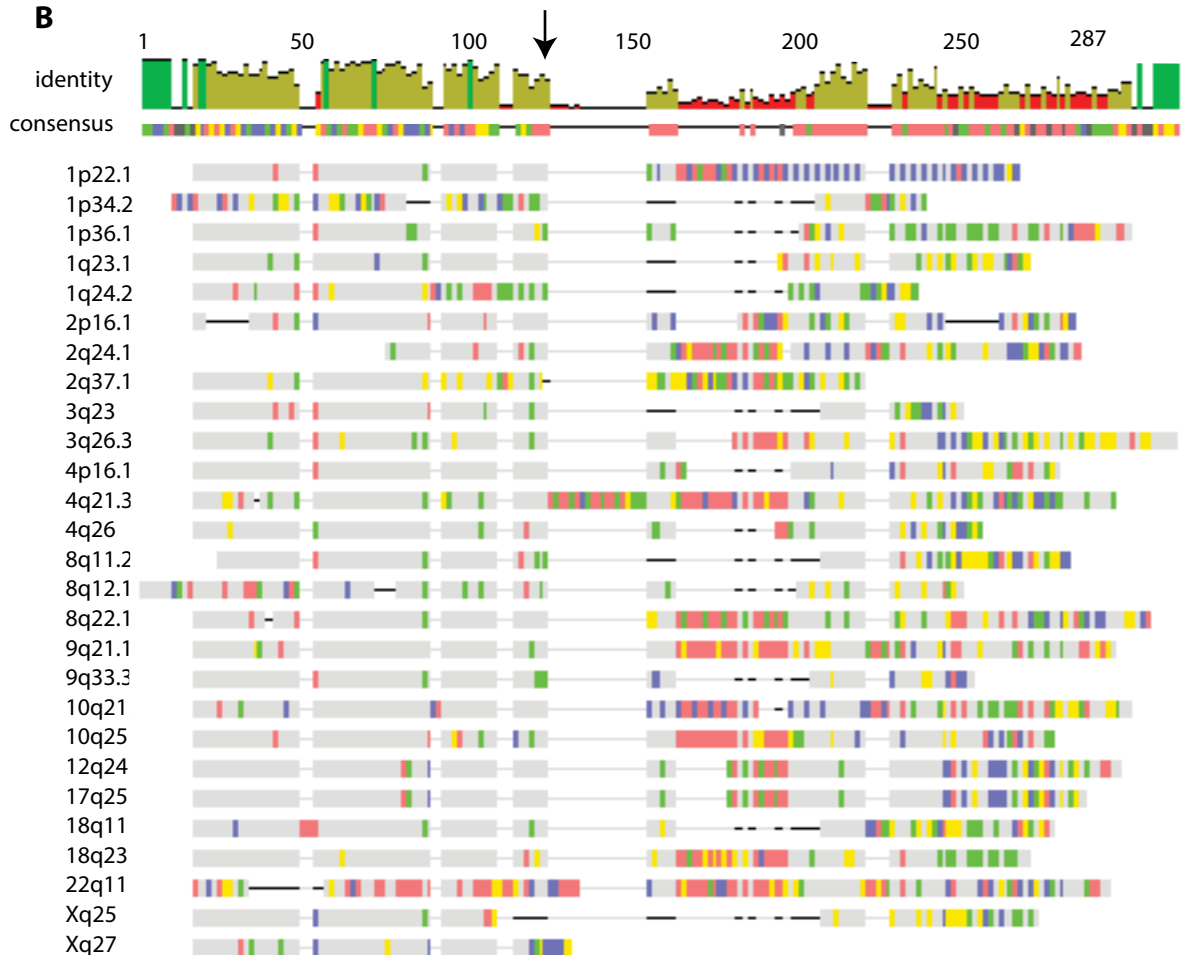
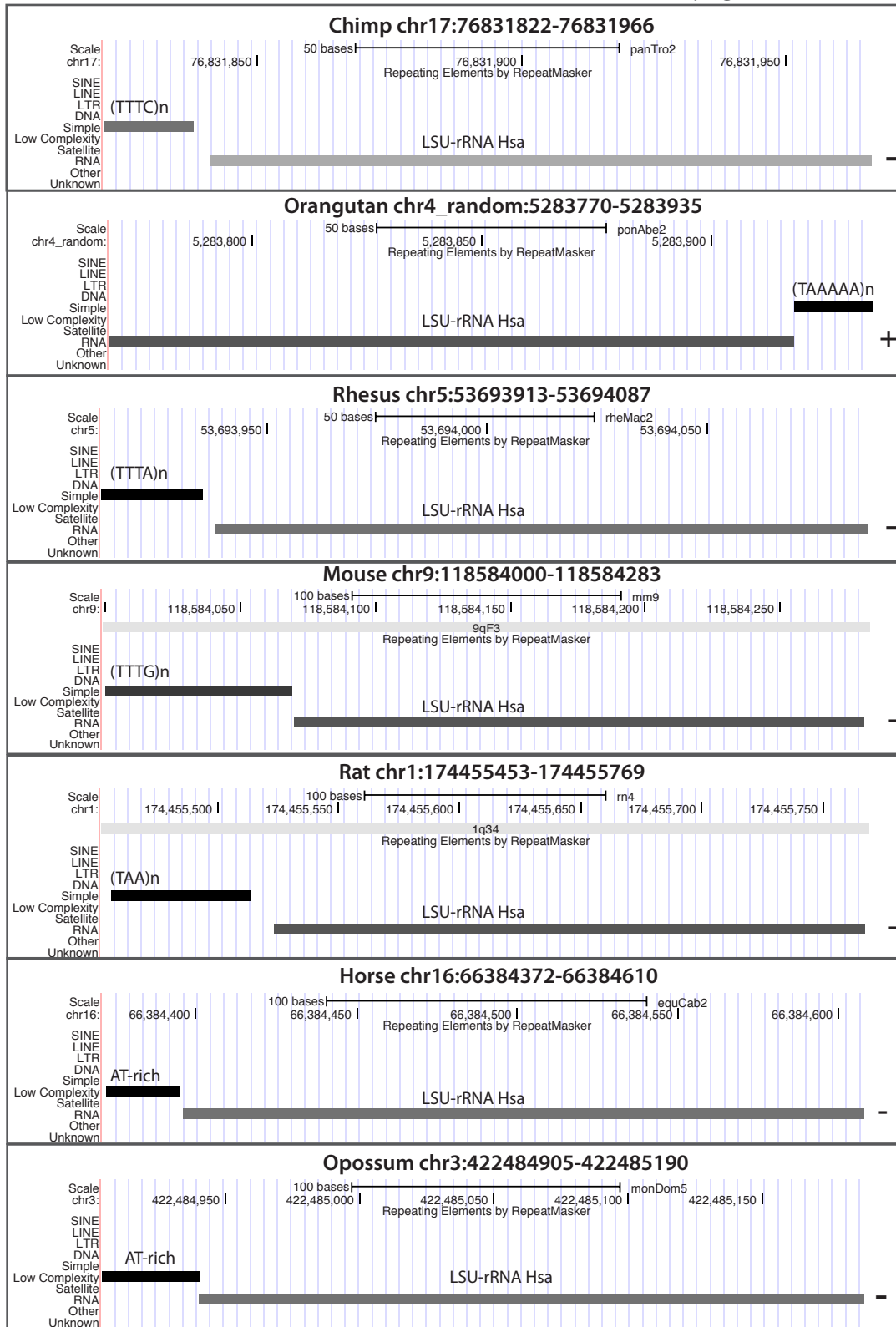


A



B





Longo_Supplemental Table 1. BLAT (Kent 2002) results from aligning the Repbase (Jurka, et al. 2005) LSU consensus sequence to the human genome (hg19). Most SINE28s have adjacent simple repeat tails (red), while very few lack tails (orange). The segmental duplications of LSU/Alu clusters (sd28alu) do not typically have tails (green) (with one exception). The majority of other fragments do not carry a simple repeat tail (white) and only a minority of non-3' fragments carry tails (yellow).

KEY	SINE28 with tail
	SINE28 without tail
	sd28alu
	Other LSU fragment with tail
	Other LSU fragment without tail
	LSU fragment too large for typical SINE

Tail?	strand	Q size	Q start	Q end	Hit location	start position	end position
n	+	5035	834	872	chr1	226736634	226736677
n	+	5035	1447	1688	chr1	145277300	145277489
n	-	5035	1447	1688	chr1	120543889	120544074
n	-	5035	2022	2792	chr1	108113078	108113729
C-rich	-	5035	3600	3705	chr1	28242609	28242680
n	+	5035	3637	4115	chr1	237766308	237766761
n	-	5035	4132	4495	chr1	91852782	91853148
n	+	5035	4192	4283	chr1	1815107	1815204
AT-rich	-	5035	4803	5025	chr1	169878693	169878875
A-rich/T-rich	+	5035	4819	4991	chr1	42779320	42779416
AT-rich	-	5035	4922	5034	chr1	25483508	25483621
A-rich	+	5035	4960	5034	chr1	157139814	157139887
A-rich	-	5035	4960	5034	chr1	92972727	92972785
n	-	5035	1	385	chr10	32699237	32699591
n	-	5035	3636	4026	chr10	50911089	50911384
n	-	5035	3663	3956	chr10	61447835	61448118
(CAA)n/AT-rich	-	5035	4711	5034	chr10	68805210	68805481
T-rich	-	5035	4871	5034	chr10	115463429	115463530
n	+	5035	254	576	chr11	85195011	85195304
n	+	5035	3635	3962	chr11	77597473	77597801
n	-	5035	3666	4306	chr11	111691917	111692715
n	+	5035	4495	5035	chr12	38531461	38531863
n	-	5035	4501	5028	chr12	34372315	34372728
n	-	5035	4508	4572	chr12	85562703	85562757
(TTTTTA)n	-	5035	4611	5034	chr12	127650452	127650987
n	-	5035	834	875	chr14	93185602	93185643
n	-	5035	1976	2115	chr14	23987017	23987089
n	+	5035	1303	1977	chr16	47538680	47539269
T-rich	-	5035	3639	3967	chr16	81902149	81902314
n	-	5035	1	158	chr17	33478112	33478275
n	+	5035	2257	2887	chr17	19355607	19356096
(T)n	-	5035	2400	2533	chr17	55868624	55868752
T-rich	-	5035	4616	5034	chr17	75158031	75158387
n	-	5035	2570	2656	chr18	9296921	9297002

A-rich	+	5035	4819	5033	chr18	77772846	77773065
A-rich	+	5035	4884	5034	chr18	21191694	21191827
n	+	5035	1734	1792	chr19	36066690	36066743
n	+	5035	1911	2080	chr19	36066505	36066673
n	-	5035	2224	4368	chr19	24182398	24186182
(TTTA)n	-	5035	4253	5024	chr19	42070022	42070800
n	+	5035	4480	5035	chr19	43910344	43912093
n	-	5035	254	513	chr2	115695095	115695274
acro1	-	5035	1736	4968	chr2	133036257	133039236
n	+	5035	1955	2025	chr2	230015773	230015823
n	-	5035	2396	2546	chr2	230045487	230045630
n	-	5035	2609	2712	chr2	230045630	230045740
n	+	5035	3866	4125	chr2	78928721	78928981
n	-	5035	3909	4307	chr2	63435936	63436042
n	+	5035	4039	4092	chr2	117770019	117770073
n	+	5035	4164	4314	chr2	223100349	223100443
T-rich	-	5035	4201	4309	chr2	203210991	203211097
n	-	5035	4500	4638	chr2	191402649	191402786
A_rich	-	5035	4596	5034	chr2	159440393	159440797
A_rich	-	5035	4805	5034	chr2	59830730	59830956
(TTAGGG)n	+	5035	4909	5027	chr2	232110115	232110175
n	-	5035	2576	2653	chr20	46687595	46687667
n	-	5035	4591	4698	chr20	45112785	45112876
n	-	5035	4501	5028	chr21	15456972	15457490
n	+	5035	4	137	chr22	21265394	21265503
A_rich	+	5035	4823	5034	chr22	22210671	22210856
n	+	5035	1950	2030	chr3	119248242	119248323
AT-rich	+	5035	2561	2836	chr3	108250257	108250341
(T)n	+	5035	4858	5034	chr3	179497378	179497506
(T)n	-	5035	4946	5034	chr3	142384658	142384747
n	-	5035	1956	2126	chr4	70296618	70296753
n	+	5035	2326	2462	chr4	167063724	167063839
AT_rich	+	5035	4162	4594	chr4	134710172	134710410
A-rich	+	5035	4501	5033	chr4	76807186	76807697
n	-	5035	4512	4725	chr4	107924734	107924904
AT-rich	+	5035	4857	5034	chr4	7584195	7584364
AT_rich	+	5035	4868	5033	chr4	120158524	120158690
T-rich	-	5035	4879	5034	chr4	86968257	86968396
n	+	5035	9	409	chr5	122990601	122991006
n	+	5035	2560	2757	chr5	71146741	71146942
n	-	5035	4040	4226	chr5	65424940	65425073
A-rich	+	5035	4171	4411	chr5	154788122	154788302
n	-	5035	4535	4693	chr5	65425190	65425223
n	+	5035	1707	1925	chr6	133593948	133594107
n	+	5035	2362	2467	chr6	120583444	120583537
n	-	5035	2596	2653	chr6	70183052	70183166
n	+	5035	3636	3900	chr6	134995824	134996048
n	-	5035	3870	4281	chr6	144692022	144692284
n	-	5035	2576	2653	chr7	22130020	22130092
n	+	5035	3625	4042	chr7	68527377	68527788

n	+	5035	4588	4631	chr7	48672360	48672403
n	+	5035	4597	4876	chr7	104966297	104966455
n	-	5035	145	211	chr8	136786015	136786081
n	-	5035	1365	1719	chr8	69218468	69218676
n	-	5035	1474	1791	chr8	56755032	56755322
n	-	5035	1951	2021	chr8	109304591	109304662
n	-	5035	3637	4011	chr8	70602252	70602616
n	+	5035	4037	4362	chr8	21455974	21456316
n	-	5035	4501	5034	chr8	46951586	46952097
AT-rich	-	5035	4614	5033	chr8	56821700	56822036
(TAAA)n	+	5035	4816	5034	chr8	98577765	98577938
A-rich	-	5035	4958	5034	chr8	49375958	49376029
n	+	5035	4185	4343	chr9	32069822	32069979
A-rich	+	5035	4695	5034	chr9	79186653	79186950
A-rich	-	5035	4959	5034	chr9	126813737	126813809
n	-	5035	25	379	chrX	47749438	47749802
n	-	5035	838	875	chrX	74966904	74966941
n	+	5035	3636	4116	chrX	108297348	108297826
n	+	5035	4545	4632	chrX	88735818	88735906
n	+	5035	4821	5035	chrX	139858400	139858607
A-rich	+	5035	4963	5028	chrX	121106711	121106776
n	+	5035	4501	5028	chrY	20487333	20487859
n	-	5035	4501	5028	chrY	19691913	19692439
n	+	5035	4545	4635	chrY	3005050	3005138

Longo_Supplemental Table 2. Genomes screened for SINE28. Genomes where SINE28 was found are denoted in red.

Species	Build	Date
Cat	felCat4	Dec.2008
Chimp	panT	Oct. 2010
Cow	bosTau6	Nov. 2009
Dog	canFam2	May 2005
Elephant	loxAfr3	July 2009
Gibbon	nomLeu1	Jan. 2010
Guinea Pig	cavPor3	Feb. 2008
Horse	equCab2	Sept. 2007
Human	hg19	Feb. 2009
Marmoset	calJac3	Mar. 2009
Mouse	mm9	July 2007
Opossum	monDom5	Oct. 2006
Orangutan	ponAbe2	July 2007
Panda	ailMel2	Dec. 2009
Pig	susScr2	Nov. 2009
Platypus	ornAna1	Mar. 2007
Rabbit	oryCun2	Apr. 2009
Rat	rn4	Nov. 2004
Rhesus	rheMac2	Jan. 2006
Sheep	oviAri1	Feb. 2010
<i>A. gambia</i>	anoGam1	Feb. 2003
<i>A. melifera</i>	apiMel2	Jan. 2005
<i>C. brenneri</i>	caePb2	Feb. 2008
<i>C. briggsae</i>	cb3	Jan. 2007
<i>C. elegans</i>	ce6	May 2008
<i>C. instestinalis</i>	di2	Mar. 2005
<i>C. japonica</i>	caeJaap1	Mar. 2008

Species	Build	Date
<i>C. remanei</i>	caeRem3	May 2007
Chicken	galGal3	May 2006
<i>D. ananassae</i>	droAna2	Aug. 2005
<i>D. erecta</i>	droEre1	Aug. 2005
<i>D. grimshaw</i>	droGri1	Aug. 2005
<i>D. melanogaster</i>	dm3	Apr. 2006
<i>D. mojavensis</i>	droMoj2	Aug. 2005
<i>D. persimilis</i>	droPer1	Oct. 2005
<i>D. pseudoobscura</i>	dp3	Nov. 2004
<i>D. sechelia</i>	droSec1	Oct. 2005
<i>D. simulans</i>	droSim1	Apr. 2005
<i>D. virilis</i>	droVir2	Aug. 2005
<i>D. yakuba</i>	droYak2	Nov. 2005
Fugu	fr2	Oct. 2004
Lamprey	petMar1	Mar.2007
Lancelet	braFlo1	Mar. 2006
Lizard	anoCar2	May 2010
Medaka	oryLat2	Oct. 2005
<i>P. pacificus</i>	priPac1	Feb. 2007
<i>S. cerevisiae</i>	sacCer3	Apr. 2011
<i>S. purpuratus</i>	strtPur2	Sept. 2006
Stickleback	gasAcu1	Feb. 2006
Tetraodon	tetNig2	Mar. 2007
Turkey	melGal1	Dec. 2009
<i>X. tropicalis</i>	xenTro3	Nov. 2009
Zebra finch	taeGut1	July 2008
Zebrafish	danRer7	July 2010

Longo_Supplemental Table 3. RNA Pol III motifs. RNA polymerase III sequence elements used to screen SINE28 sequences using Elph motif finder.

PromoterPiece-Element-Species_AuthorYear	Motif Sequence
Abox-7SL-SceSpo_Willis1993	TRGCNNAGTGG
Abox-CBL3-Hsa_Mrazek2007	AGCTCAAGCGG
Abox-Cp1-Cpa_Rovira1996	TAGCTCAGTGG
Abox-DASla-Dno_Churakov2005	TRGCTCRAGYRG
Abox-ERI-Mda_Borodulina2001	TGGCGCAGTGGA
Abox-OAX-Xtr_Okada1991	AGCTCAGTTGG
Abox-SINE3-Dre_Kapitonov2003	AGCTAAGCAGGGCTG
Abox-SNR52-Sce_GuffantiJune2006	TTGGGCTAGCGGTAAAGG
Abox-SNR52-Sce_Harismendy2003	TAGCGGTAAAGG
Abox-SOR-Sar_Borodulina2001	TAGCACAGTGGG
Abox-TAL-Mro_Borodulina2001	TGGCGCAGTGGT
Abox-TAN-newt_Okada1991	GGTCCAGTGG
Abox-U6-Mmu_Das1988	AGATTAGCATGG
Abox-ZOD1-Sce_GuffantiJuly2006	TGGCGCTTTGG
Abox-consensus_Galli1981	TGGCNNAGTGG
Abox-sRNA85-Tryp_Liu2003	TAGTCTAGCTGG
Abox-tDNAs-Hsa_Oler2010	TRGYBBARBBGD
Abox-tRNA-Ath_Kruszka2003	TRGCNNAGTGG
Abox-tRNA-Consensus_Geiduschek1988	TRGCNNAGY
Abox-tRNAala-Dno_Churakov2005	TAGATCAGTGG
Abox-tRNAglu-Sce_Geiduschek1988	TGTGTAACGG
Abox-tRNAleu-Sce_Geiduschek1988	TGGCCGAGCGG
Abox-tRNAlys-Tryp_Liu2003	TAGCTCAGTCGG
Abox-tRNAmet-Xtr_Geiduschek1988	TGGYRCAGNNR
Abox-tRNAmir-consensus_Smit1995	TGGCNNAGTGG
Abox-tRNAser-Spo_Geiduschek1988	TGTYCGNGTGR
Abox-tRNAtyr-Sce_Geiduschek1988	TRGNCRNGTTGR
Abox-vRNA1-Hsa_Mrazek2007	AGCTCAGCGG
Abox-vRNA2-Hsa_Mrazek2007	AGCTCAGCGG
Abox-vRNA3-Hsa_Mrazek2007	AGCTCAGCGG
Bbox-7SL-SceSpo_Willis1993	GGTTCGANTCC
Bbox-CBL3-Hsa_Mrazek2007	GTTCGAGAC
Bbox-Cp1-Cpa_Rovira1996	GGTTCGATTCC
Bbox-DASla-Dno_Churakov2005	GGTTCRRTYCC
Bbox-ERI-Mda_Borodulina2001	GAGTTCAATCCC
Bbox-OAX-Xtr_Okada1991	GTTCAAATC
Bbox-SNR52-Sce_GuffantiJune2006	CGTTCGAAAC
Bbox-SNR52-Sce_Harismendy2003	GTTCGAAAC
Bbox-SOR-Sar_Borodulina2001	GGGTTTCGATTCC
Bbox-TAL-Mro_Borodulina2001	CGGTTTCGATCCC
Bbox-TAN-newt_Okada1991	GTTCAAATC
Bbox-ZOD1-Sce_GuffantiJuly2006	GGTTCGAACT
Bbox-consensus_Galli1981	GGTTCGANNCC
Bbox-sRNA85-Tryp_Liu2003	GTTCGAGTC
Bbox-snoRNAtrRNA-Dme_Isogai2007	GHTCRWDHB
Bbox-snoRNAtrRNADominantSeq-Dme_Isogai2007	GTTCGANYC
Bbox-tDNAs-Hsa_Oler2010	DGWTTCRADHCBN
Bbox-tRNA-Ath_Kruszka2003	GGTTCGANTCC
Bbox-tRNA-Consensus_Geiduschek1988	GGTTCGANTCC

Bbox-tRNAala-Dno_Churakov2005	GGTTCATCCC
Bbox-tRNAglu-Sce_Geiduschek1988	AGACCGGGGTTCCGACTCCCCG
Bbox-tRNAleu-Sce_Geiduschek1988	ATGCAAGAGTTCGAATCTCTT
Bbox-tRNAmet-Xtr_Geiduschek1988	GGTNGATRGATYGAAAYN
Bbox-tRNAmir-consensus_Smit1995	GGTTCGANNCC
Bbox-tRNApro-Cel_Geiduschek1988	RRTYCCGGGWYCWNTCCC
Bbox-tRNAser-Spo_Geiduschek1988	NNGHSCAGRTWSAWATCC
Bbox-tRNAtyr-Sce_Geiduschek1988	RRTCNNNNGWTBSACNCS
Bbox-vRNA1-Hsa_Mrazek2007	TGTTCGAGAC
Bbox-vRNA2-Hsa_Mrazek2007	GTTCGAGAC
Bbox-vRNA3-Hsa_Mrazek2007	TGTTCGAGAC
Bbox2-tRNAlys-Tryp_Liu2003	GTTCGATC
Cbox-SINE3-Dre_Kapitonov2003	TGGATGGGAGACCACTAG
DSE1-H1-Hsa_Pagano2007	ATTTGCAT
DSE2-H1-Hsa_Pagano2007	ATGCAAAT
ICE-SINE3-Dre_Kapitonov2003	CCTGGT
ICR-5S-Hsa_Gogolevsky2009	AGCTAAGCAGGGTCTGGGCCTGGTTAGTACTTGGATGGGAGA
ICR-AmnSINE1-Hsa_Gogolevsky2009	AGCTAAGCAGGGTCTGGGCCTGGTCAATDCTTGGATGGAAGM
ICR-MEGall-batSINE_Gogolevsky2009	AGCTAAGCAGGGTCTGGGCCTGGTTAGTACTTGGATGGGAGA
ICR-Ped1-Rca_Gogolevsky2009	AGCTAAGCAGGGTCTGGGCCTGGTTAGTACTTGGATGGGAGA
ICR-SINE3-Dre_Gogolevsky2009	AGCTAAGCAGGGCTGAGCCTGGTCAGTACCTGGATGGGAGA
ICR-osSINE1-Ssa_Gogolevsky2009	AGCTAAGCAGGGTTGGTCCTGGTCAGTCCCTGGATGGGAGA
PSE-7SK-Hsa_Lescure1991	TTGACCTAAGTG
PSE-H1-Hsa_Myslinski2001	TCACCATAAACGTGAAAT
PSE-H1-Mmu_Myslinski2001	TCACCCTAAACGTAAAT
PSE-U6-Hsa_Lescure1991	CTTACCGTAACTT
PSE-U6-Hsa_Lobo1989	TACCGTAACTTGAAAGT
PSE-U6-Hsa_Pagano2007	CTTACCGTAACTTGAAAGT
PSE-U6-Mmu_Das1988	CTCACCTAACTGTAAAGTA
PSE-U6-Mmu_Lescure1991	CTCACCTAACTT
PSE-U6-Xtr_Lescure1991	CTCTCCTTAAGTT
PSE1-H1a-Hsa_Pagano2007	nCACCATAAAnGTGAAAn
PSE2-H1b-Hsa_Pagano2007	nTTTCACnTTTATGGTGn