Probe	Number of Hits Genome-Wide (>/=250 bp, >/=85% Sequence	Highest Percent Identity Genome-wide, non-
	Identity)	BRCA1
SINE1	143	89.6
SINE2	3693	89.4
SINE3	121,177	91.3
SINE4	95,322	91.5
SINE5	105,706	92.1
SINE6	178,780	96.1

Table S1. Characteristics of SINE probes used in nuclear run-on and end-labeled smallRNA blot experiments.

Table S2. Primers used in qRT-PCR experiments.

Primer Name	Sequence
BRCA1 Region 1 F	GTGGTGCTTCTGTGGTGAAG
BRCA1 Region 1 R	ACAGGTGCCTCACACATCTG
BRCA1 Region 2 F	GGCTATCCTCTCAGAGTGACATT
BRCA1 Region 2 R	CTGATGTGCTTTGTTCTGGA
PIWIL4 F	TGCTGATGTGAGTTACAAAGTCC
PIWIL4 R	CCTATTAGCTGCTTCTCACACG
Drosha F	GGAATTAGGCACAGCATTTATCC
Drosha R	GAGACTGTGATCCGGTAGTGG
Dicer F	AAATACATGCACTATGTGAAGAGC
Dicer R	TCGCTCATATGGTTTATATTTGC
BAZ2B F	GTGGCTTCAGTAGTTTCAAAAGG
BAZ2B R	GACACTGTGGACAGGTTAAACG
DYNC1H1 F	GACTCTGGATATCTTGAAACATGG
DYNC1H1 R	AGAGGGAAATCTTTCATCAGAGG
SETD1A F	AAGATCGAGAGGAAGCTGTGG
SETD1A R	AGTCTGATGTGCTGTCATTTTCG
USP48 F	TAGTGACTACATGCTGGGAGACG
USP48 R	AATGTATCGCCTATTACTGTTTTGC
β-Actin F	TGGTGATGGAGGAGGTTTAGTAAGT
β-Actin R	AACCAATAAAACCTACTCCTCCCTTAA

 Table S3. Primers used in sequenom experiment.

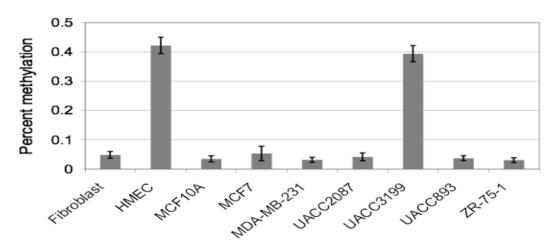
Primer Name	Sequence
BRCA1 Promoter A F	AGGAAGAGAGTAATTGGAAGAGTAGAGGTTAG
	AGGGT
BRCA1 Promoter A R	CAGTAATACGACTCACTATAGGGAGAAGGCTC
	CAAAACAAAAAATAAAAACCTCCT

Primer Name	Sequence
Actin F	TCATGTTTGAGACCTTCAACACC
Actin R	CACAGCTTCTCCTTAATGTCACG
BRCA1 Exon 9 F	AATGTCACCTGAAAGAGAAATGG
BRCA1 Exon 9 R	TGCTAAAAACAGCAGAACTTTCC
BRCA1 Exon 11 F	ATAAAGCTCCAGCAGGAAATGG
BRCA1 Exon 11 R	TTTCTGATGTGCTTTGTTCTGG
BRCA1 Exon 13 F	ATAGGTGGTACATGCACAGTTGC
BRCA1 Exon 13 R	AGTAAGATGTTTCCGTCAAATCG
BRCA1 3' UTR F	GAAGCTGTTGCTTTCTTTGAGG
BRCA1 3' UTR R	GAAGACCTAGTCCTTCCAACAGC
Pseudo F	AGCTCGCTGAGACGTTCTGG
Pseudo R	CCCCTCGCATAAGAATACCC
NBR2 F	TTCACTGTGTTAGCCAGGATGG
NBR2 R	TGAGAAAACAGGCTTAAAAGAGG
AA622657/DB546216 F	TGAGACAGTTTGGCTCATTCC
AA622657/DB546216 R	CTTTGGGATCACTTGAGGTTAGG
BRCA1 Promoter A F	ACGGAAACCAAGGGGCTACC
BRCA1 Promoter A R	ATGCAATAAGCCGCAACTGG
BRCA1 Promoter B F	GAATCCTCGTGATAGGAACTGG
BRCA1 Promoter B R	GGTCCCATCCTCTCATACATACC
14-1 F	CAGAAGAGGCATGATGTTTGG
14-1 R	GCACCTTGACCTATCTGATGTCC
14-2 F	TAATCCAGGAGGATATTGTAGGG
14-2 R	CAATGTTGGTACAAGTTATCTCAGG
14-3 F	AAGTCAAAAATGTTATAGTCATAGG
14-3 R	TGTTTTCTAGCTTTTCATTGC
14-4 F	GGTAATTTTAAATAGGGGTGTCTGG
14-4 R	TTCATTGGAACAGAAAGAAATGG
14-5 F	AAACCAAAGAACTAATGACAACG
14-5 R	GACTGCTTTGGACAATAGGTAGC
14-6 F	CAGCTAATCGTTTTAGTGACAGG
14-6 R	CAAGAACAAAGCTTCCACAGC
14-7 F	TAGCATTTATCTGCTGGTAACAC
14-7 R	ATCGAGGTCATCCTGGCTAACG
14-8 F	TATTTCAGAATATCCGCTTTTGG
14-8 R	AAATGAGCTAAGCGCTACTTCG
BRCA1 SINE 1 F	GCAGACACTGACAATGAAGAGACC
BRCA1 SINE 1 R	ACCCAAGTCTCCTTGACCTTACC
BRCA1 SINE 2 F	TTTTTCTATTTCCAATCCCTTCTGC
BRCA1 SINE 2 R	CTGTTCCCTCCCAGACAACC
BRCA1 SINE 3 F	TTTTTAAAAGATAGGGTCTCAGTCACC
BRCA1 SINE 3 R	TATAGGCCCTTCCTTGCTTTGG
BRCA1 SINE 4 F	CCAAATGCTTAAAAATATCAGTGACC
BRCA1 SINE 4 R	CCAAAGTATGGGCTACAGAAACC
BRCA1 SINE 5 F	GCTAACACTCAGTGATGAGGATGC

Table S4. Primers used to create templates for riboprobes synthesis for dot blot membranes.

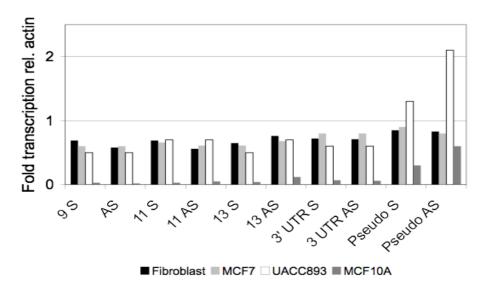
BRCA1 SINE 5 R	TAAGTGCTATGTTCCCAGTCTGC
BRCA1 SINE 6 F	AGGCAGGAGAATCTTTTGAATGC
BRCA1 SINE 6 R	TTACTTGCTTTTGGTCAAATCTCTGC

Figure S1. BRCA1 promoter methylation in all cell lines. BRCA1 promoter methylation levels as analyzed by Sequenom. A total of 18 CpG residues across 317 bp spanning the promoter and first exon of BRCA1 and the first exon of NBR2 were analyzed; average and standard error for each cell line is shown. The analyzed sequence contains the 267 bp region shown to have optimal promoter activity in transient transfection experiments [28].



BRCA1 promoter methylation

Figure S2. BRCA1 transcription rate is reduced in MCF10A cells. BRCA1 transcription in fibroblast cells, non-cancerous mammary epithelial cells MCF10A, and two sporadic breast cancer cell lines MCF7 and UACC893, as analyzed by nuclear run-on. Probes 9, 11, 13, and 3' UTR are from BRCA1 exons and the probes assay transcription solely from the BRCA1 gene; the pseudo probes assay transcription from exon one of both the BRCA1 gene and pseudogene. See Figure 1A for location of probes. S: sense; AS: antisense.



BRCA1 exon transcription

Figure S3. Transcription and small RNA from 14 kb segmental duplication. (A) Diagram of the BRCA1 genomic region and relative location of 14 kb and 24 kb intrachromosomal segmental duplications. Regions are not presented to scale. Location of 14 kb segmental duplication tiling probes and relative location of probes shown in Figure 1A that fall within this region are depicted. (B) Transcription of regions within the 14-kb segmental duplication as analyzed by nuclear run on in control and breast cancer cell lines. S: sense; AS: antisense. (C) Small RNA abundance from 14 kb segmental duplication in control and breast cancer cell lines. Small RNA levels are normalized to miR-16.

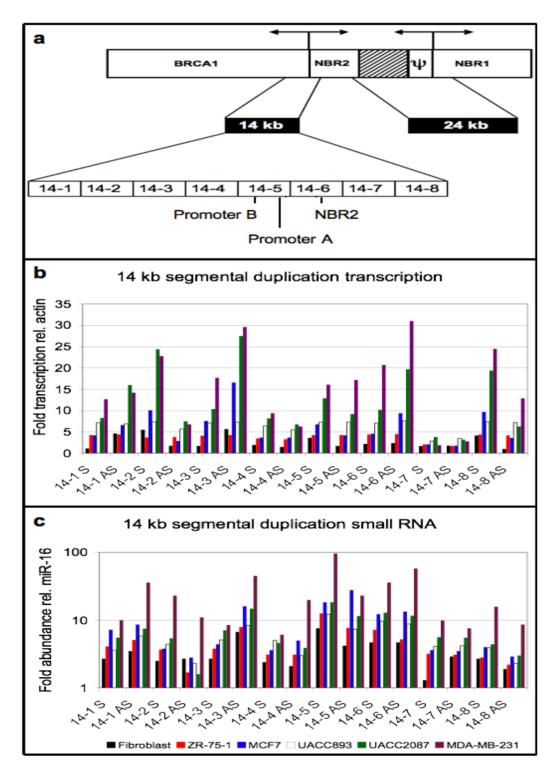
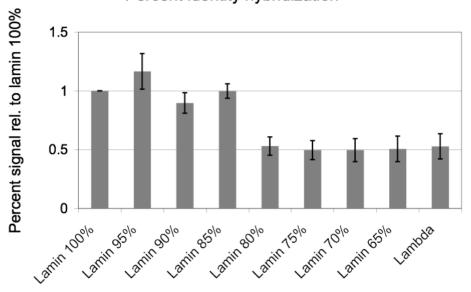


Figure S4. Stringency of hybridization conditions in nuclear run-ons. Oligos with a range of sequence identity to lamin B1 were engineered to contain sequence changes relative to the consensus sequence of lamin B1 and spotted onto nuclear run-on dot blot membranes as described. Six replicates of nuclear run-on experiments were performed, as described, and signal intensity of each probe was quantified by Quantity One Software (BioRad). Data presented is average and standard error of percent signal relative to the 100% identity probe for all six replicates. Percent signal of lambda DNA, present as a negative control, is presented to represent background signal levels.



Percent identity hybridization

Figure S5. Small RNA levels of GAPDH and ubiquitin messages. Small RNA abundance of GAPDH and ubiquitin mRNAs were analyzed by small RNA fractionation followed by radioactive end-labeling and hybridization to dot blot membranes. Small RNA levels are normalized to microRNA 16 (miR-16).

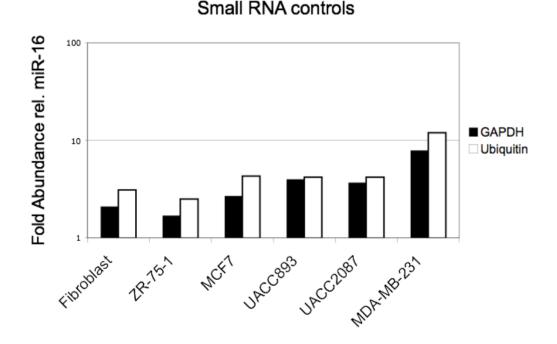


Figure S6. Expression of RNA processing enzymes in non-infected cells and GFP shRNA infected cells. Expression of indicated RNA processing enzymes in non-infected cells and GFP shRNA infected cells for (A) MDA-MB-231 cell line and (B) MCF7 cell line. Expression was analyzed by qRT-PCR and transcript levels for all enzymes were normalized to β -actin. Expression of all genes in the GFP shRNA infected cells was set to 1. Data shown is average and standard error for two technical replicates of each primer pair.

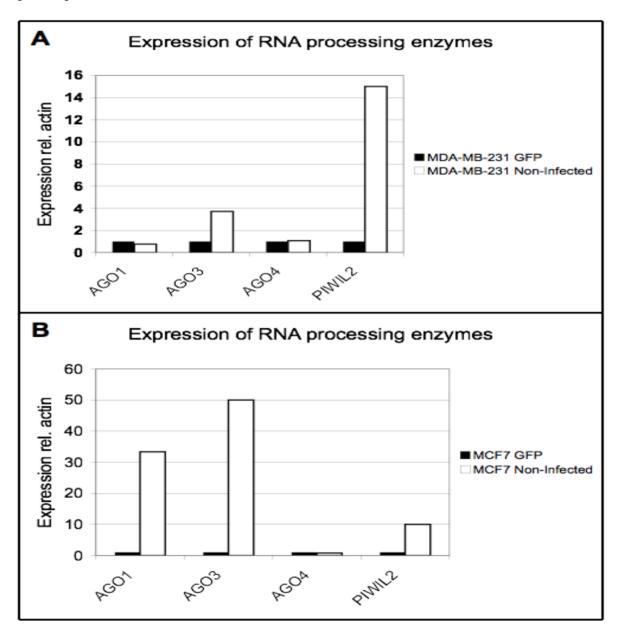


Figure S7. Effects of SINE sense/antisense expression on BRCA1 transcript levels. Expression of BRCA1 was determined 6 and 24 hours post-transfection of plasmids containing the SINE1 probe in either the sense or antisense orientation relative to the BRCA1 coding transcript. Non-transfected, mock and empty vector transfections serve as controls. Expression was normalized to β -actin and data is depicted relative to the mock treatment control, which was set to 1. Data depicted is average and standard error of two biological replicates.

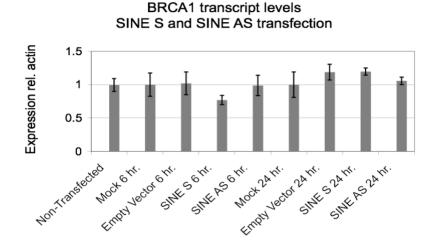
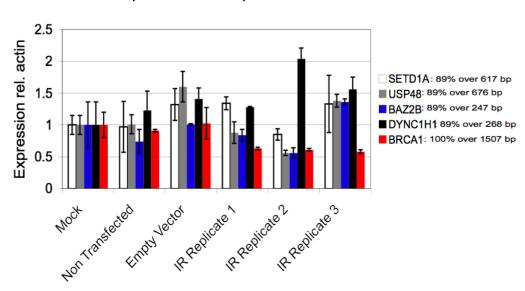


Figure S8. Expression of candidate genes for SINE dsRNA induced knockdown. Four genes containing SINE sequences homologous to the SINE 1 construct used in the transfection experiments within their introns were selected by a bioinformatics search and were defined as candidate genes for SINE 1 IR mediated knockdown. Expression level of all genes 6 hours post transfection was analyzed by qRT-PCR and normalized to β -actin. 3 independent replicates of inverted repeat transfection were compared to an empty vector transfection, a mock treatment, and non-transfected cells. Results for expression of BRCA1, as is shown in Figure 5, are depicted for the purpose of comparison. Data shown is average and standard error for two technical replicates of each primer pair.



Transcript level 6 hours post nucleofection

Figure S9. Alignment of BAZ2B to SINE1. SINE 1 sequence is cropped to show only relevant region BAZ2B. Identity is indicated by shading.

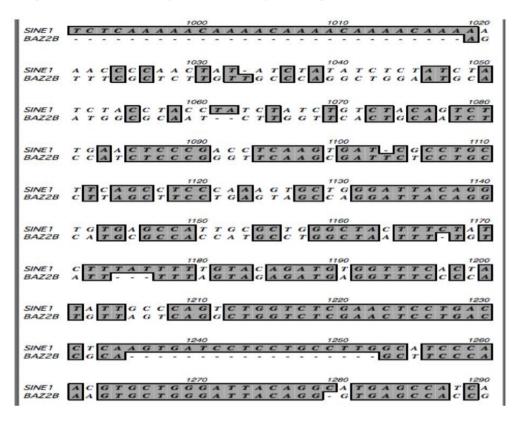


Figure S10. Alignment of DYNC1H1 to SINE1. SINE 1 sequence is cropped to show only relevant region DYNC1H1. Identity is indicated by shading.

SINE1 TCTCAAAA DYNC1H1	ACAAAACAAA	1010 1020 1010 1020 1020 1020 1020 1020	2
SINET DYNC1H1 AGTTTCGC	1030 C T A T A T C T A C T C T T G T T G C C	1040 TATCTCTATCTAT CAGGCTGGAGTA	,
SINE1 CTACCTAC DYNC1H1 CAATGGCG	1060 CTATCT-ATC CGATCTCGGC	1070 TGTCTACAGTCT TCACTGCACCCT)
SINE1 TGAACTCC DYNCIHICTGCCTCC	1090 CGACCTCAAC CGGGTTCAAC	T G A T - C G C C T G C C A A T T C T C C T G C	,
SINE1 T T C A G C C T DYNC1H1 C T C A G C C T	1120 CCCAAAGTGC TCCAGGTAGC	1130 T G G G A T T A C A G G T G G G A T T G C A G A	2
SINE1 TGTGAGCC DYNC1H1 CATGCACC	1150 A T T G C G C T G C G C C A C G C C T C	1160 F G C T A C T T T C T A T F G C T A A T T - C T G T	2
SINE1 C TTTATTT DYNCIHI ATTTT	1180 TTGTACAGA TAGTAGAGA	1190 T G T G G G T T T C A C T A C A G G G G T T T C T C C G	2
SINE1 DYNC1H1 TGTTGCCC DYNC1H1 TGTTGCTC	1210 A G T C T G G T C T A G G C T G G T C T	1220 1230 1230 1230 1230 C G A A C T C C T G A C C G A A C T C C T G A C	7
SINE 1 C T C A A G T G DYNCIHI C T C A G G A G	1240 A T C C T C C T G C A T C C A C C C G C	1250 T C T T G G C A T C C C A C T C A G C C T C C C A	,
SINE1 A C G T G C T G DYNC1H1 A G T G C T A	1270 G G A T T A C A G G G G A T T A C A G A	1280 T G A G C C A T C A C G T G A G C C A C C G	2

Figure S11. Alignment of SETD1A to SINE1. SINE 1 sequence is cropped to show only relevant region SETD1A. Identity is indicated by shading.

SINE 1 GAGGCTGAGGCATGAGGCATGAGAACCC SETDIA
640 SINE 1 A G G A T G G C A G A G G T T G C A A T G A G C T G A G A A SETDIA G T T G T T G T T G T T G T T G T G T
SINE 1 GAGTGCCACTGCACTCCACCCTGGGCAACA SETDIA GAGTTTCGCT-CTTGGCCC-AGGCTGGA SETDIA GAGTTTCGCT-CTTGTCGCCC-AGGCTGGA
700 SINE 1 GAGGAA - GACTCTGTCTCAAAAAAAAAAAAAAAAAAAAAAAAAAAA
750 SINET A A A A A A A A A A A A A A G G C C A G G A G C G G T G G C SETDIA C C T C C G C C T C C T G A G T T C A A G A G A T T - C
SINE 1 T C A C G C C T G T A A T C C C A A C A C T T G G G A G A SETDIA T C C T G C C T - C A G C C T C - C G T T G T A G C T G G
SINE 1 G G T - C A G G A G T T C A A G A C C A G C C T G A C C A SETDIA G A T T A C A G G C A C C C A C C A C C A C C C C
SINE 1 A C A C G G A G A A A C C C T G T C T C T A C T A A A A A T T T T T T G T A T T T T T
SINE 1 A C A A A A T T A G C C G G G C A T G G C A T A T G G T SETDIA A C A G G G T T T C A C T A T G T T G G C C A G A C T G G T A T G T A T G T T G G C C A G A C T G G T A T G T A T G T T G G C A G A C T G G T G G A C T G G T A T G T A T G T T G G C A G A C T G G T G G A C T G G T A T G T A T G T T G G C A G A C T G G T G G A C T G G T A T G T A T G T T G G C A G A C T G G T G G A C T G G T G G A C T G G T G G A C T G G T G G A C T G G A C T G G T G G A C T G G T G G A C T G G T G G A C T G G T G G A C T G G A C T G G T G G A C T G G A C T G G T G G A C T G G A C T G G A C T G G T G G A C T G G T G G A C T G G A C T G G A C T G G G T T C C A C T G G A C T G G A C T G G C A C T G G G T G G A C T G G C A C T G G A C T G G C A C T G G G T C A C T G G A C T G G C A C T G G G T C A C T G G C A C T G A C T C A C T C A C T C A C T C A C T C A C T C A C T C A C T C A C T
SINE1 C C T G T A A T C C T G G C T A C T C A G G A G G C T G A G SETDIA C T T G A A C T C C T G A G C T C A G G C G A T C C A C
SINE 1 G C A G G A G A A T C T C T T G A A C C C G G G A G G C G G G G G G G
940 SINE 1 A G G T T G C A G T A A A C T G A G A T C G C G C C A T T G SETDIA G G A T T A C A G
990 SINE 1 CACTCCAGCC SETDIA CACCCAGATT T
1000 SINE I C C A T C T C A A A A A C A A A A C A A A A
1050 SINE 1 A A A A A C C C C A A C T A T A T C T A T A
1060 SINE 1 TATCTACCTACCTATCTATCTGTCTACAGT SETDIA TG A TCT TG GCTCACTGCAAC
1090 1100 1100 1110 SINE1 C T C G A A C T C C C G A C C T C A A G T G A T C G - C C T G G T T C C A G C G A T T T T C C T 1100 1110 SETD1A C T C C G C C T C C C T G G T T C C A G C G A T T T T C C T 1100 1100 1110
SINE 1 G C T C A G C C T C C C A A A G T G C C T G G G A T T A C A SETD 1 A G C C C T C A G C C T C C C A A A G C G C T G G G A T T A C A
1150 1160 1160 1170 SINE1 G G T G T G A G C C A T T G C G C T G G G C T A C T T T C T SETDIA G G T G T G A G C C A C A C A C A C C T G G C T A A T T T T
1180 1190 1200 SINE I A T C T T T A T T T T G T A C A G A T G T G G T T T C A C A G A G A T G G G G T T T C A C SETDIA G T A T T T T T A G C A G A G A T G G G G T T T T A C
SINE 1 TATATTGCCCAGTCTGGTCTCGAACTCCTG SETDIA CATGTTGGCCAGGCTGGTCCCAAACTCATG
1240 1250 1260 SINE 1 A C C T C A A G T G A T C C T C C T G C C T T G G C A T C C T C C T C G G C C T C C SETDIA A C C T C A A G T G A T C C A C C C A C C T C G G C C T C C
SINE 1 CAACGTGCTGGGGATTACAGGCATGAGCCAT SETDIA CAAAGTGCTAGGCTTTCACCTGTGAGCCAC
1300 1310 1320 SINE1 C A A G C C T G G C C T T A T C T T C A T T T T A T A G A SETD1A T G C A C

Figure S12. Alignment of USP48 to SINE1. SINE 1 sequence is cropped to show only relevant region USP48. Identity is indicated by shading.

SINE1 <u>GTGCACTGT</u> AGTCCCAGTTACTCGGGAGGC USP48 <u>AGTCCCCAG</u> CTACTTGGGAAGGC
SINE 1 T G A G G C A T G A G A A T A A G T T G A A C C T G T G A G USP48 T G A G G C A G G A G A A T C G C C T G A A C C C G G G A G
SINET GCGGGGGTTACAGTGAGCCAATATTGCACC USP48 ATGGAGGCTGAAGTGAGCAAAGATCGTGCC
SINE 1 A T T G C A T T C C A G C C T G G G C A A C A G A G T G A G USP48 A C T G C A A T C T A G C C T G G G C G A C A G A G C A A G
340 SINET A TTG T G C C A A A A A A A A A A A A A A G G C A T T G G USP48 A C T C C A T C T C A A A A A A A A A A
SINET GAATTGTGAGAAAAG - ACACTTA GA GAATA USP48 TAATTATTAATTAATTAATTAATTAATTAATTAATTA
400 410 410 410 410 410 420 420 420 420 420 420 420 42
SINET TTACCCTCTCTCGCCGGGTGCAG-TGGCT USP48 TTATATATATATATTATTATAATTATT
ADD SINE 1 CATGTCTGCAGCAATCC-CAGGAGGCTGAGGCA USP48 TATTATTATTTTTTCATGAGAACCATG-A
SINE1 G C A G C T C A G C T T G A G A G A G A G A A G T T A G G A A A G T T A A G A A A A G T T A A G A
SINE 1 C C A G C C T G G C T A C A T G G G T G A A A C C C T G T C T USP48 C T T G C C T A T C A G A T A T T
SINE 1 C T A C T A A A A A A T A C A A C A A A A
580 590 590 590 590 590 590 600 059 059 059 059 059 059 059 0
610 SINE 1 T G G G A G G C T G A G G C A T G A G A A T C A C T G G A USP 48 T C A A C G A A A T T A A T C A A T G G -
SINE 1 A C C C A G G A T G G C A G A G G T T G C A A T G A G C T G USP48 - C A A A A A A A A C A G - G G T T G A G A G G C T G
670 SINET A G A A G A G T G C C A C T G C A C T C C A C C C T G G G G C USP48 G G T G C A G T G G C - T T A C A C - C T A T A A T C C - C
700 710 720 SINE1 A A C A G A G G A A G A C T C T G T C T C A A A A A A A A 720 USP48 A A C A
730 740 750 SINE1 A A A A A A A A A A A A A A A A A A A
760 770 780 SINE1 G G C T C A G C T G
790 800 810 SINE1 A G A G G T C A G G A G T T C A A G A C C A G C C T G A C C USP48 - A G G T C G A C A G T T C A A G A A C A G C C T G A C C
820 830 840 SINE1 A A C A C G G A G A A A C C C C T G T C T C T A C T A A A A A USP48 A A C A C G G A G A A A C C C C C G T C T C T A C T A A A A A A A A A A A A A A
850 SINE 1 T A C A A A A A T T A G C C G G G C A T G G T G G C A T A T USP48 C A C A A A A - T T A G C C G G G C G T G G T G G C A C A T
SINE 1 G C C T G T A A T C C T G G C T A C T C A G G A G G C T G A USP48 G C C T G T A A T C C C A G C T G • T C A G G G G G A T G A
SINE 1 G G C A G G A G A A T C T C T T G A A C C C G G G A G G C G USP48 G A C A G G A G A A T C G C T T G A A C C C A G G A G G T G
SINE 1 GAGGTTGCAGTAAACTGAGATCGCGCCATT USP48 GAAGTTGTGGGGGGGGGGGGGGGCCATT
900 SINE 1 G C A C T C C A G C C T G G G C A A C A A G A G C G A A A C USP 48 G C A C T C C A G C C T G G G C A A C A A G A G C G A A A C
1000 1010 1020 1020 SINET TCCATCTCAAAAACAAAACAAAACAAAACA
USP48 TCCATCTCAAAAAAAAAAAA