Supplementary information for

The Microprocessor controls the activity of mammalian retrotransposons

Sara R. Heras, Sara Macias, Mireya Plass, Noemí Fernandez, David Cano, Eduardo Eyras, José L. Garcia-Perez & Javier F. Cáceres*

Supplementary Figures 1 to 8

Supplementary Table 1. Distribution of reads obtained from DGCR8 HITS-CLIP mapping to repetitive elements.

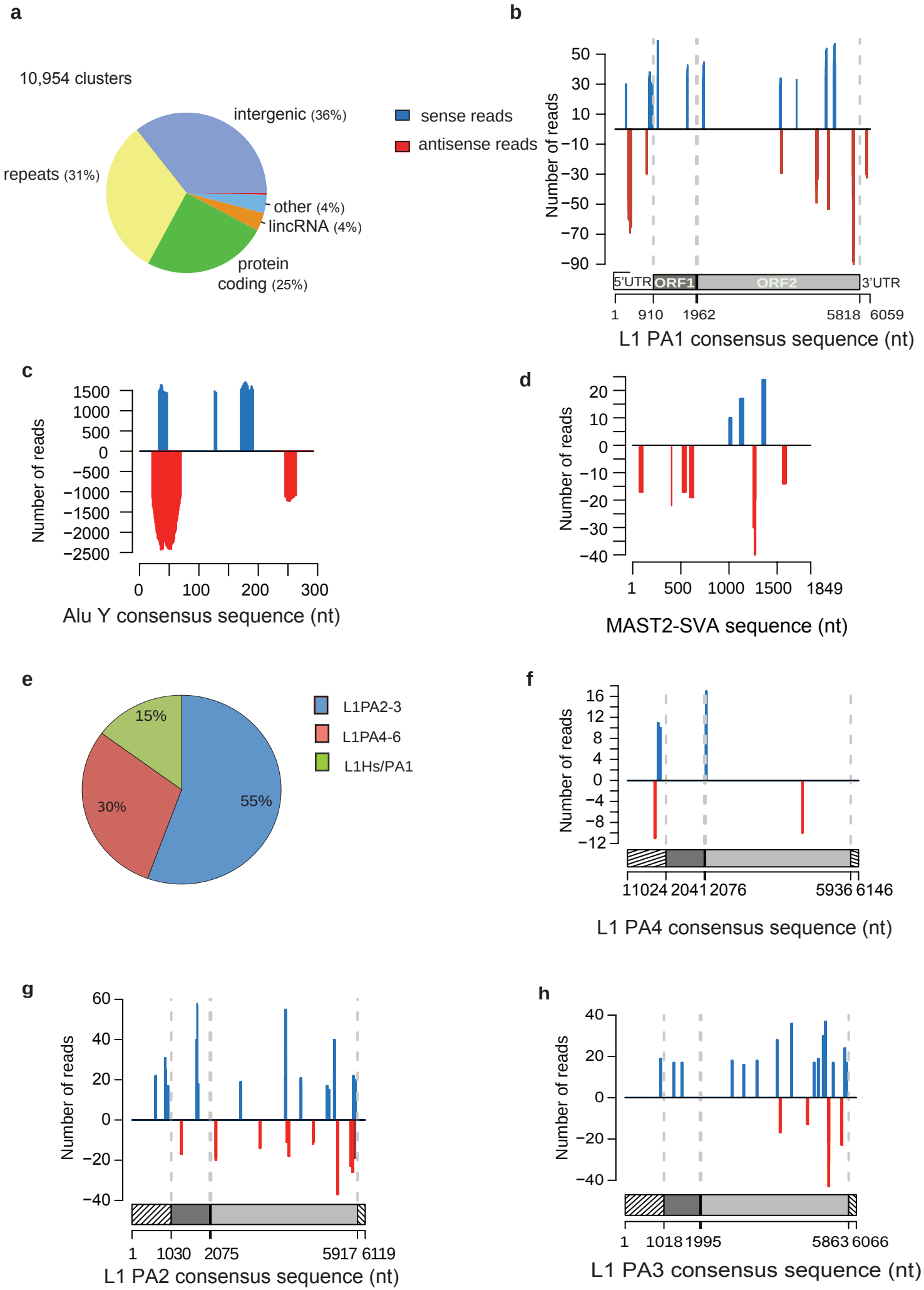
Supplementary Table 2. Human repeats sequences identified bound by DGCR8 by HITS-CLIP method

Supplementary Table 3. Description of significant DGCR8 binding sites on human retrotransposon sequences (Excel file)

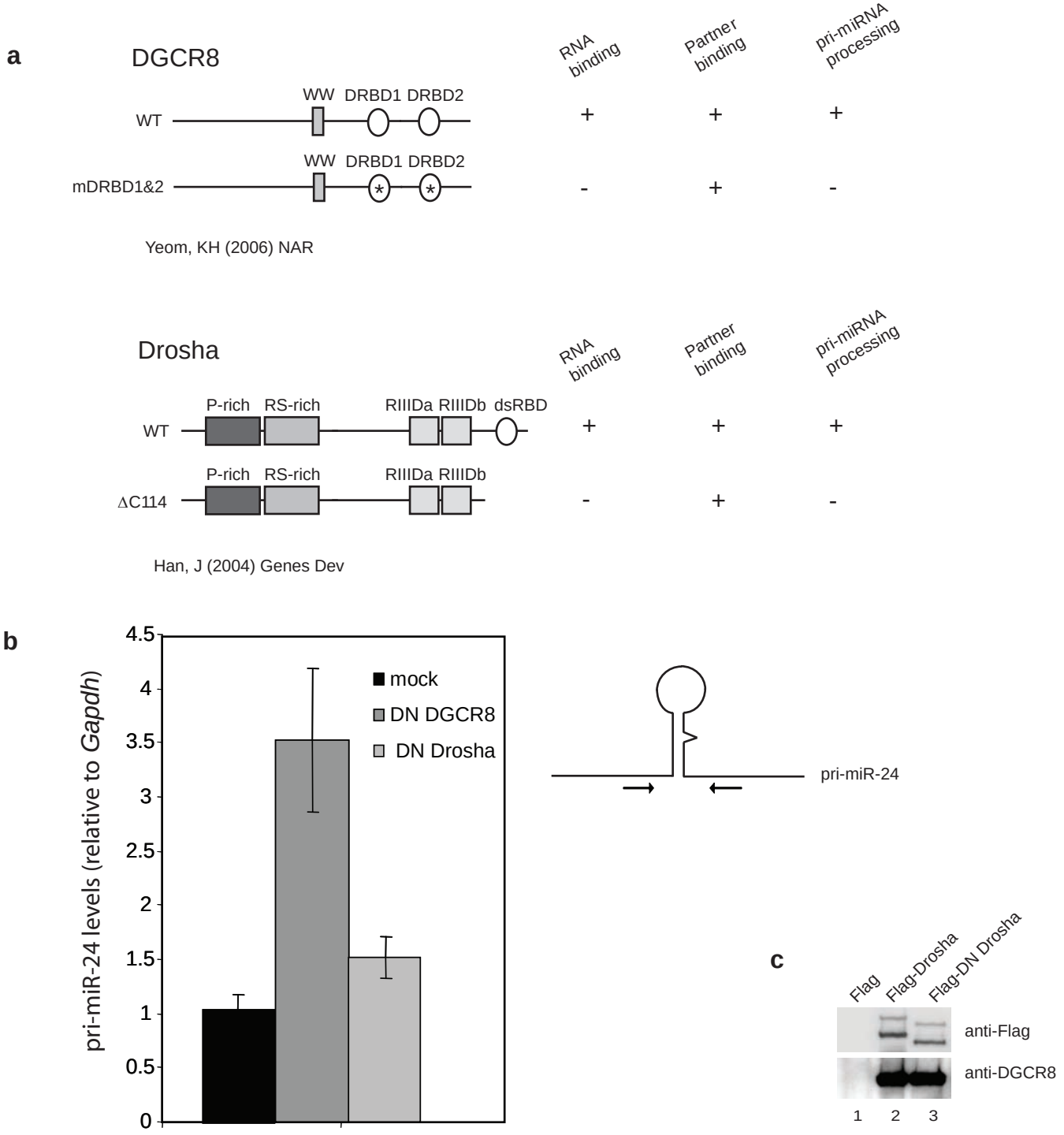
Supplementary Table 4. List of oligonucleotides used in this study.

Supplementary Table 5. Description of the plasmids used in this study.

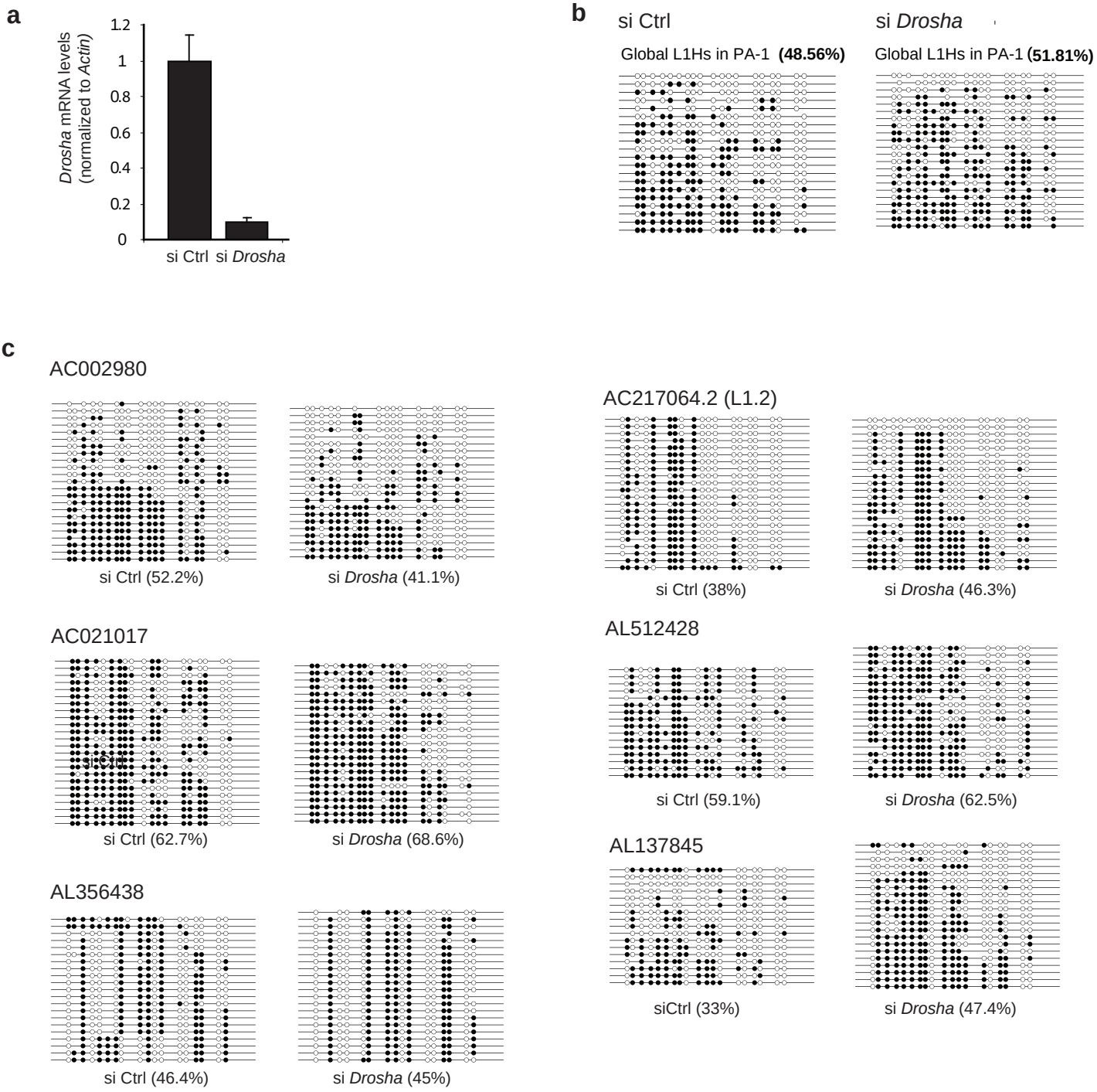
Supplementary figure 1



Supplementary Figure 1. Distribution of DGCR8 HITS-CLIP reads mapping to human Transposable Elements (**a**) Distribution of reproducible DGCR8 significant clusters (FDR<0.01) at the genomic level, with 35.5% mapping to intergenic regions (3,904 clusters), 31.3% to repetitive elements (3,445 clusters), 25.2% to protein-coding genes (2,758 clusters) and 4% to long intergenic non-coding RNAS (lincRNAs) (402 clusters). (**b**-**f**) Distribution of DGCR8 HITS-CLIP reads mapping to human Transposable Elements. Only reads that mapped to annotated Transposable Elements with extreme similarity to consensus sequence (identity >= 99% and coverage>= 90% are shown for L1PA1, Alu Y, SMAT2-SVA and L1PA4 (**b**, **c**, **d** and **f**, respectively), whereas an identity >= 90% and coverage>= 90% was considered for older L1 subfamilies (**g**,**h**). (**b**) Distribution of significant DGCR8 binding sites in a human RC-L1Hs consensus sequence. Only sense (blue) and antisense (red) peaks with a FDR <0.01 and a minimum of 29 reads are represented. Schematic representation of the RC-L1 element (bottom). UTR, untranslated region; ORF, open reading frame. (**c**) Distribution of significant DGCR8 binding sites in a human Alu Y consensus sequence. Only sense (blue) and antisense (red) peaks with a FDR <0.01 are represented. (**d**) Reads mapping to a MAST2-SVA sequence. (**e**) Analysis of cloned ORF1 cDNA sequences (256bp) from H293T cells revealed expression of several L1 subfamilies (LHs, L1PA2-3 and L1PA4-6). Analyses were performed using RepeatMasker. (**f**) Reads mapping to an L1PA4 consensus sequence. (**g**) Reads mapping to an L1PA2 consensus sequence. (**h**) Reads mapping to an L1PA3 consensus sequence.

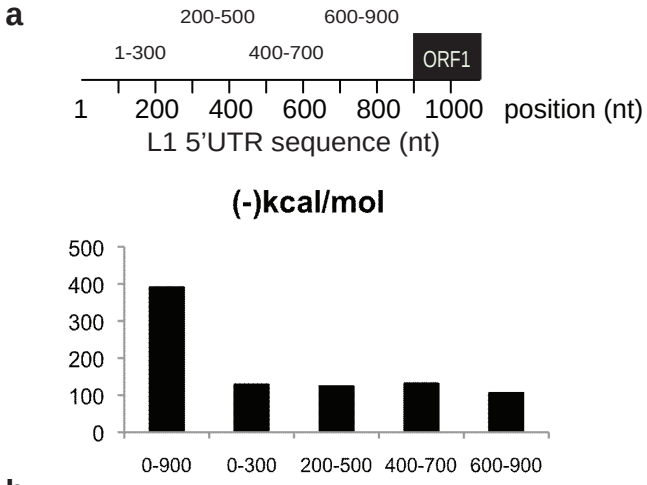


Supplementary Figure 2 Dominant negative (DN) forms of DGCR8 and Drosha. (**a**) Schematic representation and functional description of the dominant negative forms of DGCR8 and Drosha. (**b**) Accumulation of pri-miR-24-2 upon overexpression of DN forms of DGCR8 and Drosha in HeLa cells. The levels of unprocessed pri-miR-24-2 were quantified with respect to mock-transfected cells (black bar). A schematic representation of the qRT-PCR assay is shown on the right, with arrows representing primers used in the qRT-PCR analysis. GAPDH was used as an internal control. (**c**) Overexpressed Flag-DNDrosha binds endogenous DGCR8 to a similar extent (lane 3) when compared to Flag WT Drosha (lane 2), as shown by Flag immunoprecipitation experiments.

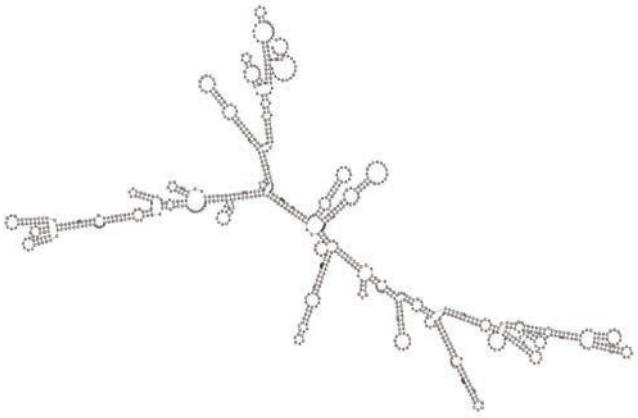


Supplementary Figure 3. Knock-down of Drosha in PA-1 cells does not result in significant changes in the DNA methylation status of the endogenous LINE-1 promoter. (**a**) Confirmation of the knock-down of Drosha at the mRNA level by qRT-PCR. (**b**) Global LINE-1 promoter methylation analysis upon Drosha knock-down (si*Drosha*). In the panel, each line corresponds to a sequenced clone with the highest sequence similarity to a consensus RC-L1 and circles correspond to each of the 20 CpG residues present in the 5´UTR of human RC-L1s (Open circles= unmethylated, closed circles= methylated CpGdinucleotides.). The percentage of methylated residues is also indicated for each RC-L1 in Microprocessor depleted cells or control. (**c**) The status of DNA methylation for six full-length RC-L1 promoters in control PA-1 cells and Drosha knocked-down was examined upon bisulfite conversion. As in panel (**b**), the percentage of DNA methylation for the LINE-1 promoter is indicated.

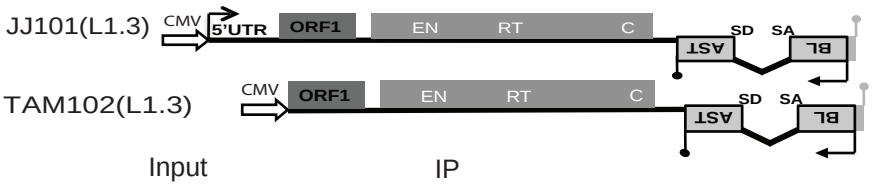
Supplementary figure 4

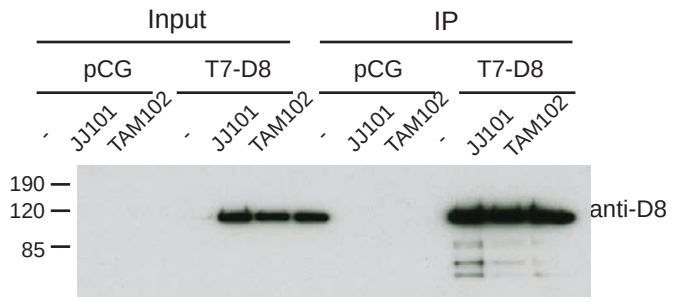


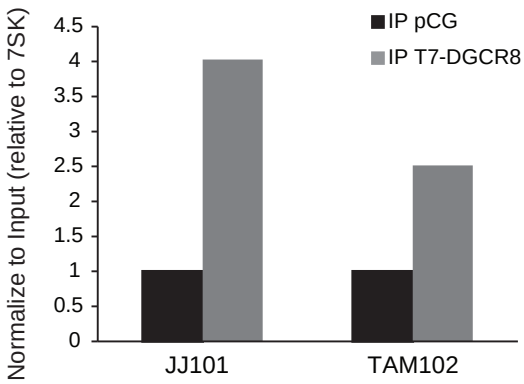
b



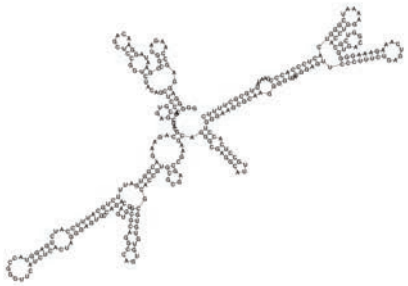
c



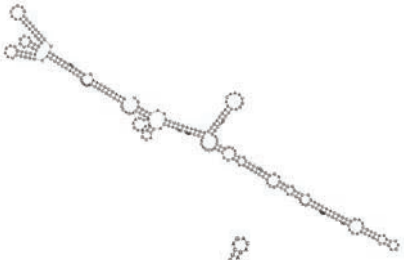




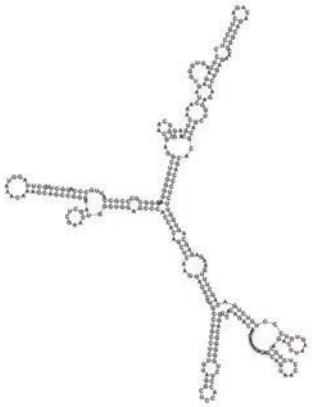
d



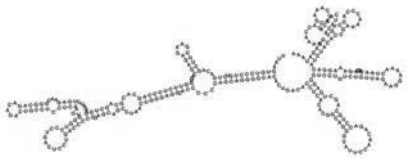
e



f



g



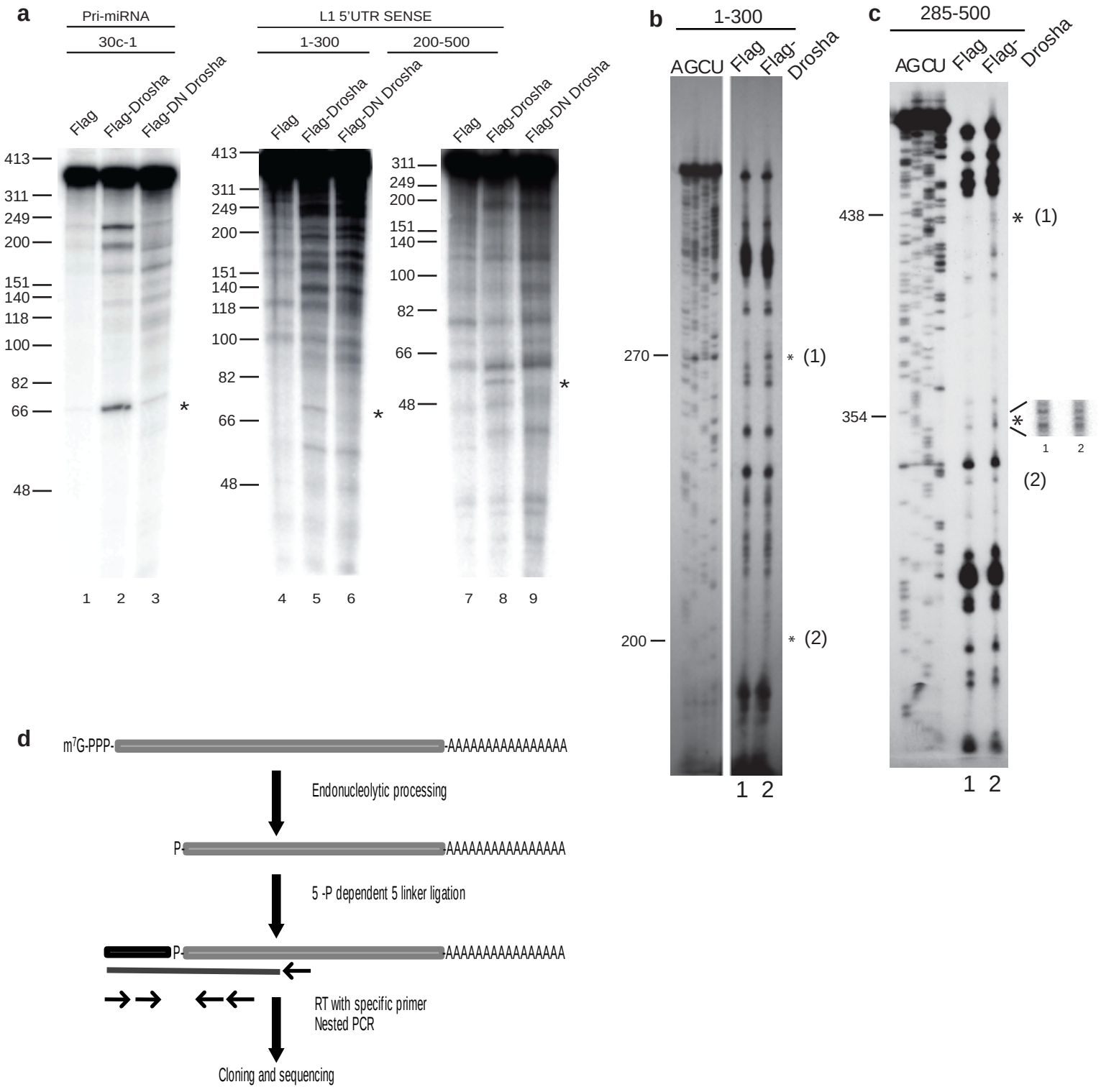
h

```
1     GAATATTGCGCTTTTCAGACCGGCTTAAGAAACGGCGCACCACGAGACTATATCCCACAC  60
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1     GAATATTGCGCTTTTCAGACCGGCTTAAGAAACGGCGCACCACGAGACTATATCCCACAC  60

61    CTGGCTCGGAGGGTCCTACGCCCACGGAATCTCGCTGATTGCTAGCACAGCAGTCTGAGA  120
      |||||||                 ||||||||||||||||||||||||||||||||||
61    CTGGCTCCCACCCTGGTAGCGGGAGCCAATCTCGCTGATTGCTAGCACAGCAGTCTGAGA  120

121   TCAAACTGCAAGGCGGCAACGAGGCTGGGGGAGGGGCGCCCGCCATTGCCCAGGCTTGCT  180
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
121   TCAAACTGCAAGGCGGCAACGAGGCTGGGGGAGGGGCGCCCGCCATTGCCCAGGCTTGCT  180

181   TAGGTAAACAAAGCAGCCGGGAAGCTCGAACTGGGTGGAGCCCACCACAGCT  232
      ||||||||||||||||||||||||||||||||||||||||||||||||||
181   TAGGTAAACAAAGCAGCCGGGAAGCTCGAACTGGGTGGAGCCCACCACAGCT  232
```

i

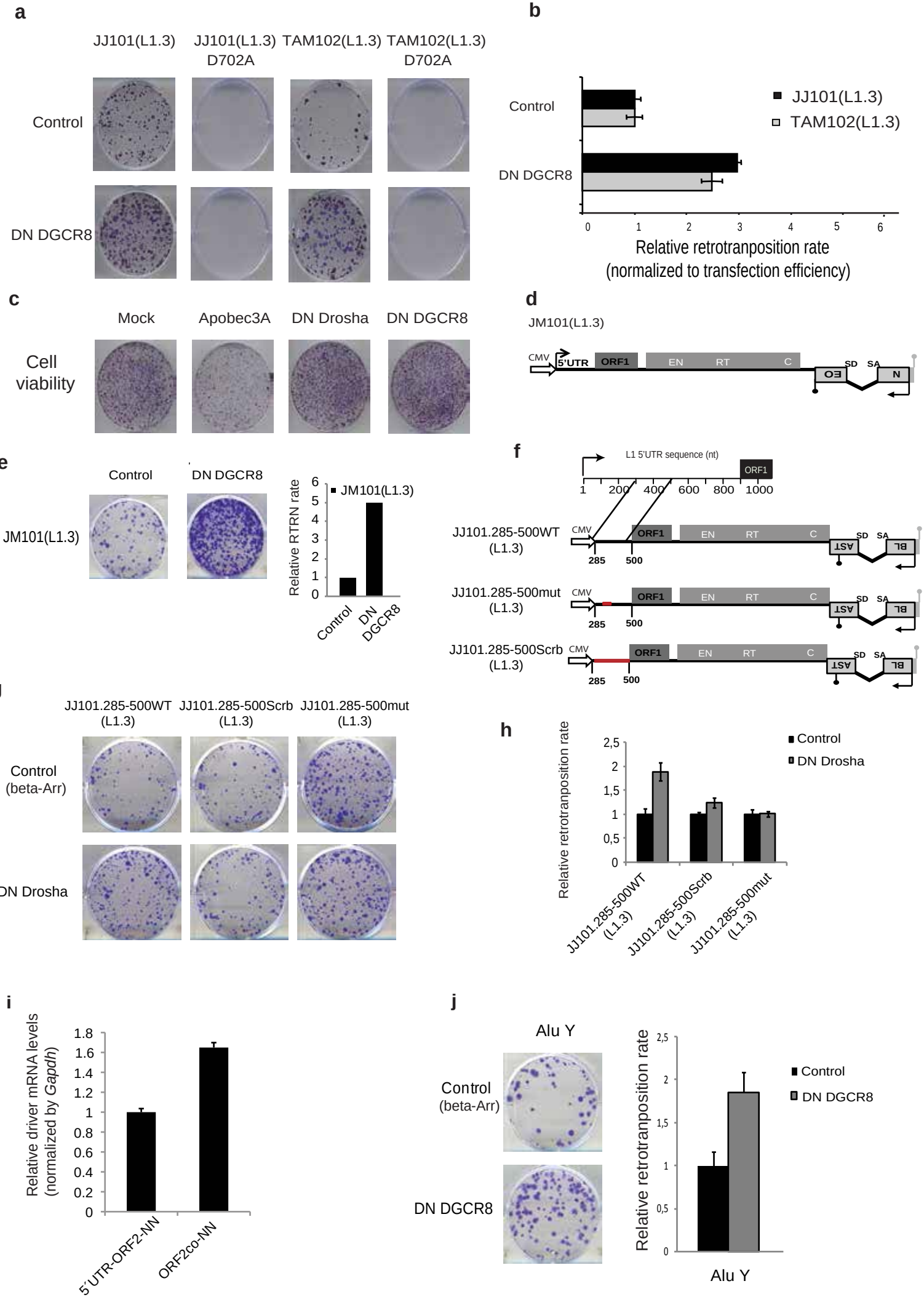
Supplementary Figure 4 The 5´UTR of human RC-L1s is predicted to form a stable secondary structure that is recognized and bound by DGCR8 (**a**) The sense L1 5'UTR region was divided in four overlapping 300 nucleotide regions: 1-300, 200-500, 400-700, and 600-900 (top panel). These four fragments as well as the full-length 5´UTR region from a human RC-L1 were folded using the mfold server at http://rna.tbi.univie.ac.at/cgi-bin/RNAfold.cgi. The bottom graph shows the free energy (as (-)kcal/mol) of the four overlapping RNAs and the full-length 5´UTR RNA of a human RC-L1 (indicated in the X-axis). (**b**) Predicted secondary structure for the full-length 5´UTR RNA from a human RC-L1. (**c**) Constructs expressing a full-length copy of the human LINE-1 element (JJ101(L1.3)) or a truncated version lacking the 5'UTR (TAM102(L1.3)) were co-transfected in HEK293T cells with a plasmid expressing a tagged version of DGCR8 (T7-DGCR8). Overexpressed DGCR8 was immunoprecipitated and LINE-1 RNA association was monitored by qRT-PCR with primers spanning the ORF1. RNA co-immunoprecipitation was normalized to the Input levels and was expressed relative to the co-immunoprecipitated 7SK RNA, that serves as a negative control (see **Fig. 1e**). Deletion of the 5'UTR reduced by 2-fold the efficiency of LINE-1 RNA immunoprecipitation by DGCR8 (compare JJ101 levels, full length, with TAM102). (**d**) Predicted secondary structure for the 1-300 RNA from a human RC-L1. (**e**) Predicted secondary structure for the 200-500 RNA from a human RC-L1. (**f**) Predicted secondary structure for the 400-700 RNA from a human RC-L1. (**g**) Predicted secondary structure for the 600-900 RNA from a human RC-L1. (**h**) Alignment of second stem from 200-500 L1.3 region (top) and mutated second stem (bottom) sequences. (**i**) Predicted secondary structure of the 1786 bp 5´UTR region from a mouse TF RC-L1 (L1spa, GenBank: AF016099.1; -843 kcal/mol(19).

Supplementary figure 5

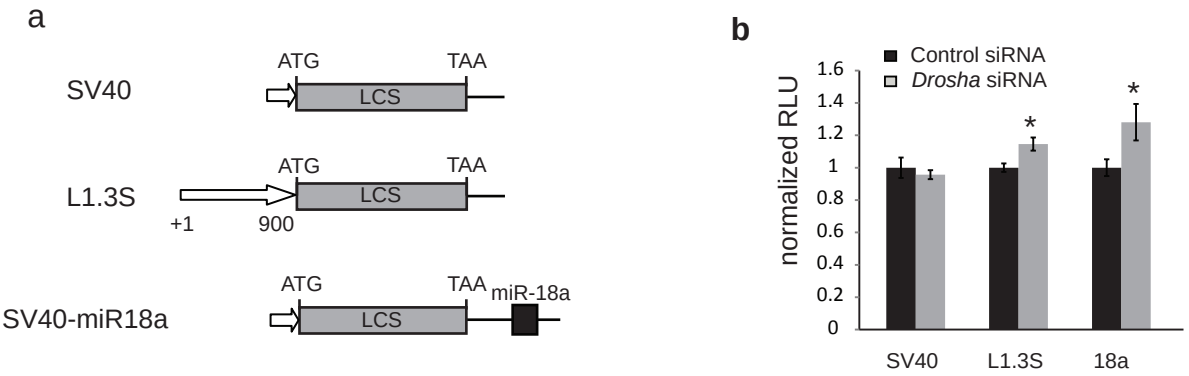


Supplementary Figure 5 Drosha-dependent cleavage sites at the 5´UTR of human RC-L1. (**a**) The regions spanning nucleotides 1-300 and 200-500 from human RC-L1 5'UTR were incubated in vitro with WT Drosha (lanes 2, 5 and 8) or DN-Drosha (lanes 3, 6 and 9) Flag immunoprecipitates. The formation of the cleavage products is dependent on the RNA binding activity of Drosha, as processing is abolished when adding immunopurified DN-Drosha (lanes 3, 6 and 9, marked by asterisks). Pri-miR-30c-1 was used as a positive control. (**b**,**c**) A transcript spanning nucleotides 1-300 (b) or 285-500 (**c**) from the 5´UTR a human RC-L1 was incubated with Flag (lane 1) or Flag-Drosha (lane 2) immunoprecipitates. RNA was purified and subjected to reverse transcriptase (RT) extension with a 5'end-labeled primer annealing at the end of the RNA fragment. cDNA products were analyzed on a 6% denaturing acrylamide gels parallel to a DNA sequence generated with the same primer (left panel). Asterisks and numbers on the left depict the positions of the cleavage sites as RT stops at these positions dependent on the addition of Flag-Drosha immunoprecipitate (numbers are relative to full-length sequence). Panel c shows the same experiment but using a transcript spanning nucleotides 285-500 from L1 5'UTR. The asterisks depict the positions of the cleavage sites, and consistently, the same RT stops were obtained when a longer fragment was used, spanning positions to 200-500 (data not shown). On the right, a magnification of the cleavage number (2) is shown, from a replicate experiment. (**d**) A modified 5-RACE was used to in vivo map processing sites in the 285-500 region. Briefly, an RNA linker (black bar) is ligated to total RNA isolated from cultured cells. Only RNAs that have undergone an endonucleolytic processing event (i.e., containing a 5´Phosphate) could be ligated to the linker. Upon reverse transcription with an L1 specific primer, nested PCR is used to amplify processed substrates. (**e**) Summary of *in vivo* processing sites on the 285-500 region. The nucleotide sequence of the 285-500 region from an L1Hs consensus sequence is shown. Highlighted in red are the characterized in vitro processing sites (see **Supplementary Fig. 5c**). Asterisk and bars denote the observed *in vivo* processing sites in RNA derived from PA-1 and HEK293T cells, respectively.

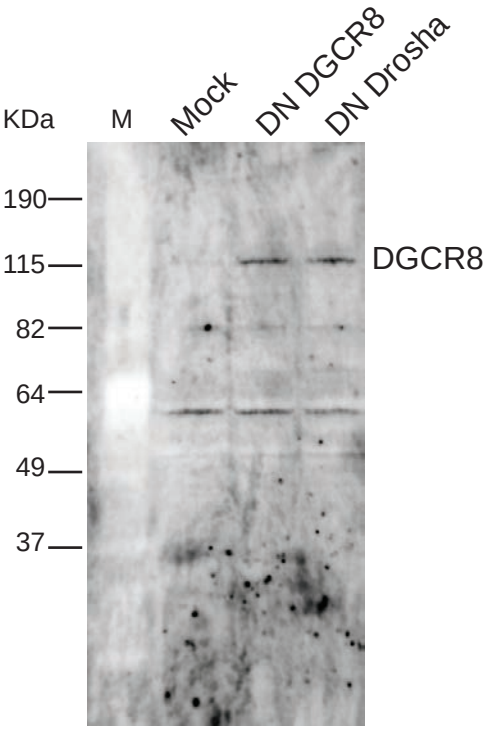
Supplementary figure 6



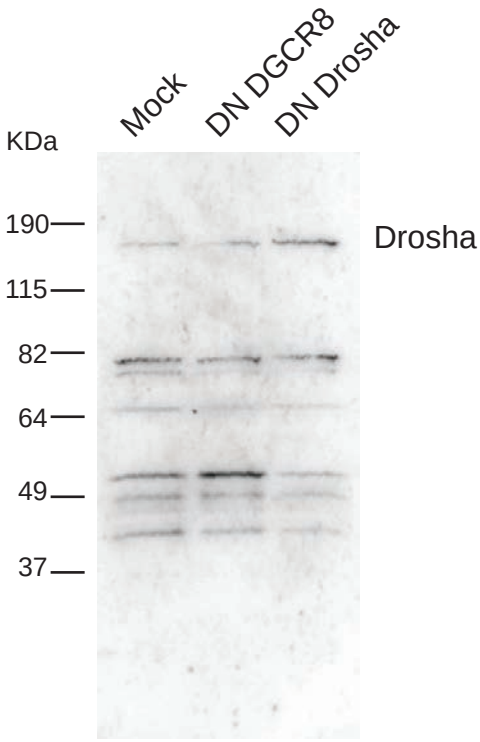
Supplementary Figure 6. The rate of LINE-1 retrotransposition increases upon Microprocessor depletion. (**a**) Each image shows representative data from assays conducted in triplicate. In these assays, cells were co-transfected with two plasmids: JJ101(L1.3) or TAM102(L1.3) (indicated on the top) and an expression plasmid for DN-DGCR8, or β-arrestin (control) as a normalization control. (**b**) Quantitation of LINE-1 retrotransposition (from panel (**a**)). Blasticidin-resistant foci were manually counted. In the graph, data is presented as the proportion of the activity seen in cultures co-transfected with the plasmid expressing the negative control β-arrestin (control) and normalized using transfection efficiency and toxicity. (**c**) Toxicity control experiments for the expression of dominant negative alleles of Drosha or DGCR8 HeLa cells. HeLa cells were co-transfected with two plasmids: a linear blasticidin resistant expressing plasmid (pCDNA-6-myc-His) and the indicated DN-DGCR8, DN-Drosha expression plasmids or control plasmids expressing β-arrestin (mock) or Apobec3A proteins. Notably, expression of DN-Drosha or DN-DGCR8 results in an increased number of blasticidin resistant foci when compared with the control beta-arrestin (mock). Expression of Apobec3A results in mild toxicity (as manifested by the reduced number of blasticidin resistant foci), as previously described. (**d**) Scheme of the plasmid used tagged with an mneoI retrotransposition indicator cassette (following the nomenclature used in **Fig. 4**). (**e**) Each image shows representative data from assays conducted in triplicate. In these assays, cells were co-transfected with two plasmids: JM101(L1.3) and an expression plasmid for DN-DGCR8, or β-arrestin (control) as a normalization control. The right graph shows a quantitation of the rate of LINE-1 retrotransposition. (**f**) The 5'UTR 285-500 region can render the engineered retrotransposition construct reporter sensitive to Microprocessor depletion. The cartoon depicts the LINE-1 based retrotransposition constructs used in cultured cells. (**g**) Representative data of engineered LINE-1 retrotransposition assays is shown (assays conducted in triplicate), and the quantification is shown on the right side. (**h**) As in **Figure 5**, data is presented as the proportion of the activity seen in cultures co-transfected with the plasmid expressing the negative control (β-arrestin) and normalized using transfection efficiency and toxicity (**i**, **j**) Alu retrotransposition is regulated by the Microprocessor (**i**) Expression level of mRNAs produced from transfected 5´UTR-ORF2NN and ORF2coNN drivers in cultured HeLa cells. 48h after transfection total RNA was isolated from transfected cells, DNase I treated, reverse transcribed and used in a quantitative RT-PCR using primers directed to the SV40 late polyadenylation sequence present in both mRNAs (See **Supplementary Tables 3 and 4**). The graph indicates the relative expression level of each driver normalized to GAPDH. (**j**) Trans-retrotransposition experiments using ORF2co-NN as a driver. Each image shows representative data from Alu trans-retrotransposition assay conducted in duplicate. β-arrestin is used as a control that does not affect Alu retrotransposition. The right side graph indicates the proportion of the activity seen in cultures co-transfected with the indicated plasmid for Alu. Data was normalized using transfection efficiency and toxicity. Averages ± s.d are shown in **(b, h, i, j**).



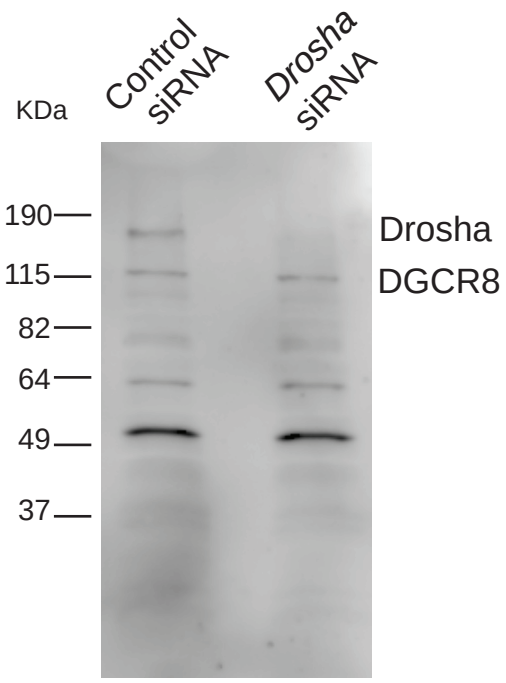
Supplementary Figure 7. The Microprocessor negatively regulates L1 retrotransposition in vivo, by targeting the 5'UTR region (**a**,**b**) A construct comprising the 5'UTR from a RC-L1 (L1.3S) or a non-L1 sequence comprising the SV40 promoter region (SV40) fused to a Firefly Luciferase ORF (LCS) were transfected into HeLa cells in the presence of a siRNA against Drosha or a non-targeting control. Luciferase levels were analyzed relative to a cotransfected Renilla Luciferase reporter. An SV40 construct containing a single miR-18a target site in its 3´UTR (SV40-miR18a) was used to monitor the miRNA processing activity of the Microprocessor. RLU, Relative Light Unit. Error bars indicate standard deviation (n=3) and *P< 0.05 (t test).



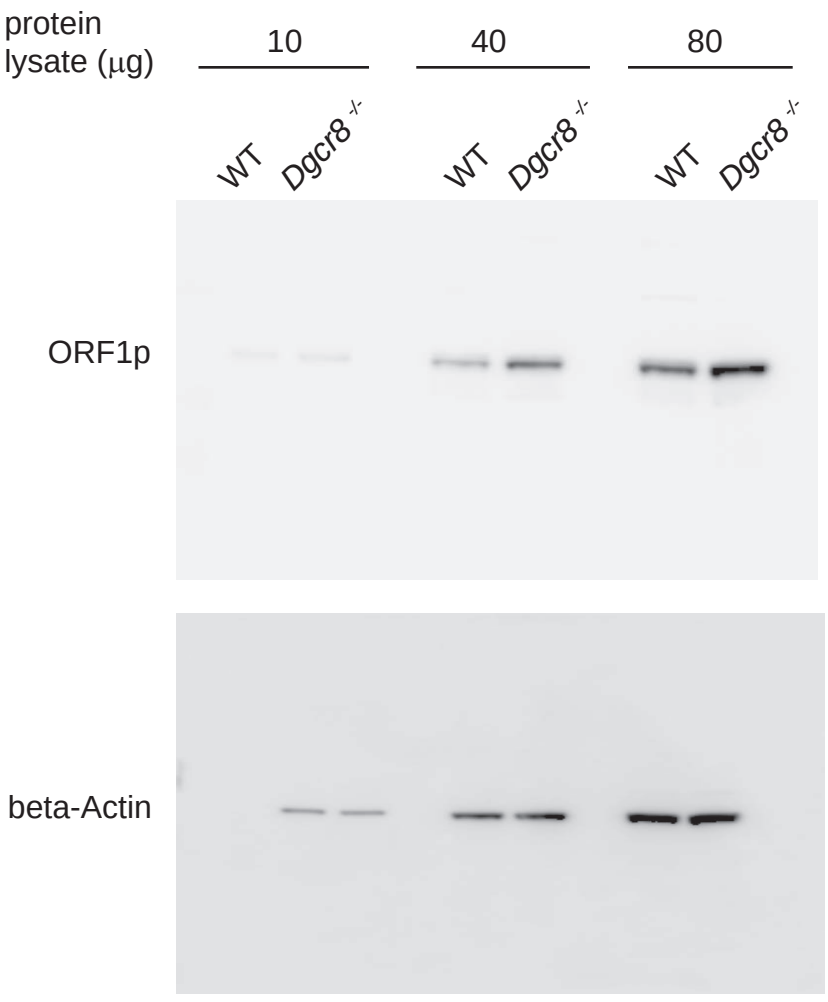
Uncropped version of the western blot shown in Fig. 2b



Uncropped version of the western blot shown in Fig. 2b



Uncropped version of the western blot shown in Fig. 2a



Uncropped version of the western blot shown in Fig. 2d

Supplementary Figure 8. Uncropped versions of images shown in the article

Supplementary Table 1. Distribution of reads obtained from DGCR8 HITS-CLIP mapping to repetitive elements.

	number of unique reads	distribution of total reads	Distribution of reads mapping repeats sequences
total CLIP1+CLIP2	5496250	100	
human repeats sequences	2850320	51.85935865	100
TEs (LINE, LTR, SINE, DNA)	2071807	37.69491926	72.68682113
LINE	911216	16.57886741	31.96890174
rRNA	537014	9.77055265	18.84048107
LTR	499025	9.079372299	17.50768335
SINE	471202	8.573154423	16.53154734
DNA	190364	3.463525131	6.678688709
tRNA	89760	1.633113486	3.149120099
Low_complexity	59388	1.080518535	2.083555531
Simple_repeat	44277	0.805585627	1.55340453
Other	48074	0.874669093	1.686617643

Supplementary Table 2. Human repeats sequences identified by DGCR8 HITS-CLIP method

human repeats sequences identified		
repeat class	**matching repeat**	**number of unique reads**
LINE	L1	579775
rRNA	rRNA	537014
LINE	L2	295959
SINE	Alu	257720
SINE	MIR	210523
LTR	ERVL-MaLR	205020
LTR	ERV1	138412
LTR	ERVL	131029
DNA	hAT-Charlie	96067
tRNA	tRNA	89760
Low_complexity	Low_complexity	59388
DNA	TcMar-Tigger	47652
Simple_repeat	Simple_repeat	44277
LINE	CR1	25204
DNA	hAT-Tip100	20693
snRNA	snRNA	19919
LTR	ERVK	10981
LINE	RTE	9871
DNA	hAT-Blackjack	7525
Satellite	centr	7047
LTR	Gypsy	6474
DNA	TcMar-Mariner	4777
Satellite	Satellite	4669
LTR	Gypsy	4583
RNA	RNA	4172
DNA	hAT	3365
Unknown	Unknown	3156
srpRNA	srpRNA	2779
Other	Other	2648
DNA	MuDR	2631
DNA	TcMar-Tc2	2449
scRNA	scRNA	2019
SINE	tRNA	1988
DNA	TcMar	1561
LTR	LTR	1242

DNA	DNA	1077
DNA	PiggyBac	947
LTR	ERVL	929
DNA	hAT	898
SINE	Deu	643
DNA	TcMar	616
RC	Helitron	569
Satellite	telo	467
SINE	SINE	418
LTR	ERV	355
DNA	DNA	293
LINE	RTE-BovB	213
LINE	Dong-R4	157
SINE	SINE	111
RC	Helitron	82
DNA	PiggyBac	78
LTR?	LTR	47
Unknown	Unknown	45
LINE	L1	37
Satellite	acro	36
DNA	Merlin	28
LINE	Penelope	15

Supplementary Table 4. Oligonucleotides used in this study

GAATGATTTTGACGAGCTGAGAGAA	FORWARD 5'UTR from L1.3, 67bp of amplicon from pos 1017-1083 (HUMAN) qRTPCR
GTCCTCCCGTAGCTCAGAGTAATT	REVERSE 5'UTR from L1.3, 67bp of amplicon from pos 1017-1084 (HUMAN) qRTPCR
taatacgactcactatagggCTTCGGCTCGCGCACGGTGC	FORWARD for antisense probe to detect human L1.3 (0-150nt)
CCGGTCTACAGCTCCCAGCGTGAGC	REVERSE for antisense probe to detect human L1.3 (0-150nt)
GAGTCAACGGATTTGGTCGT	GAPDH forward RT-PCR
TTGATTTTGGAGGGATCTCG	GAPDH reverse RTPCR
CATCCCCGATAGAGGAGGAC	FORWARD human 7SK
GCGCAGCTACTCGTATACCC	REVERSE human 7SK
taatacgactcactatagggGAATAGGAACAGCTCCGGTCTACAG	FORWARD T7 coupled for SENSE L1.3 transcription 0-300
TTAAGCCGGTCTGAAAAGCGCAAT	REVERSE 0-300 L1.3 human for in vitro transcription.
taatacgactcactatagggCGCAGGCCAGTGTGTGTGCGCAC	FORWARD 200-500 L1.3 human for in vitro transcription (SM772).
AGCTGTGGTGGGCTCCACCCAG	REVERSE 200-500 L1.3 human for in vitro transcription (SM393).
taatacgactcactatagggGATCAAACTGCAAGGCGGCAAC	FORWARD 400-700 L1.3 human for in vitro transcription.
TGCTCGGGGGTCAGGAGTCAGG	REVERSE 400-700 L1.3 human for in vitro transcription.
taatacgactcactatagggCTTAAGTGTCCCTGTCTGACAG	FORWARD 600-900 L1.3 human for in vitro transcription.
CCTTCTAACAGACAGGACCCTCAGCT	REVERSE for L1.3 transcription 600-900
AAGGGGTTAGGGAGTTTTTTT	FORWARD CpG L1 human for bisulfite sequencing (Coufal et al Nature 2009)
TATCTATACCCTACCCCCAAAA	REVERSE CpG L1 human for bisulfite sequencing (Coufal et al Nature 2009)
ATTGTATATTGGAGTTGTTTTTTTT	FORWARD for nested PCR of the L1 locus Ac002980
GAGGATTTTTGGTGATAAAGGATTT	FORWARD for nested PCR of the L1 locus AC021017
TTTTTGGTTAAGGAAGTGTTTGATAA	FORWARD for nested PCR of the L1 locus AL356438
GGTTTTTGTTGTTGTATTTGTGTGT	FORWARD for nested PCR of the L1.2 locus (AC217064.2)M80343.1
AAATTAGTTGGGTATGATGGTAGGT	FORWARD for nested PCR of the L1 locus AL512428
GGAAATTGAATAATTTGTTTTTGAAT	FORWARD for nested PCR of the L1 locus AL137845
CAAAAACCCACTTAAAAAAACAATC	REVERSE for nested PCR L1-HS
GAAATTAGTCTGAACAGGTGAGAGG	FORWARD qRTPCR mouse (223-247) TF L1
TCCTCTGGTCCGGAAGGT	REVERSE qRTPCR mouse (336-353) TF L1
ACCGAGCGCGGCTACAG	FORWARD actin B
CTTAATGTCACGCACGATTTCC	REVERSE actin B
AGCTGAGGCGCTGCTTCT	FORWARD amplifying the beginning the stem of pre-miR-24-1 (human)
CCTCGGGCACTTACAGACA	REVERSE amplifying the end the stem of pre-miR-24-1 (human)
CCGGTCTACAGCTCCCAGCGTGAGC	REVERSE for RT of the antisense L1 transcript
CTTCGGCTCGCGCACGGTGC	FORWARD for PCR antisense L1 transcript
taatacgactcactatagggGAATATTGCGCTTTCAGACCGGC	FORWARD for amplification of stem 350-500 fragment L1.3 human (SM394) (to PCR with SM393)
ATTGGCTCCCGCTACCAGGGTGGGAGCCAGGTGTGGGATATAG	REVERSE to mutate 2ndary st in 350-500 stem to PCR with SM394 (1st step PCR)
CCACCCTGGTAGCGGGAGCCAATCTCGCTGATTGCTAGCACAGC	FORWARD to mutate 2ndary st in 350-500 stem to PCR with SM393 (1st step PCR)
TAGCGGCCGCGGGGGAGGAGCCAAGTG	FORWARD with *Not I* site for amplification of 1-300 fragment L1.3 human.
TAGCGGCCGCGAAAAGCGCAATATTC	REVERSE with *Not I* site for amplification of 1-300 fragment L1.3 human.
TAGCGGCCGCGAAGCGCAAGGGGTCAG	FORWARD with *Not I* site for amplification of 200-500 fragment L1.3 human.
TAGCGGCCGCACCCAGTTCGAGCTTC	REVERSE with *Not I* site for amplification of 200-500 fragment L1.3 human.
TAGCGGCCGCGAGATCAAACTGCAAGGC	FORWARD with *Not I* site for amplification of 400-700 fragment L1.3 human.
TAGCGGCCGCACCCACTTGAGGAGGCAG	REVERSE with *Not I* site for amplification of 400-700 fragment L1.3 human.
TAGCGGCCGCTTAAGTGTCCCTGTCTGAC	FORWARD with *Not I* site for amplification of 400-700 fragment L1.3 human.
TAGCGGCCGCTTTGTGGTTTTATCTAC	REVERSE with *Not I* site for amplification of 400-700 fragment L1.3 human.
AGGAAATACAGAGAACGCCACAA	FORWARD ORF1 LINE-1, amplicon of 256bp from pos 1615 - 1851
GCTGGATATGAAATTCTGGGTTGA	REVERSE ORF1 LINE-1, amplicon of 256bp from pos 1615 - 1851

GCAGGCAGGCCTCCTTGAGCTG	5' UTRL1as primer used for RT to perform 5´ Phosphate-dependent 5´ RACE
TTGGAGCTCCACCGCGGTGG	FORWARD for first PCR in the 5' RACE (upSK primer)
GGCTCCACCCCAGTTCGAGCTTCC	REVERSE for first PCR in the 5' RACE (484as primer).
CGCTCTAGAACTAGTGGATC	FORWARD for nested PCR in the 5' RACE (KS primer).
ACCTAAGCAAGCCTGGGCATGG	REVERSE for nested PCR in the 5' RACE (447as primer).
TGGACAAACCACAACTAGAATGC	sv40pA FORWARD for RT-qPCR to determine drivers mRNA levels.
TTGCAGCTTATAATGGTTAC	sv40pA REVERSE for RT-qPCR to determine drivers mRNA levels.

Supplementary Table 5. Plasmids used in this study

Plasmid name	Description
pCGT7-DGCR8	A plasmid overexpressing human DGCR8 ORF (ENST00000351989, Ensembl) was cloned in a pCG backbone [1] containing a T7 epitope-tag at its N-terminus, using *Xba*I/*Dpn*II sites.
Flag-Drosha	was fully described in [2].
mDRBD1&2 DGCR8 (or DN-DGCR8)	contains a Dominant Negative (DN) form of DGCR8 [3].
ΔC114 Drosha (or DN-Drosha)	contains a DN form of Drosha [2]. Both were kindly provided by Narry Kim (Seoul).
Pk-β-Arrestin (or b-Arr) and **pk-α-Apobec3A** (or A3A)	contain the human β-Arrestin ORF or the human APOBEC3A ORF, respectively cloned in vector pK as described [4].
SV40-FF	contains an SV40 promoter cloned in pGL3-Basic (Promega) and
SV40-FF-miR18	**SV40-FF-miR18** was previously described [5].
L1.3S-FF (or L1.3S)	contains a 909-bp long 5′UTR from a human L1PA1/L1Hs element (L1.3) [6] cloned in pGL3-Basic (Promega) [7].
L1PA2S-FF, L1PA3S-FF and **L1PA4S-FF**	have been previously described in [7].
mL1spa-FF	contains the 1,787-bp long 5′UTR from a mouse $L1T_F$ element (L1spa) [8] cloned in pGL3-Basic (Promega).
pRL-TK-Renilla	encoding Renilla luciferase under the control of a HSV-TK promoter was obtained from Promega.
JJ101(L1.3)	contains a full-length human RC-L1 (L1.3) [6] tagged with a *mblasI* retrotransposition indicator cassette [9] and was cloned in pCEP4 (Invitrogen) [10].
JJ101(L1.3) D702A	is a derivative of JJ101(L1.3) that carries a missense mutation (Asp702Ala) [11] in the RT domain of L1-ORF2.
TAM102(L1.3)	is a derivative of JJ101(L1.3) that lacks the 5′UTR within the LINE-1 [9].
TAM102(L1.3) D702A	is a derivative of TAM102(L1.3) that contains a missense mutation (Asp702Ala) [11] in the RT domain of L1-ORF2.
JJ101.1-300(L1.3)/ JJ101.200-500(L1.3) JJ101.400-700(L1.3) and JJ101.600-900(L1.3)	are derivatives of TAM102(L1.3) that contain the 1-300, 200-500, 400-700 or 600-900 regions from the 5'UTR of a human RC-L1 (L1.3) [6] cloned into the *Not* I site, respectively.
pcDNA-6-myc-His	contains a blasticidin reporter gene and it was obtained from Invitrogen. It was linearized with *Eco*RI.
pCEP-EGFP	contains the coding sequence of the humanized enhanced green fluorescent protein (hEGFP) cloned in pCEP4 (Invitrogen) [12].
JM101(L1.3)	contains a full-length human RC-L1 (L1.3) [6] tagged with a *mneoI* retrotransposition indicator cassette [13] and was cloned in pCEP4 (Invitrogen) [14].
pCEP-5′UTR-ORF2NN	contains the 5′UTR and ORF2 coding sequence from of L1.3[6] cloned in pCEP4[15].
pCEP-ORFco-NN	contains the ORF2 coding sequence from plasmid L1PA1ch[16] cloned in pCEP4.
Q19-33-AluYa5	contains a 2.1-kb fragment containing a 7SL promoter, a copy of the NF1 Alu[17], the *neo*^{III} self-splicing indicator cassette[18], a thirty-three bp poly(A) tail, and a BC1 transcription termination sequence cloned in pBSKS-II (Invitrogen).
Q19-33-AluSx3	contains a 2.1-kb fragment containing a 7SL promoter, a copy of an Sx Alu (A-20, see[7]), the *neo*^{III} self-splicing indicator cassette[18], a thirty-three bp poly(A) tail, and a BC1 transcription termination sequence cloned in pBSKS-II (Invitrogen).

1 Caceres, J. F., Misteli, T., Screaton, D. R., Spector, D. L. & Krainer, A. R. Role of the modular domains of SR proteins in subnuclear localization and alternative splicing specificity. *J Cell Biol.* **138**, 225-238 (1997).

2 Han, J. *et al.* The Drosha-DGCR8 complex in primary microRNA processing. *Genes & Development* **18**, 3016-3027 (2004).

3 Yeom, K. H., Lee, Y., Han, J., Suh, M. R. & Kim, V. N. Characterization of DGCR8/Pasha, the essential cofactor for Drosha in primary miRNA processing. *Nucleic Acids Res* **34**, 4622-4629 (2006).

4 Bogerd, H. P. *et al.* Cellular inhibitors of long interspersed element 1 and Alu retrotransposition. *Proc Natl Acad Sci U S A* **103**, 8780-8785 (2006).

5 Guil, S. & Caceres, J. F. The multifunctional RNA-binding protein hnRNP A1 is required for processing of miR-18a. *Nat Struct Mol Biol* **14**, 591-596 (2007).

6 Sassaman, D. M. *et al.* Many human L1 elements are capable of retrotransposition. *Nat Genet* **16**, 37-43 (1997).

7 Macia, A. *et al.* Epigenetic control of retrotransposon expression in human embryonic stem cells. *Mol Cell Biol* **31**, 300-316 (2011).

8 Naas, T. P. *et al.* An actively retrotransposing, novel subfamily of mouse L1 elements. *Embo J* **17**, 590-597 (1998).

9 Morrish, T. A. *et al.* DNA repair mediated by endonuclease-independent LINE-1 retrotransposition. *Nat Genet* **31**, 159-165 (2002).

10 Beck, C. R. *et al.* LINE-1 retrotransposition activity in human genomes. *Cell* **141**, 1159-1170 (2010).

11 Wei, W. *et al.* Human L1 retrotransposition: cis preference versus trans complementation. *Mol Cell Biol* **21**, 1429-1439 (2001).

12 Garcia-Perez, J. L. *et al.* LINE-1 retrotransposition in human embryonic stem cells. *Hum Mol Genet* **16**, 1569-1577 (2007).

13 Freeman, J. D., Goodchild, N. L. & Mager, D. L. A modified indicator gene for selection of retrotransposition events in mammalian cells. *Biotechniques* **17**, 46, 48-49, 52 (1994).

14 Moran, J. V. *et al.* High frequency retrotransposition in cultured mammalian cells. *Cell* **87**, 917-927 (1996).

15 Alisch, R. S., Garcia-Perez, J. L., Muotri, A. R., Gage, F. H. & Moran, J. V. Unconventional translation of mammalian LINE-1 retrotransposons. *Genes Dev* **20**, 210-224 (2006).

16 Wagstaff, B. J., Barnerssoi, M. & Roy-Engel, A. M. Evolutionary conservation of the functional modularity of primate and murine LINE-1 elements. *PloS one* **6**, e19672, doi:10.1371/journal.pone.0019672 (2011).

17 Wallace, M. R. *et al.* A de novo Alu insertion results in neurofibromatosis type 1. *Nature* **353**, 864-866 (1991).

18 Esnault, C., Casella, J. F. & Heidmann, T. A Tetrahymena thermophila ribozyme-based indicator gene to detect transposition of marked retroelements in mammalian cells. *Nucleic Acids Res* **30**, e49 (2002).