The Microprocessor controls the activity of mammalian retrotransposons

Sara R Heras1,2,6, Sara Macías1,6, Mireya Plass3,4, Noemí Fernández1, David Cano2, Eduardo Eyras3,5, José L García-Perez2 & Javier F Cáceres1

More than half of the human genome is made of transposable elements whose ongoing mobilization is a driving force in genetic diversity; however, little is known about how the host regulates their activity. Here, we show that the Microprocessor (Drosha-DGCR8), which is required for microRNA biogenesis, also recognizes and binds RNAs derived from human long interspersed element 1 (LINE-1), Alu and SVA retrotransposons. Expression analyses demonstrate that cells lacking a functional Microprocessor accumulate LINE-1 mRNA and encoded proteins. Furthermore, we show that structured regions of the LINE-1 mRNA can be cleaved in vitro by Drosha. Additionally, we used a cell culture–based assay to show that the Microprocessor negatively regulates LINE-1 and Alu retrotransposition in vivo. Altogether, these data reveal a new role for the Microprocessor as a post-transcriptional repressor of mammalian retrotransposons and a defender of human genome integrity.

Transposable elements have been a major force in shaping mammalian genomes. At least four distinct classes of mobile DNA elements are dispersed through the human genome, comprising half of its genomic mass1. Most transposable elements are defective ‘fossil’ copies accumulated through human evolution, although a small fraction of non-long terminal repeat (non-LTR) retrotransposons are currently active in the germline and somatic human genomes. LINE-1 (or L1) is a family of non-LTR retroelements in mammals that represents one-fifth of the human genome1. Although >99% of LINE-1 copies are inactive, an average human genome contains 80–100 retrotransposition-competent L1s (RC-L1s) capable of mobilization1,2. In addition, RC-L1–encoded proteins are responsible for the mobilization of non-autonomous non-LTR retrotransposons (short interspersed elements (SINEs) including Alu and SVA)3–5, cellular mRNAs6,7 and selected non-coding cellular RNAs8, which together comprise more than one-tenth of the human genome. Thus, the activity of RC-L1s, Alus and SVAs continues to impact the human genome (as new mobilization events are a constant source of human DNA variation)9–11) and contribute to human disease1.

All characterized active or RC-L1s belong to the youngest subfamily of L1 elements, denoted L1Hs or L1-PAs (ref. 12). RC-L1s are 6-kb-long elements that contain a 5′ untranslated region (UTR) with an internal promoter followed by two non-overlapping open reading frames (ORFs) and end in a short 3′ UTR containing a poly(A) tail1. LINE-1-encoded proteins bind in cis to the same mRNA from which they were translated, thus generating a ribonucleoprotein particle (RNP) that is a proposed retrotransposition intermediate13. Both encoded proteins are required for LINE-1 retrotransposition14, although little is known about how cellular host factors affect and regulate LINE-1 retrotransposition. By contrast, Alu retrotransposons are 0.3-kb-long elements derived from the 7SL RNA and are transcribed by RNA polymerase III15,16. Primate-specific Alu elements contain two monomers separated by an A-rich tract and end in a poly(A) tail required for their trans-retrotransposition by LINE-1–encoded proteins6. An average human genome contains >1 million Alu elements, and there are >6,000 active Alu core sequences per average genome (which belong to the Y and S subfamilies)16.

Owing to the potential negative impact of newly inserted transposable elements, it is highly likely that restriction mechanisms that operate at different levels in the gene expression cascade may have evolved to control a high rate of transposition-mediated mutagenesis. Indeed, several restriction cellular factors have been described that affect mobilization of transposable elements in mammals1. Among the factors known to modulate retrotransposition, several RNA-mediated processes have been previously identified17,18. Notably, piwi-interacting RNAs have been shown to control the expression of LINE-1s specifically in the germ line of mammals and insects1,18.

MicroRNAs (miRNAs) are small RNAs (21–24 nt) that act as key regulators of many biological processes19. They are generated from longer RNA transcripts known as primary miRNAs (pri-miRNAs), which are characterized by the presence of hairpin RNA secondary structures. A complex known as the Microprocessor, comprising the RNase III–type enzyme Drosha and its partner DGCR8, catalyzes the nuclear step of miRNA biogenesis20–22. DGCR8 recognizes the RNA

1Medical Research Council Human Genetics Unit, Institute of Genetics and Molecular Medicine, University of Edinburgh, Edinburgh, UK. 2Centre for Genomics and Oncological Research, Pfizer, University of Granada, Andalusian Regional Government (GENYO), Granada, Spain. 3Computational Genomics Group, Universitat Pompeu Fabra, Barcelona, Spain. 4Bioinformatics Centre, Department of Biology, University of Copenhagen, Copenhagen, Denmark. 5Catalan Institution for Research and Advanced Studies (ICREA), Barcelona, Spain. 6These authors contributed equally to this work. Correspondence should be addressed to J.L.G.-P. (jlgp@genyo.es) or J.F.C (javier.caceres@igmmed.ac.uk).

Received 24 April; accepted 29 July; published online 1 September 2013; doi:10.1038/nsmb.2658
RESULTS

The Microprocessor binds RNAs from transposable elements

To characterize potential new RNA targets of the Microprocessor, we recently performed high-throughput sequencing of RNAs isolated by cross-linking immunoprecipitation (HITS-CLIP) of endogenous DGCR8 protein in human HEK293T cells. We found that DGCR8 binds not only to pri-miRNAs but also to many cellular RNAs. Additionally, we observed that 25% of the binding sites mapped to miRNAs, whereas 4% of binding sites corresponded to long noncoding RNAs (lncRNAs), thus suggesting an expanded role for DGCR8 in controlling the abundance of transposable-element mRNAs. Binding to antisense sequences might be caused by the presence of Alu- and LINE-1-derived sequences within 3' UTRs of human genes27 and the presence of a conserved antisense promoter within the LINE-1 5' UTR26,28, and it is also likely that permissive transcription could lead to the generation of sense and antisense RNA sequences derived from transposable elements. We next analyzed whether DGCR8 directly binds sense RNAs derived from transposable elements, by immunoprecipitation of endogenous DGCR8 protein and analysis of the associated RNAs by semiquantitative RT-PCR. Notably, we confirmed binding of endogenous DGCR8 to sense RNA sequences derived from L1Hs, Alu Y and pri-miR-24-1 transcripts (Fig. 1d,e). In addition, we observed minor binding to Alu S-, Alu J- and SVA-derived RNAs (Fig. 1e). Additional controls revealed a lack of DGCR8 binding to 7SK- and β-actin–derived RNAs. Altogether, these data revealed that DGCR8, the RNA-binding component of the Microprocessor, binds LINE-1–, Alu- and SVA-derived RNAs. These data suggest that the Microprocessor may have a role in controlling the abundance of transposable-element miRNAs.

Figure 1

DGCR8 binds a constellation of transcripts from repetitive elements. (a) Pie chart showing distribution of reads mapping to repetitive elements in sense and antisense orientation. 73% of the clusters map to transposable elements including DNA transposons, LINE-1s, LTR-containing retrotransposons and SINES. (b) Top, distribution of DGCR8-binding sites in a human RC-L1Hs consensus sequence (false discovery rate (FDR) < 0.01). Only sense peaks with a minimum of 29 reads are represented. Bottom, schematic representation of the RC-L1 element. (c) Distribution of major DGCR8-binding sites in a human Alu Y consensus sequence in the sense orientation (FDR < 0.01). (d) Amplification of Alu Y, LINE-1 and 7SK mRNAs by RT-PCR upon immunoprecipitation (IP) of endogenous DGCR8 protein. (e) Real-time RT-PCR analysis of several transposable elements (LINE-1, Alu Y, Alu S, Alu J and SVA) upon immunoprecipitation of endogenous DGCR8 of one representative experiment. Grey bars represent the relative enrichment over control IgG immunoprecipitation (black bars), and data are expressed as a percentage of input. A well-known target of the Microprocessor (pri-miR-24-1) was used as an internal positive control, whereas 7SK and ACTB were used as negative controls.
The Microprocessor regulates the abundance of L1 mRNAs

In order to study the physiological effect of this interaction, we determined LINE-1 mRNA steady-state levels upon a reduction of Microprocessor components in human teratocarcinoma PA-1 cells, which are known to naturally overexpress L1 mRNAs and L1-encoded ORF1 protein. Notably, we observed a two-fold increase in the levels of human RC-L1 mRNA upon short interfering RNA (siRNA)-mediated depletion of DROSHA (Fig. 2a). We further confirmed this by overexpression of dominant negative (DN) forms of both DROSHA and DGCR8 proteins in PA-1 cells, which resulted in a strong accumulation of sense full-length L1 mRNA (Fig. 2b and Supplementary Fig. 2a) while inducing accumulation of pri-miRNAs, as expected (Supplementary Fig. 2b). The detected smaller L1 RNA transcripts (Fig. 2b) probably arise from trans-splicing or by the use of premature polyadenylation signals, as previously described. We also observed that, in agreement with previous studies, expression of DN DROSHA induces an increase in DGCR8 protein levels (Fig. 2b), as the DGCR8 mRNA is itself a substrate of the Microprocessor. Previous studies have demonstrated that DNA methylation of the mammalian L1 promoter (which contains a canonical CpG island) negatively correlates with L1 expression levels. Therefore we analyzed L1 promoter methylation upon transient DROSHA depletion in PA-1 cells by bisulfite DNA conversion, but we found no significant changes at a genome-wide level (Supplementary Fig. 3a,b) or at specific RC-L1s loci (Supplementary Fig. 3c). Thus, the observed changes in L1 mRNA levels upon Microprocessor inactivation are likely not to be caused by alterations in the levels of the LINE-1 promoter. We next examined LINE-1 expression levels in mouse DGCR8-deficient embryonic stem (ES) cells (Dgcr8+/−) and parental ES cells. Notably, Dgcr8−/− ES cells displayed a two-fold increase in the amount of L1 mRNA for all three known active LINE-1 subfamilies (T₅, A₂ and G₂ elements) (Fig. 2c and S.R.H. and S.M., unpublished data). Consistently with this, we also observed elevated levels of L1-ORF1p from T₅ elements in Dgcr8−/− ES cells (Fig. 2d), which correlated with elevated levels of T₅ LINE-1 mRNA levels (Fig. 2e). These observations have led us to hypothesize that DGCR8 binds the L1 mRNA and that the activity of Microprocessor negatively regulates L1 RNA abundance at a post-transcriptional level.

The Microprocessor cleaves the 5’ UTR of LINE-1 mRNAs

We next tested whether LINE-1–derived RNAs could be directly processed by the endonucleolytic activity of Drosha in vitro. We focused on the 5’-UTR region of human RC-L1s, on the basis of the number of CLIP reads mapping to this region (Fig. 1b), its predicted secondary structure (Supplementary Fig. 4a,b) and its relevance for DGCR8 binding (Supplementary Fig. 4c). We generated four overlapping radio-labeled RNA transcripts corresponding to the L1Hs 5’ UTR (Fig. 3a, 300 nt each, and Supplementary Fig. 4d–g). Remarkably, in vitro incubation of these RNAs with immunopurified Microprocessor from cultured cells (with Flag-Drosha) revealed specific cleavage of the regions ranging from 1 to 300 and from 200 to 500 nt within the L1 5’ UTR (Fig. 3b, lanes 4 and 6), as well as of pri-miR-30c-1 radio-labeled RNA that was used as a positive control for the activity of the Microprocessor (Fig. 3b, lane 2). By contrast, we failed to detect processing of the other 5’ UTR–derived RNAs, despite the presence of a predicted stable secondary structure (nucleotides 400–700 and 600–900, Fig. 3b, lanes 8 and 10, and Supplementary Fig. 4f,g). We speculate that processing of these regions by the Microprocessor requires additional sequences that might affect RNA folding and/or recognition. Alternatively, it is possible that DGCR8 binds these regions in a Drosha-independent manner, thus avoiding processing. Importantly, in vitro cleavage of the 1–300 and 200–500 L1-derived RNAs was abolished when a mutant immunopurified Microprocessor unable to bind RNA was used (Flag–DN Drosha, Supplementary Fig. 5a, lanes 3, 6 and 9). Controls revealed that the tagged proteins Flag-Drosha and Flag–DN Drosha still bind endogenous DGCR8 to a similar extent, as revealed by co-immunoprecipitation analyses (Supplementary Fig. 2c). Further analyses revealed that the 285–500 region of the RC-L1 5’ UTR could adopt a predicted structure that could be efficiently processed by Drosha in vitro (Fig. 3c). Notably, disruption of this structure by mutagenesis abolished processing (Fig. 3c, lane 4, and Supplementary Fig. 4h). In sum, these data revealed that the Microprocessor binds and processes RC-L1–derived RNAs.
Figure 3 The 5′ UTR of L1 mRNA is cleaved by immunopurified Drosha in vitro. (a) Schematic representation of four 300-nt in vitro-transcribed fragments spanning the sense L1 5′ UTR region used in b. (b) In vitro processing. Transcripts were incubated (+) or not (−) with immunopurified Microprocessor (with Flag-Drosha). Cleavage products are indicated with asterisks (lanes 2, 4 and 6). An RNA ladder marker indicates sizes in bp on the left. (c) Top, predicted pri-miRNA–like structure of this region. Bottom, in vitro processing of 285–500 L1 5′ UTR region (lane 2), showing abolished cleavage in mutants (mut) that disturb this structure (lane 4 and top).

characterize the sites of Drosha-mediated in vitro processing within the 1–300 L1-derived RNA. Notably, we detected processing at positions +200 to +270 in L1.3 (ref. 42) (Supplementary Fig. 5b). Additionally, we mapped the in vitro processing site of the 285–500 L1-derived RNA at positions +354 to +438 in L1.3 (Supplementary Fig. 5c). To confirm these processing sites in vivo, we used a 5′-phosphate–dependent 5′ rapid amplification of cDNA ends (RACE) assay with total RNA derived from cultured cells (Supplementary Fig. 5d and Online Methods). Notably, we detected in vivo processing sites near position +438 (±1/2 bp) in the RNA fraction derived from PA-1 and HEK293T cells (66% and 77% of sequenced clones, respectively), thus confirming our in vitro results (Supplementary Fig. 5e). Altogether, these data indicate that the Microprocessor recognizes and processes at least two regions of the 5′ UTR of the RC-L1 mRNA and suggest that this activity may interfere with the ability of the RC-L1 mRNA to mobilize.

The Microprocessor regulates L1 retrotransposition in vivo

We next used a retrotransposition assay to analyze the role of the Microprocessor in LINE-1 mobility in cultured human cells. This assay is based on the activation of a reporter gene upon a round of L1 retrotransposition 14. We used a retrotransposition indicator cassette that consists of a blasticidin-resistance gene (BSD) in the antisense orientation relative to a full-length LINE-1 element that is disrupted by an intron in the sense orientation. Upon transfection into cells, the BSD can only be activated after a round of LINE-1 retrotransposition, thus allowing the selection and quantification of cells with new retrotransposition events in culture (Fig. 4a)14,43. Because the LINE-1 5′ UTR has internal promoter activity 1, we included an exogenous strong promoter (CMV) to correct for differences in the expression of the tagged L1 construct, as this could affect the overall rate of mobilization (Fig. 4b). Notably, the retrotransposition rate of a full-length L1 construct (JJ101(L1.3)) increased by five-fold upon overexpression of DN Drosha (Fig. 4c,d). Interestingly, removal of the 5′ UTR from the engineered construct (TAM102(L1.3)) partially abolished the increase in retrotransposition upon Microprocessor depletion (Fig. 4c,d). Furthermore, we observed severely reduced retrotransposition levels with constructs containing a reverse transcriptase (RT)-mutated engineered LINE-1 allelic construct (Fig. 4b,c). We observed an inhibitory effect upon APOBEC3A overexpression in this assay (Fig. 4c), in agreement with a proposed role for APOBEC proteins in inhibiting both LINE-1 and Alu retrotransposition 44. Additionally, co-transfection of a DN DGCR8 expression vector also increased retrotransposition, thus confirming the results observed with DN Drosha (Supplementary Fig. 6a,b). Importantly, expression of either DN Drosha or DN DGCR8 was not toxic to cultured cells (Supplementary Fig. 6c). Finally, we confirmed that engineered L1 retrotransposition increased five-fold upon overexpression of a DN DGCR8 expression vector, using a different retrotransposition assay based on the activation of a neomycin-resistance gene (mneo, also termed aphA1 14, Supplementary Fig. 6d,e). In sum, these data correlate with the processing data and further suggest that the Microprocessor controls LINE-1 retrotransposition.

To further explore this regulation, we next tested LINE-1 constructs that lack a full-length 5′ UTR but contain the four overlapping regions assayed for in vitro processing (Fig. 3a). Notably, L1 retrotransposition increased with constructs containing the 1–300 and 200–500 regions upon Microprocessor depletion, consistent with the in vitro processing data (Fig. 4e,f). Accordingly, when we removed the 5′ UTR and included the predicted stem-loop structure contained in region 285–500 in the engineered construct (Supplementary Fig. 6f and Fig. 3c), we observed a two-fold increase in the retrotransposition rate upon DN Drosha co-transfection (Supplementary Fig. 6g,h). Importantly, a mutated version of this predicted stem-loop structure (Fig. 3c and Supplementary Fig. 4h) or a scrambled version did not significantly change retrotransposition efficiency (Supplementary Fig. 6g,h). Thus, these data agree with the in vitro and in vivo processing data and further suggest that the 5′ UTR of RC-L1s is a target of the Microprocessor during the retrotransposition cycle.

The Microprocessor regulates Alu trans-retrotransposition

We next tested whether Alu-derived RNAs could be processed in vitro by the Microprocessor. To do so, we first generated a 290-nt radiolabeled
RNA transcript corresponding to the core sequence derived from an active AluYa5 element\(^1\)\(^2\). Remarkably, upon in vitro incubation of this RNA with immunopurified Microprocessor (using Flag-Drosha), we observed processing of the AluYa5-derived RNA, but this was not as effective as that observed with the pri-miR-30c-1 radiolabeled RNA\(^4\). Thus, these data further support that the Microprocessor can bind and process RNAs derived from active Alu transposable elements.

Next, in order to determine whether the Microprocessor can control the mobilization of Alu in cultured cells, we used a previously established Alu trans-retrotransposition assay\(^5\)\(^6\)\(^7\) (Fig. 5b). To do this, we tagged the AluYa5 core with a neo\(\_\)Tet cassette\(^8\). As Alu are RNA poly-Amerase III-derived transcripts, the neo\_Tet cassette is based on a self-spliceable intron that activates a neomycin-resistance reporter epitope upon insertion into genomic DNA. In this assay, cells are co-transfected with a LINE-1-derived construct lacking ORF1p (‘ORF2 driver’) that is able to produce ORF2p mediating the mobilization in trans of the tagged spliced Alu RNA (Fig. 5b). The number of G418-resistant foci is used to determine the level of trans-retrotransposition of the Alu core. Notably, the retrotransposition rate of the Alu Ya5 construct increased by 3.5-fold upon overexpression of DN Drosha in cells co-transfected with a driver that contains the 5’ UTR and ORF2 sequences of a RC-L1 (5’ UTR-ORF2-NN, Fig. 5c). Controls revealed severely reduced trans-retrotransposition in cells transfected with only the tagged Alu, as expected (S.R.H. and S.M., unpublished data). As the Alu mobilization assay requires the presence of a driver L1 and a tagged Alu, it is likely that the observed increase in trans-retrotransposition upon Microprocessor depletion reflects both Alu and L1-driver in vivo RNA processing, thus making these assays difficult to quantify. To overcome these limitations, we removed the L1 5’ UTR region and used a RC-L1 ORF2 codon-optimized sequence in the driver construct\(^4\)\(^5\)\(^6\) (plasmid ORF2co-NN, Fig. 5d). As expected, the codon-optimized LINE-1 driver produced more mRNA upon transfection into cultured cells, as determined by quantitative RT-PCR\(^4\)\(^5\)\(^6\) (Supplementary Fig. 6i). Remarkably, we observed a 2.5-fold increase in the mobilization rate of an Alu Ya5 construct upon co-transfection with the ORF2co-NN driver and DN Drosha (Fig. 5d). Notably, assays conducted in the presence of DN DGCR8 produced similar results (Supplementary Fig. 6j). Thus, our data suggest that the Microprocessor negatively regulates the mobilization of Alu retrotransposons in cultured cells. These results are consistent with the existence of RNA polymerase III-derived pri-miRNAs and with previously characterized miRNAs encoded within Alu sequences\(^7\). Additionally, these data further support that DGCR8 can bind a constellation of transposable element–derived RNAs despite having different sequences or changes in sequences within distinct subfamilies (Fig. 1 and Supplementary Tables 2 and 3).

DGCR8-dependent regulation of LINE-1s is miRNA independent

To further analyze the regulation of transposable elements by the Microprocessor, we used luciferase-based reporter constructs containing the 5’ UTR of an RC-L1 element to monitor the in vivo processing
Figure 5 Alu is processed in vitro, and its retrotransposition is regulated by the Microprocessor. (a) In vitro processing of Alu Ya5. Cleavage of in vitro–transcribed Alu Ya5 core RNA upon addition of Drosha (+) using Flag-Drosha immunoprecipitates (lane 4). Processing of pri-miR-30c-1 is shown as a positive control (lane 2). (b) Rationale of the Alu trans-mobilization assay. In the scheme, a cartoon depicts the ORF2 driver containing an exogenous promoter (white arrow, CMV) and the coding sequence for LINE-1 ORF2 (gray rectangle). In addition, a tagged Alu (gray arrow) with a neoTet retrotransposition indicator cassette (black-boxed backwards Neo) containing an encoded poly(A) tail (A3β) is depicted. The neoTet cassette contains a self-splicable group I intron (black curvy lane). Lollipops, polyadenylation signals. With this configuration, expression of a neomycin-resistance gene can only occur upon a round of trans-retrotransposition (bottom). (c,d) Trans-retrotransposition experiments using 5′ UTR-ORF2-NN (c) and ORF2co-NN (d) as a driver (scheme shown). Each image shows representative data from Alu trans-retrotransposition assays conducted in duplicate. β-arrestin is used as a control, as it does not affect Alu retrotransposition. Relative retrotransposition activity is quantitated on the right. Data were normalized by transfection efficiency and toxicity. Shown is average of three independent biological replicates ± s.d.

of this region in cultured cells. We focused on the 5′ UTR region of human RC-L1s, as we showed that it could be processed by the Microprocessor in vitro and in vivo (Figs. 3 and 4). Interestingly, reporters containing the human RC-L1 5′ UTR element displayed a small but significant increase in luciferase activity upon Drosha depletion (P = 0.04), thus suggesting in vivo processing by the Microprocessor (Supplementary Fig. 7a,b). Parallel assays with a luciferase reporter containing a canonical miR-18a–binding site in its 3′ UTR served as a control of miRNA levels and thus of Drosha function. As expected, the miR-18a–containing reporter showed increased levels of luciferase activity upon Drosha depletion (Supplementary Fig. 7a,b). Transfection of reporters containing the 5′ UTR of a human RC-L1 element with plasmids expressing DN DGCR8 or DN Drosha produced similar results (S.R.H. and S.M., unpublished data). Additionally, evolutionarily older LINE-1 5′ UTRs generated similar results upon Microprocessor depletion (S.R.H. and S.M., unpublished data), consistent with the presence of CLIP reads mapping to evolutionary older LINE-1s (L1PA2–L1PA4, Supplementary Fig. 1). Similarly, when we used a luciferase construct containing the 5′ UTR of an active mouse LINE-1 (Tc class, L1spa29), which also adopts a potential secondary ordered structure (Fig. 6a and Supplementary Fig. 4i), we observed a three-fold increase in reporter activity in Dgcr8−/− ES cells compared to parental cells (Fig. 6b). These data also suggest that the Microprocessor can negatively regulate RC-L1

Figure 6 LINE-1 regulation by the Microprocessor is Dicer and miRNA independent. (a) Schematic representation of firefly luciferase reporters containing the 5′ UTR from a mouse RC-L1 L1spa– (ml1spa) and controls (simian virus 40 promoter (SV40) and SV40-miR18a, containing a target site for miR-18a in the 3′ UTR). LCS, firefly luciferase ORF. (b) Luciferase activity in relative light units (RLU) for Dgcr8−/− Dicer1+/− ES cells and parental wild-type cells (v6.5 and f/f Dicer, respectively) transfected with vectors described in a. Error bars, s.d. (n = 3 biological replicates). For L1 levels, *P = 0.0007 in Dgcr8−/− and *P = 0.001 in Dicer1−/−; for miR-18a levels, *P = 2.17 × 10−17 in Dgcr8−/− and *P = 0.001 in Dicer1−/− (one-tailed t test). (c) Cartoon depicting the LINE-1 engineered constructs assayed in cultured cells (following nomenclature in Fig. 4). Construct JJ101(L1.3)Δ′ UTR is a derivative of plasmid JJ101(L1.3) that lacks LINE-1 3′ UTR. (d) Retrotransposition assays in HeLa cells co-transfected with JJ101(L1.3) or JJ101(L1.3)Δ′ UTR vectors plus the indicated expression plasmids (β-arrestin (β-arr) or DN Drosha), as indicated at left. Each image shows representative data from L1 retrotransposition assays conducted in triplicate. The graph shows a quantification of the assay, and data are presented as the proportion of the activity in cultures co-transfected with the plasmid expressing the negative control (β-arrestin) and are normalized by transfection efficiency and toxicity. Shown are averages of three independent biological replicates ± s.d.
expression in mammalian cells at the post-transcriptional level. Controls with a luciferase reporter containing a miR-18a-binding site in Dicer8−/− ES cells revealed a significant increase in luciferase activity when compared to wild-type cells. On the other hand, previous studies have shown that Dicer depletion results in increased levels of LINE-1 mRNAs27. As both the Microprocessor and Dicer are involved in miRNA biogenesis, we next tested whether changes in gene expression of luciferase constructs containing the 5′ UTR of an active mouse LINE-1 are a consequence of lacking canonical miRNAs. To do this, we transfected mouse Dicer1−/− ES cells49 with the same reporter containing the 5′ UTR of an active mouse LINE-1. Importantly, the same reporter displayed a small increase in luciferase activity in the absence of Dicer1 in knockout cells compared to parental cells (Fig. 6a,b), thus confirming a modest role for Dicer-dependent small RNAs in controlling LINE-1 mRNA abundance47. Notably, the effect in luciferase expression was greater in the absence of DGC88 than in the absence of Dicer, a result further suggesting a direct regulation of transposable element–derived RNAs by the Microprocessor rather than regulation due to defects in small-RNA biogenesis. As expected, parallel experiments revealed a significant (P < 0.05 (one-tailed t test), n = 3 biological replicates) increase in luciferase activity produced by the reporter containing the miR-18a in Dicer1−/− ES cells. Additional northern blot controls confirmed lack of canonical miRNAs in Dicer1−/− and Dicer8−/− ES cells (S.R.H. and S.M., unpublished data). These data suggest a miRNA-independent role for DGCR8 in regulating LINE-1 expression and retrotransposition, rather than an indirect effect due to the absence of miRNAs.

To further support these findings, we generated a human RC-L1 retrotransposition construct lacking the 3′ UTR region (Fig. 6c, construct JJ101(L1.3)A′ UTR) and determined its retrotransposition rate upon Microprocessor depletion. We reasoned that, as most miRNAs interact with 3′ UTRs19,50, this construct would allow testing for whether the effects of Microprocessor depletion on engineered retrotransposition were mediated by cellular miRNAs. Importantly, previous studies have demonstrated that the 3′ UTR of RC-L1s is not required for engineered LINE-1 retrotransposition44, and we confirmed these findings (Fig. 6d, samples co-transfected with β-arrestin). Notably, we observed a four-fold increase in the rate of engineered LINE-1 retrotransposition for the construct that lacks the 3′ UTR, upon DN Drosha co-transfection (Fig. 6d, JJ101(L1.3)A′ UTR). Parallel control experiments revealed a similar increase in retrotransposition for the wild-type construct that contains the 3′ UTR region (Fig. 6d, JJ101(L1.3)). In sum, these data are consistent with L1 retrotransposition being regulated by the direct processing of RC-L1 mRNAs rather than by the absence of inhibitory miRNAs upon Microprocessor depletion. Altogether, our data suggest that the Microprocessor can negatively regulate RC-L1 expression in mammalian cells at the post-transcriptional level, probably in an miRNA independent manner.

DISCUSSION

The activity of LINE-1 retrotransposons may pose a risk for the genome of an individual. Previous studies have demonstrated that heritable LINE-1 and SINE insertions accumulate mostly during embryogenesis34,51,52. Additionally, recent evidence has also demonstrated that human LINE-1s and Alus are active in selected somatic cells and in several types of tumors9,37,53–56. Thus, different mechanisms may have evolved to reduce retrotransposition in somatic cells. Indeed, DNA methylation of mammalian LINE-1 promoters is a well-known mechanism to downregulate expression and subsequent retrotransposition of LINE-1 elements56. Failure to transcriptionally repress LINE-1 expression may negatively impact the human genome in a myriad of ways. Besides DNA methylation36, few host factors have been previously described to negatively regulate L1 or Alu retrotransposition, including APOBEC proteins44, proteins associated with the L1 ORF1 protein and its ribonucleoprotein57, the exonuclease Trex 1 (ref. 58), RNA interference–like mechanisms driven by LINE-1 (ref. 17) or piwi-interacting RNAs18 and MOV10 (refs. 59,60).

Here, we describe a new role for the Microprocessor in controlling the activity of mammalian LINE-1s and SINEs. We have shown that the core components of the Microprocessor directly bind a constellation of mammalian transposable element–derived RNAs, including known active LINE-1s, Alu and SVAs. Additionally, we have demonstrated that the Microprocessor can process structures within the 5′ UTR of RC-L1s in vitro and probably in vivo, and potentially, other factors that co-purify with the Microprocessor might be able to modulate this activity in vivo21. Similarly, we have demonstrated that the core from an active Alu Ya5 is a substrate for the Microprocessor. Furthermore, the use of luciferase-based reporters and engineered retrotransposition constructs lacking 3′-UTR sequences suggests that Microprocessor regulation of LINE-1 mRNAs is both miRNA and Dicer independent (Fig. 6). Additionally, we have demonstrated that the mobilization of full-length engineered L1s in cultured cells can be controlled by the Microprocessor in vivo, probably by binding and processing of 5′ UTRs. However, we also speculate that other sites of Microprocessor processing may exist within coding LINE-1 sequences, on the basis of identified binding sites. It is worth mentioning that additional binding sites for DGCR8 were found in the antisense strand of RC-L1, Alu and SVA sequences, although their role remains to be determined. As the whole genome is transcribed and LINE-1s are dispersed through the genome56, we propose that these binding sites might be present within antisense LINE-1 and/or SINE sequences located in gene transcripts (probably within 3′ UTR sequences) and/or lncRNAs29. However, additional experiments are required to determine whether these derived antisense RNAs are processed by the Microprocessor and whether they have a role in transposable-element biology or genome biology. Although LINE-1 retrotransposons are transcriptionally silenced by DNA methylation in germ cells and in most somatic human tissues3,36, we propose that the Microprocessor restricts L1 retrotransposition at a post-transcriptional level, acting against endogenous retrotransposons that escape transcriptional silencing in mammalian cells. Furthermore, the ability of the Microprocessor to restrict L1 retrotransposition in somatic cells may also have an impact on other transposable elements that require L1-encoded proteins for their mobilization; such is the case for Alu and SVA elements.
In sum, we propose a model whereby the Microprocessor complex binds the RC-L1 mRNA within the nucleus, probably cotranscriptionally, and cleaves hairpin structures contained in the L1 mRNA (Fig. 7). We speculate that this cleavage results in LINE-1 mRNA destabilization, with a concomitant decrease in LINE-1-encoded proteins, ultimately reducing retrotransposition rates. This mRNA destabilization, with a concomitant decrease in LINE-1–mammalian retrotransposition at a post-transcriptional level.

**METHODS**

Methods and any associated references are available in the online version of the paper.

**Accession codes.** Sequencing raw data for endogenous and overexpressed DGCR8 HITS-CLIP have been deposited in the Gene Expression Omnibus database, under accession code GSE39086.

**ACKNOWLEDGMENTS**

We thank N. Hastie and J.V. Moran for comments and critical reading of the manuscript. We also are grateful to M. Madej, J. Reddington and R. Meehan for advice on DNA methylation assays and to I. Adams for discussions. We thank R. Blelloch (University of California San Francisco, San Francisco, California, USA), V.N. Kim (Seoul National University, Seoul, Korea), S.L. Martin (University of Colorado School of Medicine, Aurora, Colorado, USA), A. Roy-Engel (Tulane Cancer Center, New Orleans, LA USA), T. Heidmann (Institut Gustave Roussy, Villejuif, France and Université Paris-Sud, Orsay, France) and J.V. Moran (Howard Hughes Medical Institute, University of Michigan Medical School, Ann Arbor, Michigan, USA) for their generous gifts of reagents. S.M. was supported by a long-term European Molecular Biology Organization postdoctoral fellowship. S.R.H. was supported by a Marie Curie Intra-European Fellowship and a Marie Curie CIG-Grant (PCIG10-GA-2011-303812). M.P. and E.E. were supported by the Spanish Ministry of Science (BIO2011-23920) and by the Sandra Ibarra Foundation (IICS-55007420). J.L.G.-P. was supported by Core funding from the Medical Research Council and by the Wellcome Trust (grant 095518/B/11/2).

**AUTHOR CONTRIBUTIONS**

S.R.H., S.M., J.L.G.-P. and J.F.C. conceived of, designed and interpreted the experiments. S.R.H., S.M., D.C. and N.F. performed the experiments and data analysis. S.R.H., S.M., J.L.G.-P. and J.F.C. interpreted the experiments. M.P. and E.E. provided advice on DNA methylation assays and to I. Adams for discussions. We thank N. Hastie and J.V. Moran for comments and critical reading of the manuscript. We also are grateful to M. Madej, J. Reddington and R. Meehan for advice on DNA methylation assays and to I. Adams for discussions. We thank R. Blelloch (University of California San Francisco, San Francisco, California, USA), V.N. Kim (Seoul National University, Seoul, Korea), S.L. Martin (University of Colorado School of Medicine, Aurora, Colorado, USA), A. Roy-Engel (Tulane Cancer Center, New Orleans, LA USA), T. Heidmann (Institut Gustave Roussy, Villejuif, France and Université Paris-Sud, Orsay, France) and J.V. Moran (Howard Hughes Medical Institute, University of Michigan Medical School, Ann Arbor, Michigan, USA) for their generous gifts of reagents. S.M. was supported by a long-term European Molecular Biology Organization postdoctoral fellowship. S.R.H. was supported by a Marie Curie Intra-European Fellowship and a Marie Curie CIG-Grant (PCIG10-GA-2011-303812). M.P. and E.E. were supported by the Spanish Ministry of Science (BIO2011-23920) and by the Sandra Ibarra Foundation (IICS-55007420). J.L.G.-P. was supported by Core funding from the Medical Research Council and by the Wellcome Trust (grant 095518/B/11/2).

**COMPELLING FINANCIAL INTERESTS**

The authors declare no competing financial interests.
44. Bogerd, H.P. et al. Cellular inhibitors of long interspersed element 1 and Alu retrotransposition. *Proc. Natl. Acad. Sci. USA* **103**, 8780–8785 (2006).
45. 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 (2011).
46. Han, J.S. & I Brooke, J.D. A highly active synthetic mammalian retrotransposon. *Nature* **429**, 314–318 (2004).
47. Borchert, G.M., Lanier, W. & Davidson, B.L. RNA polymerase III transcribes human microRNAs. *Nat. Struct. Mol. Biol.* **13**, 1097–1101 (2006).
48. Gui, S. & Cáceres, J.F. The multifunctional RNA-binding protein hnRNP A1 is required for processing of miR-18a. *Nat. Struct. Mol. Biol.* **14**, 591–596 (2007).
49. Babiarz, J.E., Ruby, J.G., Wang, Y., Bartel, D.P. & Blelloch, R. Mouse ES cells express endogenous shRNAs, siRNAs, and other Microprocessor-independent, Dicer-dependent small RNAs. *Genes Dev.* **22**, 2773–2785 (2008).
50. He, L. & Hannon, G.J. MicroRNAs: small RNAs with a big role in gene regulation. *Nat. Rev. Genet.* **5**, 522–531 (2004).
51. Garcia-Perez, J.L. et al. LINE-1 retrotransposition in human embryonic stem cells. *Hum. Mol. Genet.* **16**, 1569–1577 (2007).
52. Kano, H. et al. L1 retrotransposition occurs mainly in embryogenesis and creates somatic mosaicism. *Genes Dev.* **23**, 1303–1312 (2009).
53. Musti, A.R. et al. Somatic mosaicism in neuronal precursor cells mediated by L1 retrotransposition. *Nature* **435**, 903–910 (2005).
54. Lee, E. et al. Landscape of somatic retrotransposition in human cancers. *Science* **337**, 967–971 (2012).
55. Evory, G.D. et al. Single-neuron sequencing analysis of L1 retrotransposition and somatic mutation in the human brain. *Cell* **151**, 483–496 (2012).
56. Solyom, S. et al. Extensive somatic L1 retrotransposition in colorectal tumors. *Genome Res.* **22**, 2328–2338 (2012).
57. Goodier, J.L., Cheung, L.E. & Kazazian, H.H. Mapping the LINE1 ORF1 protein interactome reveals associated inhibitors of human retrotransposition. *Nucleic Acids Res.* doi:10.1093/nar/gkt512 (2013).
58. Stetson, D.B., Ko, J.S., Heidmann, T. & Medzhitov, R. Trex1 prevents cell-intrinsic initiation of autoimmunity. *Cell* **134**, 587–598 (2008).
59. Arjan-Odedra, S., Swanson, C.M., Sherer, N.M., Wolinsky, S.M. & Malim, M.H. Endogenous MOV10 inhibits the retrotransposition of endogenous retroelements but not the replication of exogenous retroviruses. *Retrovirology* **9**, 53 (2012).
60. Goodier, J.L., Cheung, L.E. & Kazazian, H.H. MOV10 RNA helicase is a potent inhibitor of retrotransposition in cells. *PLoS Genet.* **8**, e1002941 (2012).
ONLINE METHODS
Cross-linking immunoprecipitation. High throughput sequencing and cross-linking immunoprecipitation (HITS-CLIP) was performed as previously described24. HITS-CLIP data available at the GEO database under the accession number GSE39086 was used to profile the distribution of DGCGR8 CLIP reads on retrotransposons24. Reads were mapped to the genome with bowtie24.

Small-RNA read mapping. To profile the distribution of reads on retrotransposons, only those overlapping annotated retrotransposons were used. This was done to maximize the number of reads considered, as there may be indels and/or nucleotide substitutions in the genome, compared to the consensus sequence, that may reduce the number of mapped reads. In the case of L1Hs and Alu elements, the annotation of the repetitive element was taken from the repeatmasker table of hg18 from UCSC (http://genome.ucsc.edu/)25 and the consensus sequences from Repbase25. In the case of older LINE-1 elements (L1PA2-PAS), the consensus sequences from ref. 12 were used to annotate L1PA2-5 elements in the human genome. In the case of SVA elements, we used two known active SVA sequences as queries (SVA-MAST2 and SVA.2)25. For each retrotransposon sequence, the overlap between the mapped reads and the annotated retrotransposons was calculated with fjoin26. Next, each of the annotated retrotransposons in the genome that had overlapping reads was aligned to its corresponding consensus sequence with exonerate (parameters -m a:l –score 1 –n 1)65. For L1Hs and Alu elements, only those reads mapping to annotated retrotransposons that had an alignment coverage ≥90% and a sequence identity ≥90% with its corresponding consensus sequence were further used. Owing to the lower conservation among older LINE-1 subfamily members, those reads that had an alignment coverage ≥90% and a sequence identity ≥90% were also considered in the analysis. Finally, the position of HITS-CLIP reads was transferred on the consensus sequence with the alignment as a guide. If a read mapped to more than one position in the consensus sequence and/or in the genome, one was selected at random to build the profile. To identify significant positions in the consensus sequences covered by reads, we calculated the false discovery rate (FDR) as described in ref. 66 by computing the background frequency after randomly placing the same reads on the consensus sequences for 1,000 iterations. Only those positions with an FDR <0.01 were considered significant and displayed in the final plots. This procedure was performed independently for sense and antisense reads. (Supplementary Fig. 1 and Supplementary Table 3).

Analysis of expressed L1 mRNAs. Total RNA was extracted from HEK293T cells with Trizol (Invitrogen), and 1 µg was treated with DNase I (Invitrogen), reverse transcribed with M-MLV RT (Promega) and used in RT-PCR reactions with LINE-1 ORF1 primers as described26,24. Amplified products were cloned in pGEMT-Easy (Promega) and at least 25 independent clones sequenced. Sequences were analyzed with RepeatMasker (http://repeatmasker.org/).

RNA extraction and quantitative RT-PCR. Total RNA was isolated from cells with Trizol (Invitrogen) and treated with DNase (RNQ1 DNase, Promega, M601A) and checked for DNA contamination by PCR. 500 ng of total RNA was used for quantitative RT-PCR analyses with the SuperScript III Platinum SYBR Green One-Step qRT-PCR Kit (Invitrogen, 11736-051). Data were analyzed with Bio-Rad CFX Manager software. All experiments show the average and s.d. of at least three independent biological replicates. Primers for qRT-PCR analysis are listed in Supplementary Table 4.

Northern blot analyses. Hybridization was performed at 42 °C with an antisense riboprobe to the first 150 nt of a human RC-L1 (L1.3). For normalization purposes, the membrane was stripped and rehybridized with a GAPDH probe. Radioactive signals were analyzed with a Fuji FLA-5100 Phosphorimager.

Immunoprecipitations and RT-PCR. For immunoprecipitation of endogenous DGCGR8, anti-DGCGR8 antibody (Abcam, ab90579) and Dynabeads (10001D, Invitrogen) were used. For Flag-Drosha and Flag–DN Drosha, anti-Flag M2 affinity (Sigma, A220) was used. For T7-DGCR8, T7Tag Antibody Agarose (Millipore, 69026) was used. Primers used for these analyses are listed in Supplementary Table 4.

Cell lines and antibodies. HEK293T, PA-1 and HeLa cells were purchased from ATCC. Cytogenetic authentication of HeLa and PA-1 cell lines was performed by spectral karyotyping (SKY)-FISH. Cell lines used in this study were tested for mycoplasma contamination at least monthly. HEK293T and HeLa cells were grown under standard conditions in Dulbeccos Modified Eagle Medium (DMEM). Human teratocarcinoma PA-1 cells were cultured in Minimum Essential Media (MEM) supplemented with 10% heat-inactivated FBS. Mouse embryonic stem cells (mES) were grown on gelatin-coated plates (Sigma) without feeders in DMEM-high glucose supplemented with 15% (v/v) FBS (GIBCO-Invitrogen). Dgcr8-/-, mES cells were purchased from Novus Biologicals (NRA1-19349) and the parental strain (v6.5) from Thermo Scientific (M1402). Dicer1-/- and If Dicer1 were kindly provided by R. Bl loaf (UCSF)25. The rabbit anti-DGCR8 was from Abcam (ab90579), the rabbit anti-Drosha antibody was from Upstate (07-717) and the mouse/goat anti-β-actin antibody was from Sigma (T4026), whereas the anti-mouse ORF1p antibody was kindly provided by S. Martin (U of Colorado)13. The working dilutions were 1:250, 1:500, 1:5,000 and 1:2,000, respectively. Uncropped forms of western blots are shown in Supplementary Figure 8.

siRNA depletions. Knockdown of Drosha in PA-1 cells was achieved with two rounds of siRNA transfection with Dharmafect 4 solution (Dharmacon) according to the manufacturer’s instructions. Briefly, cells were seeded in six-well plates to 40% confluency and after 24 h were transfected with 25 nM of each siRNA (DROSHA or control, from Dharmacon) and 10 µl of the transfection reagent. The transfection medium was replaced after 24 h, and cells were grown for another 24 h. Cells were then retransfected according to the same protocol. siRNAs against Drosha and non-targeting siRNAs (control) were purchased from Dharmacon (L-016996-00 and D-001810-02, respectively). Cells were collected 24 h after the second transfection for analyses.

Radioactive RNA labeling and in vitro processing reactions. Templates for RNA synthesis and radiolabeling were obtained by PCR from a plasmid containing the 5’UTR from L1.3 (L1.3S-FF; primers listed in Supplementary Table 4). The forward or 5’-end primer included the sequence of a T7 promoter. Additionally, the pri-miR-30c-1 was cloned in pGEMT-easy (Promega) and NdeI digested before transcription. Transcription reactions were performed with T7 RNA polymerase (Ambion, AM2082) in the presence of 40 µmol of [α-32P]UTP. Probes were gel purified, phenol extracted and ethanol precipitated according to standard procedures. Approximately 50,000 c.p.m. of each probe was incubated with 15 µl of unpurified Flag-Drosha, Flag–DN Drosha beads or control immunoprecipitates in the presence of buffer A (0.5 mM ATP, 20 mM creatine phosphate and 3.2 mM MgCl2). Reactions were incubated for 30 min at 30 °C, and then standard phenol/chloroform extraction and ethanol precipitation were performed. RNAs were resolved in an 8–10% 1× TBE polyacrylamide urea gel. Gels were analyzed with a Phosphorimager (FLA-5100 Phosphorimager, Fuji).

Cell transfection and dual luciferase assays. HEK 293T cells were co-transfected with a 5’ UTR L1 luciferase construct (L1.3S-FF) and dominant-negative forms of DGCGR8 or Drosha. A complete list of plasmids used in this study can be found in Supplementary Table 5. Mouse ES cells (Dgcr8-/-, Dicer1-/- and their wild-type counterparts, v6.5 and If/dicer) were transfected with mL1spa construct and SV40 and SV40-miR18a plasmids as controls, with Lipofectamine 2000 (Invitrogen). Cells were collected 48 h post-transfection. For siRNA studies in HeLa cells, DNA plasmids were transfected during the second round of siRNA transfection with Dharmafect DUO (Dharmacon), and cells were collected 24 h after transfection. Efficient Drosha knockdown was monitored by co-transfection of an miR-18a reporter system (SV40-miR18a). In all assays, a plasmid expressing the Renilla luciferase gene was used as an internal control. Cells were lysed with passive lysis buffer (Promega), and the levels of firefly and Renilla luciferase activity were measured with the Promega Dual Luciferase Reaction system. The data are expressed as a ratio of firefly luciferase activity to Renilla luciferase activity and normalized to mock value (or wild-type value). Luminescence was measured with a Monolight 3010 luminescence (PharMingen).

Mapping cleavage sites by primer extension. 170 nM of unlabeled RNA was subjected to in vitro processing with or without immunopurified Drosha. RNA was phenol purified, precipitated and denatured for 3 min at 95 °C. Primer extension was carried out with 5’-end-labeled oligonucleotides pairing to the 3’ end of each of the RNA templates used. Primer annealing and extension was carried out in the
Toxicity of DN Drosha, DN DGCR8, and other RNA-binding proteins in cultured human cells can accommodate multiple LINE-1 retrotransposition events. For this reason, retrotransposition assays were carried out in HeLa-HA cells and a previously established transient Line-1 retrotransposition assay was used. 

We used HeLa-HA cells and a previously established transient LINE-1 retrotransposition assay was used. The retrotransposition efficiency assay was performed as previously described. With this configuration, the LINE-1 retrotransposition efficiency assay was performed as previously described.

A slightly modified version of a previously published protocol was used. Briefly, HeLa-HA cells were co-transfected with a driver vector containing a blasticidin-resistance cassette and selecting BSD-resistant foci were counted manually.

Bisulfite DNA sequencing. Genomic DNA from mock siRNA– and Drosha siRNA–transfected PA-1 cells was extracted and purified with a Wizard DNA genomic purification kit (Promega). Bisulfite conversion of genomic DNA was performed with EZ DNA methylation-Gold Kit (Zymo Research D5005) according to the manufacturer's instructions. Bisulfite DNA sequencing was performed globally, as described. The methylation status of the LINE-1 5' UTR was analyzed with mfold, as described. Bisulfite DNA sequencing was performed globally, as described. The methylation status of the LINE-1 5' UTR was analyzed with mfold, as described.

Mapping cleavage sites in vivo with 5'-phosphothioate-labeled siRNA. Line-1 retrotransposition activity in human genomes was analyzed by 5'-phosphothioate-labeled siRNA. The retrotransposition activity was analyzed by 5'-phosphothioate-labeled siRNA. The retrotransposition activity was analyzed by 5'-phosphothioate-labeled siRNA. The retrotransposition activity was analyzed by 5'-phosphothioate-labeled siRNA. The retrotransposition activity was analyzed by 5'-phosphothioate-labeled siRNA. The retrotransposition activity was analyzed by 5'-phosphothioate-labeled siRNA. The retrotransposition activity was analyzed by 5'-phosphothioate-labeled siRNA.