Retrotransposons jump into alternative-splicing regulation via a long noncoding RNA

Reini F Luco

A conserved long noncoding RNA expressed at the 5S rDNA ribosomal locus has acquired a novel function in alternative-splicing regulation in primates, owing to the insertion of a mobile Alu element. This discovery opens new perspectives regarding the roles of transposable elements in expanding the human transcriptome and may be applied as a biotechnology tool to drive gene-specific changes in alternative splicing.

It was originally thought that most of the human genome is functionally irrelevant because it does not encode any known proteins. However, intronic RNA, repetitive and transposable elements, small RNAs and long noncoding RNAs (lncRNAs) are increasingly being shown to perform a variety of important regulatory functions essential for genome organization, stability and evolution[1,2].

Shan and colleagues[3] have now expanded the roles of noncoding RNAs in gene-expression regulation by identifying a dual function of the human lncRNA h5S-OT in transcription and precursor-mRNA (pre-mRNA) splicing, owing to the integration in primates of an antisense Alu element in the 5S-OT locus. 5S-OT is a highly conserved lncRNA that is expressed within the 5S rDNA locus and is important for expression of the 5S ribosomal subunit (**Fig. 1a**). The 5S rDNA locus is the only ribosomal subunit clustered apart in tandem repeats. Interestingly, only approximately one-tenth of the repeat is actually transcribed, and in the early 1990s, it was shown that removal of a GC-rich DNA sequence upstream of the 5S rDNA locus impairs 5S rRNA transcriptional activity *in vitro*[4,5]. Shan and colleagues[3] have now shown, in mouse and human cells, that this sequence corresponds to the CpG-rich, RNA polymerase II–dependent promoter of the lncRNA 5S-OT. By physically interacting with its own promoter, 5S-OT promotes its own expression and the transcription of the RNA polymerase III–dependent 5S rRNA (**Fig. 1a**). Surprisingly, knockdown of 5S-OT also changed the splicing of more than 200 exons (approximately 4% of all alternatively spliced exons) only in human cells, and there was no preference toward more inclusion or exclusion of the regulated exons in the absence of the lncRNA.

When investigating what the human h5S-OT-dependent exons might have in common, the authors[3] found that all exons shared an Alu element within a 2-kb window upstream or downstream of the regulated exon. Alu elements are the most abundant primate-specific retrotransposable elements in the human genome and belong to the short interspersed nuclear elements (SINE) family of repetitive elements[6] Even though they may be deleterious when they are inserted in a coding region, these elements are also known to expand the repertoire of regulatory elements by adding or deleting transcriptional and post-transcriptional signals, thereby reshaping the human transcriptome[6]. In particular, when they are inserted at intronic regions in an antisense orientation, Alu elements contain several cryptic splice sites that can be recognized by the splicing machinery, thereby leading to incorporation of Alu elements into mature transcripts[7].

Shan and colleagues[3] have now found a new mechanism of alternative-splicing regulation by Alu elements that is independent of the insertion of cryptic splice sites. In its 3′ end, the human h5S-OT lncRNA contains an antisense Alu element, in which the poly(A) tail has evolved into a pyrimidine-rich (Py) sequence similar to the polypyrimidine tract (PPT) necessary for recruitment of splicing regulators, such as U2AF or PTB (**Fig. 1b**). Protein pulldown assays of biotinylated h5S-OT identified U2AF65 as the cofactor involved in h5S-OT-dependent regulation of alternative splicing, and U2AF65 knockdown affected ~50% of 5S-OT-sensitive exons, thus supporting a U2AF65-based mechanism of alternative-splicing regulation by h5S-OT. U2AF65 is a core splicing regulator that binds at intron-exon junctions and is necessary for assembly of the splicing machinery and intron excision[8]. The authors propose that by Alu–anti-Alu pairing, the Py sequence of the h5S-OT lncRNA is targeted to an Alu-containing intron, thereby inducing recruitment of U2AF65 and assembly of the splicing machinery at the target point, with subsequent inclusion of the downstream exon (**Fig. 1b**). Surprisingly, even though most of the alternatively spliced exons analyzed showed the presence of Alu elements within their flanking introns, only 4% were affected by down-regulation of h5S-OT, thus suggesting that the specificity of this novel regulatory mechanism is not entirely understood.

An interesting twist in the story is the potential use of h5S-OT as a targeting tool to modulate the alternative splicing of specific genes. Shan and colleagues[3] took advantage of the newly identified mechanism of splicing regulation to change the inclusion levels of a specific exon by replacing the anti-Alu sequence of the h5S-OT lncRNA with an antisense sequence complementary to the intronic region flanking the targeted alternatively spliced exon[3] (**Fig. 2**). In concordance with the polar effect of U2AF65 in alternative-splicing regulation[8], U2AF65 induced exon inclusion when it was targeted upstream of the regulated exon but

Reini F. Luco is at the Institute of Human Genetics, CNRS UPR1142, University of Montpellier, Montpellier, France.
e-mail: reini.luco@igh.cnrs.fr

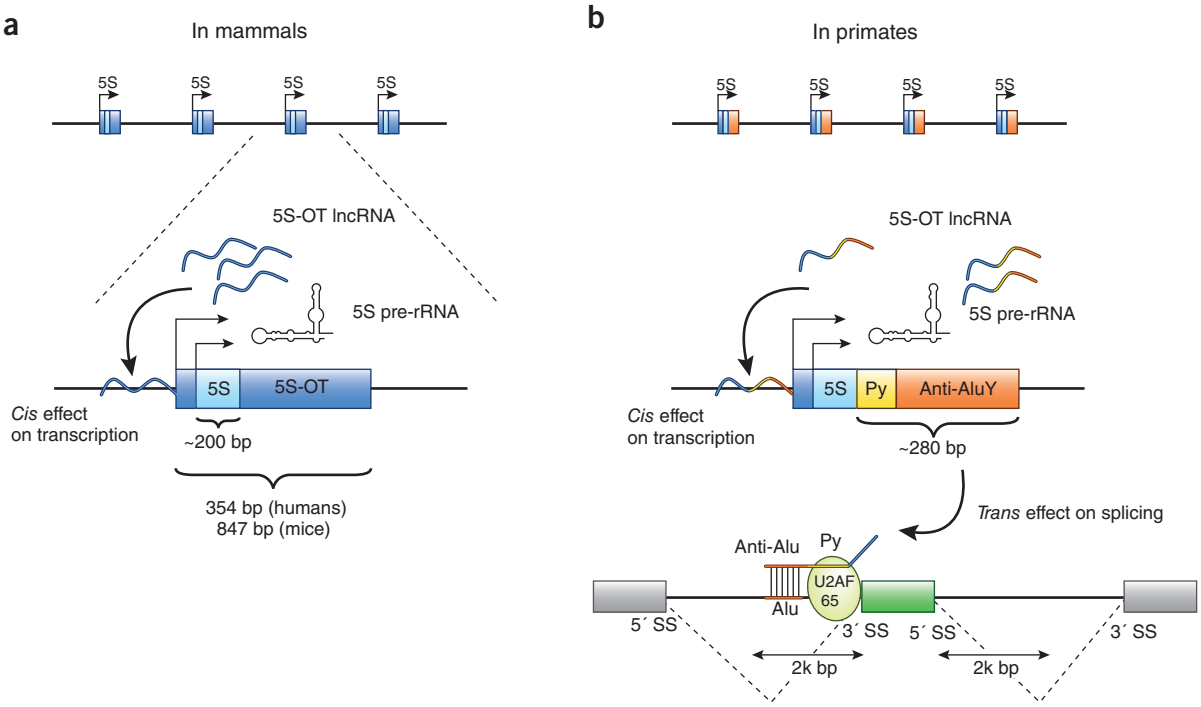

Figure 1 A primate-specific role of the lncRNA 5S-OT in transcription and alternative splicing. (**a**) Mice and humans express the lncRNA 5S-OT (dark blue) at the 5S rDNA locus (light blue). 5S-OT promotes expression of its own promoter and the 5S ribosomal subunit in *cis*. (**b**) In humans, insertion of an Alu element in the antisense direction (orange) in the 5S-OT locus has expanded the lncRNA function to a *trans* regulatory effect on alternative splicing. Via Alu–anti-Alu pairing with Alu elements within 2 kb upstream or downstream of the regulated exon, a pyrimidine sequence (Py, yellow) present in the human h5S-OT lncRNA targets the splicing factor U2AF65 to alternatively spliced exons, thereby inducing a change in exon-inclusion levels.

induced exon exclusion when it was targeted to the downstream intron (**Fig. 2**). The authors successfully used this method to change the alternative splicing of 9 out of 11 tested exons with variable efficiencies (25–60% change in exon-inclusion levels), thus suggesting that the system can be adapted as a molecular tool to drive gene-specific changes in alternative-splicing patterns without affecting the overall protein levels[3]. Of note, similar systems have already been used to switch alternative-splicing patterns. In those systems, a bifunctional antisense oligonucleotide that contains a portion complementary to the target pre-mRNA is linked to a nonhybridizing tail that contains a specific RNA-binding motif necessary to recruit a splicing repressor or activator to the target region[9,10]. Such mechanisms have successfully been used to decrease levels of disease-specific splicing variants, and antisense oligonucleotides are under clinical trials for treating spinal muscular atrophy[9,11]. However, bifunctional and antisense oligonucleotides require previous knowledge of the relevant regulatory sequences to be targeted, thus restricting their application to only well-characterized genes. The h5S-OT-based mechanism of splicing regulation allows targeting of any sequence upstream or downstream of an exon of interest to increase or decrease exon-inclusion levels, thus making the system much more flexible and adaptable to any gene. An alternative, highly promising mechanism to modulate splicing is the innovative CRISPR–dCas9 system, which can also be targeted to specific RNAs[12,13] and has successfully been used to target transcriptional regulators to a gene of interest in order to change the gene's transcriptional levels[14], thus suggesting that such a system could also be adapted to target specific spicing regulators to an alternatively spliced exon.

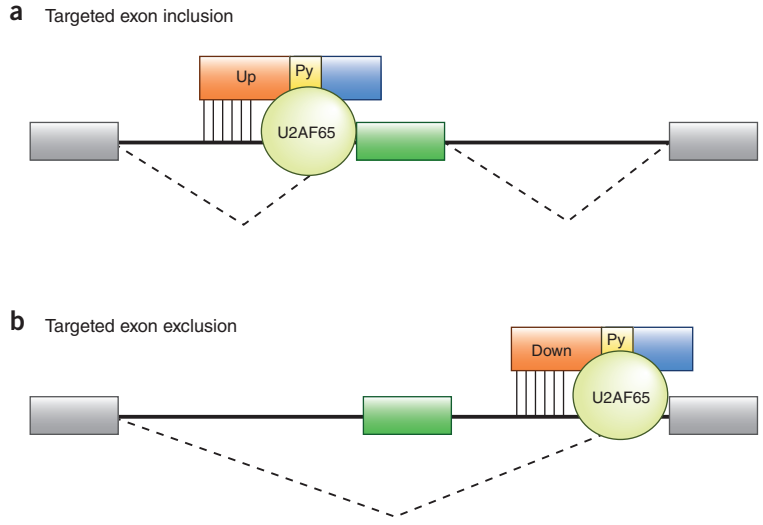

Figure 2 Targeted modulation of alternative splicing by adapting the h5S-OT sequence. (**a**,**b**) By replacing the anti-Alu sequence in h5S-OT with a sequence complementary to the intronic region upstream or downstream of the target exon, it is possible to increase (**a**) or decrease (**b**) exon-inclusion levels, respectively.

In the past few years, lncRNAs have increasingly been implicated in the regulation of gene expression at multiple levels, including alternative splicing[15–17]. Shan and colleagues[3]

now propose a novel mechanism of lncRNA-mediated splicing regulation dependent on Alu retrotransposable elements. Because Alu elements are exclusive to primates, it is possible that this Alu-dependent regulation of splicing may have contributed to the transcriptome divergence of primates from the rest of the mammals and to the appearance of human-specific alternative-splicing events. Along these lines, it would be interesting to address whether h5S-OT-dependent exons play important roles in human morphology and/or physiology. Finally, the adaptability of this endogenous mechanism to target any sequence of interest to increase or to decrease exon-inclusion levels opens new perspectives into the development of much more flexible, simpler biotechnology tools to modulate alternative-splicing outcomes in normal cells or to cause reversion of disease-specific splice variants[18].

COMPETING FINANCIAL INTERESTS
The author declares no competing financial interests.

1. Wilusz, J.E., Sunwoo, H. & Spector, D.L. *Genes Dev.* **23**, 1494–1504 (2009).
2. Ule, J. *Biochem. Soc. Trans.* **41**, 1532–1535 (2013).
3. Hu, S.W. *Nat. Struct. Mol. Biol.* **23**, 1011–1019 (2016).
4. Nielsen, J.N., Hallenberg, C., Frederiksen, S., Sørensen, P.D. & Lomholt, B. *Nucleic Acids Res.* **21**, 3631–3636 (1993).
5. Sørensen, P.D. & Frederiksen, S. *Nucleic Acids Res.* **19**, 4147–4151 (1991).
6. Häsler, J. & Strub, K. *Nucleic Acids Res.* **34**, 5491–5497 (2006).
7. Sorek, R., Ast, G. & Graur, D. *Genome Res.* **12**, 1060–1067 (2002).
8. Shao, C. *et al. Nat. Struct. Mol. Biol.* **21**, 997–1005 (2014).
9. Osman, E.Y., Yen, P.F. & Lorson, C.L. *Mol. Ther.* **20**, 119–126 (2012).
10. Brosseau, J.P. *et al. Nucleic Acids Res.* **42**, e40 (2014).
11. Hua, Y. *et al. Genes Dev.* **29**, 288–297 (2015).
12. Nelles, D.A. *et al. Cell* **165**, 488–496 (2016).
13. Abudayyeh, O.O. *et al. Science* **353**, aaf5573 (2016).
14. Gilbert, L.A. *et al. Cell* **154**, 442–451 (2013).
15. Tripathi, V. *et al. Mol. Cell* **39**, 925–938 (2010).
16. Gonzalez, I. *et al. Nat. Struct. Mol. Biol.* **22**, 370–376 (2015).
17. Wang, K.C. & Chang, H.Y. *Mol. Cell* **43**, 904–914 (2011).
18. Wang, G.S. & Cooper, T.A. *Nat. Rev. Genet.* **8**, 749–761 (2007).

mRNA decapping in 3D

Jeff Coller

The degradation of mRNAs involves removal of the 5′ protective cap via a decapping-enzyme complex, in a largely irreversible process that commits the transcript for destruction. Understanding how the decapping reaction is catalyzed and regulated are major goals in the field. New data suggest how the chemistry of decapping is controlled and orchestrated within the cell.

Eukaryotic mRNAs are post-transcriptionally modified, so that each receives a 7-methylguanosine cap at its 5′ end. The cap is joined to the mRNA 5′ end via a unique 5′-to-5′ triphosphate linkage. The cap distinguishes mRNAs from other RNA transcripts within the cell and promotes splicing, export, translation and mRNA stability[1,2]. Removal of the cap is absolutely required before the mRNA can be digested by 5′-to-3′ exoribonucleases[3]. Thus, decapping is a critical step in controlling mRNA half-life and therefore gene expression.

In studies of mRNA stability, it was first suggested in the late 1970s that an activity must exist within cells that catalytically removes 5′ caps and commits mRNAs for destruction[4,5]. Biochemical hints of a decapping activity were first reported in the 1980s in *Saccharomyces cerevisiae*, when Stevens isolated an activity that specifically cleaved m^7GDP from the 5′ ends of *in vitro*–derived transcripts[6]. The Parker laboratory made a series of discoveries in the 1990s identifying two genes absolutely required for removing the 5′ caps from mRNA transcripts; the corresponding proteins are now referred to as Dcp1 and Dcp2 (refs. 7,8). Dcp2 and Dcp1 are conserved[9,10] and are now known to be crucial regulators of mRNA stability, playing roles in normal, microRNA-mediated and nonsense-mediated mRNA-decay pathways[11].

The discovery of two genes being absolutely required for mRNA decapping has been enigmatic. Dcp2, a member of the Nudix hydrolase family of proteins, has clearly been shown to contain the catalytic activity and to be sufficient to bind and decap mRNAs *in vitro*. However, *in vivo*, Dcp1 is essential for Dcp2's decapping function and thus serves a crucial but unknown role in promoting the decapping reaction. The C terminus of Dcp2 contains the Nudix domain and an RNA-binding motif referred to as the Box-B helix (**Fig. 1a**). The N terminus of Dcp2 contains the so-called N-terminal regulatory domain, which binds the EVH1 domain of Dcp1. In addition to Dcp1 and Dcp2, there are several known coactivators of mRNA decapping, including Edc1–4, Lsm14a (SCD6), Pat1, Lsm1–7, Dhh1 (DDX6), PNRC2 and UPF1 (ref. 11).

Over the past ten years, many structural studies of Dcp2 have investigated how decapping is catalyzed and possibly regulated by coactivators[12–15]. The picture that has emerged indicates that Dcp2 is a bilobed protein with a flexible connector between its N and C termini (**Fig. 1a**) that allows Dcp2 to adopt multiple distinct conformations and to oscillate between open and closed forms[12]. In a closed, compact form of Dcp2, the N-terminal regulatory domain interacts closely with the Nudix domain. There is growing evidence that this compact form is the active conformation of Dcp2 (ref. 16), thus suggesting that the active site is a composite between the C-terminal Nudix domain and the N-terminal regulatory domain (which interacts with Dcp1 (ref. 16)). However, it remains unclear how Dcp1 contributes to this active state of Dcp2, because no contacts have been identified that might explain how, for instance, Dcp1 might stabilize a compact conformation. Moreover, the composite active site formed by the Nudix domain and the N-terminal regulatory domain was hypothetical and had not been observed to coordinate cap binding within the enzyme itself[16].

Three recent independent studies (two in this issue and one in the June issue of *Nature Structural & Molecular Biology*)[17–19] provide some of the most detailed understanding of the eukaryotic decapping complex to date. These studies have generated high-resolution structures of a heterotrimeric complex composed of Dcp2, Dcp1 and an important decapping regulator: Edc1 (ref. 17), Edc3 (ref. 18) or PNRC2 (ref. 19). In addition, two of the structures are bound to cap-nucleotide analogs, thus providing the first snapshots of the decapping enzyme with its active site loaded. Together, the three studies provide crucial new insights into how

Jeff Coller is at the Center for RNA Molecular Biology, Case Western Reserve University, Cleveland, Ohio USA.
e-mail: jmc71@case.edu