Cell Research

www.nature.com/cr
www.cell-research.com

RESEARCH HIGHLIGHT

Nascent RNA m^{6}A modification at the heart of the gene–retrotransposon conflict

Victor Billon[1] and Gael Cristofari[1]

Cell Research (2021) 31:829–831; https://doi.org/10.1038/s41422-021-00518-5

Retrotransposons are highly abundant mobile genetic elements in mammalian genomes and can interfere with many genomic processes, particularly the transcription of genes. A study by Xiong et al. reveals that segments of nascent RNAs containing retrotransposons are enriched in *N*^{6}-methyladenosine and act as transcriptional roadblocks for elongating RNA polymerase II, but that cellular proteins capable of binding to these RNAs, SAFB/SAFB2, can mitigate this effect.

Long interspersed element 1 (L1) retrotransposons are transposable elements capable of propagating within genomes by a copy-and-paste mechanism through an RNA intermediate and a reverse transcription step. They are present in many eukaryotic lineages, but have been, and still are, particularly active in mammals, acting as potent endogenous mutagens. They are strongly silenced by several layers of transcriptional and post-transcriptional mechanisms, both in the nucleus and in the cytoplasm, that restrict their expression and mobilization to the germ cells, the early embryos, and a very narrow set of adult somatic cells. Nevertheless, some of these elements manage to break free from these locks and insert into new genomic locations, often landing in introns, and occasionally leading to genetic disease.[1]

Inserting a large piece of DNA into a gene is potentially more damaging than single nucleotide mutations. A full-length L1 copy is ~6 kb in length, carries its own promoter sequence, as well as two open reading frames encoding the retrotransposition machinery, and ends with a polyadenylation signal.[1] L1s also contain multiple cryptic splice sites, and binding sites for transcription factors and RNA-binding proteins. Surprisingly, when insertion occurs into introns, transplanting so many *cis*-acting regulatory elements does not appear, in most cases, to prevent polymerase II from producing functional transcripts. This can be due in part to purifying selection that eliminates the most detrimental L1 insertions, particularly those in sense orientation relative to genes, or to specific protein–RNA complexes that can mask L1 sequences embedded in introns and prevent them from becoming functional.[2]

In mammals, several DNA and RNA base modifications have been implicated in the regulation of transposable elements or in preventing damaging consequences of their activation. Well-known examples of such modifications include 5-methylcytosine in DNA, an epigenetic mark associated with transcriptional repression,[3] and adenosine-to-inosine (A-to-I) RNA editing that inhibits the activation of inflammatory pathways that could arise in response to double-stranded RNA structures formed by sense and antisense retrotransposon sequences.[4]

However, the most abundant mRNA modification, *N*^{6}-adenosine methylation (m^{6}A), has only recently been detected in RNA from retrotransposons.[5–7] First, a CRISPR screen performed in mouse embryonic stem cells and designed to identify regulators of retroviral-like retrotransposons, the IAP elements, discovered that the m^{6}A writers METTL3 and METTL14 are key regulators of retrotransposon RNA stability. Notably, they exert opposing effects on different retrotransposon families: while m^{6}A destabilizes intracisternal A particle RNA, it appears to stabilize L1 RNA.[5] Furthermore, m^{6}A RNA modifications may regulate the formation or maintenance of retrotransposon-associated heterochromatin, at least at specific retrotransposon loci, possibly through chromatin-associated RNA.[6,8]

In a new study published in *Cell Research*, Xiong et al.[9] explore m^{6}A RNA modifications in nascent transcripts of human cells using a newly developed method, termed MINT-seq, that provides much higher intronic coverage than previous studies. They observe that RNA segments derived from L1 elements and embedded in introns in the sense orientation relative to genes are particularly enriched in m^{6}A modifications. Such m^{6}A-rich domains, called MILs for m^{6}A-rich intronic L1s, are characterized by decreased transcriptional elongation downstream of L1 (Fig. 1a). The levels of m^{6}A enrichment in these intronic regions correlate with the extent of transcription block, and deletion or inversion of MILs by genome editing rescues efficient polymerase II progression. Although it is still unclear whether m^{6}A is directly responsible for transcription attenuation, or whether other unappreciated characteristics of MILs are involved, a first hint of the underlying mechanisms arose from the systematic analysis of publicly available RNA–protein crosslinking immunoprecipitation (CLIP) data. Several RNA-binding proteins are enriched at MILs, particularly SAFB and SAFB2. The authors demonstrate biochemically that these two poorly characterized proteins can specifically bind L1 RNA in the sense orientation, a property enhanced by m^{6}A modifications. Upon depletion of SAFB/SAFB2, transcription blockade at MILs is exacerbated suggesting that these two factors limit MIL-induced transcriptional attenuation and thereby contribute to fine-tuning transcription (Fig. 1a).

MINT-seq, as most short-read-based approaches, cannot discriminate transcripts produced from the internal L1 promoter and serving as intermediates of L1 replication, from genic transcripts that may contain L1 sequences in their introns.[10] However, looking at carefully chosen L1 loci, known to be autonomously transcribed, and using synthetic reporter constructs, the authors show that m^{6}A is deposited both on passenger L1 sequences embedded in intronic nascent RNAs and on L1s transcribed from their own promoters.

[1]Université Côte d'Azur, Inserm, CNRS, IRCAN, Nice, France
Correspondence: Gael Cristofari (Gael.Cristofari@univ-cotedazur.fr)

Published online: 9 June 2021

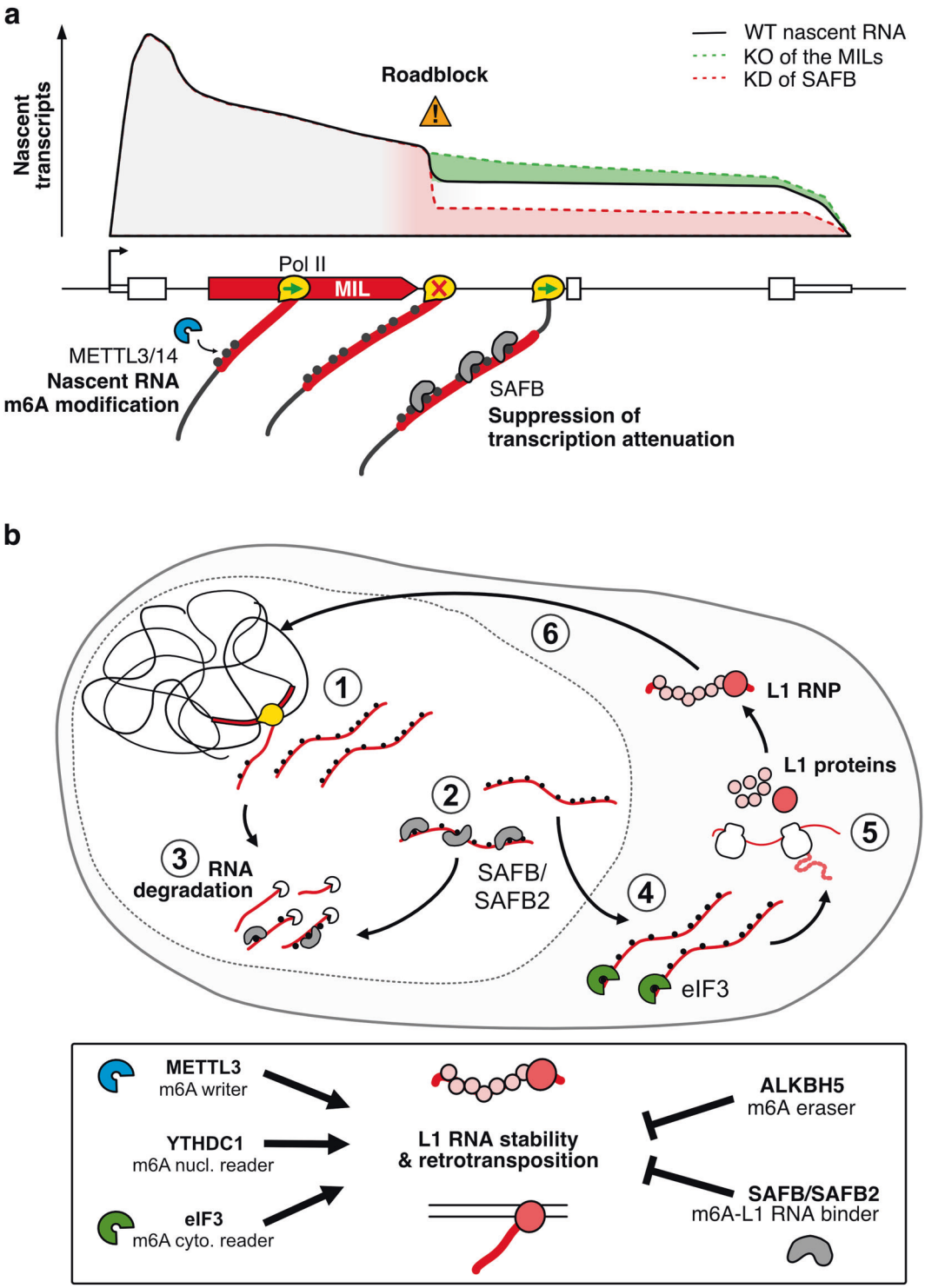

Fig. 1 A dual role for SAFB/SAFB2 in protecting the genome against L1 retrotransposons. a SAFB limits transcriptional attenuation caused by intronic L1s. RNA polymerase II (Pol II) transcribes a gene containing a sense-oriented intronic L1 (red). The nascent RNA segment containing the L1 sequence is m^{6}A-modified (black circles) by METTL3/14 (blue) and bound by SAFB/SAFB2 (grey). The intronic L1 causes a transcription roadblock (dark line). Deletion (KO, green dashed line) or inversion of the L1 sequence prevents attenuation while depletion of SAFB increases it (KD, red dashed line). **b** A model for the influence of m^{6}A on L1 mobilization. When transcribed from its own promoter, L1 RNA can also be N^6A-methylated by METTL3/14 (1). SAFB/SAFB2 proteins preferentially bind m^{6}A-modified L1 RNA (2). Unmodified L1 RNAs and SAFB-bound L1 RNAs are destabilized (3). In the cytoplasm, m^{6}A-modified L1 RNA can hijack the translation initiation factor eIF3 (green), enhancing L1 translation (4). L1 ribonucleoprotein particles (RNPs) formed by L1 proteins and RNA (5) drive reverse transcription and genomic integration (6). Inset, m^{6}A pathway effectors regulating L1 mobilization.

They could also address the effect of m^6A methylation and its readers on L1 RNA stability and retrotransposition. m^6A methylation stabilizes L1 RNA and increases L1 mobility, a process promoted by the m^6A reader YTHDC1 and antagonized by SAFB/SAFB2 (Fig. 1b). In agreement with these results, another recent study concluded that m^6A positively regulates L1 expression and retrotransposition, and further identified a cytoplasmic effector of this pathway, eIF3, which stimulates L1 RNA translation.[7]

Xiong et al. have uncovered the role of SAFB/SAFB2 proteins as peacekeepers that safeguard gene function against potent endogenous mutagens. Not only do these proteins limit the ability of L1 to propagate, but also once new L1 insertions occur, they alleviate L1-mediated transcription attenuation. Future work will decipher the full set of m^6A writers, readers and erasers, and how they stand in the gene–retrotransposon conflict (Fig. 1b, inset).

ACKNOWLEDGEMENTS

Work in the laboratory of GC is supported by the French government, through the Agence Nationale de la Recherche (Idex UCAJEDI, ANR-15-IDEX-01; Labex SIGNALIFE, ANR-11-LABX-0028-01; RetroMet, ANR-16-CE12-0020; ImpacTE, ANR-19-CE12-0032), the Fondation pour la Recherche Médicale (FRM, DEQ20180339170), Inserm (GOLD cross-cutting programme on genomic variability), CNRS (GDR 3546), and the University Hospital Federation (FHU) OncoAge.

ADDITIONAL INFORMATION

Competing interests: G.C. is an unpaid associate editor of the journal *Mobile DNA* (Springer-Nature) and an unpaid review editor of the Genetics, Genomics and Epigenomics of Aging section within the *Frontiers in Aging* journal (Frontiers).

REFERENCES

1. Kazazian, H. H. & Moran, J. V. *N. Engl. J. Med.* **377**, 361–370 (2017).
2. Attig, J. et al. *Cell* **174**, 1067–1081 (2018).
3. Deniz, Ö., Frost, J. M. & Branco, M. R. *Nat. Rev. Genet.* **20**, 417–431 (2019).
4. Eisenberg, E. & Levanon, E. Y. *Nat. Rev. Genet.* **19**, 473–490 (2018).
5. Chelmicki, T. et al. *Nature* **591**, 312–316 (2021).
6. Liu, J. et al. *Nature* **591**, 322–326 (2021).
7. Hwang, S.-Y. et al. *Nat. Commun.* **12**, 880 (2021).
8. Xu, W. et al. *Nature* **591**, 317–321 (2021).
9. Xiong, F. et al. *Cell Res*. https://doi.org/10.1038/s41422-021-00515-8 (2021).
10. Lanciano, S. & Cristofari, G. *Nat. Rev. Genet.* **21**, 721–736 (2020).