Title

Alu transposable elements rewire enhancer-promoter network through RNA pairing.

Journal

Molecular Cell, 83(18)

Authors

Wen, Xingzhao

Zhong, Sheng

Publication Date

2023-09-21

DOI

10.1016/j.molcel.2023.08.026

Peer reviewed

HHS Public Access

Author manuscript

Mol Cell. Author manuscript; available in PMC 2024 September 21.

Published in final edited form as:

Mol Cell. 2023 September 21; 83(18): 3234–3235. doi:10.1016/j.molcel.2023.08.026.

Alu transposable elements rewire enhancer-promoter network through RNA pairing

Xingzhao Wen[1], **Sheng Zhong**[1,2,*]

[1]Program in Bioinformatics and Systems Biology, University of California San Diego, La Jolla, CA 92093, USA

[2]Shu Chien-Gene Lay Department of Bioengineering, University of California San Diego, La Jolla, CA 92093, USA

Abstract

A recent study[1] revealed that interacting enhancer RNAs (eRNAs) and promoter-transcribed upstream antisense RNAs (uaRNAs), can identify enhancer-promoter interactions. Complementary sequences within the interacting eRNAs and uaRNAs, predominantly Alu sequences, confer the specificity for eRNA-uaRNA pairing and hence enhancer-promoter recognition.

The human genome contains hundreds of thousands of enhancers and tens of thousands of promoters. It remains to be resolved what enhancers communicate with what promoters in a given cell type and more importantly, what determines the specificity of enhancer-promoter communication[2]? Two general rules emerged. First, communicating enhancer-promoter pairs (cE-Ps) often exhibit spatial proximity in the 3-dimensional nucleus and co-localize in the same topologically associating domain (TAD). Second, the chromatin regions of some promoters are more open and thus more available for communicating with enhancers. However, these rules cannot fully explain the extent of specificity of cE-Ps. Emerging evidence indicates that the DNA sequence of the promoters contributes to specifying its cognate enhancers. What remains to be resolved is how the DNA sequence determines such specificity[2].

One possibility involves the binding of specific proteins at enhancers and promoters and specific protein-protein interactions (PPIs) between enhancer-associated and promoter-associated proteins (PPI-rendered specificity). While this model is supported in several cases, it seems unlikely that PPI-rendered specificity alone can explain numerous cE-Ps.

RNAs can be produced from enhancers (eRNA) and promoters (upstream antisense RNA (uaRNA)). The second, less explored possibility involves eRNAs and uaRNAs. Early studies

* szhong@ucsd.edu .

Conflict of Interests

S.Z. is a founder and board member of Genemo, Inc.

of individual eRNAs revealed that an eRNA can promote chromatin looping between the enhancer and the target promoter via Mediator complex recruitment[3] or assistance of cohesin[4], and promote the transcription of target gene. Genome-wide analysis of chromatin-associated RNAs (caRNAs) by MARGI (mapping RNA-genome interactions) revealed that most promoters are enriched with caRNAs that are transcribed from distal genomic sequences (distal caRNAs)[5]. Furthermore, the amount of promoter-associated distal caRNAs correlates with the promoter's transcriptional activity[5], and MARGI-identified promoter-associated distal caRNAs can regulate the promoters' transcription activity[6]. These genome-wide data allude to a possible generalization of the mode of action of the individually studied eRNAs, in that most eRNAs can attach to the target gene's promoter to regulate its activity.

Yet unknown is how eRNAs are recruited to specific promoters? A plausible hypothesis is that chromatin-associated eRNAs and uaRNAs are brought together by protein-mediated RNA-RNA interactions (RRIs). A technology to globally map protein-mediated RRIs is required to test this hypothesis. The MARIO (Mapping RNA interactome *in vivo*) technology, based on proximityligation of RNAs within the same crosslinked protein-RNA complexes, revealed tens of thousands of RRIs from cultured cells and mouse brain[7]. RIC-seq (RNA *in situ* conformation sequencing) improved on MARIO to carry out the RNA ligations *in situ* and reduce the linker sequence between the ligated RNAs to a single nucleoside[8]. Applied to seven human cell lines, RIC-seq revealed in each cell line an RRI map of eRNAs and uaRNAs[8,1]. To test if these enhancer–promoter RNA interaction (EPRI) maps reflect cE-Ps, Cai et al. knocked down the eRNAs of a set of typical enhancers and super enhancers. Most target genes defined by the EPRI map exhibited reduced transcription. In contrast, the transcription of non-target genes in the genomic neighborhoods of these enhancers was unaffected[8], indicating cE-Ps can be inferred by EPRIs.

Motif analysis revealed an enrichment of Alu sequences in the interacting eRNAs and uaRNAs, where the eRNA-contained Alu sequences often complement those embedded in their interacting uaRNAs[1]. Alu is a group of transposable elements specific to the primate genomes[9]. More than 10% of the human genome comprises Alu elements, totaling approximately one million copies[9]. The enriched Alu sequences in EPRIs suggest a model where complementary RNA sequences between chromatin-associated eRNA and uaRNA specify enhancer-promoter pairing. Supporting this model, RNA FISH and DNA FISH revealed spatial co-localization of an eRNA-uaRNA pair that contain complementary Alu sequences, as well as the enhancer and the promoter[8]. A subset of eRNA-uaRNA pairs form duplexes by PARIS ((psoralen analysis of RNA interactions and structures)[10]. Deleting a single Alu element from an enhancer or the promoter of the target gene reduced long-range chromatin interactions and the transcription of the target gene[1]. Reversely, insertion of an Alu element to an unregulated promoter promoted chromatin looping with an Alu-containing enhancer and increased the promoter's transcription activity[1]. To tease apart the functions of the DNA and the RNA, an Alu-containing RNA was tethered to an uaRNA of an unregulated promoter, which again increased chromatin looping and the promoter's transcription activity[1]. These data establish the role of caRNA in determining the cE-Ps.

A small fraction (~7%) of the EPRI-supported enhancer–promoter pairs do not contain Alu elements. These eRNA-uaRNA pairs nevertheless exhibit sequence complementarity and can form RNA duplexes[1]. Thus, the chromatin-associated eRNA and uaRNA are likely responsible for specifying cE-Ps regardless of whether they contain Alus.

Alu elements are evolving with the human genome. Inter-species comparisons revealed that Alu sequences are evolutionarily selected, in that (1) they are evolved to acquire enhancers' epigenetic signatures and (2) the Alu elements closer to promoters are more evolutionarily conserved than those further away from the promoters[9]. For any genomic sequence to be evolutionarily selected, this genomic sequence must benefit the survival of the host species. These evolutionary analyses indicate that Alu elements, especially those near promoters, should benefit the survival of humans. It begs the question, how do Alus provide such benefit(s)?

The discovery of Alu's role in determining cE-Ps suggests that the transposition of Alu elements may actively rewire the E-P network in the human genomes. The intersection of the cancer-related genomic variations identified by the International Cancer Genome Consortium (ICGC) and EPRIs led to an "Alu variant-to-function" map in each of the seven analyzed cell lines[1]. These maps comprise 17,231 Alu-related insertions or deletions (ICGC Alu variants) and 8,623 target genes. These ICGC Alu variants are significantly depleted in known oncogenes but enriched in tumor suppressor genes[1]. These data suggest that Alu transposition-mediated innovation of the E-P network is poorly tolerated in oncogenes but is expanding in tumor suppressor genes. It highlights a possible route for Alus to benefit the human species by diversifying the E-P network of tumor suppressor genes. Finally, a personal genomic variation that overlaps with an oncogene-linked ICGC Alu variation will likely increase this person's cancer risk.

Acknowledgments

This work is funded by NIH grants DP1DK126138, R01GM138852, U01CA200147, R01HD107206, and a Kruger Research Grant.

References

1. Liang L et al. Complementary Alu sequences mediate enhancer-promoter selectivity. Nature 619, 868–875 (2023). [PubMed: 37438529]
2. Galouzis CC & Furlong EEM Regulating specificity in enhancer-promoter communication. Curr. Opin. Cell Biol 75, 102065 (2022). [PubMed: 35240372]
3. Ørom UA et al. Long noncoding RNAs with enhancer-like function in human cells. Cell 143, 46–58 (2010). [PubMed: 20887892]
4. Li W et al. Functional roles of enhancer RNAs for oestrogen-dependent transcriptional activation. Nature 498, 516–520 (2013). [PubMed: 23728302]
5. Sridhar B et al. Systematic Mapping of RNA-Chromatin Interactions In Vivo. Curr. Biol 27, 602–609 (2017). [PubMed: 28132817]
6. Calandrelli R et al. Stress-induced RNA-chromatin interactions promote endothelial dysfunction. Nat. Commun 11, 5211 (2020). [PubMed: 33060583]
7. Nguyen TC et al. Mapping RNA-RNA interactome and RNA structure in vivo by MARIO. Nat. Commun 7, 12023 (2016). [PubMed: 27338251]

8. Cai Z et al. RIC-seq for global in situ profiling of RNA-RNA spatial interactions. Nature 582, 432–437 (2020). [PubMed: 32499643]
9. Su M, Han D, Boyd-Kirkup J, Yu X & Han J-DJ Evolution of Alu elements toward enhancers. Cell Rep 7, 376–385 (2014). [PubMed: 24703844]
10. Lu Z et al. RNA Duplex Map in Living Cells Reveals Higher-Order Transcriptome Structure. Cell 165, 1267–1279 (2016). [PubMed: 27180905]

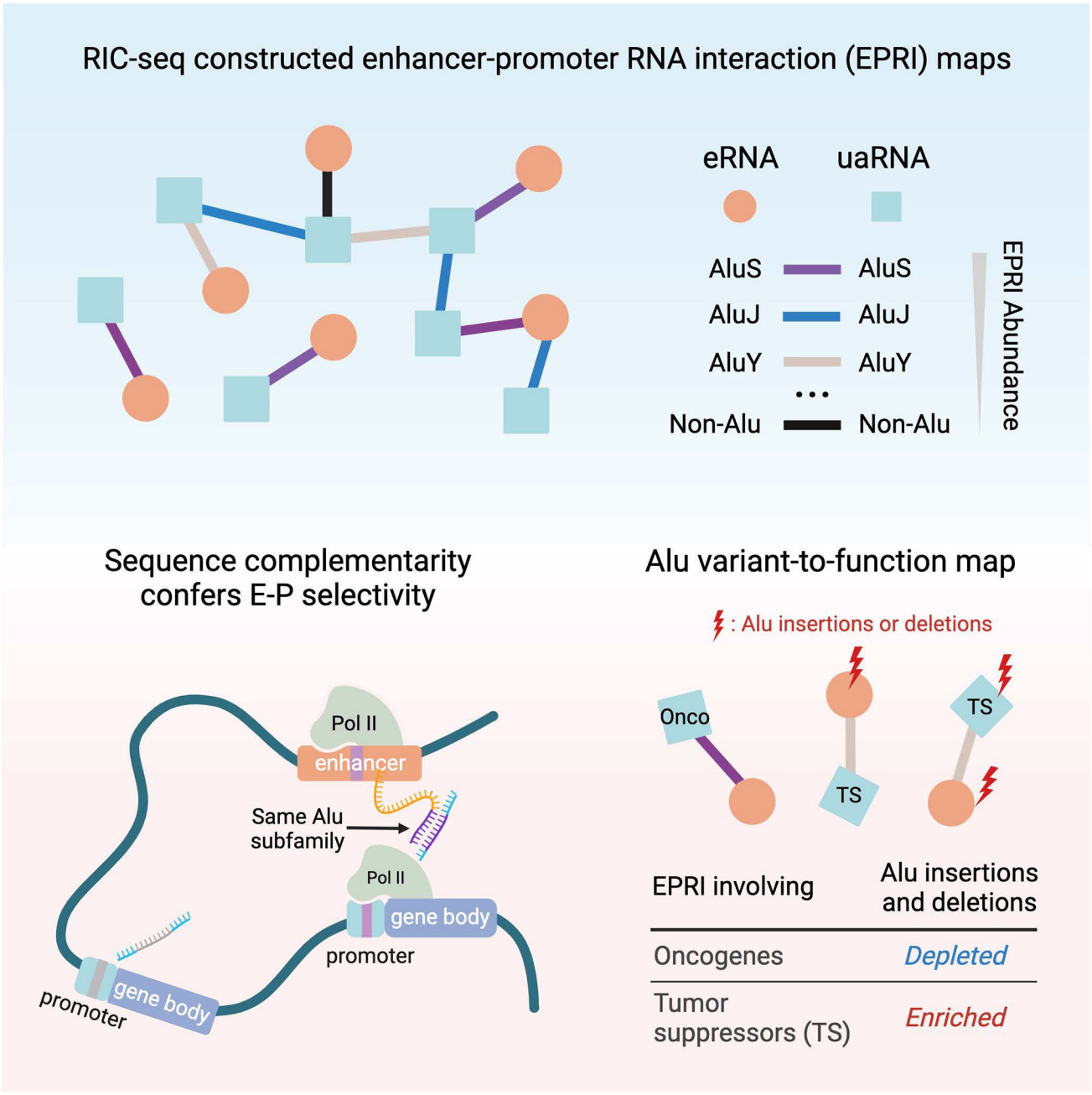

Figure 1.

Complementary Alu RNA sequences confer specificity in enhancer-promoter communication. Top: RIC-seq is applied to generate enhancer-promoter RNA interaction (EPRI) maps. The interacting eRNA-uaRNA pairs often contain complementary RNA sequences. Lower left: EPRIs reflect the communicating enhancer-promoter (cE-P) pairs. Complementary Alu sequences confer the specificity of cE-Ps. Lower right: Cancer-related Alu insertions and deletions in the human genomes are depleted in the EPRI subnetwork involving oncogenes and enriched in the subnetwork involving tumor suppressor genes.