



OPEN ACCESS

EDITED BY

Chunhui Hou,
Chinese Academy of Sciences (CAS), China

REVIEWED BY

Craig John Benham,
University of California, Davis, United States
Ansuman Lahiri,
University of Calcutta, India

*CORRESPONDENCE

Rakesh K. Mishra,
✉ mishra@ccmb.res.in

RECEIVED 24 October 2023

ACCEPTED 27 December 2023

PUBLISHED 22 February 2024

CITATION

Pathak RU, Phanindhar K and Mishra RK (2024),
Transposable elements as scaffold/matrix
attachment regions: shaping organization and
functions in genomes.
Front. Mol. Biosci. 10:1326933.
doi: 10.3389/fmolb.2023.1326933

COPYRIGHT

© 2024 Pathak, Phanindhar and Mishra. This is
an open-access article distributed under the
terms of the [Creative Commons Attribution
License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or
reproduction in other forums is permitted,
provided the original author(s) and the
copyright owner(s) are credited and that the
original publication in this journal is cited, in
accordance with accepted academic practice.
No use, distribution or reproduction is
permitted which does not comply with these
terms.

Transposable elements as scaffold/matrix attachment regions: shaping organization and functions in genomes

Rashmi Upadhyay Pathak¹, Kundurthi Phanindhar¹ and
Rakesh K. Mishra^{1,2,3*}

¹CSIR-Centre for Cellular and Molecular Biology, Hyderabad, India, ²Academy of Scientific and Innovative Research (AcSIR), Ghaziabad, India, ³Tata Institute for Genetics and Society, Bengaluru, India

The hierarchical structure of eukaryotic genomes has regulatory layers, one of them being epigenetic “indexing” of the genome that leads to cell-type-specific patterns of gene expression. By establishing loops and defining chromatin domains, cells can achieve coordinated control over multi-locus segments of the genome. This is thought to be achieved using scaffold/matrix attachment regions (S/MARs) that establish structural and functional loops and topologically associating domains (TADs) that define a self-interacting region of the genome. Large-scale genome-wide mapping of S/MARs has begun to uncover these aspects of genome organization. A recent genome-wide study showed the association of transposable elements (TEs) with a significant fraction of S/MARs, suggesting that the multitude of TE-derived repeats constitute a class of anchorage sites of chromatin loops to nuclear architecture. In this study, we provide an insight that TE-driven dispersal of S/MARs has the potential to restructure the chromosomes by creating novel loops and domains. The combination of TEs and S/MARs, as elements that can hop through the genome along with regulatory capabilities, may provide an active mechanism of genome evolution leading to the emergence of novel features in biological systems. The significance is that a genome-wide study mapping developmental S/MARs reveals an intriguing link between these elements and TEs. This article highlights the potential of the TE–S/MAR combination to drive evolution by restructuring and shaping the genome.

KEYWORDS

nuclear matrix, nuclear matrix attachment region, transposable element, genome, topologically associated domain

Introduction

Genomic function is regulated by the hierarchical organization of chromatin in a 3D nuclear context. The current idea of the functional organization of chromatin in an interphase nucleus is dominated by the view that the chromatin folds into clusters of loops and segregates into two distinct types of compartments, transcriptionally repressive and permissive. At the chromosome level, the self-associating chromatin loop clusters, termed TADs, ensure cell-type specific expression of associated genes. The chromatin loops topologically couple the *cis*-regulating elements (CREs), such as the enhancers and promoters of target genes. The CREs themselves are also involved in the folding of the

chromatin fiber by virtue of networking of *trans*-acting protein factors (TFs) bound to them (Szabo et al., 2019).

S/MARs (scaffold/matrix attachment regions) are DNA elements instrumental in the formation of chromatin loops as they attach to the nuclear matrix (NuMat) in the interphase nucleus. Traditionally, S/MARs have been viewed as regions of DNA that anchor chromatin loops to a proteinaceous scaffold and are thought to have a structural role. Taking cues from experimentally characterized S/MARs from the eukaryotic genome, several functions, including the initiation of DNA replication, transcriptional activation, chromatin remodeling, and insulation of domains, are attributed to S/MARs. These elements, therefore, not only act as structural loop anchors but also regulate the functional aspects of chromatin (Liebich et al., 2002). However, the functional heterogeneity, lack of conservation of DNA sequences, and difficult and method-specific experimental output have put S/MARs on the backstage of recent scientific advances. We present here an argument that these important determinants of structural chromatin loops and functional domains need to be investigated at the genomic scale in a variety of eukaryotes as they may also be drivers of genome evolution.

S/MARs are an intriguing class of CREs. Typically, CREs are non-coding DNA that are responsible for the activation and sustenance of the transcriptional status of a locus and contain binding sites for TFs and other regulatory molecules. Promoters, enhancers, and insulators are the best-understood types of CREs. Occurring exclusively in the eukaryotic genome, S/MARs are DNA elements that are defined biochemically as opposed to other CREs that are defined functionally. They are demarcated based on their attachment to a ribo-proteinaceous scaffolding structure in the nucleus known as the NuMat. Identified for the first time in *Drosophila*, these sequences have heterogeneous functions and lack sequence conservation, although they have a functionally relevant conservation of the secondary structure of DNA. The reason for the sequence heterogeneity is attributed to their involvement in multiple independent processes (Mirkovitch et al., 1984; Bode et al., 1992; Laemmli et al., 1992).

A recent study indicated that TE-derived sequences identify as a significant portion of S/MARs (Sureka et al., 2022). The observation that TEs can act as S/MARs has been reported earlier as well. As TEs are mobile DNA elements with sequence properties uniquely suited to gene regulatory functions, their combination with S/MARs presents a tantalizing possibility of driving evolution through changes in domain architecture. It is well known that through diverse invasion strategies, TEs occupy a substantial fraction of all eukaryotic genomes. Environmental and genetic factors further modulate and produce dramatic variation in TE content. These factors present TEs as a major source of genetic variation and a driver of evolution (Rebollo et al., 2012; Wells and Feschotte, 2020). In the present article, we focus on the TE-S/MAR combination as powerful mobile regulatory sequences that can connect with nuclear architecture and significantly impact the evolution of genomes.

Genome-wide distribution of S/MARs: dynamic and constitutive S/MARs

As S/MARs are implicated in the structural delimitation of chromatin loops, it is reasonable to assume that they would

correlate with TADs to organize chromatin in discrete neighborhoods with coordinated gene expression. In recent years, with the advent of conformation capture techniques, there has been significant advancement in the knowledge of genome-wide chromatin topology and domain organization. However, S/MARs have never been placed in the context of chromosome folding at the genome level, mainly because information on *in vivo* S/MARs in an identical context where TADs have been mapped is not available. A couple of studies where S/MARs and TADs have been mapped in an identical setup do show that the two features coincide. The overlap of TAD borders with S/MARs in pericentromeric heterochromatin was found to be highly significant, indicating their plausible role in genome organization along with TADs (Saha et al., 2020; Verma et al., 2021). It is pertinent here to note that while TADs are generally erased during mitosis, chromatin loop anchors are thought to persist. This raises the intriguing possibility that a fraction of S/MARs (those that coincide with major TAD borders) are retained on condensed mitotic chromosomes to transmit the memory of TADs through the cell cycle. However, this aspect remains to be investigated.

Most experimentally identified S/MARs have been reported on a case-by-case basis, and genomic approaches are only beginning to emerge (Table 1). A handful of studies investigate S/MARs on the genomic scale (whole genome or whole chromosome), and only few among them experimentally map S/MARs to validate. Most studies use the bioinformatics approach as experimental mapping of S/MARs requires substantial efforts and can generate method-specific results. The bioinformatics approaches, although informative, are based on the occurrence of certain motifs (ORI sequences, TG-richness, Topo II sites, and AT-richness) or physical properties of the DNA sequence (duplex destabilization, triple helix, and curving/kinking of DNA). Severe discordance exists between sequence-based S/MAR prediction and their *in vivo* mapping, as the tools detect only a fraction of the elements experimentally mapped (Pathak et al., 2014). This is understandable as a predicted S/MAR might not be functional *in vivo* and *vice versa*. However, these observations underline the importance of mapping S/MARs *in vivo* to refine bioinformatics approaches for the purpose. In this context, robust biochemical methods for the isolation of S/MARs have proved to be very useful. Most protocols employed to isolate the NuMat are optimized to remove soluble proteins and the bulk of chromatin without aggregation to reveal the non-chromatin nuclear skeleton. The initial protocol used to isolate the NuMat employed high-salt (1 M and 2 M NaCl) extraction (Berezney and Coffey, 1974). In 1982, a modified method was introduced that used ammonium sulfate at lower molarity [0.25 M (NH₄)₂SO₄] but nearly the same ionic strength as 1 M NaCl (Capco et al., 1982). Later on, concerns about the high ionic strengths involved in extraction protocols led to two alternative procedures. While Mirkovitch et al. (1984) exposed the nuclei to a low concentration of lithium 3,5-diiodosalicylate (25 mM LIS) to obtain the NuMat preparation, Jackson and Cook agarose-encapsulated the cells and electrophoretically removed the non-NuMat material (Jackson and Cook, 1986). All the alternative methods yielded NuMat preparations that ultra-structurally resembled the conventional one. However, subtle differences were noted; for instance, the non-chromatin scaffold that resists low-salt extraction (using LIS) showed better preservation of components of the lamina, nucleolus, and fibro-granular internal meshwork. In

TABLE 1 Genome-wide studies on S/MARs.

S. no.	Year	Organism	Approach	Key feature	Reference
1	2004	<i>Arabidopsis</i>	<i>In silico</i> mapping using SMARTest	First genome-wide investigation of S/MARs	Rudd et al. (2004)
2	2007	<i>C. elegans</i>	<i>In silico</i> computational prediction	Genome-wide analysis of the distribution of matrix recognition signature	Anthony and Blaxter (2007)
3	2007	Human	Bioinformatics computational prediction	Genome-wide prediction of MARs	Girod et al. (2007)
4	2009	Human HeLa cells	Array-based hybridization	Tiling microarray to map S/MARs on chr14 to chr18	Linneman et al. (2009)
5	2009	Human primary cell AoAF	Array-based hybridization	Tiling microarray to map S/MARs on chr14 to chr18	Linneman and Krawetz (2009)
6	2010	<i>Giardia</i>	<i>In silico</i> computational prediction	First report of S/MARs from a unicellular eukaryote, the protozoan parasite <i>Giardia lamblia</i>	Padmaja et al. (2010)
7	2014	<i>Arabidopsis</i>	Array-based genomic hybridization	Tiling microarray to map S/MARs on chr4	Pascuzzi et al. (2014)
8	2014	<i>Drosophila</i>	Biochemical isolation and NGS	First genome-wide <i>in vivo</i> mapping of MARs by sequencing	Pathak et al. (2014)
9	2019	Human	Bioinformatic data mining approach	Genome-wide mapping of S/MARs based on the ChIP-Seq data of S/MARBP	Narwade et al. (2019)
10	2022	<i>Bombyx mori</i>	Biochemical isolation and NGS	Genome-wide mapping and dynamics of S/MARs during silk gland development	Chunduri et al. (2022)
11	2022	<i>Drosophila</i>	Biochemical isolation and NGS	Structural and developmental dynamics of S/MARs	Sureka et al. (2022)

recent years, S/MARs isolated by different methods have been sequenced and compared. One such study shows that though distinct groups of elements are isolated by low-salt and high-salt extraction protocols, there is considerable overlap among them. Both methods of extraction uncover interactions that are the components of the same network, validating either the low-salt (LIS) or high-salt (NaCl) method as good enough to isolate S/MARs (Linnemann et al., 2009). Another recent study found that interventions like heat stabilization strengthen the NuMat, where the basic molecular components remain unchanged (Sureka et al., 2022).

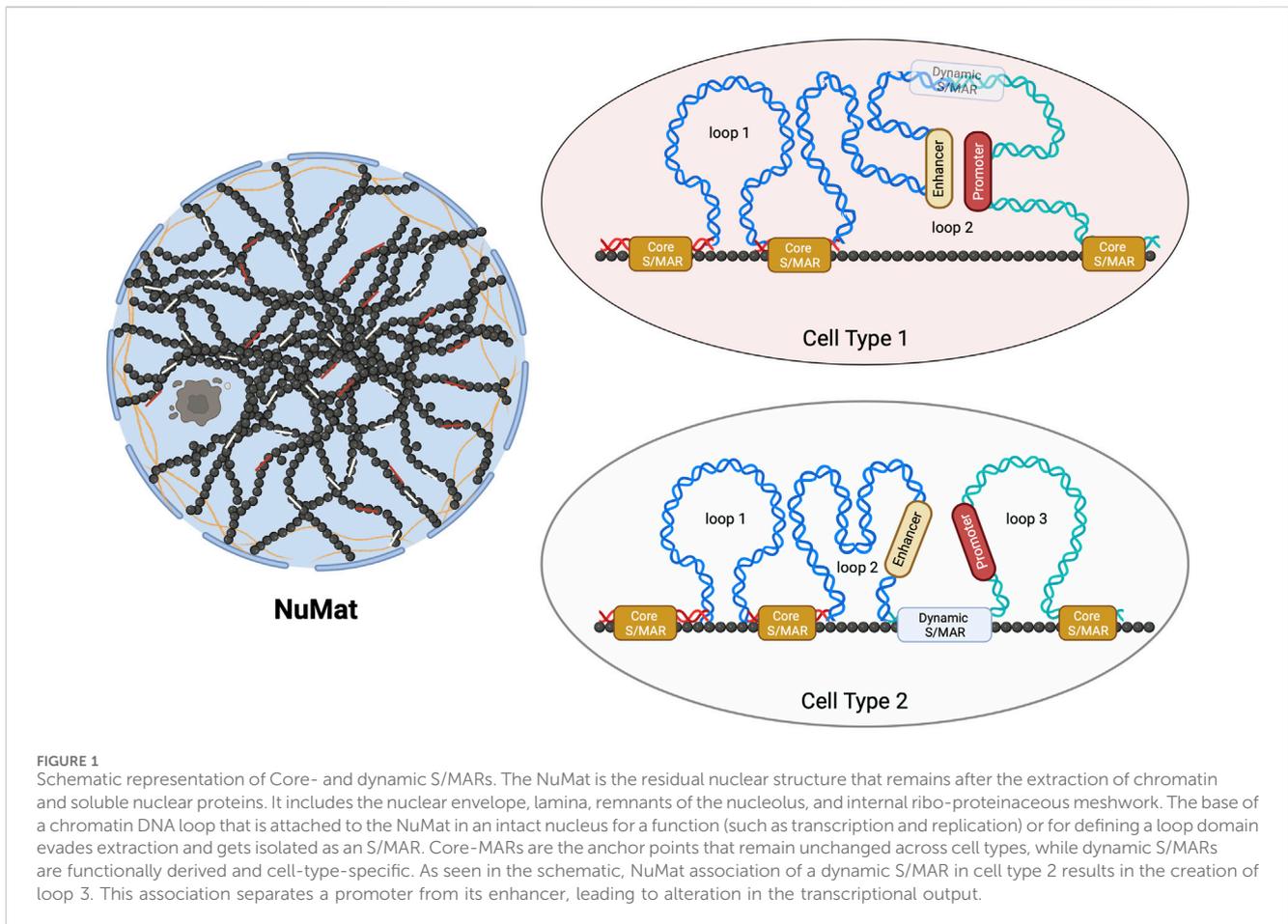
Only two studies of genome-wide mapping of S/MAR investigate their dynamics during embryonic development or cellular differentiation (Chunduri et al., 2022; Sureka et al., 2022). These studies, for the first time, uncovered that these attachment points can be categorized as dynamic or stable (Core-MARs) (Figure 1). The idea of categorization of S/MARs is not novel and has been suggested in several previous studies. Avramova et al. (1995), while studying these elements in maize, stated that “strong” S/MARs delimit the structural domains containing the transcriptional units and the “weak” S/MARs cause dynamic folding of larger loops into subdomains. Tikhonov et al. (2000) divided these elements into two groups: the “durable” S/MARs, mapping outside genes and defining borders of chromatin loops, and the “unstable” S/MARs with regulatory roles, mapping mainly within introns. They conclude that S/MARs possess both domain-defining and regulatory roles.

With the genome-wide study conducted during *Drosophila* embryonic development (Sureka et al., 2022), it is clear that the functional contribution of S/MARs to gene regulation is driven by both dynamic and stable attachments. The dynamic attachment sites are reversible (likely to be protein factor-induced), functionally

derived, closer to or within genes, and differ in various cell types. The stable “Core-MARs” are constitutive structural domains defining anchor points that remain unchanged across cell types. The interesting finding that this study brought to light is that the majority of Core-MARs are associated with repetitive DNA and that the TEs of LINE and LTR types are predominant among them. This study, for the first time, provided a glimpse into how dynamic and stable S/MARs can differ in their sequence composition and that TEs form a major class of Core-MARs. Thus, the study has identified a class of S/MARs stable across cell types with sequence characteristics different from those of the AT-rich or TG-rich classes of S/MARs containing TE-derived repeats.

Conservation of MARs through evolution

In the absence of sequence conservation, it is difficult to study the status of S/MARs as regulatory elements across species. During evolution, segments of genomes are shuffled (by recombination), placing orthologous genes from closely related species in very different genomic environments. Such genes still show a highly similar level of transcriptional output. A hypothesis to explain this paradox is based on the location of S/MARs, as these elements would stably anchor and appropriately position genes within the nuclear space. To gain insight, S/MARs from the genomes of closely related species have been investigated. Although not conducted on the genomic scale, these studies provide valuable details of the mechanistic paradigm of S/MARs. In one such study, the *Sh2/A1*-homologous regions of two grass species, rice and sorghum (about 50 kb of sorghum and 30 kb of rice DNA), were screened for the location and distribution of S/MARs. The position of S/MARs



relative to the neighboring genes was maintained, indicating the preservation of structural organization and folding of the chromatin in the region, but no conservation of sequences or motifs was found. The gene composition, orientation, order, and placement of S/MARs were found to be evolutionarily conserved in these closely related species. Interestingly, however, S/MARs in the region co-localize with TEs (Avramova et al., 1998).

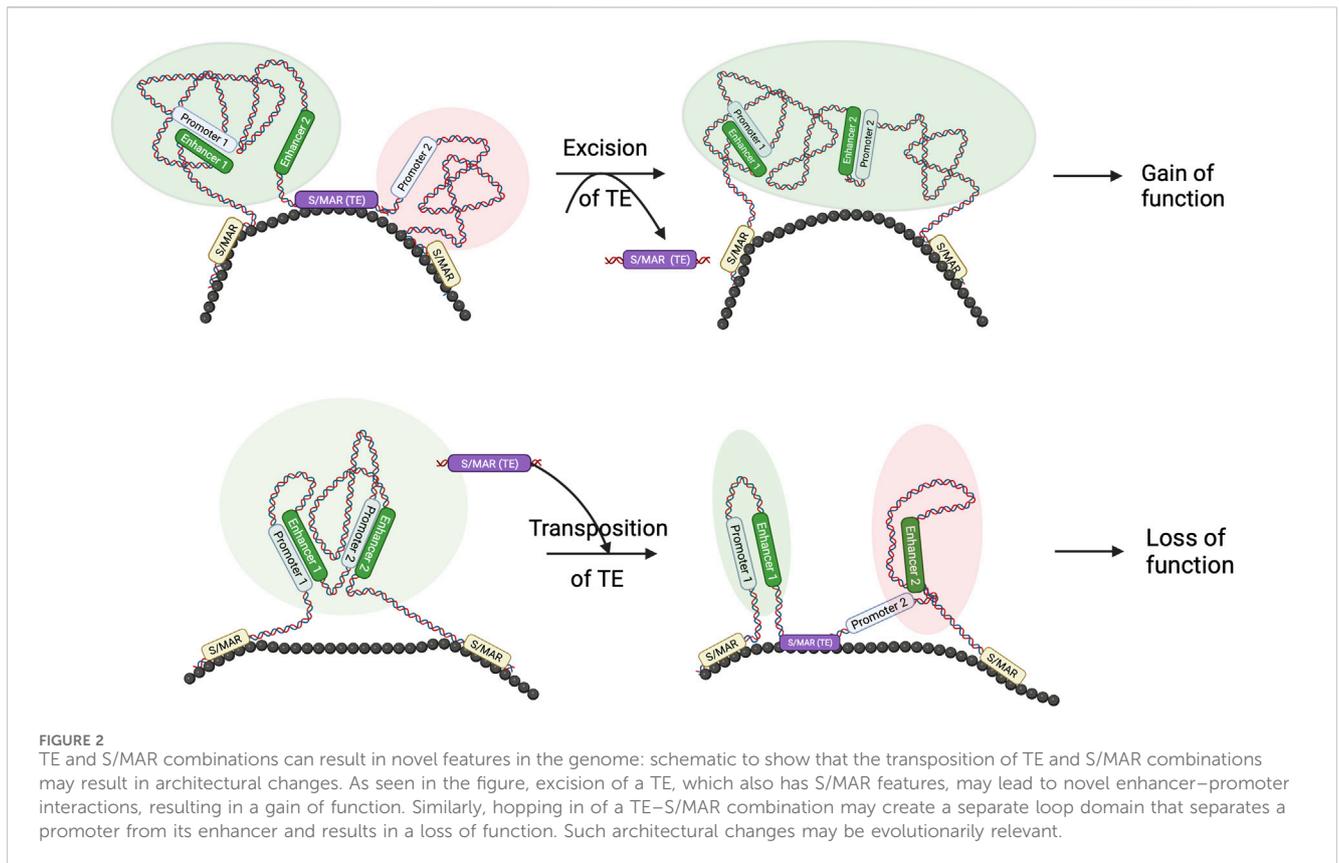
In a similar study, collinear chromosomal segments of maize and sorghum at the alcohol dehydrogenase (*Adh1*) locus carrying four genes were analyzed, and S/MARs were found to be at comparable positions in the two species, often flanking individual genes. Here too, TEs were found on the same fragments as S/MARs, suggesting that TEs may themselves be the S/MARs (Tikhonov et al., 2000). Another comparison between maize and Arabidopsis *Adh* locus found similar organizational conservation, and the authors conclude that stable S/MARs create large organizational domains that are conserved, and these are distinct from the smaller functional domains created by dynamic S/MARs (Paul and Ferl, 1998).

S/MAR organization has been shown to be conserved at the orthologous loci in animal genomes as well. A study examining the interspecies conservation of looping elements at the important human imprinting center at 15q11-q13 found that despite extensive divergence of the DNA sequence, the homologous mouse locus had a similar organization of S/MARs (Greally et al., 1999). Non-coding DNA in the human–mouse orthologous intergenic regions contain “islands” of conserved sequences, a

significant fraction of which were predicted S/MARs (Glazko et al., 2003). Taken together, these studies indicate that S/MAR organization is strikingly similar around an orthologous gene or group of genes, although sequence similarity may be minimal. These underline the importance of S/MARs in the evolutionary conservation of a functional genomic arrangement.

Significance of TEs as Core-MARs

Although more genome-wide studies in related species are required, it is highly probable that Core-MARs are conserved to convey architectural information, and TEs form a sizeable proportion of such elements. We hypothesize that changes in architecturally placed TEs, viz., Core-MARs, are events that establish specific new genome architectures that are important from the evolutionary perspective (Figure 2). The concept that genome architecture is formatted by TEs is not new, and it has been known for a while that transpositional changes cause evolutionary diversification by conferring adaptive benefits or reproductive isolation. Comparative whole-genome sequencing has provided evidence for TEs as organizers of functional genomic architecture, as these elements often map in a similar position and identical orientation in aligned syntenic segments of related genomes, indicating positive selection (Zhu et al., 2003). To get an integrated view of S/MARs in evolution, it may be worthwhile



to search for cases where TEs that have been major engines of speciation are also Core-MARs.

It is also important to note here that non-coding repetitive DNA is far more taxonomically discriminating than coding sequences. For example, each order of mammals has its own characteristic set of TEs (Platt et al., 2018). Approximately 50% of the human genome is composed of TE-derived sequences compared with ~1% of protein-coding sequences. This figure might be higher as many TE sequences in the genome would have evolved beyond recognition. The staggering evolutionary success of TEs, as evidenced by the sheer abundance of these elements in the genomes of highly evolved species, as well as the variety of changes induced by their mobility, suggest that these elements are “domesticated” to serve the evolutionary interests of their hosts. The property of TEs that enables them to provide ready-to-use mobile CREs is exploited by the host genome, and thus, TE-derived promoters, enhancers, and insulators are common (Jordan et al., 2003). TEs are also known to facilitate the partitioning of the genome into distinct domains, and this is precisely the same function envisaged for S/MARs. Notably, binding sites for CTCF (architecturally important protein, also known to bind to S/MARs) were propagated through TE expansion (Schmidt et al., 2012). Several conserved, robust, and tissue-independent CTCF sites were identified in the study that spanned six mammalian species. These were ChIP-seq data and did not investigate TADs or S/MARs in the context. However, given the properties of CTCF-binding sites, the data can be extrapolated to evolutionarily conserved TADs and Core-MARs, highlighting the link between TEs, S/MARs, and TADs through evolution.

TEs, as Core-MARs, gain more power when they land in a region of the genome that is potentially important for the regulation of a

cluster of genes. Such locus control regions (LCRs) contain a complex collection of enhancer and silencer elements and confer epigenetic regulation of the linked genes in the locus. Many times, the LCRs contain abundant TE-derived sequences and S/MARs. Some of the best-characterized LCRs include the ones located at the human β -globin locus, MHC locus, and serpin gene locus. Importantly, each of these loci harbors several S/MARs that contain predominantly TE-derived repetitive DNA sequences, pointing to a link between the evolution of these loci and S/MARs (Rollini et al., 1999; Zhang and Qian, 2002; Ottaviani et al., 2008).

Put together, genetic signals, such as S/MARs, format the DNA for interaction with the nuclear architecture and ensure proper access to coding sequences. Among these, the Core-MARs that are lineage-specific confer a characteristic genome architecture independent of the coding sequence. In addition, as they are located in the non-coding regions, they are amenable to meaningful evolutionary changes. A subset of these elements, the combination of TEs and Core-MARs, may provide a specific mechanism for the generation of evolutionarily significant variations in the genome via changes in the domain architecture.

MAR/TE combinations as hotspot of evolutionary events

Apart from single-nucleotide changes, large-scale rearrangements such as inversions, translocations, and fusion

drive chromosome evolution. Studies reveal that chromosome breakpoints cluster in an area rich in repeats, S/MARs, and Topo II binding sites, indicating that the loop anchorage regions work as recombination hotspots (de la Paz et al., 2012). Mysterious clustering of fragile sites around S/MARs indicates that the chromatin structure around loop anchorage regions induces breakpoints that may be mediated by Topo II present on the site (Razin and Iarovaia, 2005). Interestingly, S/MARs often are hotspots for retroelement insertions (Narwade et al., 2019). As TEs naturally possess CREs, their insertions result in genome-wide dispersal of regulatory elements. At times, TEs may also carry cell-type-specific elements, adding a new regulatory property to the domain where they get inserted. The combination of TE and S/MAR insertions provides ready-to-use regulatory elements that can alter the properties of a chromatin domain, increasing the genomic reservoir that can be acted upon by natural selection. Any nonfatal transposition event retained in the genome could increase the possibilities of genetic drift and evolution. Thus, TE–S/MAR combinations may act as regions important for the restructuring of the genome during evolution. By studying TE–S/MAR in related species, it may also be possible to decipher the evolutionary record of architectural transformations.

Conclusion

S/MARs, as dispersed regulatory sequences, offer several potential functional attributes. Not only do they connect unlinked coding sequences into co-ordinately controlled subsystems by virtue of attachment to nuclear architecture but they can also act as drivers of chromosomal changes that are evolutionarily significant. The combination of S/MARs and TEs allows TE-mediated hopping of S/MARs that brings about the reformatting of topological and functional domains, a potentially valuable component of the evolutionary process. In this context, genome-wide sequencing of *in vivo* S/MARs from various sources is highly desirable. Understanding the dynamics of these elements through the cell cycle, tissue differentiation, and organismal development would help better define their role in the organization of chromatin. Placing them alongside other relevant chromatin landmarks such as TAD borders, enhancers, and promoters will help us to understand the variety of structural and functional roles played by them. Extensive genome-wide mapping of S/MARs from several closely related species will help us to better understand the functional and evolutionary relevance of these unique regulatory elements.

References

- Anthony, A., and Blaxter, M. (2007). Association of the matrix attachment region recognition signature with coding regions in *Caenorhabditis elegans*. *BMC Genom.* 8, 418. doi:10.1186/1471-2164-8-418
- Avramova, Z., Sanmiguel, P., Georgieva, E., and Bennetzen, J. L. (1995). Matrix attachment regions and transcribed sequences within a long chromosomal continuum containing maize *Adh1*. *Plant Cell* 7, 1667–1680. doi:10.1105/tpc.7.10.1667
- Avramova, Z., Tikhonov, A., Chen, M., and Bennetzen, J. L. (1998). Matrix attachment regions and structural colinearity in the genomes of two grass species. *Nucleic Acids Res.* 26, 761–767. doi:10.1093/NAR/26.3.761
- Berezney, R., and Coffey, D. S. (1974). Identification of a nuclear protein matrix. *Biochem. Biophys. Res. Commun.* 60, 1410–1417. doi:10.1016/0006-291X(74)90355-6
- Bode, J., Kohwi, Y., Dickinson, L., Joh, T., Klehr, D., Mielke, C., et al. (1992). Biological significance of unwinding capability of nuclear matrix-associating DNAs. *Sci.* 255, 195–197. doi:10.1126/SCIENCE.1553545
- Capco, D. G., Wan, K. M., and Penman, S. (1982). The nuclear matrix: three-dimensional architecture and protein composition. *Cell* 29, 847–858. doi:10.1016/0092-8674(82)90446-9

Data availability statement

The original contributions presented in the study are included in the article/Supplementary material; further inquiries can be directed to the corresponding author.

Author contributions

RP: writing—original draft. KP: writing—review and editing. RM: writing—review and editing, conceptualization, funding acquisition, project administration, and resources.

Funding

The authors declare that financial support was received for the research, authorship, and/or publication of this article. KP was supported by the UGC fellowship. Work in the RKM lab was supported by CSIR, Government of India grant (MLP0139), and SERB grant (GAP0466), a statutory body under the Department of Science and Technology, Government of India.

Acknowledgments

The authors thank RKM lab members for the continuous inputs and critical discussions.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The authors declare that they were editorial board members of *Frontiers*, at the time of submission. This had no impact on the peer review process and the final decision.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors, and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

- Chunduri, A. R., Rajan, R., Lima, A., Ramamoorthy, S., and Mamillapalli, A. (2022). Dynamics of nuclear matrix attachment regions during 5th instar posterior silk gland development in *Bombyx mori*. *BMC Genomics* 23, 247–318. doi:10.1186/s12864-022-08446-3
- dela Paz, J. S., Patti, E. S., Scott, J. D., Sandy, S., Clare, A. H., and Riggs, C. D. (2012). Chromosome fragile sites in Arabidopsis harbor matrix attachment regions that may be associated with ancestral chromosome rearrangement events. *PLoS Genet.* 8. doi:10.1371/journal.pgen.1003136
- Girod, P. A., Nguyen, D. Q., Calabrese, D., Puttini, S., Grandjean, M., Martinet, D., et al. (2007). Genome-wide prediction of matrix attachment regions that increase gene expression in mammalian cells. *Nat Methods.* 4, 747–753. doi:10.1038/nmeth1076
- Glazko, G. V., Koonin, E. V., Rogozin, I. B., and Shabalina, S. A. (2003). A significant fraction of conserved noncoding DNA in human and mouse consists of predicted matrix attachment regions. *Trends Genet.* 19, 119–124. doi:10.1016/S0168-9525(03)00016-7
- Greally, J. M., Gray, T. A., Gabriel, J. M., Song, L., Zemel, S., and Nicholls, R. D. (1999). Conserved characteristics of heterochromatin-forming DNA at the 15q11-q13 imprinting center. *Proc. Natl. Acad. Sci. U. S. A.* 96, 14430–14435. doi:10.1073/pnas.96.25.14430
- Jackson, D. A., and Cook, P. R. (1986). Replication occurs at a nucleoskeleton. *EMBO J.* 5, 1403–1410. doi:10.1002/j.1460-2075.1986.tb04374.x
- Jordan, I. K., Rogozin, I. B., Glazko, G. V., and Koonin, E. V. (2003). Origin of a substantial fraction of human regulatory sequences from transposable elements. *Trends Genet.* 19, 68–72. doi:10.1016/S0168-9525(02)00006-9
- Laemmli, U. K., Käs, E., Poljak, L., and Adachi, Y. (1992). Scaffold-associated regions: cis-acting determinants of chromatin structural loops and functional domains. *Curr. Opin. Genet. Dev.* 2, 275–285. doi:10.1016/S0959-437X(05)80285-0
- Liebach, I., Bode, J., Reuter, I., and Wingender, E. (2002). Evaluation of sequence motifs found in scaffold/matrix-attached regions (S/MARs). *Nucleic Acids Res.* 30, 3433–3442. doi:10.1093/nar/gkf446
- Linnemann, A. K., and Krawetz, S. A. (2009). Silencing by nuclear matrix attachment distinguishes cell-type specificity: association with increased proliferation capacity. *Nucleic Acids Res.* 37, 2779–2788. doi:10.1093/nar/gkp135
- Linnemann, A. K., Platts, A. E., and Krawetz, S. A. (2009). Differential nuclear scaffold/matrix attachment marks expressed genes. *Hum. Mol. Genet.* 18, 645–654. doi:10.1093/hmg/ddn394
- Mirkovitch, J., Mirault, M. E., and Laemmli, U. K. (1984). Organization of the higher-order chromatin loop: specific DNA attachment sites on nuclear scaffold. *Cell* 39, 223–232. doi:10.1016/0092-8674(84)90208-3
- Narwade, N., Patel, S., Alam, A., Chattopadhyay, S., Mittal, S., and Kulkarni, A. (2019). Mapping of scaffold/matrix attachment regions in human genome: a data mining exercise. *Nucleic Acids Res.* 47, 7247–7261. doi:10.1093/nar/gkz562
- Ottaviani, D., Lever, E., Mitter, R., Jones, T., Forshew, T., Christova, R., et al. (2008). Reconfiguration of genomic anchors upon transcriptional activation of the human major histocompatibility complex. *Genome Res.* 18, 1778–1786. doi:10.1101/gr.082313.108
- Padmaja, S. S., Lakshmanan, J., Gupta, R., Banerjee, S., Gautam, P., and Banerjee, S. (2010). Identification of Scaffold/Matrix Attachment (S/MAR) like DNA element from the gastrointestinal protozoan parasite *Giardia lamblia*. *BMC Genomics* 11, 386. doi:10.1186/1471-2164-11-386
- Pascuzzi, P. E., Flores-Vergara, M. A., Lee, T. J., Sosinski, B., Vaughn, M. W., Hanley-Bowdoin, L., et al. (2014). In vivo mapping of arabidopsis scaffold/matrix attachment regions reveals link to nucleosome-disfavoring poly(dA:dT) tracts. *Plant Cell* 26, 102–120. doi:10.1105/tpc.113.121194
- Pathak, R. U., Srinivasan, A., and Mishra, R. K. (2014). Genome-wide mapping of matrix attachment regions in *Drosophila melanogaster*. *BMC Genomics* 15, 1022. doi:10.1186/1471-2164-15-1022
- Paul, A. L., and Ferl, R. J. (1998). Higher order chromatin structures in maize and Arabidopsis. *Plant Cell* 10:1349–1359. doi:10.1105/tpc.10.8.1349
- Platt, R. N., Vandewege, M. W., and Ray, D. A. (2018). Mammalian transposable elements and their impacts on genome evolution. *Chromosom. Res.* 26, 25–43. doi:10.1007/S10577-017-9570-Z
- Razin, S. V., and Iarovaia, O. V. (2005). Spatial organization of DNA in the nucleus may determine positions of recombination hot spots. *Mol. Biol.* 39:543–548. doi:10.1007/S11008-005-0070-4
- Rebollo, R., Romanish, M. T., and Mager, D. L. (2012). Transposable elements: an abundant and natural source of regulatory sequences for host genes. *Annu. Rev. Genet.* 46, 21–42. doi:10.1146/annurev-genet-110711-155621
- Rollini, P., Namciu, S. J., Marsden, M. D., and Fournier, R. E. K. (1999). Identification and characterization of nuclear matrix-attachment regions in the human serpin gene cluster at 14q32.1. *Nucleic Acids Res.* 27, 3779–3791. doi:10.1093/NAR/27.19.3779
- Rudd, S., Frisch, M., Grote, K., Meyers, B. C., Mayer, K., and Werner, T. (2004). Genome-wide *in silico* mapping of scaffold/matrix attachment regions in Arabidopsis suggests correlation of intragenic scaffold/matrix attachment regions with gene expression. *Plant Physiol.* 135, 715–722. doi:10.1104/pp.103.037861
- Saha, P., Sowpati, D. T., Soujanya, M., Srivastava, I., and Mishra, R. K. (2020). Interplay of pericentromeric genome organization and chromatin landscape regulates the expression of *Drosophila melanogaster* heterochromatic genes. *Epigenetics Chromatin* 13, 41–19. doi:10.1186/s13072-020-00358-4
- Schmidt, D., Schwalie, P. C., Wilson, M. D., Ballester, B., Gonçalves, A., Kutter, C., et al. (2012). Waves of retrotransposon expansion remodel genome organization and CTCF binding in multiple mammalian lineages. *Cell* 148, 335–348. doi:10.1016/J.CELL.2011.11.058
- Sureka, R., Avvaru, A. K., Sowpati, D. T., Pathak, R. U., and Mishra, R. K. (2022). Structural and developmental dynamics of Matrix associated regions in *Drosophila melanogaster* genome. *BMC Genomics* 23, 725–818. doi:10.1186/s12864-022-08944-4
- Szabo, Q., Bantignies, F., and Cavalli, G. (2019). Principles of genome folding into topologically associating domains. *Sci. Adv.* 5, eaaw1668. doi:10.1126/sciadv.aaw1668
- Tikhonov, A. P., Bennetzen, J. L., and Avramova, Z. V. (2000). Structural domains and matrix attachment regions along colinear chromosomal segments of maize and sorghum. *Plant Cell* 12, 249–264. doi:10.1105/tpc.12.2.249
- Verma, S., Pathak, R. U., and Mishra, R. K. (2021). Genomic organization of the autonomous regulatory domain of eyeless locus in *Drosophila melanogaster*. *G3 Genes, Genomes, Genet.* 11, jkab338. doi:10.1093/G3/JOURNAL/JKAB338
- Wells, J. N., and Feschotte, C. (2020). A field guide to eukaryotic transposable elements. *Annu. Rev. Genet.* 54, 539–561. doi:10.1146/annurev-genet-040620-022145
- Zhang, S. B., and Qian, R. L. (2002). The interaction between the human beta-globin locus control region and nuclear matrix. *Cell Res.* 12, 411–416. doi:10.1038/sj.cr.7290144
- Zhu, L., Swergold, G. D., and Seldin, M. F. (2003). Examination of sequence homology between human chromosome 20 and the mouse genome: intense conservation of many genomic elements. *Hum. Genet.* 113, 60–70. doi:10.1007/s00439-003-0920-x