

Review

Satellite DNAs rising from the transposon graveyards

Eva Šatović-Vukšić^{1,*}, Patrik Majcen^{1,2}, and Miroslav Plohl¹

¹Division of Molecular Biology, Ruđer Bošković Institute, 10000 Zagreb, Croatia

²Elementary School Trnovec, 42202 Trnovec, Croatia

*Corresponding author. Division of Molecular Biology, Ruđer Bošković Institute, Bijenička 54, 10000 Zagreb, Croatia (Email: esatovic@irb.hr).

Abstract

Repetitive DNA sequences, as transposable elements (TEs) and satellite DNA (satDNA) spread and diversify within host genomes, impacting genome biology in numerous ways. In the first part of this review, we emphasize the evolutionary importance of satDNAs and TEs, providing a short summary of their roles and the mechanisms by which they influence the structure and function of genomes. We also discuss the broad, complex, and extensive relationships between TEs and satDNAs. Following that, we bring together different mechanisms on the generation of satDNA from TE, as it has been demonstrated that almost any part of any type of TE can undergo tandemization and produce novel satDNAs. Importantly, we here present a hypothesis that would explain the existence of particular types of monomers, namely composite satDNA monomers which display multiple subsequent stretches of similarity to various TEs, for which the explanation was lacking so far. We propose that even highly shuffled and degraded TE remnants residing in heterochromatin 'TE graveyards' can give rise to new satDNA sequence monomers, transforming these genomic loci into DNA 'recycling yards'. Furthermore, we emphasize important evolutionary questions regarding the causes, mechanisms, and frequency of these occurrences.

Keywords: satellite DNAs; transposable elements; heterochromatin; genome evolution.

1. Introduction

Repetitive DNA components build a significant portion of eukaryotic genomes. They are traditionally divided into 2 major groups, satellite DNAs (satDNAs), comprised of arrays of sequences repeated in tandem, and transposable elements (TEs), interspersed throughout the genome.^{1–5} According to conventional concepts, satDNAs are dominant in the pericentromeric, subtelomeric, and interstitial chromosomal regions, where they constitute blocks of constitutive heterochromatin.⁶ The most basic classification of TEs divides them into 2 main groups according to their transposition intermediates: Class I (Retrotransposons), and Class II (DNA transposons). In the 'copy-and-paste' mechanism employed by Class I elements, an RNA intermediate is present. The majority of Class II elements are mobilized by 'cut-and-paste' process, in which the transposon is excised and transferred to a new genomic site.⁷

SatDNAs and TEs are often called the 'dark matter' of the genome, as their functions were initially unknown. In addition, the repetitive nature of these sequences causes significant technical problems in sequencing and assembly, resulting in their general underrepresentation in outputs of genome projects.^{8,9} New sequencing methodologies, in particular long-read sequencing supported by specialized bioinformatics tools, are on the way to solve these problems and to shed more light on the genomic composition of the 'dark matter' repeats.^{8,10} As a result, scientific literature has extensively covered numerous topics related to the structure, organization, function, and evolution of these sequences in diverse model and non-model organisms, enabling new insights into

repetitive DNAs.^{6,11–19} Both satDNAs and TEs have tremendous impact on genome architecture and evolution, making them crucial players in the process of comprehending the overall structure and function of the genome.

In this article, we bring the basic notions on satDNAs and TEs, summarize their evolutionary significance and their crucial contributions to the constitution and evolution of eukaryotic genomes. We discuss the intricate and extensive connections between TEs and satDNAs, as new data are enhancing our understanding of these relationships. We summarize various mechanisms regarding the generation of satDNAs from TEs. Notably, we present a hypothesis that seeks to explain the existence of satDNA sequences that share short stretches of similarity with multiple TEs, suggesting that these satDNA sequences may originate from 'TE graveyards'.

2. Evolutionary significance of satDNAs and TEs

TEs and satDNAs play crucial roles in various processes that significantly influence the structure and function of the genome. While TEs and satDNAs differ in their structures, mechanisms of dissemination, and organizational patterns, there are numerous connections between them, which will be explored in the following sections. On one hand, they often complement each other in some of the genomic functions (Fig. 1). On the other hand, due to substantial differences in their structure and organization, many of the roles that TEs and satDNAs fulfil in the genome are specific to certain subclasses of these repetitive elements.

Received 24 April 2025; revised 2 September 2025; accepted 15 September 2025

© The Author(s) 2025. Published by Oxford University Press on behalf of Kazusa DNA Research Institute.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

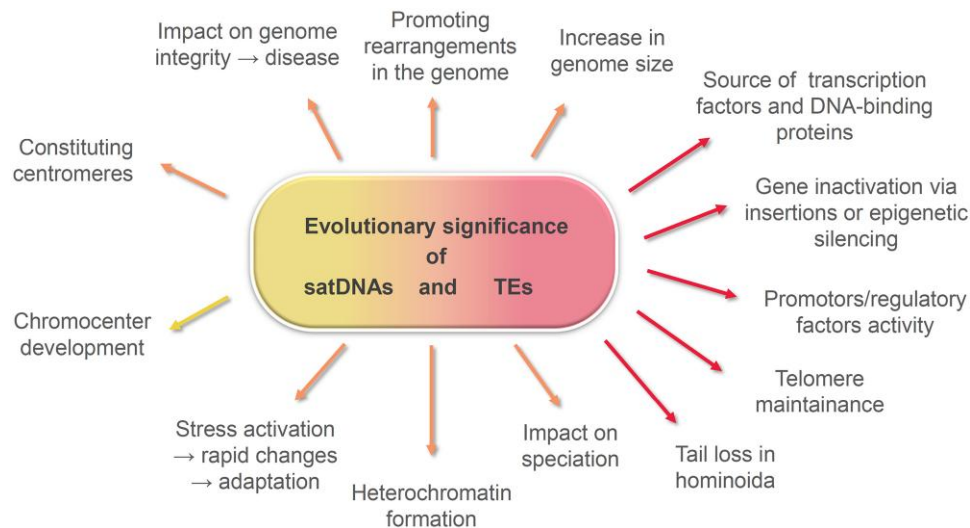


Fig. 1. Schematic depiction of the major contributions of satDNAs (yellow) and TEs (red) to genome structure, function, and evolution. Shared roles are denoted by orange arrows.

As the main common role, TEs and satDNAs shape the overall genomic landscape by inducing structural rearrangements through illegitimate recombinations, deletions, inversions, translocations, and segmental duplications. Both TEs and satDNAs can make significant changes to the genome size.^{20–25}

The domestication of TEs has yielded numerous proteins, which are co-opted for functions in physiology of many species. For instance, numerous DNA-binding proteins and transcription factors are derived from transposases, and the majority of primate-specific regulatory sequences are derived from TEs (7). In addition, TE machinery serves as a mediator in the formation of intron-free gene copies (retrogenes), which can lead to the evolution of a novel trait through neofunctionalization.²⁶ TE can also serve as telomeric sequences at the ends of chromosomes, such as in *Drosophila*.²⁷

TEs extensively affect regulatory networks that are involved in processes like dosage compensation, immunity, and early embryonic development. There are numerous examples of changes in phenotype driven by TE activity, some of them being beneficial for the organism.¹² The colour polymorphism of the peppered moth is a well-known case of such adaptive, TE-induced change.²⁸ Recently, their involvement in the mechanism that facilitated tail-loss process in hominoids has been proposed.²⁹ In addition to the direct gene inactivation, TE insertions can also affect the expression of nearby genes by alterations at the epigenetic level (eg histone modifications and chromatin packing).³⁰ TE Helitrons were shown to be powerful genome shuffling agents with wide-reaching biological consequences.³¹ TE movement, regulatory activities, and effects on genome integrity can also cause (and intensify) the effects of many diseases.⁷ In addition, differences in TEs distribution can influence speciation through the formation of reproductive isolation.³²

Centromeres are essential for proper segregation of chromosomes, and satDNAs and TEs are the most common DNA component in centromeres of animals and plants.³³ Typically, satDNAs and TEs extend to the pericentromeric regions much more than it would be necessary for the centromeric function alone. Pericentromeric satDNA repeats were

shown to be the main contributors to large-scale nuclear organization that supports general transcription.³⁴ In addition, satDNA sequences have been continuously associated to evolutionary breakpoint areas and fragile sites in a variety of taxa.¹³ There are numerous neurological, congenital, and developmental disorders caused by short tandem repeat expansions,³⁵ and changes in the repeat copy number can also affect social behaviour.³⁶

SatDNAs have key roles in heterochromatin formation and maintenance,³⁷ and contribute to heterochromatin organisation in embryonic stem cells.³⁸ TEs were also suggested to be mediators and facilitators of heterochromatin formation, through recruitment of Heterochromatin Protein 1 and repressive chromatin marks.³⁹ Different studies revealed that satDNA transcription and satDNA transcripts are involved in various cellular processes, while the improper regulation of satDNA transcription was shown to be associated with genomic instability and human diseases.^{13,40} Pericentromeric satDNA transcription is significantly elevated across many cancers, and these transcripts have a variety of biological functions in either aiding the cancerous state (mutation induction, disruption of epigenetic regulation, tumour cell proliferation, inflammation, resistance to cancer treatment, destabilizing genome integrity), or opposing it (innate immune system activation). Elevated transcription of pericentromeric satDNAs is typically triggered by various environmental stressors, initializing the mechanism that can affect regulation of many genes through reversible changes in the chromatin state.⁴⁰ Under stress conditions, TEs are also activated. This activation may modify genes structure and activity of genes, aiding in the process of adaptation and survival.^{41,42}

SatDNAs are involved in other numerous processes that are crucial for the cells. For example, it has been demonstrated that the formation of the chromocenter and the maintenance of the entire genome within the nucleus rely on the presence of satDNA repeats located across multiple chromosomes. The chromocenter contains DNA-binding proteins and physically links different chromosomes by bringing together their corresponding pericentromeric satDNA repeats.⁴³

3. Relationships between satDNAs and TEs

An increasing number of reports shows that satDNAs and TEs are interconnected in various ways (reviewed in Šatović-Vukšić and Plohl¹⁶ and Zattera and Bruschi⁴⁴). The heterochromatin is the most frequent site of interaction, as both types of repetitive sequences are particularly prevalent in this genomic region. There, they not only coexist but also form complex and dynamic networks. It was observed that 1 satDNA array can be directly followed by satDNA of another type, and different types of TEs can be simultaneously found in the immediate vicinity of a satDNA. TEs can be inserted into satDNA arrays or inserted into other TEs.⁴⁵ Moreover, cases of multiple insertions have been reported, such as a TE being inserted within another TE, both of which are located within a satDNA array.^{46,47}

Neither satDNA nor TEs are exclusively limited to heterochromatin.^{16,48} SatDNA sequences outside of the heterochromatin can be found in different organizational forms: as single monomers or monomer fragments, in arrays of various (usually short) length, or incorporated into TEs.^{45,49–55}

TEs significantly contribute to satDNA evolution by generating repeats that can be dispersed through the genome, and in some cases, amplified into long arrays of novel satDNAs. SatDNA repeats of various species were formed by tandemization of a complete TE or its subsegments and structural components.^{56–60} This way, satDNA can arise from the sequence segment of a long-terminal repeat (LTR), gag or pol domains of LTR retrotransposons; untranslated regions of LINE; terminal inverted repeats or sequence segments from the central parts of DNA transposons.^{44,61} Additionally, satDNAs can be formed through the expansion of short-internal arrays found within TEs.^{50,52,62,63} A common example of such expansions is observed in Helitron/Helitron TEs, which have conserved sequence segments at their ends, while their central regions often contain arrays of satDNA-like repeats.

TEs are also proposed to facilitate and contribute to genomic dispersal of satDNA repeats.^{45,51,62,64–66} In line with that, it was shown that TEs are responsible for novel, highly dispersed organization of numerous satDNAs across the whole-genome, completely contrasting canonical concepts of compartment-localized satDNA organization.⁵⁴ In extreme cases, complete satelitomes (entirety of satDNAs in 1 organism), comprised of numerous satDNAs, can be based on TEs (being TE-derived, TE-incorporated, or TE-propagated).⁵⁵

Hybrid forms between satDNAs and TEs also exist. For example, ‘transitional’ 154TR sequence is at the same time a tandem repeat embedded in a TE, but also found as large expanded arrays within constitutive heterochromatic loci, similar to classical satDNAs.⁶⁷ The Cg170/HindIII sequence displays fluctuation between 3 forms, TE-incorporated, standalone satDNA arrays, and an ‘intermediate’ form. In the case of the latter, tandem repeats were found to be associated with TE Helitron components only on one side of the array.⁵⁵ These intermediate arrays may result from recombination events between element-incorporated and standalone arrays.⁵⁴ Alternatively, they may be generated by aberrant rolling circle replication (RCR), as will be discussed later.

4. Mechanisms that generate satDNAs (from TEs)

Regarding the mechanisms of satDNA generation, general models of satDNA evolution suggest 2 stages in their

emergence: amplification processes which generates small number of tandem repeats, followed by their expansion into longer arrays.⁶⁸ Formation of tandem repeats can occur through different mechanisms. For example, in a process not necessarily related to TEs, such DNA replication. During replication, the displacement of the DNA strand can happen, resulting in mispairing of the complementary bases and loop formation. Staggered mispairing in the leading DNA strand results in a duplication of the template sequence⁶⁹ (Fig. 2a). These simple repeats can expand into satDNA arrays via unequal exchange.⁶⁸

Different hypotheses exist, providing potential explanations for the formation of satDNA sequences from TEs or from repeats present within TEs. Some of them are based on the aforementioned concept of loop formation.

Izsvák et al.⁷⁰ propose the model explaining generation of tandem repeats from a TE, based on the construction of a stem-and-loop structure. During DNA replication, after passing the TE, the newly synthesized DNA strand can fold back on itself, creating a loop structure. Subsequently, this looped strand may disassociate from the replication complex. DNA synthesis then reinitiates at the 3' end of the loop, using the nascent strand as a new template to replicate the TE again, and forming the stem. The duplicated segment, containing 2 copies of the TE, is then released as an extrachromosomal stem-and-loop structure (Fig. 2b). This structure can be incorporated into a new location in the genome, facilitated by local homology between the target DNA sequence and the amplified extrachromosomal fragment. The mechanism is based on studying the *Angel* MITE in zebrafish, where the ability of intrastrand base pairing of single-stranded *Angel* molecules was demonstrated in vitro.⁷⁰

The mechanism for the formation of satDNAs from Miniature Inverted-repeat TEs (MITEs) was put forward by Hikosaka and Kawahara,⁷¹ based on studying *Xstir* sequences in the genomes of several *Xenopus* species. Three types of *Xstir*-related structures were observed, including MITE, tandem array, and a composite structure of MITE and tandem array. The alignment analyses revealed that tandem repeats may be derived from internal sequences of the MITE. The proposed mechanism involves the formation of a stem-and-loop structure by the 2 elements in close proximity. This structure forms during DNA replication due to the delay on 1 DNA strand. During the delay, nucleases excise the loop portion of the stem-and-loop structure, and the remaining segments are joined together. As a result, a ‘hybrid’ element is formed, containing segments from both of the original elements. Repeating this process contributes to the additional extension of the construct. The resulting sequences still represent interspersed repeats; however, with the involvement of recombination processes, they could progress into arrays of tandem repeats, satDNAs (Fig. 2c).

The mechanism proposed by Scalvenzi and Pollet⁶² explains the evolution of MITE and related satDNAs. It is based on the Tc1/mariner MITE in *Xenopus*, named miDNA4. MiDNA4 possesses a satellite DNA that exists as a single monomer or as an array of a variable number of copies. They suggest that the ancestral MITE captured a pre-existing tandem repeat. The satDNA-like sequence they described is flanked by AT-rich sequences that form short direct repeats of 5 or more base pairs. They propose that these microhomologies can lead to internal deletions and integrations during the processes of DNA replication or repair. This can result in longer

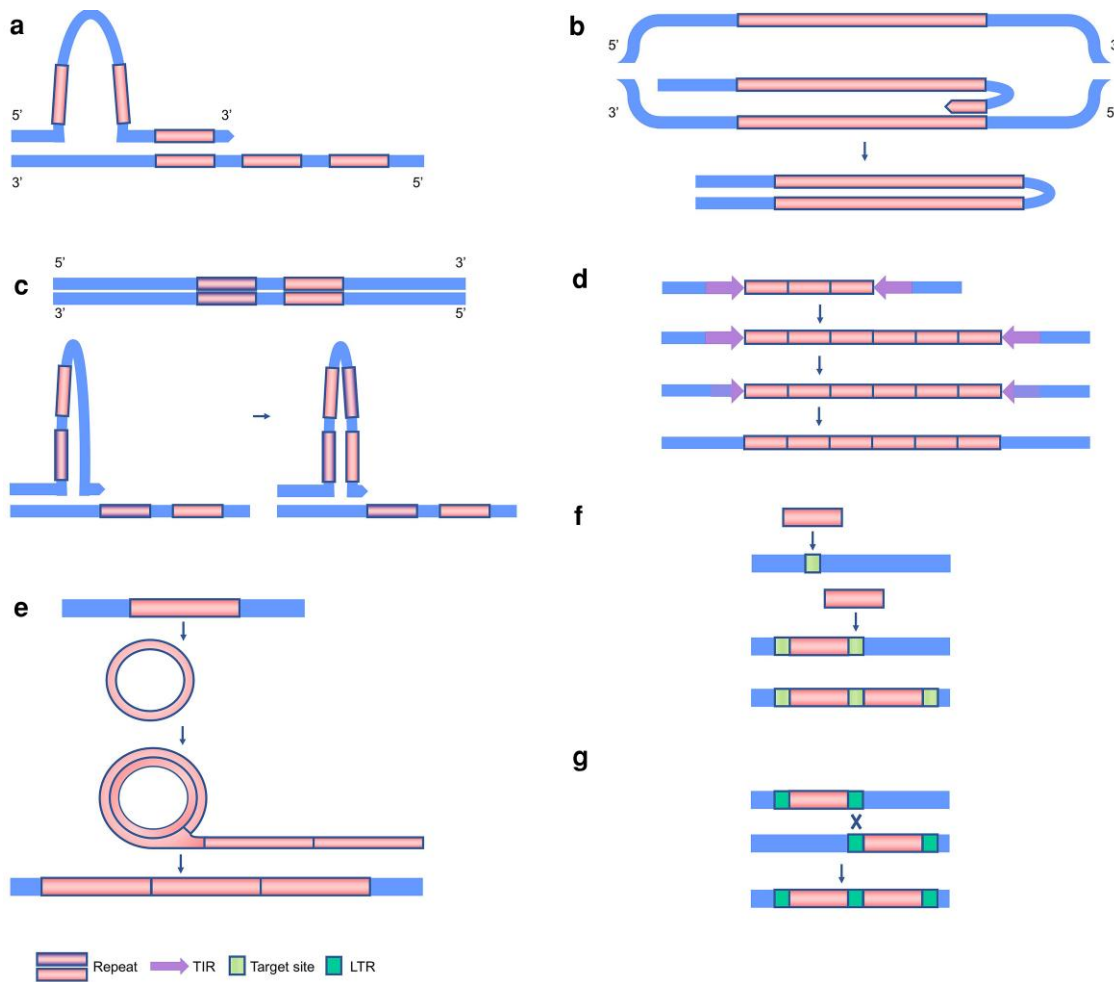


Fig. 2. Schemes illustrating different mechanisms for the generation of satDNAs or tandem repeats. a) Staggered mispairing in the leading DNA strand results in a duplication of the template sequence. Adapted from Moran and Morish.⁶⁹ b) Formation of the duplicated segment after newly synthesized DNA strand folds back on itself creating a loop structure. The newly formed extrachromosomal stem-and-loop structure can be incorporated into a new location in the genome. Adapted from Izsvák *et al.*⁷⁰ c) Molecular mechanisms for creating tandem repeats from TEs in close proximity. Stem-and-loop structure is formed during replication by the 2 elements in close proximity. Nucleases excise the loop portion of the stem-and-loop structure, and the remaining segments are joined. Repeating this process contributes to the additional extension of the construct. Adapted from Hikosaka and Kawahara.⁷¹ d) The expansion of tandem repeats within MITE and formation of satDNA arrays. Tandem repeats within TEs can experience copy number changes. Subsequently, TEs containing longer tandem arrays can give rise to classical satDNA arrays after accumulating mutations in the terminal segments. Adapted from Scalvenzi and Pollet.⁶² e) RCR mechanism contributing to the formation of tandem repeats. Terminator sequence can be removed from the circular DNA template resulting in aberrant replication, followed by insertion of the resulting concatemer. Adapted from McGurk and Barbash⁶⁸ and Xiong *et al.*⁷² f) Multiple insertion of the same type of TE into a single genomic site, forming a tandem. Adapted from McGurk and Barbash.⁶⁸ g) Generation of tandem sequences via recombination between LTRs of 2 different elements. Adapted from McGurk and Barbash.⁶⁸

internal arrays. Over time, MITEs containing longer tandem arrays can give rise to classical satDNA arrays devoid of inverted repeats at element ends, by accumulating mutations in the terminal segments (Fig. 2d).

TE Helitrons have the ability to capture various fragments of the host genome,⁷³ which often include tandem repeats, as shown in multiple studies.^{65,74–78} DNA motifs that are promoting the formation of stem-and-loop structures, like direct and inverted repeats or palindromes, can also be found within the structure of Helitron/Helitron elements.⁷³ They enable the Helitrons to participate in the above-described processes. In addition, Helitrons employ several mechanisms for their propagation and for the amplification or tandemization of their segments. They utilize an RCR mechanism during transposition.⁷² The RCR initiates at the 5' end and advances towards the 3' end, where a terminal hairpin structure serves as a recognition site for termination and subsequent DNA

cleavage. Through an intramolecular recombination event involving internally repeated sequence, the 3' terminator sequence can be removed from the circular DNA template. If that occurs, the subsequent cycle of replication generates a tandem array of truncated TEs. This causes the sequences at the Helitron's 5' end to become amplified more frequently than those at the 3' end. Such formation of incomplete Helitron templates may contribute to further partial RCR, resulting in additional multiplication of the internal sequences (Fig. 2e). Finally, the tandemized, truncated Helitron copies are integrated into the new genomic location.⁷² The mechanism proposed by Xiong *et al.*⁷² is based on their analysis of 27 plant genomes, which revealed numerous tandem arrays of partially decayed, truncated Helitrons. Many of the detected arrays had multiple 5' but single 3' Helitron end, while the number of repeats in arrays ranged from several to several hundreds.

By analysing numerous genomes from 5 populations of *Drosophila melanogaster*, McGurk and Barbash⁶⁸ revealed that TEs commonly form dimers. Their results suggest that insertion site preference is the major mechanism by which dimers are formed and that their formation is related to the periods of active transposition. They believe that the abundance of TE dimers has the potential to provide source material for expansion into satDNA arrays, based on their discovery of copy number expansion of the DNA transposon hobo to 16 tandem copies. These authors present multiple potential mechanisms for generating tandem repeats from TEs, 1 of which is the previously described RCR. The second mechanism relies on a double or multiple insertion of the same type of element into a single genomic site, forming a tandem. This is possible for TEs that create target site duplications (TSDs) upon insertion, as this allows for subsequent insertion(s) into the same target site (Fig. 2f). The third mechanism focuses on satDNAs derived from TEs containing repetitive segments such as LTRs and/or tandemly repeated regulatory motifs. These segments serve as substrates for expansion by unequal exchange (Fig. 2g). For example, ectopic recombination between LTRs of 2 different elements could generate tandemized retrotransposons that share 1 LTR.⁶⁸ Similar mechanism was proposed by Wong and Choo,⁷⁹ based on similarities between TE components and satDNAs reported in various organisms, such as wheat, *Arabidopsis*, *Drosophila*, and the Cetaceans. They suggested that satDNA repeats originate from the duplication of a portion of a TE sequence. This duplication occurs through unequal crossing-over between homologous TE elements, which may be located on the same chromosome or on different chromosomes, presumably in a similar way as in Fig. 2g.

Zattera and Bruschi⁴⁴ update and summarize recorded cases of TEs that have given rise to the tandem repeat sequences. Based on that, they propose that Non-Homologous End-Joining (NHEJ) and Non-Allelic Homologous Recombination (NAHR) DNA repair mechanisms may contribute to the expansion of satDNAs from TE. If microhomology-initiated NHEJ occurs between sister chromatids, it can result in a variety of events, including sequence duplications. This mechanism is known to play a dominant role in gene duplications and is significant in completing some TE-related instability events.⁴⁴ The NAHR mechanism offers important insights into the expansion of repeat arrays. It relies on the location of the paralogous TE copies, which serve as the template for the repair of the chromatid that has suffered the double-strand break. This may lead to tandem duplications of TEs (in the same way as depicted in Fig. 2g). This way, intra- or inter-chromatid NAHR may contribute to the expansion of the initial repeat in a way similar to the unequal crossovers.⁴⁴

5. SatDNAs emerging from TE heterochromatin graveyards

As discussed above, it has previously been observed that particular satDNA could be derived from a particular TE. Here, we would like to discuss 1 specific type of composite satDNA monomers, which has not been addressed previously. In several instances, we have observed that the sequence of a satDNA monomer can contain multiple short segments that resemble various types of TEs.^{54,55} In these cases, whether the satDNAs originated from TEs was debatable as they could

not be assigned to any specific TE. The origin and the characteristics of monomer sequences structured in this way are currently an open question, as they appear to result from a series of complex events. In this chapter, we propose a scenario that aims to explain their origin.

TEs are frequently enriched in constitutive heterochromatin, part of the genome with a reduced number of functional genes.³⁹ TEs can insert themselves next to the other repetitive sequences via shared TSD, as described above. Additionally, they can integrate into other repetitive sequences, including previously inserted TEs of the same or different type. The insertion of TEs into other TEs produces what is known as nested insertion.⁸⁰ Therefore, multiple insertions are common, and such loci are known to serve as hotspots and target sites for further TEs insertions.^{46,81} Sometimes, due to the imperfect process of transposition, truncated versions of TEs may be integrated,⁷² increasing the diversity of sequences at the insertion sites.

Many of the inserted copies of TEs progressively accumulate mutations and deletions over time.⁸² Genomic loci that are rich in numerous mutated and/or truncated elements in close proximity were observed.^{47,55} This can result in genomic segments that present short stretches of similarity with different TEs, and over time, their remnants eventually diverge to the point of almost being unrecognizable as TEs.⁸² Related to this, the term 'graveyard of dead transposons' has often been used for heterochromatin because it harbours numerous remnants from ancient TE insertions,⁸³ features a 'clustered-scrambled' organization and a high density of repetitive sequences,⁸⁴ and contains piRNA clusters that are particularly enriched in TE relics.⁸⁵

We have reported the presence of the complex and 'shuffled' loci in the genome of the bivalve *Crassostrea gigas*.⁴⁷ These loci appear to be generated by insertion, deletion, tandemization, and recombination events, involving satDNAs and structural components of the Helitron TE. In these instances, different parts of the elements were truncated, or tandemized, and such structures were found inserted within other (complete or truncated) elements. We found central arrays of tandem repeats within Helitron/Helitron elements that were oriented in different directions or organized in arrays of different lengths. The same type of arrays was observed both within and outside of the elements, and different types of tandem arrays were identified within a single element, as well as the same type of arrays being found across different Helitron elements.⁴⁷

While monomer sequences of different satDNAs have very little or nothing in common, they frequently contain distinct sequence features, such as conserved motifs, inverted repeats, and palindromes.⁸⁶ It has been proposed that these structural features may play a role in providing signals that aid mechanisms responsible for the fast proliferation of satDNA repeats, both within arrays and throughout the genome.⁴⁵

In the genomic sites described above, DNA segment tandemization could occur through different mechanisms, aided by such sequence motifs which often accompany TEs. We propose that the tandemization and subsequent propagation of DNA segments from these genomic regions would result in novel satDNA repeats. These repeats are likely to include short stretches of similarity to different types of TEs within the monomer sequence (Fig. 3). Therefore, such satDNA sequences would indeed have their origin in what we refer to as 'TE graveyards'. These composite satDNA sequences reveal that

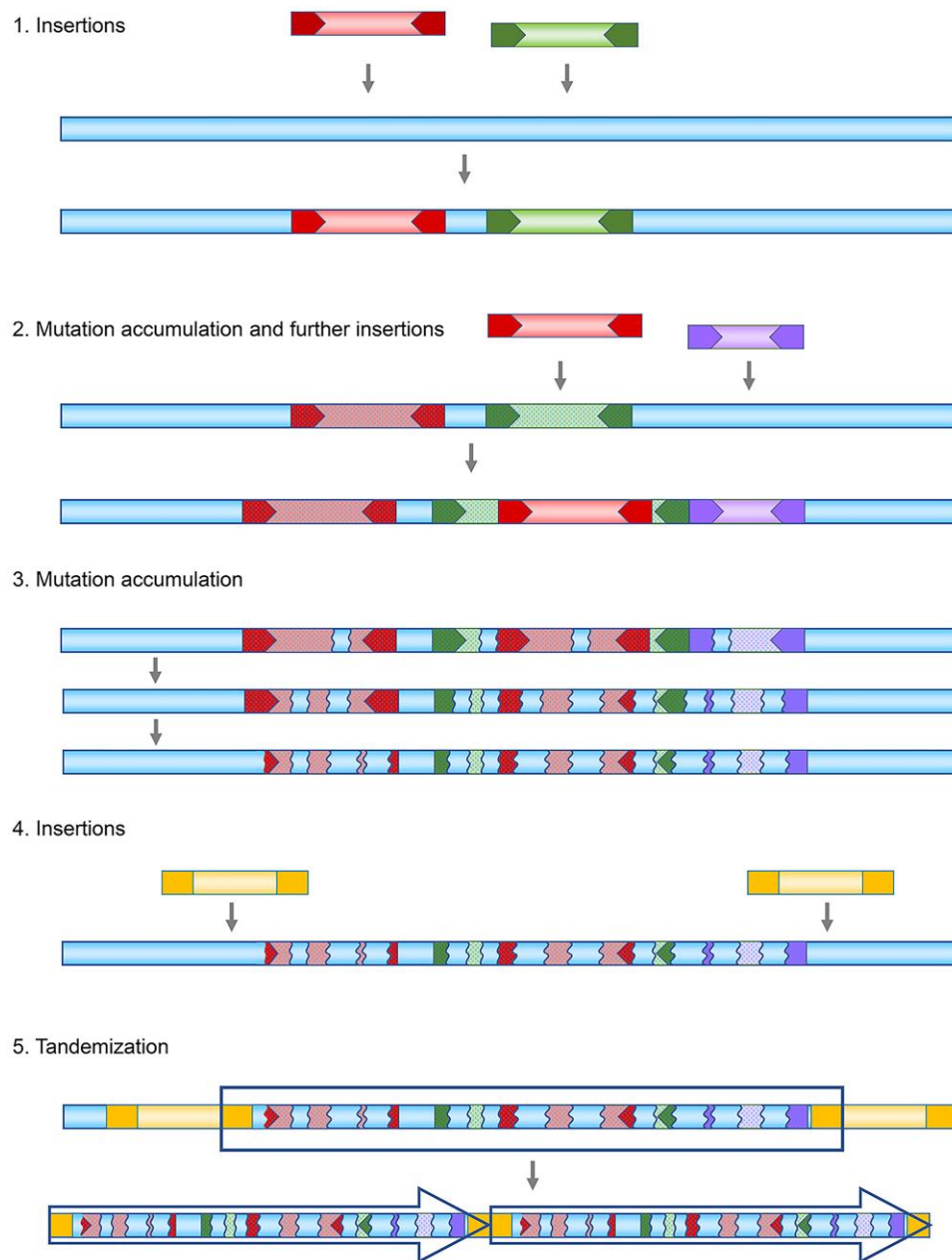


Fig. 3. The proposed mechanism for forming composite satDNA monomers with short stretches of similarity to multiple TEs through a series of interlacing events, including insertions, mutation accumulation, and sequence degeneration.

the DNA found in ‘graveyards’ is recycled, transforming heterochromatin TE graveyards into dynamic DNA ‘recycling yards’.

It has been proposed that degraded copies of interspersed repeats may constitute a significant fraction of unassigned genome sequences.⁸⁷ The potential roles of these deteriorated elements as resources for host genomes have also been discussed. Truncated TE copies can modulate host gene expression by serving as new regulatory sequences, alternative splice sites, polyadenylation signals, and new transcription factor binding sites.⁸⁸ We would add the formation of novel satDNA sequences to the potential usage of these remnants.

The considerations presented raise compelling evolutionary questions: How does selection operate on the newly generated satDNA sequences? Do the newly generated composite satDNA sequences acquire functional roles?

Whether these sequences obtain important roles in the genome and participate in functional interactions or not, this underscores a fascinating reality of genomes utilizing the available DNA with remarkable efficiency. Composite satDNA sequences show us that even the DNA residing in ‘graveyards’ is being recycled, transforming heterochromatin TE graveyards into productive ‘recycling yards’.

6. Conclusions and future perspectives

In this article, we have outlined the importance of satDNA sequences and TEs in genome evolution and architecture, and their involvement in various genomic processes and functions. It is also evident that TEs and satDNAs form a complex network of sequences that significantly impact the structure and, ultimately, the functionality of every eukaryotic genome.

The intricate and wide-ranging connections between TEs and satDNAs show that these 2 types of sequences exist in a genome in various (and transient) forms. Here, we put forward the idea that even heavily shuffled and degraded TE remnants residing in the ‘heterochromatin graveyards’ give rise to the novel composite satDNA sequences, turning such genomic loci into DNA ‘recycling yards’.

This observation offers a foundation for various future studies that will explore the causes, mechanisms, and possible functional significance of TE-based composite repetitive components in eukaryotic genomes. Gaining insights into these aspects will deepen our understanding of the broader implications of such sequences in genome architecture and function.

Funding

None declared.

Conflict of interest

None declared.

References

- Charlesworth B, Sniegowski P, Stephan W. The evolutionary dynamics of repetitive DNA in eukaryotes. *Nature*. 1994;371:215–220. <https://doi.org/10.1038/371215a0>
- Schmidt T, Heslop-Harrison JS. Genomes, genes and junk: the large-scale organization of plant chromosomes. *Trends Plant Sci*. 1998;3:195–199. [https://doi.org/10.1016/S1360-1385\(98\)01223-0](https://doi.org/10.1016/S1360-1385(98)01223-0)
- Jurka J, Kapitonov VV, Kohany O, Jurka MV. Repetitive sequences in complex genomes: structure and evolution. *Annu Rev Genomics Hum Genet*. 2007;8:241–259. <https://doi.org/10.1146/annurev.genom.8.080706.092416>
- López-Flores I, Garrido-Ramos MA. The repetitive DNA content of eukaryotic genomes. *Genome Dyn*. 2012;7:1–28. <https://doi.org/10.1159/000337118>
- Biscotti MA, Olmo E, Heslop-Harrison JS. Repetitive DNA in eukaryotic genomes. *Chromosome Res*. 2015;23:415–420. <https://doi.org/10.1007/s10577-015-9499-z>
- Garrido-Ramos MA. Satellite DNA: an evolving topic. *Genes (Basel)*. 2017;8:1–41. <https://doi.org/10.3390/genes8090230>
- Wells JN, Feschotte C. A field guide to eukaryotic transposable elements. *Annu Rev Genet*. 2020;54:539–561. <https://doi.org/10.1146/annurev-genet-040620-022145>
- Sedlazeck FJ, Lee H, Darby CA, Schatz MC. Piercing the dark matter: bioinformatics of long-range sequencing and mapping. *Nat Rev Genet*. 2018;19:329–346. <https://doi.org/10.1038/s41576-018-0003-4>
- Wallrath LL, Rodriguez-Tirado F, Geyer PK. Shining light on the dark side of the genome. *Cells*. 2022;11:330. <https://doi.org/10.3390/cells11030330>
- Šatović E, Tunjić-Cvitanović M, Plohl M. Tools and databases for solving problems in detection and identification of repetitive DNA sequences. *Period Biol*. 2020;121–122:7–14. <https://doi.org/10.18054/pb.v121-122i1-2.10571>
- Slotkin RK, Martienssen R. Transposable elements and the epigenetic regulation of the genome. *Nat Rev Genet*. 2007;8:272–285. <https://doi.org/10.1038/nrg2072>
- Chuong EB, Elde NC, Feschotte C. Regulatory activities of transposable elements: from conflicts to benefits. *Nat Rev Genet*. 2017;18:71–86. <https://doi.org/10.1038/nrg.2016.139>
- Louzada S et al. Decoding the role of satellite DNA in genome architecture and plasticity—an evolutionary and clinical affair. *Genes (Basel)*. 2020;11:72. <https://doi.org/10.3390/genes11010072>
- Thakur J, Packiaraj J, Henikoff S. Sequence, chromatin and evolution of satellite DNA. *Int J Mol Sci*. 2021;22:4309. <https://doi.org/10.3390/ijms22094309>
- Liao X et al. Repetitive DNA sequence detection and its role in the human genome. *Commun Biol*. 2023;6:954. <https://doi.org/10.1038/s42003-023-05322-y>
- Šatović-Vukšić E, Plohl M. Satellite DNAs—from localized to highly dispersed genome components. *Genes (Basel)*. 2023;14:742. <https://doi.org/10.3390/genes14030742>
- Flynn JM, Yamashita YM. The implications of satellite DNA instability on cellular function and evolution. *Semin Cell Dev Biol*. 2024;156:152–159. <https://doi.org/10.1016/j.semcdb.2023.10.005>
- Fonseca-Carvalho M et al. Answering the cell stress call: satellite non-coding transcription as a response mechanism. *Biomolecules*. 2024;14:1–14. <https://doi.org/10.3390/biom14010124>
- Cabral-de-Mello DC, Palacios-Gimenez OM. Repetitive DNAs: the ‘invisible’ regulators of insect adaptation and speciation. *Curr Opin Insect Sci*. 2025;67:101295. <https://doi.org/10.1016/j.cois.2024.101295>
- SanMiguel P et al. The paleontology of intergene retrotransposons of maize. *Nat Genet*. 1998;20:43–45. <https://doi.org/10.1038/1695>
- Naville M et al. Massive changes of genome size driven by expansions of non-autonomous transposable elements. *Curr Biol*. 2019;29:1161–1168.e6. <https://doi.org/10.1016/j.cub.2019.01.080>
- Lehmann R et al. DNA transposon expansion is associated with genome size increase in mudminnows. *Genome Biol Evol*. 2021;13:1–14. <https://doi.org/10.1093/gbe/evab228>
- Mora P et al. Making the genome huge: the case of *Triatoma del-ponte*, a triatominae species with more than 50% of its genome full of satellite DNA. *Genes (Basel)*. 2023;14:371. <https://doi.org/10.3390/genes14020371>
- Majid M, Yuan H. Comparative analysis of transposable elements in genus *Calliptamus* grasshoppers revealed that satellite DNA contributes to genome size variation. *Insects*. 2021;12:1–18. <https://doi.org/10.3390/insects12090837>
- Oppert B et al. The genome of the yellow mealworm, *Tenebrio molitor*: it’s bigger than you think. *Genes (Basel)*. 2023;14:2209. <https://doi.org/10.3390/genes14122209>
- Schrader L, Schmitz J. The impact of transposable elements in adaptive evolution. *Mol Ecol*. 2019;28:1537–1549. <https://doi.org/10.1111/mec.14794>
- George JA et al. Genomic organization of the *Drosophila* telomere retrotransposable elements. *Genome Res*. 2006;16:1231–1240. <https://doi.org/10.1101/gr.5348806>
- van’t Hof AE et al. The industrial melanism mutation in British peppered moths is a transposable element. *Nature*. 2016;534:102–105. <https://doi.org/10.1038/nature17951>
- Xia B et al. The genetic basis of tail-loss evolution in humans and apes. *Nature*. 2024;626:1042–1048. <https://doi.org/10.1038/s41586-024-07095-8>
- Horvath JE, Slotte T. Short tandem repeat variation and epigenetic control. *Genome Biol*. 2017;18:180. <https://doi.org/10.1186/s13059-017-1339-3>
- Grabundzija I et al. A Helitron transposon reconstructed from bats reveals a novel mechanism of genome shuffling in eukaryotes. *Nat Commun*. 2016;7:10716. <https://doi.org/10.1038/ncomms10716>
- Serrato-Capuchina A, Matute D. The role of transposable elements in speciation. *Genes (Basel)*. 2018;9:254. <https://doi.org/10.3390/genes9050254>
- Plohl M, Meštrović N, Mravinac B. Centromere identity from the DNA point of view. *Chromosoma*. 2014;123:313–325. <https://doi.org/10.1007/s00412-014-0462-0>
- Podgornaya OI. Nuclear organization by satellite DNA, SAF-A/hnRNPU and matrix attachment regions. *Semin Cell Dev Biol*. 2022;128:61–68. <https://doi.org/10.1016/j.semcdb.2022.04.018>
- Chintalapathi SR, Pineda SS, Deveson IW, Kumar KR. An update on the neurological short tandem repeat expansion disorders and the emergence of long-read sequencing diagnostics. *Acta Neuropathol Commun*. 2021;9:98. <https://doi.org/10.1186/s40478-021-1201-x>

36. Hammock EAD, Young LJ. Microsatellite instability generates diversity in brain and sociobehavioral traits. *Science*. 2005;308:1630–1634. <https://doi.org/10.1126/science.1111427>
37. Martienssen RA. Maintenance of heterochromatin by RNA interference of tandem repeats. *Nat Genet*. 2003;35:213–214. <https://doi.org/10.1038/ng1252>
38. Novo CL et al. Satellite repeat transcripts modulate heterochromatin condensates and safeguard chromosome stability in mouse embryonic stem cells. *Nat Commun*. 2022;13:1–16. <https://doi.org/10.1038/s41467-022-31198-3>
39. Marsano RM, Dimitri P. Constitutive heterochromatin in eukaryotic genomes: a mine of transposable elements. *Cells*. 2022;11:761. <https://doi.org/10.3390/cells11050761>
40. Ugarković Đ, Sermek A, Ljubić S, Feliciello I. Satellite DNAs in health and disease. *Genes (Basel)*. 2022;13:1154. <https://doi.org/10.3390/genes13071154>
41. Negi P, Rai AN, Suprasanna P. Moving through the stressed genome: emerging regulatory roles for transposons in plant stress response. *Front Plant Sci*. 2016;7:1448. <https://doi.org/10.3389/fpls.2016.01448>
42. Ahmad SF, Singchat W, Panthum T, Srikulnath K. Impact of repetitive DNA elements on snake genome biology and evolution. *Cells*. 2021;10:1–27. <https://doi.org/10.3390/cells10071707>
43. Jagannathan M, Cummings R, Yamashita YM. The modular mechanism of chromocenter formation in *Drosophila*. *Elife*. 2019;8:1–16. <https://doi.org/10.7554/eLife.43938>
44. Zattera ML, Bruschi DP. Transposable elements as a source of novel repetitive DNA in the eukaryote genome. *Cells*. 2022;11:3373. <https://doi.org/10.3390/cells11213373>
45. Šatović E et al. Adjacent sequences disclose potential for intra-genomic dispersal of satellite DNA repeats and suggest a complex network with transposable elements. *BMC Genomics*. 2016;17:997. <https://doi.org/10.1186/s12864-016-3347-1>
46. Palomeque T, Antonio Carrillo J, Muñoz-López M, Lorite P. Detection of a mariner-like element and a miniature inverted-repeat transposable element (MITE) associated with the heterochromatin from ants of the genus *Messor* and their possible involvement for satellite DNA evolution. *Gene*. 2006;371:194–205. <https://doi.org/10.1016/j.gene.2005.11.032>
47. Šatović-Vukšić E, Plohl M. Classification problems of repetitive DNA sequences. *DNA (Basel)*. 2021;1:84–90. <https://doi.org/10.3390/dna1020009>
48. Rico-Porras JM et al. Heterochromatin is not the only place for satDNAs: the high diversity of satDNAs in the euchromatin of the beetle *Chrysolina americana* (Coleoptera, Chrysomelidae). *Genes (Basel)*. 2024;15:395. <https://doi.org/10.3390/genes15040395>
49. Brajković J, Feliciello I, Bruvo-Mačarić B, Ugarković D. Satellite DNA-like elements associated with genes within euchromatin of the beetle *Tribolium castaneum*. *G3 (Bethesda)*. 2012;2:931–941. <https://doi.org/10.1534/g3.112.003467>
50. Dias GB et al. Tetris is a foldback transposon that provided the building blocks for an emerging satellite DNA of *Drosophila virilis*. *Genome Biol Evol*. 2014;6:1302–1313. <https://doi.org/10.1093/gbe/evu108>
51. Dias GB, Heringer P, Svartman M, Kuhn GCS. Helitrons shaping the genomic architecture of *Drosophila*: enrichment of DINE-TR1 in α - and β -heterochromatin, satellite DNA emergence, and piRNA expression. *Chromosome Res*. 2015;23:597–613. <https://doi.org/10.1007/s10577-015-9480-x>
52. Luchetti A. terMITES: miniature inverted-repeat transposable elements (MITES) in the termite genome (Blattodea: Termitoidae). *Mol Genet Genomics*. 2015;290:1499–1509. <https://doi.org/10.1007/s00438-015-1010-1>
53. Feliciello I et al. Evolutionary history of alpha satellite DNA repeats dispersed within human genome euchromatin. *Genome Biol Evol*. 2020;12:2125–2138. <https://doi.org/10.1093/gbe/evaa224>
54. Tunjić-Cvitanović M et al. Satellitome analysis of the Pacific Oyster *Crassostrea gigas* reveals new pattern of satellite DNA organization, highly scattered across the genome. *Int J Mol Sci*. 2021;22:6798. <https://doi.org/10.3390/ijms22136798>
55. Tunjić-Cvitanović M, García-Souto D, Pasantes JJ, Šatović-Vukšić E. Dominance of transposable element-related satDNAs results in great complexity of “satDNA library” and invokes the extension towards “repetitive DNA library”. *Mar Life Sci Technol*. 2024;6:236–251. <https://doi.org/10.1007/s42995-024-00218-0>
56. Langdon T et al. Retrotransposon evolution in diverse plant genomes. *Genetics*. 2000;156:313–325. <https://doi.org/10.1093/genetics/156.1.313>
57. Tek AL, Song J, Macas J, Jiang J. Sobo, a recently amplified satellite repeat of potato, and its implications for the origin of tandemly repeated sequences. *Genetics*. 2005;170:1231–1238. <https://doi.org/10.1534/genetics.105.041087>
58. Biscotti MA et al. Molecular and cytogenetic characterization of repetitive DNA in the Antarctic polyplacophoran *Nuttallochiton mirandus*. *Chromosome Res*. 2008;16:907–916. <https://doi.org/10.1007/s10577-008-1248-0>
59. Macas J, Koblízková A, Navrátilová A, Neumann P. Hypervariable 3' UTR region of plant LTR-retrotransposons as a source of novel satellite repeats. *Gene*. 2009;448:198–206. <https://doi.org/10.1016/j.gene.2009.06.014>
60. Sharma A, Wolfgruber TK, Presting GG. Tandem repeats derived from centromeric retrotransposons. *BMC Genomics*. 2013;14:142. <https://doi.org/10.1186/1471-2164-14-142>
61. Meštrović N et al. Structural and functional liaisons between transposable elements and satellite DNAs. *Chromosome Res*. 2015;23:583–596. <https://doi.org/10.1007/s10577-015-9483-7>
62. Scalvenzi T, Pollet N. Insights on genome size evolution from a miniature inverted repeat transposon driving a satellite DNA. *Mol Phylogenet Evol*. 2014;81:1–9. <https://doi.org/10.1016/j.ympev.2014.08.014>
63. Vondrak T et al. Characterization of repeat arrays in ultra-long nanopore reads reveals frequent origin of satellite DNA from retrotransposon-derived tandem repeats. *Plant J*. 2020;101:484–500. <https://doi.org/10.1111/tjp.14546>
64. Alkan C et al. Organization and evolution of primate centromeric DNA from whole-genome shotgun sequence data. *PLoS Comput Biol*. 2007;3:1807–1818. <https://doi.org/10.1371/journal.pcbi.0030181>
65. Šatović E, Plohl M. Tandem repeat-containing MITE elements in the clam *Donax trunculus*. *Genome Biol Evol*. 2013;5:2549–2559. <https://doi.org/10.1093/gbe/evt202>
66. Paço A, Freitas R, Vieira-Da-Silva A. Conversion of DNA sequences: from a transposable element to a tandem repeat or to a gene. *Genes (Basel)*. 2019;10:1014. <https://doi.org/10.3390/genes10121014>
67. Kuhn GCS, Heringer P, Dias GB. Structure, organization, and evolution of satellite DNAs: insights from the *Drosophila* repleta and *D. virilis* species groups. In: Ugarković Đ, editor. *Satellite DNAs in physiology and evolution. Progress in molecular and subcellular biology*. Springer; 2021;60. p. 27–56.
68. McGurk MP, Barbash DA. Double insertion of transposable elements provides a substrate for the evolution of satellite DNA. *Genome Res*. 2018;28:714–725. <https://doi.org/10.1101/gr.231472.117>
69. Moran JV, Morish T. Chromosomes: noncoding DNA (including satellite DNA). In: *Encyclopedia of life sciences*. John Wiley & Sons; 2005. p. 1–9.
70. Izsvák Z et al. Short inverted-repeat transposable elements in teleost fish and implications for a mechanism of their amplification. *J Mol Evol*. 1999;48:13–21. <https://doi.org/10.1007/pl00006440>
71. Hikosaka A, Kawahara A. Lineage-specific tandem repeats riding on a transposable element of MITE in *Xenopus* evolution: a new mechanism for creating simple sequence repeats. *J Mol Evol*. 2004;59:738–746. <https://doi.org/10.1007/s00239-004-2664-1>
72. Xiong W, Dooner HK, Du C. Rolling-circle amplification of centromeric Helitrons in plant genomes. *Plant J*. 2016;88:1038–1045. <https://doi.org/10.1111/tjp.13314>

73. Thomas J, Pritham EJ. Helitrons, the eukaryotic rolling-circle transposable elements. *Microbiol Spectr*. 2015;3:1–32. <https://doi.org/10.1128/microbiolspec.MDNA3-0049-2014>
74. Miller WJ, Nagel A, Bachmann J, Bachmann L. Evolutionary dynamics of the SGM transposon family in the *Drosophila obscura* species group. *Mol Biol Evol*. 2000;17:1597–1609. <https://doi.org/10.1093/oxfordjournals.molbev.a026259>
75. Gaffney PM et al. *Pearl*, a novel family of putative transposable elements in bivalve mollusks. *J Mol Evol*. 2003;56:308–316. <https://doi.org/10.1007/s00239-002-2402-5>
76. Yang H-P, Barbash DA. Abundant and species-specific DINE-1 transposable elements in 12 *Drosophila* genomes. *Genome Biol*. 2008;9:R39. <https://doi.org/10.1186/gb-2008-9-2-r39>
77. Kuhn GCS, Heslop-Harrison JS. Characterization and genomic organization of PERI, a repetitive DNA in the *Drosophila buzzatii* cluster related to DINE-1 transposable elements and highly abundant in the sex chromosomes. *Cytogenet Genome Res*. 2011;132:79–88. <https://doi.org/10.1159/000320921>
78. Thomas J, Vadrnagara K, Pritham EJ. DINE-1, the highest copy number repeats in *Drosophila melanogaster* are non-autonomous endonuclease-encoding rolling-circle transposable elements (Helitrons). *Mob DNA*. 2014;5:18. <https://doi.org/10.1186/1759-8753-5-18>
79. Wong LH, Choo KHA. Evolutionary dynamics of transposable elements at the centromere. *Trends Genet*. 2004;20:611–616. <https://doi.org/10.1016/j.tig.2004.09.011>
80. Gao C et al. Characterization and functional annotation of nested transposable elements in eukaryotic genomes. *Genomics*. 2012;100:222–230. <https://doi.org/10.1016/j.ygeno.2012.07.004>
81. Witte CP, Le QH, Bureau T, Kumar A. Terminal-repeat retrotransposons in miniature (TRIM) are involved in restructuring plant genomes. *Proc Natl Acad Sci U S A*. 2001;98:13778–13783. <https://doi.org/10.1073/pnas.241341898>
82. Kidwell MG. Transposable elements and the evolution of genome size in eukaryotes. *Genetica*. 2002;115:49–63. <https://doi.org/10.1023/a:1016072014259>
83. Vaury C, Bucheton A, Pelisson A. The β heterochromatic sequences flanking the I elements are themselves defective transposable elements. *Chromosoma*. 1989;98:215–224. <https://doi.org/10.1007/BF00329686>
84. Wensink PC, Tabata S, Pacht C. The clustered and scrambled arrangement of moderately repetitive elements in *Drosophila* DNA. *Cell*. 1979;18:1231–1246. [https://doi.org/10.1016/0092-8674\(79\)90235-6](https://doi.org/10.1016/0092-8674(79)90235-6)
85. Malone CD, Hannon GJ. Molecular evolution of piRNA and transposon control pathways in *Drosophila*. *Cold Spring Harb Symp Quant Biol*. 2009;74:225–234. <https://doi.org/10.1101/sqb.2009.74.052>
86. Plohl M, Luchetti A, Mestrovic N, Mantovani B. Satellite DNAs between selfishness and functionality: structure, genomics and evolution of tandem repeats in centromeric (hetero)chromatin. *Gene*. 2008;409:72–82. <https://doi.org/10.1016/j.gene.2007.11.013>
87. Henikoff S et al. Gene families: the taxonomy of protein paralogs and chimeras. *Science*. 1997;278:609–614. <https://doi.org/10.1126/science.278.5338.609>
88. Fernández-Medina RD et al. Losing identity: structural diversity of transposable elements belonging to different classes in the genome of *Anopheles gambiae*. *BMC Genomics*. 2012;13:272. <https://doi.org/10.1186/1471-2164-13-272>