Trends in Genetics

CellPress OPEN ACCESS

Review

Helitrons: genomic parasites that generate developmental novelties

Daniela Barro-Trastoy [1] and Claudia Köhler [1,2,*]

Helitrons, classified as DNA transposons, employ rolling-circle intermediates for transposition. Distinguishing themselves from other DNA transposons, they leave the original template element unaltered during transposition, which has led to their characterization as 'peel-and-paste elements'. Helitrons possess the ability to capture and mobilize host genome fragments, with enormous consequences for host genomes. This review discusses the current understanding of Helitrons, exploring their origins, transposition mechanism, and the extensive repercussions of their activity on genome structure and function. We also explore the evolutionary conflicts stemming from Helitron-transposed gene fragments and elucidate their domestication for regulating responses to environmental challenges. Looking ahead, further research in this evolving field promises to bring interesting discoveries on the role of Helitrons in shaping genomic landscapes.

Highlights

Helitrons are notoriously difficult to identify, due to low conservation of their internal sequences and few characteristic sequence features. New sequencing technologies and genome analysis tools open advanced possibilities for the identification of Helitrons, allowing one to monitor Helitron activity and its consequences.

Helitrons can transpose gene fragments, including regulatory elements of genes, thereby creating new circuits of gene regulation. In plants, this has been demonstrated for endosperm-specific regulatory elements and elements regulating the response to herbivores, but more examples are likely to be discovered.

Helitron-captured gene fragments can elicit evolutionary conflicts. This is because the host silencing machinery will target the Helitron and the captured gene, causing transacting silencing effects on the donor gene.

Transposable elements, catalysts of genome evolution

Genomes are the complete collection of genetic material in an organism, containing the essential instructions for its development and functionality. Each living organism, from the simplest single-celled microorganisms to the most complex multicellular beings, stores its genetic blueprint within its genome, which is composed of DNA sequences. One can think of genomes as giant puzzles meticulously constructed for millions of years, reflecting the evolutionary adaptations of living organisms to diverse environments, all in the pursuit of survival.

The concept of evolution is intricately linked to genomes. As time progresses, populations of organisms undergo modifications in their genomes, giving rise to variations and the acquisition of new traits. Numerous mechanisms facilitate such modifications, including random DNA mutations, recombination, duplication of genes, of chromosomes or of the whole genome, sequence rearrangements, etc. Among them, transposition is arguably one of the most influential drivers of genomic changes and genetic novelty [1,2].

Transposable elements (**TEs**; see Glossary) are segments of DNA with the ability to move from one location within the genome to another [3]. They are present in the genomes of virtually all living organisms, from bacteria to plants, fungi, and animals. Traditionally, eukaryotic TEs are divided into two main classes based on their transposition mechanisms [3]. Class I TEs, also known as retrotransposons, mobilize via an RNA intermediate, which is then reverse-transcribed into a DNA copy and integrated into the genome. In this case, the original template element remains unaltered; hence, class I TEs are often termed 'copy-and-paste elements'. By contrast, class II TEs mobilize via a DNA intermediate, which in the majority of the cases is excised from the original location. Therefore, most class II TEs are commonly referred to as 'DNA cut-and-paste elements'. Both TE classes have canonical structural hallmarks. For instance, some class I TEs have a pair of identical DNA sequences of several hundred base pairs' length located at each end, called 'long

[1]Max Planck Institute of Molecular Plant Physiology, Am Mühlenberg 1, 14476 Potsdam-Golm, Germany
[2]Department of Plant Biology, Swedish University of Agricultural Sciences and Linnean Center for Plant Biology, Uppsala 75007, Sweden

*Correspondence:
koehler@mpimp-golm.mpg.de
(C. Köhler).

terminal repeats' (LTRs). Non-LTR retrotransposons are either long interspersed nuclear elements (LINEs) or short interspersed nuclear elements (SINEs). Class II TEs are often flanked by terminal inverted repeats, and their insertion in the new location of the genome results in the duplication of host insertion sites, called 'target site duplications' [3].

One particular family of DNA TEs is **Helitrons** [4–6]. Although Helitrons occupy a substantial portion of many eukaryotic genomes [7], they remained undiscovered until the early 2000s, more than 50 years after the discovery of the first TEs [8]. This was due to the lack of structural features that characterize TEs known at that time, as well as the high diversity of their internal sequences [9,10]. Helitrons are classified as class II DNA TEs, because no RNA intermediates are involved in their transposition [8]. However, unlike classical DNA TEs, Helitrons lack terminal inverted repeats and do not produce target site duplications upon insertion [8].

Helitrons: discovery, structural hallmarks, and mechanism of transposition

Helitrons are a class of TEs initially identified exclusively through computational analyses of repetitive sequences in the *Arabidopsis thaliana*, *Oryza sativa*, and *Caenorhabditis elegans* genomes [8]. This seminal research already pinpointed the possible evolutionary importance of Helitrons, which quickly became a subject of further investigations that included genetic and molecular approaches. Despite the fact that the abundance of Helitrons varies substantially among different species, it has become evident that Helitrons are ubiquitous in nearly all eukaryotic genomes [7,11,12].

The first characterization of Helitrons highlighted specific structural hallmarks that distinguish them from conventional DNA transposons [8]. They possess unique transposon ends lacking tandem inverted repeats and do not lead to the duplication of the host target sites during the transposition process. Their characteristic features include 5′-TC and 3′-CTRR termini, often favoring 3′-CTAG sequences (Figure 1). Additionally, Helitrons commonly have a short palindromic sequence of 16 to 20 nucleotides that form a hairpin situated proximal to the 3′ end. Unlike typical DNA transposons, Helitrons usually maintain the original template element unaltered during transposition, earning the designation of peel-and-paste or rolling-circle elements (Box 1), although a few examples of Helitron excision were observed in maize [13,14]. Furthermore, Helitrons demonstrate a strong inclination for integrating within AT dinucleotide sites, thereby establishing a defining feature of their unique genomic integration preference [8].

The majority of Helitrons discovered thus far are nonautonomous, meaning that they lack the capability to encode the essential proteins required for self-transposition (Figure 1). Instead, they depend on trans-mobilization by the enzymatic machinery of their autonomous counterparts. However, the *in silico* reconstruction of potential autonomous Helitrons has unveiled the

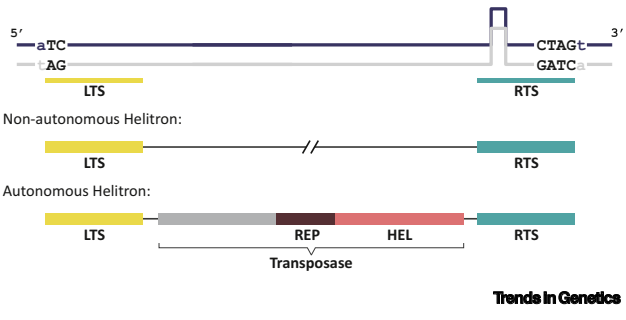

Figure 1. Model of a Helitron. Helitrons are characterized by 5′-TC and 3′-CTAG terminal motifs that are part of approximately 30-base pair long left and right terminal sequences, respectively. Nonautonomous Helitrons lack the capability to encode the essential proteins required for self-transposition, whereas autonomous Helitrons encode the RepHel transposase. Broken lines indicate that nonautonomous Helitrons can vary in length. Abbreviations: HEL, Hel domain of the RepHel transposase; LTS, left terminal sequence; REP, Rep domain of the RepHel transposase; RTS, right terminal sequence.

Glossary

Gene capture: process whereby a transposable element incorporates a host gene during its movement within the genome.
Gene collinearity: arrangement of a set of genes that are located on the same chromosome or genomic region in a conserved linear order in different species.
Helitron: a type of DNA transposon characterized by its distinctive rolling-circle replication mechanism.
Horizontal transfer: transfer of genetic material between organisms that are not directly in a lineage of descent.
Transposable element (TE): DNA sequence that has the ability to move from one location in the genome to another.

Box 1. Mechanism of Helitron transposition

The remarkable resemblance between Helitrons and prokaryotic rolling-circle transposons led to the hypothesis that Helitrons might employ a rolling-circle replication mechanism for their transposition [8]. This notion garnered acceptance, but it was not until 15 years later that experimental evidence supported the rolling-circle transposition mechanism [33,62,78]. On the basis of functional [33,62] and structural [78] analyses, it was proposed that RepHel, also referred to as 'Helitron transposase', recognizes the 5′ terminus of the Helitron element, which is essential to initiate transposition activity (Figure IA) [33,78]. This recognition leads to the introduction of a nick in the positive strand of the DNA [33], which is followed by a covalent attachment of the transposase with the newly liberated strand (Figure IB) [78], establishing a replication fork. Both cleavage and linkage are orchestrated by the Rep domain, specifically by the two tyrosines of the HUH Y2 active site, similar to many replication initiators that use HUH nucleases [33,78]. While the host's replication machinery initiates DNA replication at the site where the nick occurred (leading to the reconstitution of the Helitron donor), the strand covalently attached to the transposase is peeled off (Figure IC) [33]. This separation event is believed to be facilitated by the Hel domain, which likely plays a role in unwinding the double-stranded DNA at this juncture [33]. As the process continues, the transposase recognizes the hairpin structure near the 3′ terminus of the Helitron (Figure IC). This recognition presumably causes the pause of the transposition process (i.e., it acts as transposition terminator) and allows a second nicking at the 3′-GTAG sequence (Figure ID) [33,62]. In turn, it enables the attachment of the 3′ terminus to the 5′ terminus of the Helitron, forming a circular single-stranded intermediate (Figure IE) [33]. This circular single-stranded intermediate is subsequently or concurrently converted into a double-stranded circle, which acts as a platform for successive transposition events by inserting into the host genome (Figure IF) [62]. Notably, the insertion of the newly mobilized Helitron into the host genome tends to occur mostly preferentially between AT dinucleotides, usually located in AT-rich DNA sequences (Figure IG) [9,10,33]. Given the attributes of this transposition mechanism, Helitrons were also referred to as 'rolling-circle' or 'peel-and-paste' elements.

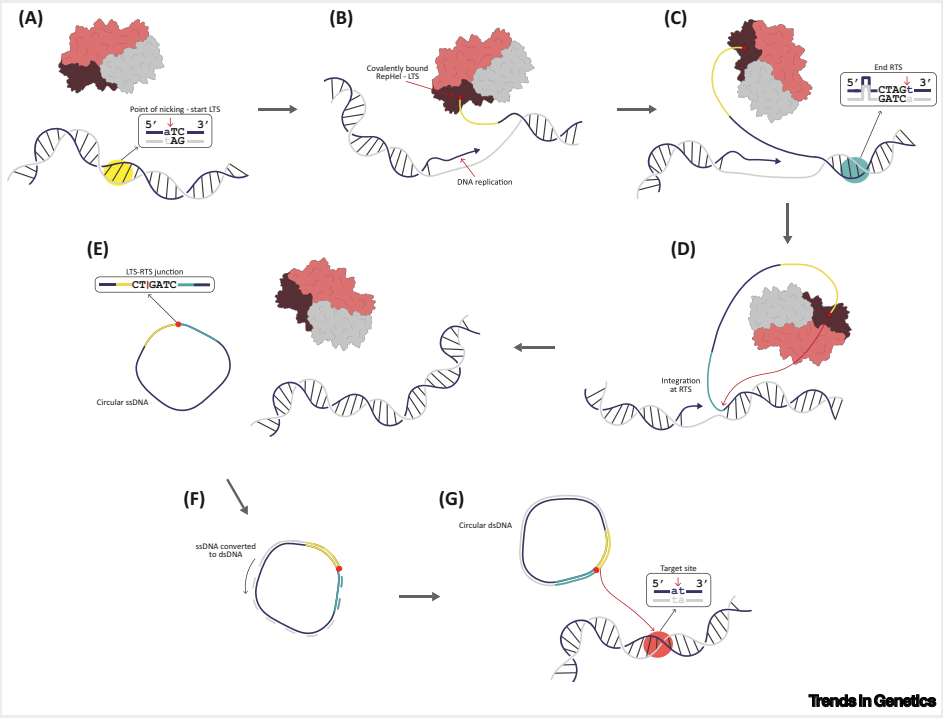

Figure I. Model of Helitron transposition. (A–G) The RepHel protein exhibits distinct colors corresponding to its N-terminal (gray), Rep (dark pink), and Hel (light pink) domains. The different colors in the DNA molecule represent the positive (blue) and negative (light gray) strands, the left (yellow) and right (light blue) terminal Helitron sequences, and the insertion point (red). Abbreviations: dsDNA, double-stranded DNA; LTS, left terminal sequence; RTS, right terminal sequence; ssDNA, single-stranded DNA.

existence of a large protein termed 'RepHel' [8], which emerges as a common characteristic among presumed autonomous Helitrons from both animal and plant genomes [7,8,15–19]. The RepHel protein includes a replication initiation domain (Rep), which shares similarity with

HUH endonucleases, coupled with a DNA helicase domain (Hel) predicted to belong to the Pif1 helicase family (SH1 helicase superfamily) (Figure 1) [8]. Some plant Helitrons additionally encode a single-stranded DNA-binding replication protein A (RPA), which may be involved in DNA replication and repair [8,20]. Although apparently less common in animals, homologs of RPA have also been identified in Helitrons from zebrafish and sea anemone [4] and in a subtype of Helitrons termed 'helentrons' in *Drosophila melanogaster* [21]. Some putative autonomous Helitrons presumably encoding intact RepHel proteins have been found in maize [20], but, to date, no *in vivo* active Helitron has been isolated. After years of discussing how Helitrons might have evolved, recent research has contributed insights into their possible origin (Box 2).

Impacts of Helitron transposition

Following their discovery and first description, Helitrons were rapidly recognized as insertional mutagens causing spontaneous mutations [22–25]. Besides the evident possible evolutionary impacts of Helitrons as insertional mutagens, it was noted that insertions that lacked coding capacity for Helitron transposases carried multiple gene fragments from different chromosomal locations. Quickly, an increasing amount of genomic data revealed that Helitrons are often prone to capture and mobilize host genomic fragments. This phenomenon leads to gene fragment duplications, genome rearrangements, exon shuffling, generation of chimeric transcripts, and dissemination of genomic regulatory elements [9–11,18,23,26–32]. These observations, coupled with the widespread occurrence of Helitrons in a wide range of organisms and their often-substantial genomic prevalence [7], highlight the relevance of Helitrons in shaping the genetic landscape, especially in plants.

Gene capture

In the context of TEs, **gene capture** refers to a process whereby TEs can acquire and mobilize genes or gene fragments from one location in the genome to another. Several hypothetical mechanisms have been proposed to explain the process of gene capture by Helitrons (Figure 2) [4,5]. Here we focus on two of these mechanisms that are substantially supported by data gained over recent years.

Gene fragment acquisition by Helitrons is considered to occur at the DNA level, because contiguous exons and introns were found within the Helitron-captured DNA [9,10,20]. One of the initially

Box 2. Evolutionary origin of Helitrons

Several hypotheses have been proposed concerning the evolutionary origin of Helitrons. Initially, due to their similarities in their mode of propagation, it was suggested that Helitrons have evolved from prokaryotic ancestral rolling-circle replication TEs, such as those belonging to the IS91 group [8]. However, despite these prokaryotic elements encode Rep proteins that share conserved motifs closely resembling the Rep domain found in Helitrons, they lack a Hel domain. A second hypothesis proposes that Helitrons may have originated as genomic parasites through the integration of ancient eukaryotic viruses such as geminiviruses, a group of single-stranded DNA viruses that infect many plant species. This hypothesis is grounded in the observation that geminiviruses employ rolling-circle replication proteins that contain both a Rep and a Hel domain and display single-stranded DNA-binding activities, similar to the features seen in some Helitrons [79]. However, the Hel domain present in the proteins encoded by geminiviruses belongs to a superfamily different from those present in the RepHel of Helitrons [80]. Because of similarities of the Hel domain with eukaryotic Pif1 helicases, it was also proposed that the Hel domain has been acquired through the capture a Pif1 gene from an ancestral eukaryotic host. However, recent phylogenetic studies indicate that the RepHel protein already had its archetypical structure with two domains before invading eukaryotic hosts [81]. This rather suggests that Helitrons are descendants of a prokaryotic plasmid element that invaded eukaryotes at early stages of eukaryotic evolution and shifted into a transposon [16,81,82]. Interestingly, while the Pif1 family of DNA helicases is conserved from bacteria to humans and has essential functions in maintaining genome integrity, it was now found that Pif1-like genes were independently lost in the Brassicales and Commelinids (a monophyletic group containing the Poales). But those plant groups contain multiple Helitron-derived Pif1-like Hel domain sequences [81], suggesting that some plants may have replaced original genomic Pif1 genes by domesticated Pif1 sequences from Helitrons.

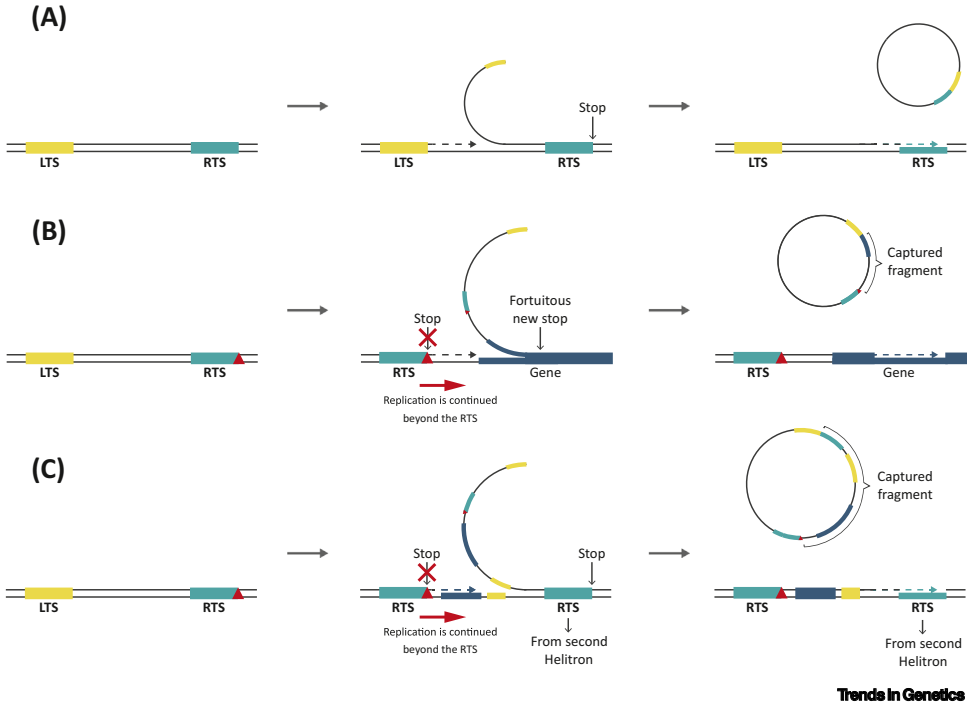

Figure 2. Mechanisms of Helitron gene capture. (A) Normal Helitron replication. (B) Malfunction or deletion of the 3′ termination signal permits the continuation of reading beyond the original 3′ terminus, extending through the neighboring sequences until a fortuitous motif downstream serves as a new termination signal. (C) Malfunction or deletion of the 3′ termination signal permits the continuation of reading beyond the original 3′ terminus, extending through the neighboring sequences until another 3′ terminus from a nearby Helitron acts as termination signal. Abbreviations: LTS, left terminal sequence; RTS, right terminal sequence.

proposed models suggests an accidental malfunction or deletion of the 3′ termination signal. Thus, as transposition occurs, the transposition machinery bypasses the original 3′ terminus and continues to read through the neighboring sequences until a fortuitous motif located downstream acts as a new termination signal (Figure 2B) [4,5]. A variation of this model proposes the creation of a chimeric Helitron when defects accumulate in two closely situated and correctly oriented Helitrons [4,5]. For example, a new Helitron can be formed if the first Helitron lacks a proper 3′ termination signal, so transposition is terminated at the 3′ end of the second Helitron, thus capturing all the intervening sequence (Figure 2C). There is supportive evidence for both models [9,33,34]. In a remarkable paper from 2016, an autonomous Helitron was reconstructed on the basis of sequences of inactive Helitrons from bat [33]. This Helitron was demonstrated to be active *in vitro* and *ex vivo*, providing the first experimental insights into Helitron transposition [33]. This breakthrough allowed the development of a Helitron transposition system that permitted functional study of structural characteristics of Helitrons [33]. The elimination of the 5′ terminus completely abolished Helitron transposition, underscoring its indispensability [33]. However, a reduced yet still noticeable transposition activity was observed when the 3′ terminus, particularly the hairpin structure near the 3′ terminus, was either eliminated or modified [33]. Further examination unveiled that reconstructed bat Helitrons devoid of the complete 3′ terminus, or solely the hairpin, made use of serendipitous downstream alternative termination signals [33]. Hairpin-lacking Helitrons or even an intact Helitron, albeit less frequently, could successfully capture an entire selection cassette placed downstream of the Helitron [33]. These findings demonstrate the ability of Helitrons to capture genome fragments when the termination signal malfunctions. Concerning the second model proposing the formation of chimeric Helitrons, it was noted that the *Arabidopsis*

genome exhibits a considerable presence of Helitrons featuring diverse combinations of termini supposedly corresponding to separate Helitron families [34]. This observation raises the possibility that these Helitrons, housing variable termini from different families, might be chimeric Helitrons resulting from truncations or defects in neighboring Helitrons.

Several Helitrons across different species, including plants and animals, have been documented to carry simultaneously a variety of fragments sourced from multiple chromosomal locations [9,10,20,23,26,27,31,35–38]. This suggests that Helitrons capture genes in a stepwise manner [19,20]. Thus, a first gene capture event occurs during a single transposition event. Then, the Helitron, together with the captured fragments, becomes mobilized and integrates into another site within the genome. With successive rounds of transposition, there is the possibility of capturing a second gene fragment. This iterative process can recur, potentially giving rise to Helitrons spanning a diverse collection of host sequences, including promoters, binding sites, exons, introns, or other genetic components from distinct genes scattered throughout the genome [20,31]. This phenomenon occasionally leads to the random recombination of exons (exon shuffling), which can result in the emergence of chimeric transcripts [31,39].

Although Helitron prevalence and gene capture vary across species, they are notable in certain genomes. The capture of host genomic fragments seems to have been particularly prevalent in maize [9,10,38,40]. Recent analyses estimate that Helitrons constitute up to 4% of the maize genome [41]. A high percentage of the unambiguously identified maize Helitrons have captured gene fragments, usually ranging from one to three fragments [9,10], but certain Helitrons have even been observed to harbor fragments from up to 12 genes [20]. In total, it is estimated that maize Helitrons have captured and amplified tens of thousands of gene fragments [9,10], and some examples of nearly full-length captured genes were reported [29,30]. Many maize Helitrons exhibit multiple and variable 3′ termini deviating from the conserved 3′-CTAG motif, possibly because the gene capture process disrupts the recognition of the transposition termination signal [20]. Multiple termini are likely a consequence of several sequential transpositions [20]. In animal genomes, especially in mammals, prevalence and activity of Helitrons appear to be scarcer than in plant genomes, with bats being a notable exception [18]. In the bat *Myotis lucifugus*, Helitrons constitute around 6% of the genome, with several of them containing one to two gene fragments [31].

Impact on transcriptomes

One of the most evident effects of Helitron mobilization is the potential for mutations resulting from their insertion within genes. Indeed, numerous instances of Helitrons acting as insertional mutagens have been documented [22,23,25,31,42–46]. Nonetheless, and probably more important, Helitron insertions can potentially contribute to enhancing transcriptome diversity. Usually, Helitrons carrying concatenated exons remain silent. Yet, in certain cases, chimeric transcripts featuring shuffled exons are detected, suggesting that these Helitrons were likely expressed, possibly because they incorporated regulatory sequences or were inserted in close proximity to such sequences (Figure 3A) [22,26,27,30,31,39,47,48]. In bats, it was observed that Helitrons tend to bear sequences corresponding to promoters, 5′ untranslated regions (UTRs), the first coding exon and first intron, and Helitrons carrying these types of sequences accumulate higher copy numbers than Helitrons carrying other genic regions [31]. It was speculated that the presence of a promoter in a Helitron can induce the transcription of adjacent regions and eventually become a functional gene [31]. For example, in *Aspergillus nidulans*, a Helitron captured and duplicated the promoter and a portion of the coding sequence of the *xanA* gene [49], and this duplicated promoter was regulated in the same manner as the original *xanA* gene [49], indicating its potential to promote the expression of genes under appropriate conditions. Thus, Helitrons foster genome diversity by changing genome structure. This potential, when subjected to

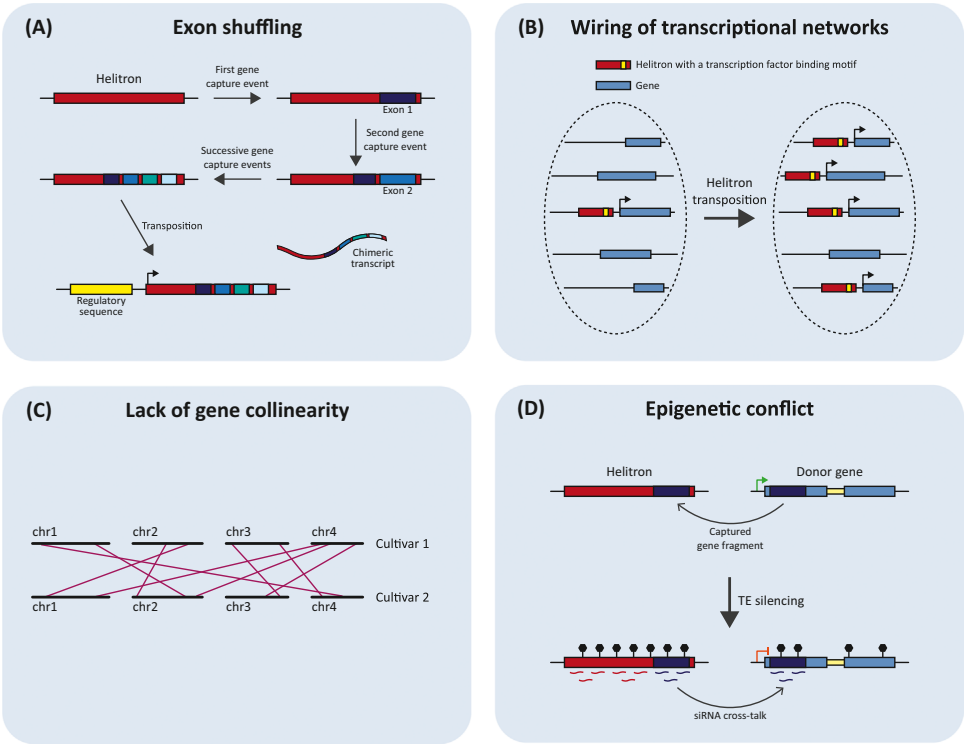

Figure 3. Evolutionary consequences of Helitron transposition. (A) Exon shuffling occurs through multiple rounds of transposition, wherein Helitrons have the potential to capture various independent gene fragments. This process may give rise to Helitrons that span a wide range of host sequences. Occasionally, this phenomenon leads to the random recombination of exons, resulting in the emergence of chimeric transcripts. (B) Wiring of transcriptional networks occurs through the dispersal and amplification of transcription factor binding motifs by Helitrons, eventually affecting the expression of nearby genes, and ultimately allowing the incorporation of new genes into regulatory networks. (C) Lack of gene collinearity between related species arises from Helitron transposition, which introduces structural variations and rearrangements in the genome, leading to the nonlinear arrangement of genes along the chromosome. (D) Epigenetic conflicts arise when Helitrons capture gene fragments. The siRNAs produced by these Helitron-captured genes can target the donor gene in *trans*, resulting in increased methylation and reduced gene expression. Abbreviation: TE, transposable element.

selective pressures, could contribute to the emergence of new genes and proteins with novel domains and functions. Interestingly, the generation of multiple transcript isoforms via alternative splicing is quite widespread in Helitron captured genes [39,47], showing that Helitrons not only interlace coding regions of different genes but also increase the diversity of transcripts by high levels of alternative splicing. Collectively, these observations pinpoint Helitrons as a potential driving force behind the evolution of genes.

Another consequence stemming from Helitron-mediated gene capture and mobility involves the dispersal and amplification of transcription factor binding motifs, thus affecting the expression of proximal genes (Figure 3B) [50–52]. As a result, new genes can be incorporated into regulatory networks [53]. In *Arabidopsis*, binding motifs for the endosperm-specific type I MADS-box transcription factor PHERES1 are notably enriched in Helitrons. It is suggested that the Helitron-mediated amplification of those binding motifs allowed the recruitment of crucial endosperm development genes into a common transcriptional network, thus contributing to endosperm evolution [50]. In *C. elegans,* a substantial number of the annotated Helitrons contain one or more heat shock elements (HSEs), with nearly two-thirds of all *C. elegans* HSEs being

associated with Helitrons. These Helitrons carrying HSEs can recruit the transcription factor heat shock factor 1 (HSF-1) and promote the expression of proximal genes in response to heat shock, thereby expanding the landscape of HSF-1 targets [51,54]. In the plant Brassicaceae family, a similar strategy was employed by members of the Copia family that carry HSEs and confer heat responsiveness to neighboring genes [55]. In the *Drosophila miranda* lineage, Helitrons played a role in the dispersion of male specific lethal (MSL) binding sites across the X chromosome, thus recruiting genes into the MSL regulatory network that regulates dosage compensation in male individuals [32]. These examples illustrate how Helitrons can contribute to evolutionary innovations through the rewiring and diversification of transcriptional networks by supplying *cis*-regulatory elements without directly affecting gene structure.

Helitrons shape genome structure

Because of their inherent ability of changing their position, along with their efficiency in capturing gene fragments and potentially achieving high copy numbers, Helitrons have a noteworthy impact on shaping genome architecture. For example, the considerable lack of **gene collinearity** observed within modern maize inbred lines is predominantly attributed to the mobilization of Helitrons transposing gene fragments (Figure 3C) [27,28,35]. It was estimated that approximately 20% of genes or gene fragments are not shared between maize lines B73 and Mo17, and a substantial portion of this divergence was found to be due to the insertion of Helitrons [27]. Similar observations were reported for the maize lines B73 and McC [35]. Also, in wheat, more than 7000 tandem duplicated Helitron regions translocated since the recent divergence of the cultivars Aikang 58 and Chinese Spring, with more than 400 regions being specific for Aikang 58 [56]. Thus, through the reorganization of gene fragments, Helitrons have played an important role in enhancing genome diversity in maize and wheat and likely contributed to the diversification of many other species. Intriguingly, both number and cumulative length of Helitrons positively correlate with genome size in plants [7,12], suggesting that plant genomes have expanded over time due to the proliferation of Helitrons.

Helitron distribution patterns have been determined so far only for a few plant species that, however, reveal some interesting commonalities and differences. In the Brassicaceae family, Helitrons are more frequently inserted in gene-poor regions, especially around centromeres [7,10,12,57]. Albeit less frequently, when Helitrons are inserted in genic regions, they are predominantly found in UTRs or introns rather than in coding regions [12]. In wheat, the centromeres of all chromosomes contain tandem duplicated Helitrons, which have probably contributed to centromere formation and centromere plasticity [56] and possibly have contributed to centromere repositioning in wheat [58]. In maize, Helitrons tend to locate in gene-rich regions rather than gene-poor regions [59]. Nevertheless, tandem Helitron repeats, consisting of truncated and disrupted Helitrons, have been identified in many plant centromeres, including those of maize [60], and in some animals, such as oysters [61], indicating a historic invasion of centromeres by Helitrons. In rice, Helitron distribution appears to lack a discernible pattern [57], but large numbers of tandem repetitive Helitrons are also found in centromeres [60]. On the basis of available studies, it thus seems that tandemly repeated Helitrons are a universal feature of plant centromeres, although the number of Helitrons per tandem array varies between species [60]. Interestingly, studies of *de novo* Helitron transposition events generated in cultured cells showed that Helitrons frequently target the promoter regions and gene bodies of highly expressed genes [62]. However, because most insertions of TEs are deleterious to the host genome, Helitron insertions into genic regions will be under strong negative selection pressure. [12,47,48,63]. Thus, the observed patterns of Helitron distributions reflect the outcome of natural selection rather than that of Helitron insertion preferences.

In animals, Helitrons appear to have undergone frequent **horizontal transfers** between distinct organisms [15,19,63–65]. This is supported by the patchy taxonomic distribution and high

sequence similarity observed in Helitron elements across distantly related lineages [19,65]. Horizontal transfer of Helitrons has been proposed for various taxonomic groups, including insects, reptiles, fishes, and bats [19,63–66]. This can explain the disparate distribution of Helitrons across mammals, with only bats showing a high prevalence of Helitrons [63]. Bats seem to be a hotspot for horizontal transfer of DNA transposons in mammals [63], and the horizontal transfer of Helitrons likely played an important role driving bat genome evolution. The mechanisms underlying the frequent occurrence of horizontal transfer of Helitrons among animals remain mysterious, but it has been suggested that these events may be linked to host–parasite interactions [65]. Not many cases of horizontal transfer of Helitrons have been reported in plants, although high differences of Helitron content are found among genera belonging to the same families or among closely related species, indicating that Helitron dynamics are independent from the phylogeny [7]. Interestingly, phylogenetic analyses of RPA proteins showed that RPA homologs from plant Helitrons constitute a distinct clade diverging prior to the separation of plants, animals, and fungi [67], suggesting an ancient origin of Helitron RPA proteins. However, within plants, RPA proteins are present only in angiosperms. The most plausible hypothesis for this observation is that the RPA encoding gene was introduced into an ancestral angiosperm by horizontal transfer [67]. RPA proteins bind to single-stranded DNA, so it is tempting to speculate that their incorporation into Helitrons may have facilitated Helitron transposition. This, in turn, could have potentially contributed to the enormous radiation of Helitrons in angiosperms.

Several studies analyzing Helitrons in different organisms discovered that most Helitrons arose relatively recently, with the majority of them being less than 6 million years old [15,18,26,29,30,38,40]. Recent comparative analyses of Helitrons among more than 300 plant genomes revealed that bryophytes have the lowest number and average density of Helitrons, while angiosperms are on the opposite end, with the highest density of Helitrons in their genomes [7]. Helitron density was highest in species of the Poaceae and Brassicaceae families, but numbers differ widely among families and even species within the same family [7], emphasizing a possible role of horizontal transfer. In plants, Helitrons likely experienced two rapid expansion phases. The first expansion occurred from 30 to 20 million years ago, coinciding with the rapid diversification of angiosperms. The second expansion occurred in the past 4 million years, partly overlapping with Quaternary glacial stage [7]. Whether Helitron expansion was promoted by climate change and facilitated adaptation by creating genetic diversity remains a fascinating subject of future investigations.

Helitrons induce epigenetic conflict

TEs are well-accepted driving forces of evolution; nevertheless, one of the most common immediate consequence of their proliferation is the detrimental impact they can exert on host fitness. In consequence, TEs are usually inactivated by epigenetic silencing mechanisms facilitated by small interfering RNAs (siRNAs). These siRNAs are generated by the highly conserved RNAi pathway found in plants, fungi, and animals, which performs silencing at the RNA level. Plants additionally employ the RNA-directed DNA methylation (RdDM) pathway, which results in silencing at the DNA level [68]. In instances where TEs capture genes, as happens with Helitrons, this can lead to epigenetic consequences for the endogenous genes. The phenomenon arises because siRNAs, which guide RdDM at TEs harboring integrated gene fragments, have the potential to selectively target the genes from which these fragments originated. This can give rise to evolutionary conflicts between TEs and genes (Figure 3D). If a TE with captured gene fragments is subjected to silencing, a siRNA crosstalk between the TE and the donor gene could drive epigenetic alterations to the gene itself, eventually affecting its function. If these affected genes play important roles, natural selection is likely to moderate the host's silencing response, thereby preventing potential harm to the gene function but inadvertently favoring the TE with the captured gene fragment. In agreement with this hypothesis, it was found that in the maize genome, when TEs,

especially Helitrons, capture fragments of genes, siRNAs derived from those TEs can act in *trans* to mediate an epigenetic response against the donor genes, increasing its methylation status [69]. Interestingly, this effect is moderated on functionally important genes and does not lead to alterations in gene expression. By contrast, it tends to be more aggressive on functionally less vital genes, resulting in reduced expression levels and potentially leading to pseudogenization [69]. Moreover, TEs harboring fragments of important genes generate apparently fewer siRNAs than TEs without captured fragments [69]. Therefore, gene capture by Helitrons bears the potential for intragenomic conflict, which in turn can influence gene evolution [69].

This intragenomic conflict plays off in the context of plant stress responses. In *Arabidopsis*, the nonautonomous Helitron *ATREP2* is enriched for sequences from genes regulating the induced-resistance response to herbivores [70]. Under normal conditions, *ATREP2* is highly methylated. However, defense responses induce the demethylation of *ATREP2* [70], causing *ATREP2* transcription and siRNA generation [70,71]. These siRNAs guide the small RNA-binding effector protein ARGONAUTE1 to *ATREP2* defense genes, inducing chromatin changes and enhanced transcription factor binding, leading to an increase of gene expression [70]. Thus, hypomethylated *ATREP2* elements establish long-lasting memory of induced resistance by enabling an enhanced transcriptional response after a previous herbivore attack.

Concluding remarks and future perspectives

Helitrons have become a captivating subject of interest due to their parasitic nature and immense potential as evolutionary toolkits (see Outstanding questions). Understanding the impact of Helitrons is important for unraveling the complexities of genome dynamics and evolution. Helitrons have been particularly resistant to automated computational identification because they lack terminal repeats, because of their propensity for gene capture leading to extreme heterogeneity in size and sequence, and because of the predominant nonautonomous nature of the vast majority of Helitrons. Their unique structural features made it challenging to definitively identify or even classify them. As a result, the classification of Helitrons has remained ambiguous and has been a subject of recurring discussion [8,10,19,34,36,57]. It is plausible that bioinformatic detection of Helitrons has led to both false-positive and false-negative findings. Nowadays, however, with the advance of machine learning algorithms and design of novel tools, the precision of Helitron identification may strongly improve and is expected to give intriguing insights into Helitron dynamics in the near future [72–75]. Furthermore, the advent of long-read sequencing now enables the accurate characterization of complex repetitive genome regions, likely providing exciting new insights into the role of Helitrons in building heterochromatic structures [76].

One important open question in the field of Helitron research is whether there are active Helitrons currently and by which trigger they may become activated. While the observation of recent insertional mutations caused by nonautonomous Helitrons in maize suggests that these elements have moved in recent times [22,23], direct evidence of Helitron activity *in vivo* is still lacking. Extrachromosomal circular DNA (eccDNA) intermediates are used as a diagnostic indicator for the activity of LTR retrotransposons [77]. Since Helitrons use a double-stranded circular DNA intermediate during their transposition cycle [62], eccDNAs may serve a similar purpose to identify active Helitrons [77].

With our increasing understanding of the impact of Helitrons in shaping genomes and facilitating the generation of developmental novelties, particularly in plants, the obvious question arises of whether and how Helitrons could be harnessed as novel tools for genetic engineering and crop improvement. Helitrons have been linked to various domestication traits [43,44,76], implying that the deliberate activation of Helitrons could potentially facilitate the creation of novel or improved traits. Testing these ideas will provide exciting avenues for future research.

Outstanding questions

How is the transposition and activity of Helitrons regulated within genomes? What factors influence their mobility and stability?

What determines the differential amplification of Helitrons in different groups of organisms?

Do functional differences exist among distinct Helitron families? If so, how do these differences contribute to genomic dynamics in diverse species?

Which genetic pathways were established in consequence of Helitrons' transposition?

How might Helitrons be used as innovative tools for genetic engineering and the improvement of crops?

CellPress
OPEN ACCESS

Acknowledgments

This work was financially supported by the Max Planck Society.

Declaration of interests

No interests are declared.

References

1. Baduel, P. and Quadrana, L. (2021) Jumpstarting evolution: how transposition can facilitate adaptation to rapid environmental changes. *Curr. Opin. Plant Biol.* 61, 102043
2. Colonna Romano, N. and Fanti, L. (2022) Transposable elements: major players in shaping genomic and evolutionary patterns. *Cells* 11, 1048
3. Wells, J.N. and Feschotte, C. (2020) A field guide to eukaryotic transposable elements. *Annu. Rev. Genet.* 54, 539–561
4. Kapitonov, V.V. and Jurka, J. (2007) Helitrons on a roll: eukaryotic rolling-circle transposons. *Trends Genet.* 23, 521–529
5. Lal, S. *et al.* (2009) Helitrons: enigmatic abductors and mobilizers of host genome sequences. *Plant Sci.* 176, 181–186
6. Thomas, J. and Pritham, E.J. (2015) Helitrons, the eukaryotic rolling-circle transposable elements. *Microbiol. Spectr.* 3, 1–32
7. Li, S.F. *et al.* (2023) Landscape and evolutionary dynamics of Helitron transposons in plant genomes as well as construction of online database HelDB. *J. Syst. Evol.* 61, 919–931
8. Kapitonov, V.V. and Jurka, J. (2001) Rolling-circle transposons in eukaryotes. *Proc. Natl. Acad. Sci. U. S. A.* 98, 8714–8719
9. Du, C. *et al.* (2009) The polychromatic Helitron landscape of the maize genome. *Proc. Natl. Acad. Sci. U. S. A.* 106, 19916–19921
10. Yang, L. and Bennetzen, J.L. (2009) Distribution, diversity, evolution, and survival of Helitrons in the maize genome. *Proc. Natl. Acad. Sci. U. S. A.* 106, 19922–19927
11. Xiong, W. *et al.* (2014) HelitronScanner uncovers a large overlooked cache of Helitron transposons in many plant genomes. *Proc. Natl. Acad. Sci. U. S. A.* 111, 10263–10268
12. Hu, K. *et al.* (2019) Helitron distribution in Brassicaceae and whole genome Helitron density as a character for distinguishing plant species. *BMC Bioinformatics* 20, 354
13. Li, Y. and Dooner, H.K. (2009) Excision of Helitron transposons in maize. *Genetics* 182, 399
14. Li, C. *et al.* (2023) Abundance of transgene transcript variants associated with somatically active transgenic Helitrons from multiple T-DNA integration sites in maize. *Int. J. Mol. Sci.* 24, 6574
15. Kapitonov, V.V. and Jurka, J. (2003) Molecular paleontology of transposable elements in the *Drosophila melanogaster* genome. *Proc. Natl. Acad. Sci. U. S. A.* 100, 6569–6574
16. Poulter, R.T.M. *et al.* (2003) Vertebrate helentrons and other novel Helitrons. *Gene* 313, 201–212
17. Fu, D. *et al.* (2013) New insights into Helitron transposable elements in the mesopolyploid species *Brassica rapa*. *Gene* 532, 236–245
18. Pritham, E.J. and Feschotte, C. (2007) Massive amplification of rolling-circle transposons in the lineage of the bat *Myotis lucifugus*. *Proc. Natl. Acad. Sci. U. S. A.* 104, 1895–1900
19. Han, G. *et al.* (2019) Characterization of a novel *Helitron* family in insect genomes: insights into classification, evolution and horizontal transfer. *Mob. DNA* 10, 25
20. Dong, Y. *et al.* (2011) Structural characterization of Helitrons and their stepwise capturing of gene fragments in the maize genome. *BMC Genomics* 12, 609
21. Thomas, J. *et al.* (2014) DINE-1, the highest copy number repeats in *Drosophila melanogaster* are non-autonomous endonuclease-encoding rolling-circle transposable elements (Helentrons). *Mob. DNA* 5, 18
22. Lal, S.K. *et al.* (2003) The maize genome contains a Helitron insertion. *Plant Cell* 15, 381–391
23. Gupta, S. *et al.* (2005) A novel class of Helitron-related transposable elements in maize contain portions of multiple pseudogenes. *Plant Mol. Biol.* 57, 115–127
24. Gallavotti, A. *et al.* (2004) The role of barren stalk1 in the architecture of maize. *Nature* 432, 630–635
25. Choi, J.D. *et al.* (2007) Spontaneous mutations caused by a *Helitron* transposon, *Hel-It1*, in morning glory, *Ipomoea tricolor*. *Plant J.* 49, 924–934
26. Brunner, S. *et al.* (2005) Origins, genetic organization and transcription of a family of non-autonomous Helitron elements in maize. *Plant J.* 43, 799–810
27. Morgante, M. *et al.* (2005) Gene duplication and exon shuffling by Helitron-like transposons generate intraspecies diversity in maize. *Nat. Genet.* 37, 997–1002
28. Wang, Q. and Dooner, H.K. (2006) Remarkable variation in maize genome structure inferred from haplotype diversity at the bz locus. *Proc. Natl. Acad. Sci. U. S. A.* 103, 17644–17649
29. Xu, J.H. and Messing, J. (2006) Maiza haplotype with a Helitron-amplified cytidine deaminase gene copy. *BMC Genet.* 7, 52
30. Jameson, N. *et al.* (2008) Helitron mediated amplification of cytochrome P450 monooxygenase gene in maize. *Plant Mol. Biol.* 67, 295–304
31. Thomas, J. *et al.* (2014) Rolling-circle transposons catalyze genomic innovation in a mammalian lineage. *Genome Biol. Evol.* 6, 2595–2610
32. Ellison, C.E. and Bachtrog, D. (2013) Dosage compensation via transposable element mediated rewiring of a regulatory network. *Science* 342, 846–850
33. Grabundzija, I. *et al.* (2016) A Helitron transposon reconstructed from bats reveals a novel mechanism of genome shuffling in eukaryotes. *Nat. Commun.* 7, 10716
34. Tempel, S. *et al.* (2007) Model-based identification of Helitrons results in a new classification of their families in Arabidopsis thaliana. *Gene* 403, 18–28
35. Lai, J. *et al.* (2005) Gene movement by Helitron transposons contributes to the haplotype variability of maize. *Proc. Natl. Acad. Sci. U. S. A.* 102, 9068–9073
36. Han, M.J. *et al.* (2013) Identification and evolution of the silkworm Helitrons and their contribution to transcripts. *DNA Res.* 20, 471–484
37. Langdon, T. *et al.* (2009) Fragments of the key flowering gene *GIGANTEA* are associated with Helitron-type sequences in the Pooideae grass *Lolium perenne*. *BMC Plant Biol.* 9, 70
38. Sweredoski, M. *et al.* (2008) A comparative computational analysis of nonautonomous Helitron elements between maize and rice. *BMC Genomics* 9, 467
39. Barbaglia, A.M. *et al.* (2012) Gene capture by Helitron transposons reshuffles the transcriptome of maize. *Genetics* 190, 965–975
40. Hollister, J.D. and Gaut, B.S. (2007) Population and evolutionary dynamics of Helitron transposable elements in *Arabidopsis thaliana*. *Mol. Biol. Evol.* 24, 2515–2524
41. Anderson, S.N. *et al.* (2019) Transposable elements contribute to dynamic genome content in maize. *Plant J.* 100, 1052–1065
42. Chuck, G. *et al.* (2007) The maize tasselseed4 microRNA controls sex determination and meristem cell fate by targeting Tasselseed6/indeterminate spikelet1. *Nat. Genet.* 39, 1517–1521
43. Liu, Q. *et al.* (2020) A Helitron-induced RabGDIα variant causes quantitative recessive resistance to maize rough dwarf disease. *Nat. Commun.* 11, 495
44. Li, X. *et al.* (2012) A large insertion in bHLH transcription factor BrTT8 resulting in yellow seed coat in *Brassica rapa*. *PLoS One* 7, e44145
45. Inagaki, S. *et al.* (2009) A link among DNA replication, recombination, and gene expression revealed by genetic and genomic analysis of *TEBICHI* gene of *Arabidopsis thaliana*. *PLoS Genet.* 5, e1000613
46. Tsukamoto, T. *et al.* (2010) Molecular and genetic analyses of four nonfunctional S haplotype variants derived from a common ancestral S haplotype identified in sour cherry (*Prunus cerasus* L.). *Genetics* 184, 411–427

47. Lynch, B.T. *et al.* (2015) Differential pre-mRNA splicing alters the transcript diversity of Helitrons between the maize inbred lines. *G3 (Bethesda)* 5, 1703–1711
48. Zhou, Q. *et al.* (2006) Helitron transposons on the sex chromosomes of the platyfish *Xiphophorus maculatus* and their evolution in animal genomes. *Zebrafish* 3, 39–52
49. Cultrone, A. *et al.* (2007) The tightly regulated promoter of the xanA gene of *Aspergillus nidulans* is included in a Helitron. *Mol. Microbiol.* 63, 1577–1587
50. Batista, R.A. *et al.* (2019) The MADS-box transcription factor PHERES1 controls imprinting in the endosperm by binding to domesticated transposons. *eLife* 8, e50541
51. Garrigues, J.M. *et al.* (2019) Diversification of the *Caenorhabditis* heat shock response by helitron transposable elements. *eLife* 8, e51139
52. Zhao, Y. *et al.* (2022) Transposable elements: distribution, polymorphism, and climate adaptation in *Populus*. *Front. Plant Sci.* 13, 814718
53. Qiu, Y. and Köhler, C. (2020) Mobility connects: transposable elements wire new transcriptional networks by transferring transcription factor binding motifs. *Biochem. Soc. Trans.* 48, 1005–1017
54. Schreiner, W.P. *et al.* (2019) Remodeling of the *Caenorhabditis elegans* non-coding RNA transcriptome by heat shock. *Nucleic Acids Res.* 47, 9829–9841
55. Pietzenuk, B. *et al.* (2016) Recurrent evolution of heat-responsiveness in *Brassicaceae COPIA* elements. *Genome Biol.* 17, 209
56. Wang, Z. *et al.* (2022) Helitron and CACTA DNA transposons actively reshape the common wheat - AK58 genome. *Genomics* 114, 110288
57. Yang, L. and Bennetzen, J.L. (2009) Structure-based discovery and description of plant and animal Helitrons. *Proc. Natl. Acad. Sci. U. S. A.* 106, 12832–12837
58. Zhao, J. *et al.* (2023) Centromere repositioning and shifts in wheat evolution. *Plant Commun.* 4, 100556
59. Schnable, P.S. *et al.* (2009) The B73 maize genome: complexity, diversity, and dynamics. *Science* 326, 1112–1115
60. Xiong, W. *et al.* (2016) Rolling-circle amplification of centromeric Helitrons in plant genomes. *Plant J.* 88, 1038–1045
61. Tunji, M. *et al.* (2020) Sequence composition underlying centromeric and heterochromatic genome compartments of the Pacific oyster *Crassostrea gigas*. *Genes (Basel)* 11, 695
62. Grabundzija, I. *et al.* (2018) Helraiser intermediates provide insight into the mechanism of eukaryotic replicative transposition. *Nat. Commun.* 9, 1278
63. Thomas, J. *et al.* (2011) The limited distribution of Helitrons to vesper bats supports horizontal transfer. *Gene* 474, 52–58
64. Thomas, J. *et al.* (2010) Pervasive horizontal transfer of rolling-circle transposons among animals. *Genome Biol. Evol.* 2, 656–664
65. Guo, X. *et al.* (2014) Evidence of horizontal transfer of non-autonomous Lep1 Helitrons facilitated by host-parasite interactions. *Sci. Rep.* 4, 5119
66. Coates, B.S. (2015) Horizontal transfer of a non-autonomous Helitron among insect and viral genomes. *BMC Genomics* 16, 137
67. Roffler, S. *et al.* (2015) The making of a genomic parasite - the Mothra family sheds light on the evolution of Helitrons in plants. *Mob. DNA* 6, 23
68. Erdmann, R.M. and Picard, C.L. (2020) RNA-directed DNA methylation. *PLoS Genet.* 16, e1009034
69. Muyle, A. *et al.* (2021) Gene capture by transposable elements leads to epigenetic conflict in maize. *Mol. Plant* 14, 237–252
70. Wilkinson, S.W. *et al.* (2023) Long-lasting memory of jasmonic acid-dependent immunity requires DNA demethylation and ARGONAUTE1. *Nature Plants* 9, 81–95
71. Tanurdzic, M. *et al.* (2008) Epigenomic consequences of immortalized plant cell suspension culture. *PLoS Biol.* 6, e302
72. Touati, R. *et al.* (2019) Distinguishing between intra-genomic Helitron families using time-frequency features and random forest approaches. *Biomed. Signal Process. Control* 54, 101579
73. Touati, R. *et al.* (2019) A combined support vector machine-FCGS classification based on the wavelet transform for Helitrons recognition in *C. elegans*. *Multimed. Tools Appl.* 78, 13047–13066
74. Touati, R. *et al.* (2020) Classification of intra-genomic Helitrons based on features extracted from different orders of FCGS. *Inform. Med. Unlocked* 18, 100271
75. Touati, R. *et al.* (2021) New intraclass Helitrons classification using DNA-image sequences and machine learning approaches. *IRBM* 42, 154–164
76. Gao, C. *et al.* (2016) Helitron-like transposons contributed to the mating system transition from out-crossing to self-fertilizing in polyploid *Brassica napus* L. *Sci. Rep.* 6, 33785
77. Peng, H. *et al.* (2022) Extrachromosomal circular DNA: a neglected nucleic acid molecule in plants. *Curr. Opin. Plant Biol.* 69, 102263
78. Kosek, D. *et al.* (2021) The large bat Helitron DNA transposase forms a compact monomeric assembly that buries and protects its covalently bound 5′-transposon end. *Mol. Cell* 81, 4271–4286
79. Feschotte, C. and Wessler, S.R. (2001) Treasures in the attic: rolling circle transposons discovered in eukaryotic genomes. *Proc. Natl. Acad. Sci. U. S. A.* 98, 8923–8924
80. Kazlauskas, D. *et al.* (2019) Multiple origins of prokaryotic and eukaryotic single-stranded DNA viruses from bacterial and archaeal plasmids. *Nat. Commun.* 10, 3425
81. Heringer, P. and Kuhn, G.C.S. (2022) Pif1 helicases and the evidence for a prokaryotic origin of Helitrons. *Mol. Biol. Evol.* 39, msab334
82. Heringer, P. and Kuhn, G.C.S. (2018) Exploring the remote ties between Helitron transposases and other rolling-circle replication proteins. *Int. J. Mol. Sci.* 19, 3079