Review

**Helitrons: Enigmatic abductors and mobilizers of host genome sequences**

Shailesh Lala,*, Matthew Oetjens, L. Curtis Hannah

---

**Abstract**

Helitrons are a recently discovered superfamily of eukaryotic transposable elements. They are known for their ability to capture and mobilize gene fragments and, in turn, significantly contribute to the lack of gene colinearity widely reported among different maize inbred lines. As judged by present evidence, Helitrons differ fundamentally from Class I and Class II families of transposable elements in both structure and the mechanism of transposition. These elements are poorly understood despite their massive abundance, their structural diversity and the important role they apparently play in the evolution of maize genome. Although evidence for recent Helitron movement has been reported in maize, autonomous Helitron activity has not been reported. Helitrons are postulated to transpose via a so-called rolling circle (RC) mechanism involving strand replacement catalyzed by helicase and replicase enzymes encoded by an autonomous element. However, no experimental evidence in support of this hypothesis has been reported. Several models for gene piece capture have been proposed based on the structural features of Helitrons and the comparison of captured gene fragments of related elements. However no supporting evidence is extant for any of the models proposed. A better understanding of these elements requires concrete evidence of their transposition in the present day genome, establishment of a system to assay their transposition and an analysis of additional indigenous Helitrons in other species. This review critically analyses the proposed mechanisms of Helitron transposition, their impact on genome evolution and the process by which these enigmatic elements capture and multiply host genes.

---

1. Introduction

Helitrons are a novel family of transposable elements that were only recently discovered despite their massive abundance in some eukaryotic genomes. Helitrons possess unique features that make them fundamentally different from other known transposable elements. The existence of Helitrons was first proposed from the analysis of repetitive sequences in a number of eukaryotic organisms [1]. These analyses also identified a hypothetical autonomous Helitron that would contain open reading frames encoding a DNA helicase and a replication initiator protein. Because these proteins are associated with some bacterial transposons that replicate via rolling circle replication, it has been suggested that Helitrons also transpose via a similar mechanism involving strand replacement and replication [1–3]. Furthermore, Helitrons unlike other known Classes I and II families of transposable elements, lack terminal repeats, contain short palindrome sequence located 15–20 nucleotides upstream from...
the 3’ terminus, do not duplicate host sequences upon insertion and usually insert between dinucleotides, A and T [1].

Helitrons have been reported to reside in the genome of species belonging to all eukaryotic kingdoms, but their abundance and diversity vary drastically even among closely related species. Helitrons very similar in sequence to those found in C. elegans genome have been described in the mosquito Anopheles gambia. Intriguingly, Anopheles Helitrons lack introns [4]. In the fruit fly Drosophila virilis, different families of Helitrons comprise approximately 5% of the genome [5]. However, the genome of Drosophila melanogaster harbors only a portion of 5’ replication initiator (Rep) protein domain similar to the Helitron in the Anopheles genome [4]. Helitrons have also been reported in the genomes of Zebrafish, Pufferfish and Platyfish. Fish Helitrons are unique in encoding a domain which closely resembles an apurinic-apyrimidinic (AP) endonuclease in the C-terminal portion of Rep/Hel [4]. Intriguingly, the sex chromosome of platyfish Xiphophorus maculates contains several copies of intact Helitrons which are apparently transcribed. These observations suggest that Helitrons are active in platyfish genome and may be involved in the evolution of sex chromosomes [6]. In mammals, the Helitrons have only been reported in the genome of the little brown bat Myotis lucifugus. Two families of Helitrons have multiplied to greater than 100,000 copies, representing more than 3% of the brown bat genome [7]. The amplification of the bat Helitron is estimated to have occurred between 30 and 36 MYA, which roughly coincides with an expansion of the vesper bats across the widest geographic range in the bat order. Comparison of Helitron sequences between different species indicates that horizontal transfer may have played a vital role in the evolution of these elements.

The presence of nearly identical Helitrons at different loci in one genome indicates that these elements have been active very recently, however unequivocal evidence of Helitron activity has not been reported in any organism [1,8]. The putative autonomous Helitron bearing an intact helicase (HEL) and replication protein A (RPA1)-like protein reported in some species, not in the fully sequenced genomes of maize. However, the discovery of two maize mutants, sh2-7527 and ba-ref1 caused by a recent insertion of Helitrons in shrunken-2 and barren stalk-1 genes indicates that the present-day maize genome harbors transposable Helitron [9,10].

Helitrons capture and mobilize gene fragments in different genomes but the frequency of gene piece capture is extraordinarily high in maize. Helitrons appear to have captured and inserted thousands of gene piece sequences; however the presence of a functional gene inside a Helitron has yet to be reported. Despite their tremendous impact on genome evolution of many eukaryotes, Helitrons are poorly understood due to the lack of a system to assay transposition. This review critically evaluates various proposals and hypotheses put forward to explain the mechanism by which Helitrons transpose and capture gene sequences.

2. The vast movement of gene sequences by Helitrons may impact hybrid vigor or heterosis in maize

The movement of gene pieces by Helitrons has apparently played a major role in causing the lack of gene sequence collinearity reported among different maize inbred lines. This process was first reported as haplotype variability at the bz (bronze) locus between the inbreds McC and B73. Variation was due entirely to the presence/absence of two Helitron insertions [11]. These Helitrons harbored fragments of a number of genes. The extent of gene movement attributed to Helitrons in these inbred lines was simultaneously analysed using hybridization array and oligonucleotide probes [12]. These analyses indicated that Helitron-mediated movement of gene pieces was primarily responsible for ~20% of the 20,656 gene pieces present in only one of the two inbred lines. Thus Helitrons appear to have significantly contributed to the magnificent diversity and dynamism of the maize genome. This, in turn, may have been pivotal in the domestication and evolution of the maize genome.

The gene pieces captured by Helitrons are sometimes transcribed into a common transcript giving birth to chimeric or mosaic transcripts with conjoined coding regions of the different captured gene pieces. This potentially impacts host gene expression in several ways [9,12,13]. First, the transcribed regions of the genes may interact and compete with the transcript of the wild type counterpart for mRNA processing factors, etc; thus affecting host gene expression at the transcriptional level. Second, the putative translational products may interact with the wild type proteins and produce biologically inactive hetero-multimers, and hence, serve as a dominant regulator of the wild type protein. Third, in rare cases, genes captured by Helitrons may be transcribed in the antisense direction. The resulting transcript could then potentially interact with wild type transcription in the sense orientation and regulate expression via siRNA-mediated silencing.

Importantly, if the chimeric transcripts contain important functional domains derived from different genes, the newly created transcript may encode novel biological functions and be subject to evolutionary selection. However, the creation of an intact functional gene attributed to Helitrons capture, awaits further investigation.

The Helitron-mediated + polymorphisms may contribute to copy number variation between different maize inbred lines and thus impact the phenomenon of heterosis or hybrid vigor [14]. Heterosis refers to the enhanced vigor of progeny compared to the two parents. Classically, one explanation for heterosis is that the progeny contains a greater number of “favorable” alleles than does either parent. If Helitrons are active in modern maize inbreds and have the ability to capture whole genes, then different inbreds could differ in the number of functional copies of particular genes. If, for example, gene A was duplicated in inbred X and gene B duplicated in inbred Y, then the hybrid between inbreds X and Y would contain multiple copies of both genes. This course would not occur in either parent. The proposed rolling circle mechanism of Helitron transposition involving copy and paste increases the copy number of the genes (Fig. 1) [15–17]. However, a significant impact on heterosis entails movement of functional genes by Helitrons. Helitron capture of an entire gene has yet to be reported.

The locations of the promoters driving the transcription of genes captured by Helitrons have been determined in only two cases [9,18]. It is postulated that gene pieces are transcribed by heterologous promoters, strategically located in the upstream flanking region, outside the Helitron. The capture and amplification of a maize cytochrome P450 monoxygenase (P450) gene into three different regions of the genome has been recently described [17]. All the three copies of the Helitron-captured P450 genes are transcribed, even though they lack the homologous 5’ end including the promoter. Intriguingly, one of the P450 bearing Helitrons was transposed into a retroelement and inserted in opposite orientation relative to the retroelement. This makes it highly unlikely that the P450 gene is transcribed by the retroelement promoter, suggesting that this P450 gene is transcribed by a heterologous promoter located within the Helitron and near its 5’ terminus.

The abundance, diversity and impact of Helitrons on the maize genome may not be apparent in the recently sequenced genome of maize inbred B73 because discovery of highly diverse maize Helitrons often requires demonstration of + polymorphisms at the sites of integration. Accordingly, the mass movement of gene piece sequences by Helitrons will probably complicate the annotation of the maize genes in the sequenced genome of B73. Thus, the total analysis of the impact of
extreme diversity of maize gene-like sequences generated by Helitrons warrants sequencing multiple diverse inbred lines.

3. Helitrons indiscriminately capture and mobilize gene sequences

Gene pieces captured by Helitrons appear to be random and differ drastically in size and sequence as compared to the wild type counterpart. It has been proposed that genes captured by Helitrons that bestow transposition advantage to the elements are maintained and thus display more similarity to their progenitor genes [1]. However, the biological relevance of any captured gene piece remains to be demonstrated. It is also quite feasible that the extent of similarity between the captured gene piece and the progenitor gene reflects the evolutionary time of capture rather than the biological importance to the element or to the host organism for that matter. Furthermore, genes detected within Helitrons represent truncated versions of their progenitors and are frequently interrupted by mutations, minimizing the possibility of encoding any biologically relevant protein. Almost intact portions of a cytidine deaminase and a cytochrome-P450-monooxygenase (P450) were recently reported in maize Helitrons [16,17]. However, the 5’ termini of both genes including the promoters were missing in the Helitron versions. The potential of Helitrons to capture and mobilize full-length functional genes awaits further elucidation.

4. Possible mechanisms of gene capture by Helitrons

Capture of gene sequences by plant transposable elements has been noted in a number of studies. For example, approximately 40% of the Basho family of non-autonomous Helitrons comprising 2% of Arabidopsis thaliana contains gene pieces derived from only five host genes [19]. However, the extent and magnitude of gene capture by maize Helitrons seem only matched by the Mutator-like transposable elements (MULEs) in rice [20,21]. The capture and mobilization of more than 1000 gene sequences by MULEs may have played a major role in the evolution and expression of rice genome. Whether MULEs employ mechanisms similar to maize Helitrons for gene capture awaits further investigation. The lack of host rice genes similar to approximately 20% of the genes abducted by MULEs, and their presence in other related species suggests that capture of genes by MULEs involves a mechanism that may cause destruction of the host gene. In contrast, the putative progenitors of genes captured by Helitrons apparently exist in the maize genome. However it remains a possibility that in some cases, progenitors may represent genes in a similar Helitron inserted in different regions of the genome [12].

Several mechanisms of gene capture by Helitrons have been proposed, based on the structure and postulated mechanism of their transposition. It was initially proposed that Helitron transposition begins at the 5’ terminus and the inefficient recognition of the palindromic sequence at the 3’ terminus leads to the capture of the flanking downstream genome sequence [22]. According to this model, random palindrome sequences located downstream of the insertion site act as new transposition terminators, giving rise to new Helitrons with variable 3’ ends. However, the highly conserved nature of the 3’ terminus compared to the 5’ terminus of the maize Helitrons makes this scenario of gene capture unlikely. This led to the contrasting proposal which argues that the transposition of Helitrons is initiated at the 3’

Fig. 1. Helitron-mediated gene amplification and its possible impact on gene dosage and heterosis. The left and right panels display Helitron capture of hypothetical genes, Gene 1 and Gene 2 in maize inbred lines X and Y, respectively. The exons and introns of the genes are displayed by color coded blocks and black lines, respectively. The copies of genes in each inbred line following homozygosis and in the hybrid are enclosed in circles. The bottom circle displays single alleles for genes 1A and 2A resulting from the cross between two inbred lines.
terminus of the element and capture of the genes occurs at the 5' end by some unknown mechanism [13]. Both of these models are displayed side by side in Fig. 2. It has also been proposed that genes trapped within Helitrons may represent fortuitous template or “filler DNA” during repair of double strand breaks presumably occurring during Helitron transposition [23]. These mechanisms of Helitron transposition fail to account for the facts that (1) the vast majority of genes captured by Helitrons are oriented in the same direction as the element and (2) in several cases, sequences of the genes originally residing in different regions of the genome are inserted one within another inside the Helitron [[20], Lal and Hannah, unpublished result].

Fig. 2. Proposed models for gene capture by Helitrons. The top panel displays the schematics of a Helitron insertion. The Helitron is displayed in green and the arrows mark the direction of the Helitron. The hairpin loops display the 3' palindrome sequence. The conserved sequence at the 5' and 3' terminus is indicated. Hypothetical genes flanking the 5' and 3' ends of the Helitron are displayed in blue and red, respectively. The left and right panels display the mechanism by which genes flanking the 5' and 3' ends of the Helitrons are captured during transposition. The purple balls represent the transposase molecule binding the 5' and 3' terminal ends of the Helitron and their target site, respectively. The schematics show that Helitrons capture 5' and 3' flanking gene sequences by inefficient recognition of the transposition termination at the 5' termini or 3' palindrome signals. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

Fig. 3. Schematic comparisons of the putative mechanisms of gene acquisition by Integrons and Helitrons. The left and right panels display the basic structure and the mechanism of gene capture by Integrons and Helitrons, respectively. The letter P denotes the position of the promoter. The integrase (int1) and sulfonamide (sol1) gene of an integron, and the 5' and 3' conserved terminal (CS) ends are indicated. The integration site of an integron and the short sequence allowing integration are indicated by attI and attC, respectively. The recombination sites are marked by letter X. The sequence and the orientation of the captured genes are displayed by color coded arrows.
Intriguingly, maize Helitrons share striking structural similarity to bacterial integrons. These elements capture gene sequences via site-specific recombination and generate circular intermediates [24]. Both Helitrons and integrons are mobile, lack terminal repeats and cause no duplication of host genome sequence upon insertion. Furthermore, both elements contain variable regions consisting of captured genes flanked by conserved regions at their termini. Of most importance, gene cassettes in integrons are oriented in the same direction, as is the case with gene insertion in Helitrons, and different captured genes are co-transcribed, encoding mRNAs containing the coding regions of different genes. Extra-chromosomal circular DNA originating from the repeated genomic sequences has been reported in plants [25]; however, it remains to be determined whether they provide the template for genes captured by Helitrons. In the absence of circular host DNA, we note that capture of a host gene sequence by a Helitron would require three recombination events, two flanking the sequences of the host gene and one at the integration site inside the Helitron. Whether Helitrons recruit a site-specific recombinase that is similar to integrase for capturing gene sequences awaits further investigation however, sequences similar to integrate have never been reported in Helitrons. Since putative non-autonomous Helitrons are likely to rely on proteins encoded by autonomous elements for their transposition, the lack of an integrate gene does not formally exclude the possibility of its involvement in Helitron function. We postulate that gene acquisition and directional integration of captured genes into Helitrons is mediated by integrate-like proteins provided in trans by either the host or other elements, such as retroelements residing in the genome. A scheme of the mechanism of gene capture by integrons and the proposed mechanism for Helitrons are displayed in Fig. 3.

Experimental validation of various hypotheses put forward to explain mechanism of gene capture by Helitrons awaits demonstration of direct molecular or genetic evidence for Helitron movement in the maize genome. And an increase in the dataset of Helitrons derived from different inbred lines should shed light on events involving gene capture. For example, similar Helitrons bearing different cadre of genes, inserted at the same position in different inbred lines would serve as good candidates to analyse whether Helitrons, like integrons, capture genes by events that are independent of transposition.

5. Conclusion and future prospects

The remarkably high degree of diversity and polymorphism, both in sequence and length among different members of Helitron family enabled these elements to remain elusive until only recently, despite their massive abundance and possible impact on evolution of many eukaryotic genomes. Their relatively recent discovery and lack of readily detectable transposition activity in a present-day genome have made these enigmatic elements subject to molecular basis of heterosis. The transposition of non-autonomous maize Helitrons may not be dependent on autonomous elements, but rather may rely on the transposition proteins encoded by the host genome. Furthermore, it has been postulated that the transposition helicase and ssDNA binding protein of an autonomous Helitron represent host genes captured by the element, as they display conservation of sequence and gene structure (exon—intron junction) with the host protein [1].

Helitrons were initially discovered through bioinformatic tools, which subsequently contributed to the current status of their knowledge. A better understanding of Helitron transposition and gene acquisition mechanisms will require establishment of an in vivo or in vitro system to assay their transposition. The discovery of Helitron induced maize mutants and the establishment of maize as a robust model system through six decades of study of transposable elements provides a valuable resource to investigate the activity of these elements.

Acknowledgements

The study was supported by grants from National Science Foundation (USA) to SL and LCH. Work in the Hannah laboratory is supported by National Science Foundation Grants 105-0815104, I0B-0444031, DBI-0077676, DBI-0606607, and I0B-9982626 and USDA Grants 2006-35100-17220 and 2008-35318-18649.

References


