Heredity (2005) 95, 421–422

www.nature.com/hdy

NEWS AND COMMENTARY

Plant genomes

Massive changes of the maize genome are caused by *Helitron*s

SK Lal and LC Hannah

Heredity (2005) **95,** 421–422. doi:10.1038/sj.hdy.6800764; published online 12 October 2005

One of the significant tenets from the Nobel prize-winning work of Barbara McClintock some 50 years ago was that genetic elements of the maize genome move from chromosome to chromosome, changing position and/or copy number in the process. Only now is the magnitude of McClintock's insight being fully realized and appreciated. In a recent issue of Nature Genetics, Morgante *et al* (2005) report the remarkable discovery that a vast majority of the sequence diversity distinguishing two well-known maize lines is due to a newly-described class of transposable elements termed *Helitrons*.

Morgante's group (2005) focused their analysis on two maize inbreds, B73 and Mo17. These inbreds were used extensively in the early days of commercial corn breeding and remain important today. Surprisingly, examining genomic clones from these inbreds with nearly 15 000 oligonucleotide probes revealed that some 20% of the 20 656 gene fragments detected were not shared by both inbreds. If there are 40 000 genes in the maize genome, one would expect that ~10 000 genes or gene pieces would be missing in one of the two inbreds.

To analyze this ±DNA polymorphism in more detail, Morgante coworkers perused five chromosomal regions for which DNA sequence data were available from both inbreds. These regions comprise only a small fraction of the total genome. They found nonshared sequences within these regions to be clustered and from comparing with the sequenced rice genome, these unique sequences arose from insertion into one inbred rather than to deletion in the other inbred.

Remarkably, in-depth analysis of nine insertions showed that eight were caused by *Helitrons*. *Helitrons* are eukaryotic transposable elements recognized only recently by computer analysis of repetitive DNA sequences of Arabidopsis, rice and *Caenorhabditis elegans* (Kapitonov and Jurka, 2001). *Helitrons* are quite large (>10 kbp) but unlike virtually all other classes of transposable elements, *Helitron*s lack terminal repeats and do not duplicate host sequences during the insertion process. They insert within the host dinucleotide, AT. The only invariant *Helitron* sequences are a TC 5′ terminus and a CTRR 3′ terminus. Most *Helitron*s so far investigated have a 10–16 bp palindrome near the 3′ end. From computer analysis, Kapitonov and Jurka (2001) proposed an autonomous *Helitron*. This hypothetical element encodes a HEL protein composed of a rolling-circle replication initiator and DNA helicase domains needed for transposition. It also contains a replication protein A (RPA)-like protein with putative single-stranded DNA-binding activity. As these protein sequences are involved in rolling circle replication and transposition of some bacterial transposable elements, it has been hypothesized that *Helitron*s also move through a rolling circle type of replication and strand replacement (Kapitonov and Jurka, 2001; Feschotte and Wessler, 2001). Following the report of the hypothetical *Helitron*, researchers showed two loss-of-function maize mutations (Lal *et al*, 2003; Gallavotti *et al*, 2004) to be caused by *Helitron* insertion.

A remarkable characteristic of *Helitron*s is their ability to incorporate host gene sequences. For example, one *Helitron* contains pieces of 12 different genes (Lal *et al*, 2003). From the limited sequence data available, *Helitron*s appear gene-sequence rich relative to the entire maize genome.

Although two had already been characterized by Lai *et al* (2005), characterization of the eight *Helitron*s by Morgante and co-workers revealed a number of seminal facts. Two *Helitron*s were interrupted by other transposable elements. In one case an LTR retrotransposon (transposable element that moves through RNA and contains long terminal repeats) had been inserted into the *Helitron* while in another case, a member of the DNA-based *Suppressor-mutator/Enhancer* family of elements was inserted. A third *Helitron* was inserted into a retro-transposon.

Sequence analysis of extant maize genomic DNA identified four additional probable *Helitron*s with sequence similarity to the termini of the eight *Helitron*s. All contained gene fragments. Interestingly, one of these *Helitron*s is interrupted by a retro-transposon in only one of the two inbreds.

Morgante's group (2005) searched for an autonomous *Helitron* and found 11 sequences within the maize genome with similarity to the HEL protein. Five of these were flanked by sequences similar to the RPA protein while in two cases the two genes were contained within sequences with *Helitron* termini. Whereas none of these sequences contained significant open reading frames, cDNA clones with high similarity to eight of the 11 putative HEL genes were isolated.

Some *Helitron*-contained sequences are expressed. Morgante and co-workers examined RNA derived from these inbreds and detected RNA arising from the gene pieces in four *Helitrons*. Significantly, and as in the case of the *Helitron* insertion in the *shrunken2* maize locus (Lal *et al*, 2003), these transcripts are chimeric. Sequences arising from the different captured gene pieces of the *Helitron* are conjoined in one transcript. Morgante *et al* (2005) speculate, as have others before them (Lai *et al*, 2005; Lal and Hannah, 2005), that *Helitrons* may act to shuffle exons and thereby create new genes. It is intriguing that a similar role was proposed for introns in gene evolution (for example, Gilbert, 1978)

The propensity of *Helitron*s to capture gene pieces has also been noted for another transposable element family. The MULE family of elements has captured and mobilized more than 1000 gene fragments in the rice genome (Jiang *et al*, 2004). However, the mechanism by which these two super families of transposable elements acquire gene fragments may differ. For example, 20% of the gene fragments captured by MULEs bear no apparent wild-type counterpart in the rice genome but display strong similarity to protein sequences from other plant species, suggesting that the function of the host gene is destroyed or lost during the process of acquisition. In contrast, Morgante *et al* (2005) demonstrate that the wild-type progenitor of four different gene fragments captured by a *Helitron* exists in the maize. These investigators speculate that *Helitron* transposition involves replication and

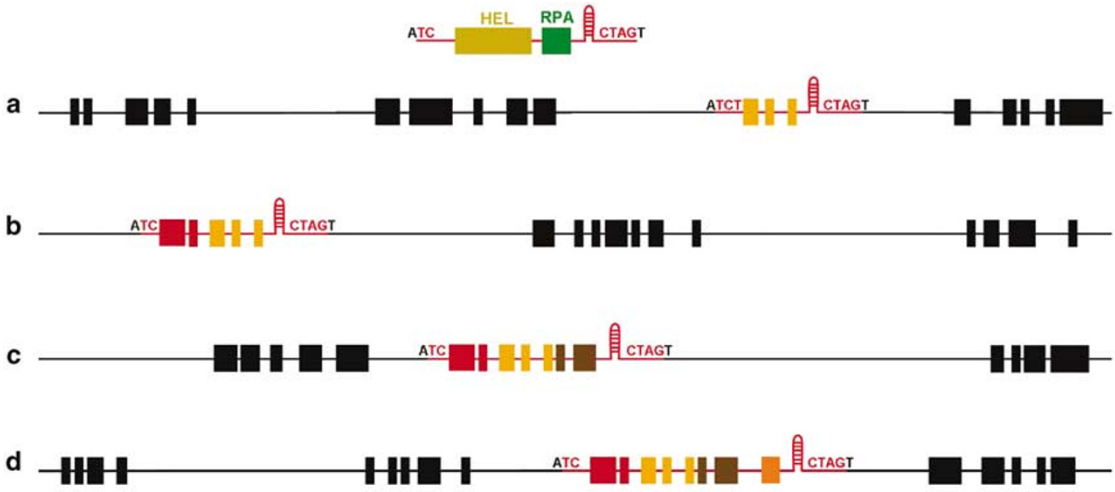

Figure 1 Hypothetical journey of a *Helitron* through the maize genome. The upper panel displays the structure of autonomous maize *Helitron* inserted between nucleotide A and T. The invariant terminal nucleotides and the coding regions of DNA helicase (HEL) and single-stranded DNA binding protein (RPA) are indicated. The loop near the 3′ end of the *Helitron* represents the palindrome sequence. The lower four panels represent a hypothetical journey of a non-autonomous *Helitron*, lacking the HEL and RPA genes, as it inserts sequentially into four different regions of the genome and grows in size by capturing gene sequences. The exons of the wild-type maize genes are represented by black boxes. The red lines represent the *Helitron* insertion and the colored boxes represent the exons of the captured genes. Exons of different genes are given different colors.

strand replacement that preserves the integrity of the donor host sequence.

While two models (Feschotte and Wessler, 2001; Gupta *et al*, 2005) can explain *Helitron* gene capture and transposition, the exact mechanism(s) is far from clear. Likely relevant is the fact that captured gene pieces are oriented in the same direction. Also, the fact that nearly identical copies of nonautonomous *Helitron*s have been found at multiple locations (Lai *et al*, 2005) demonstrates that *Helitron*s can move through the genome without altering internal DNA sequences.

Many important questions remain to be answered. We do not know whether *Helitron*s can capture whole genes. So far, only gene fragments – many spanning several exons – have been documented, although there is no apparent reason why whole genes cannot be incorporated into *Helitron*s. If whole genes are captured, then *Helitron*s may play a significant role in creating the genetic heterozygosity exploited in hybrid vigor and other genetic phenomena. As noted by Morgante and co-workers it is also possible that expression of the gene pieces contained within *Helitron*s could lead to silencing of the cognate host genes.

Presently, it is unclear whether *Helitron*s are still active in the maize genome or whether the appropriate enzymatic activities for movement are still encoded by the maize genome. Sequencing of the maize genome should shed light on these important questions.

Are gene pieces added serially during the journey of a *Helitron* through the genome? If gene-piece capture is a function of transposition, one might be able to map the history of *Helitron* movement via the order of gene capture (Figure 1).

Finally, a more in-depth understanding of the mechanism of transposition of these of these enigmatic elements may provide novel tools for maize transformation and perhaps other organisms. That DNA fragments of greater than 10 000 bp require only a few flanking sequences for transposition may allow several genes to be inserted routinely into the maize genome in a single transformation event.

SK Lal is at the Department of Biological Sciences, Oakland University, Rochester, MI 48309-4401, USA. LC Hannah is at the Department of Horticultural Sciences and Program of Plant Molecular and Cellular Biology, University of Florida, Gainesville, Fl 32611, USA.

e-mail: Hannah@mail.ifas.ufl.edu

Feschotte C, Wessler SR (2001). *Proc Natl Acad Sci USA* **98**: 8923–8924.
Gallavotti A *et al* (2004). *Nature* **432**: 630–635.
Gilbert W (1978). *Nature* **271**: 501–503.
Gupta S *et al* (2005). *Plant Mol Biol* **57**: 115–127.
Jiang N *et al* (2004). *Nature* **431**: 569–573.
Kapitonov VV, Jurka J (2001). *Proc Natl Acad Sci USA* **98**: 8714–8719.
Lai J *et al* (2005). *Proc Natl Acad Sci USA* **102**: 9068–9073.
Lal SK *et al* (2003). *Plant Cell* **15**: 381–391.
Lal SK, Hannah LC (2005). *Proc Natl Acad Sci USA* **102**: 9993–9994.
Morgante M *et al* (2005). *Nat Genet* **37**: 997–1002.