Advanced Journal of Microbiology Research Vol. 2003
Available online at http://internationalscholarsjournals.org/journal/ajmr

Plant retroviruses: Structure, evolution and future applications

Essam A. Zaki*

Genetic Engineering & Biotechnology Research Institute, GEBRI, Research Area, Borg El Arab, Post Code 21934, Alexandria, Egypt.

Accpted 20 May 2003

Retroelements, which replicate by reverse transcription, have been detected in higher plants, higher animals, fungi, insects and bacteria. They have been classified into viral retroelements, eukaryotic chromosomal non-viral retroelements and bacterial chromosomal retroelements. Until recently, retroviruses were thought to be restricted to vertebrates. Plant sequencing projects revealed that plant genomes contain retroviral-like sequences. This review aims to address the structure and evolution of plant retroviruses. In addition, it proposes future applications for these important key components of plant genomes.

Key words: Interspecies gene flow, plant genes vectors, plant retroviruses, retroelements, sequence divergence, transgenic plants.

RETROELEMENTS CLASSIFICATION

The replication of most of nucleic acids is either from DNA to DNA (chromosomal and viral nucleic acids) or from RNA to DNA (viruses and some cytoplasmic nucleic acids). However, an increasing number of nucleic acids are being found whose replication involves reverse transcription of RNA to produce DNA. This replication is driven by the enzyme reverse transcriptase (RT), which was first recognized over 30 years ago (Baltimore, 1970; Temin and Mizutani, 1970). Nucleic acids that replicate by reverse transcription are termed retroelements (Temin, 1989) and this form of replication is employed by elements in higher plants, higher animals, fungi, insects and bacteria. Retroelements have been classified into viral retroelements, eukaryotic chromosomal non-viral retroelements and bacterial chromosomal retroelements (Table 1).

RECLASSIFYING RETROELEMENTS

Eukaryotic genomes harbor mobile genetic elements known as long terminal repeat (LTR) retrotransposons. LTR retrotransposons are closely related to the infectious and endogenous retroviruses (Wilhelm and Wilhelm, 2001). LTR retrotransposons and retroviruses have two genes in common, *gag*, which codes proteins for virus particles, and *pol*, which encodes the enzymatic activities for replication (Malik et al., 2000). The viral envelope (*env*) gene of the retroviruses distinguishes them from the LTR retrotransposons. Viral envelope glycoproteins associate with cell membranes and facilitate the budding of viral core particles from infected cells. In addition, they also mediate infection by recognizing cellular receptors (Coffin et al., 1997).

Until recently, retroviruses were thought to be restricted to vertebrates. Intracellular virus- like particles, however, had been observed for several LTR retrotransposons (Malik et al., 2000). Furthermore, structural and functional data converged when it was shown that the *gypsy* element of *D. melanogaster* was able in some circumstances to function as a retrovirus (Kim et al., 1994; Song et al., 1994). This result established the convenience of classifying LTR retrotransposons as viruses. In the most recent virus taxonomy, LTR-containing retrotransposons are reclassified into two main families, Pseudoviridae (corresponding to the *copia* subgroup) and Metaviridae *(gypsy* elements). The Metaviridae are further split according to the presence of the *env* gene (genus *Errantivirus*) or its absence (genus *Metavirus*) (Hull, 2001) (Figure 1). Consequently, it is likely that the *env* gene is more widespread among invertebrates than previously thought. Furthermore, this reevaluation is not too surprising given that it is now believed that retroviruses have evolved from the *gypsy*-like retrotransposons by acquiring the *env* gene (Eickbush and Malik, 2002).

*Current Address; Department of Biological Sciences, 1392 Lilly Hall of Life Sciences, West Lafayette, IN 47907-1392, Phone: (765) 494-9837 Fax: (765) 496-1496, E-mail: ezaki@purdue.edu.

Abbreviations; INT: integrase, LTR: long terminal repeat, RT: reverse transcriptase.

Table 1. Retroelements classification.

Viral retroelements		
	I. Retroviruses (RNA in virions)	II. Pararetroviruses (DNA in virions)
Eukaryotic non-viral retroelements		
	I. Retrotransposons	II. Retroposon
LTR	+	±
RT	+	+
INT	+	+
Bacterial retroelements		
	Retron	
LTR	±	
RT	+	
INT	±	

LTR: long terminal repeat, RT: reverse transcriptase. INT: integrase. (Adapted from Hull and Will, 1989).

PLANT RETROVIRUSES

In plants, retrotransposons have been extremely successful as evident to their abundance (Kumar and Bennetzen, 1999). Their ubiquity in the plant kingdom suggests that they are of very ancient origin (Bennetzen, 2000). In addition, their abundance has played a major role in plant genome structure and evolution (Bennetzen, 2002) . In this regard, the possibility that retroviruses might exist in plants had been discussed (Kumar, 1998; Kumar and Bennetzen, 1999), but it is only very recently that plant genomes have been shown to contain retroviral- like sequences (Table 2) . In other words, the detection of *env*-like gene. It is noteworthy that the presence of an *env*-like gene that encodes a transmembrane protein is generally considered to be a predictor of a retroelement's infectious nature (Peterson-Burch et al., 2000).

Table 2. Examples of plant retroviruses.

Retrovirus	Plant	Reference
SIRE-1	Soybean	Laten et al., 1998
*Tat*1	*A. thaliana*	Wright and Voytas, 1998
Athila4	*A. thaliana*	Wright and Voytas, 1998
Cyclops	Pea	Chavanne et al., 1998
Bagy-2	Barley	Vicient et al., 2001
GM-5 and *GM*-6	*Gossypium*	Abdel Ghany and Zaki, 2002

Genome sequencing projects have enhanced our understanding of diversity and evolutionary trends among retrotransposons (Eickbush and Furano, 2002). In this regard, plant retroviruses were identified through plant genome sequencing projects. *Gossypium* retroviruses, however, were identified using a novel approach (Abdel Ghany and Zaki, 2002). These elements were isolated using specific oligonucleotides for the *gypsy env*-gene, suggesting that *env*-like genes are ubiquitous in the plant kingdom, and are evolutionary related to the *Drosophila gypsy env*-gene. In addition, it offers a simple and universal method for the isolation of *env*-like genes in plants (Abdel Ghany and Zaki, 2002).

STRUCTURE OF THE *env* GENE IN PLANTS

The retroviral envelope (*env*) gene encodes a polypeptide which is cleaved into two proteins: the surface protein (SU), which is involved in receptor recognition, and the transmembrane (TM) subunit, which anchors the entire *env* complex and is directly responsible for cell membrane fusion and virus entry (Coffin et al., 1997). The *env* genes are the most variable retroviral genes and therefore not readily identified from primary sequence data (Malik et al., 2000). Nevertheless, there are a few predictable secondary structure features including: a signal peptide, a fusion peptide, an anchor peptide, a peptide cleavage site, glycosylation sites and a C-terminal transmembrane domain (Eickbush and Malik, 2002). These conserved features in the known animal retroviral envelope proteins were compared with the

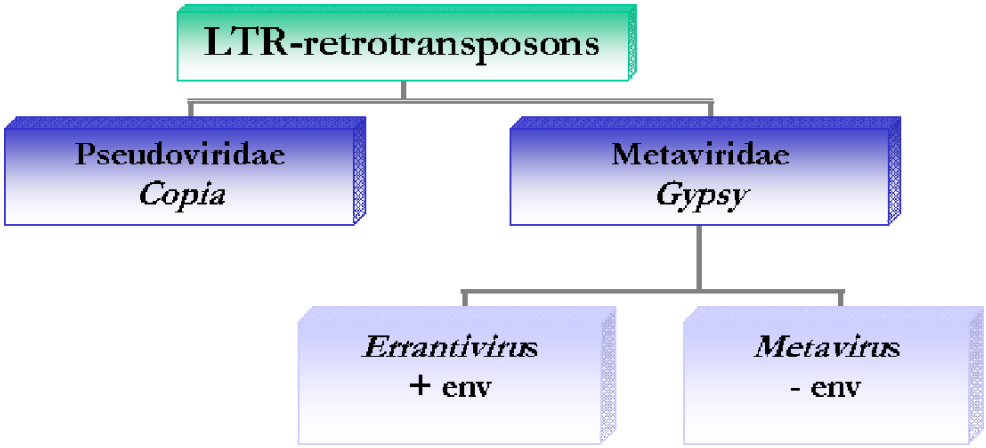

Figure 1. LTR-retrotransposons classification based on the presence of the envelope-like (*env*) gene.

deduced protein sequences of plant retroviruses (Peterson-Burch et al., 2000). This analysis revealed that plant putative *env*-like sequences possess several conserved features: a protease cleavage site, a transmembrane anchor peptide and glycosylation sites. Recently, Vicient et al. 2001 demonstrated that *Bagy*-2 transcripts undergo splicing to generate a subgenomic *env* product as do those of retroviruses. However, no unspliced, full-length transcripts were detected, suggesting the low efficiency of the splicing reaction *in vivo* (Vicient et al., 2001). Clearly, it will be necessary to demonstrate that these putative *env*-like genes encode envelope-like proteins that are capable of transferring retroviral nucleocapsids from cell-to-cell, as shown for the *Drosophila gypsy* retrotransposon (Kim et al., 1994; Song et al., 1994).

EVOLUTION OF PLANT RETROVIRUSES

What is the origin of plant retroviruses? Phylogenetic analyses of the reverse transcriptase sequences of the vertebrate retroviruses strongly suggest that vertebrate retroviruses are derivative of *gypsy*-like retrotransposons (Malik et al., 2000). As *env* genes, the principal difference between retroviruses and retrotransposons, represent antigenic sites that elicit a host immune system, segments of this gene are under strong selective pressure to diverge (Coffin et al., 1997) . Both the antiquity of the original acquisition and the rapid sequence divergence have made it difficult to ascertain the origins of the *env* gene in retroviruses (Malik et al., 2000). Indeed, it is unclear whether vertebrate env genes represent a single acquisition event or multiple events (Eickbush and Malik, 2002).

In this regard, it is intriguing to discover that the some members of the plant *copia*-like retrotransposons possess *env*-like sequences (Laten et al., 1998). Interestingly, there are no reports on the presence of *env*-like sequences in the *copia*-like retrotransposons in *D. melanogaster* or any other invertebrate or vertebrate (Eickbush and Malik, 2002). Therefore, the presence of *env*-like sequences in both *copia* and *gypsy* groups suggests that the *env* gene was acquired by these two groups of retrotransposons independently (Kumar, 1998). Alternatively, closely related relatives retroviral derivatives of *copia* and *gypsy* retrotransposons invaded the genome of plants and subsequently lost their *env* gene (Kumar and Bennetzen, 1999). Currently, it is unknown which process proceeded first. Nevertheless, the existence of plant retroviruses supports the hypothesis for an apparent horizontal transfer of plant viruses in plants (Peterson-Burch et al., 2000; Abdel Ghany and Zaki, 2002).

FUTURE APPLICATIONS OF PLANT RETROVIRUSES

The detection that some of the plant retroviruses are structurally intact and transcriptionaly active (Peterson-Burch et al., 2000; Vicient et al., 2001; Abdel Ghany and Zaki, 2002) promotes the possibility of plant retroviruses that are potent vehicles for interspecies gene flow in plants. In other words, a vector based on plant retroviruses could be an important additional tool for the production of transgenic plants with well-defined, foreign DNA inserts required for biosafety approval and commercialization. Furthermore, the development of plant retroviruses as gene vectors is of great advantage to transfer genes of interest without using currently available transformation methods which are expensive, time consuming, laborious, idiosyncratic, and therefore difficult to automate. Finally, unlike animals, plants do not sequester their germ line and infected somatic plant cells can give rise to floral organs and seeds. However, before they are used as vectors, it is imperative to understand how plant retroviruses naturally contribute to interspecies gene flow, and thus rationally evaluate

recent concerns regarding the use of genetically modified crop species.

ACKNOWLEDGEMENTS

This work was supported by a grant from the US-Egypt Science & Technology Foundation.

REFERENCES

Abdel Ghany AA, Zaki EA (2002). Cloning and sequencing of an *envelope*-like gene in *Gossypium*. Planta 216:351-353.

Baltimore D (1970). RNA -dependent DNA polymerase in virions of RNA tumor viruses. Nature. 226:1209-1211.

Bennetzen JL (2002). Mechanisms and rates of genome expansion and contraction in flowering plants. Genetica. 115:29-36.

Bennetzen JL (2000). Transposable elements contributions to plant gene and genome evolution. Plant Mol. Biol. 42:351-369.

Chavanne F, Zhang DX, Liaud MF, Cerff R (1998). Structure and evolution of T*y*3/*Gypsy* family highly amplified in pea and other legume species. Plant. Mol. Biol. 37:363-.375.

Coffin JM, Hughes SH, Varmus HE (1997). Retroviruses, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Eickbush TH, Furano AV (2002). Fruit flies and humans respond differently to retrotransposons. Curr. Opin. Genet. Dev. 12:669-674.

Eickbush TH, Malik HS (2002). Origins and evolution of retrotransposons. In: Craig et al. (eds) Mobile DNA II. ASM Press, USA, pp 1111-1144.

Hull R, Will H (1989). Molecular biology of viral and nonviral retroelements. Trends Genet. 5:357-359.

Hull R (2001). Classifying reverse transcribing elements: a proposal and a challenge to the ICTV. Arch. Virol. 146:2255-2261.

Kim A, Terzian C, Santamaria P, Pelisson A, Purd'Homme N, Bucheton A (1994). Retroviruses in invertebrate: the *gypsy* retrotransposon is apparently an infectious retrovirus of *Drosophila melanogaster*. Proc. Natl. Acad. Sci. USA. 91:1285-1289.

Kumar A, Bennetzen JL (1999). Plant retrotransposons. Annu. Rev. Genet. 33:479-532.

Kumar A (1998). The evolution of plant retroviruses: moving to green pastures. Trends Plant. Sci. 3:371-374.

Laten HM, Majumdar A, Gaucher EA (1998). *SIRE*-1, a *copia*/T*y*1-like retroelement from soybean, encodes a retroviral envelope-like protein. Proc. Natl. Acad. Sci. USA. 95:6897-6902.

Malik HS, Henikoff S, Eickbush TH (2000). Poised for contagion: evolutionary origins of the infectious abilities of invertebrate retroviruses. Genome Res. 10:1307-1318.

Peterson-Burch BD, Wright DA, Laten HM, Voytas DF (2000). Retroviruses in plants? Trends Genet. 16:151-152.

Song SU, Gerasimova M, Kurkulos M, Boeke JD, Corces VC (1994). An *env*-like protein encoded by a *Drosophila* retroelement: evidence that *gypsy* is an infectious retrovirus. Genes Dev. 8:2046-2057.

Temin HM, Mizutani S (1970). RNA-directed DNA polymerase in virions of Rous sarcoma virus. Nature. 226:1211-1213.

Temin HM (1989). Retrons in bacteria. Nature. 339:254-255.

Vicient CM, Kalendar R, Schulman AH (2001). Envelope-class retrovirus-like elements are widespread, transcribed and spliced, and insertionally polymorphic in plants. Genome Res. 11:2041-2049.

Wilhelm M, Wilhelm FX (2001). Reverse transcription of retroviruses and LTR retrotransposons. Cell. Mol. Life Sci. 58:1246-1262.

Wright DA, Voytas DF (1998). Potential retroviruses in plants: Tat1 is related to a group of *Arabidopsis thaliana* Ty3/gypsy retrotransposons that encode envelope-like proteins. Genetics. 149:703-715.