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.

Accepted 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).

Abbreviations; INT: integrase, LTR: long terminal repeat, RT: reverse transcriptase.
Table 1. Retroelements classification.

<table>
<thead>
<tr>
<th>Viral retroelements</th>
<th>II. Pararetroviruses</th>
</tr>
</thead>
<tbody>
<tr>
<td>I. Retroviruses</td>
<td>(DNA in virions)</td>
</tr>
<tr>
<td>(RNA in virions)</td>
<td></td>
</tr>
<tr>
<td>Eukaryotic non-viral retroelements</td>
<td></td>
</tr>
<tr>
<td>I. Retrotransposons</td>
<td>II. Retroposon</td>
</tr>
<tr>
<td>LTR +</td>
<td>–</td>
</tr>
<tr>
<td>RT +</td>
<td>+</td>
</tr>
<tr>
<td>INT +</td>
<td>+</td>
</tr>
<tr>
<td>Bacterial retroelements</td>
<td></td>
</tr>
<tr>
<td>Retron</td>
<td></td>
</tr>
<tr>
<td>LTR –</td>
<td></td>
</tr>
<tr>
<td>RT +</td>
<td></td>
</tr>
<tr>
<td>INT –</td>
<td></td>
</tr>
</tbody>
</table>


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.

<table>
<thead>
<tr>
<th>Retrovirus</th>
<th>Plant</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>SIRE-1</td>
<td>Soybean</td>
<td>Laten et al., 1998</td>
</tr>
<tr>
<td>Tat1</td>
<td>A. thaliana</td>
<td>Wright and Voytas, 1998</td>
</tr>
<tr>
<td>Athila4</td>
<td>A. thaliana</td>
<td>Wright and Voytas, 1998</td>
</tr>
<tr>
<td>Cyclops</td>
<td>Pea</td>
<td>Chavanne et al., 1998</td>
</tr>
<tr>
<td>Bagy-2</td>
<td>Barley</td>
<td>Vicient et al., 2001</td>
</tr>
<tr>
<td>GM-5 and GM-6</td>
<td>Gossypium</td>
<td>Abdel Ghany and Zaki, 2002</td>
</tr>
</tbody>
</table>

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