The *Journal of Cotton Science 8:179–185 (2004)*
http://journal.cotton.org, © The Cotton Foundation 2004

MOLECULAR BIOLOGY AND PHYSIOLOGY

Ty3/*gypsy* Retro-Transposons in Egyptian cotton (*G. barbadense)*

Essam A. Zaki* and A. A. Abdel Ghany

ABSTRACT

The objective of this research was to isolate and characterize Ty3/*gypsy* sequences in cultivated allotetraploid cotton, *Gossypium barbadense* L. Using a degenerate oligonucleotide designed to amplify these sequences by PCR, part of the reverse transcriptase gene of Ty3/*gypsy* retrotransposons was isolated, cloned, and sequenced. These retro-elements exist in the cotton genome in high copy number, and their encoded transcripts were detected. Phylogenetic analysis shows that these retro-elements are heterogeneous and closely related to elements present in other plant species. These results indicate that Ty3/*gypsy* retro-transposons are standard component of the *Gossypium* genome, and support the fact that Ty3/*gypsy* retro-transposons represent a major component of the plant genome.

Retro-elements and their derivatives are ubiquitous and abundant components of plant genomes (Kumar and Bennetzen, 1999). Based on their structure, retro-transposons are divided into two groups: long terminal repeat (LTR) retro-transposons and non-LTR-retro-transposons (Eickbush and Malik, 2002). LTR-retro-transposons are further divided into two major groups, Ty1/*copia* and Ty3/*gypsy*. The major structural difference between the Ty1/*copia* and Ty3/*gypsy* groups is the order of the reverse transcriptase (RT) and integrase (IN) domains in their *pol* genes (Wilhelm and Wilhelm, 2001). The RT genes have conserved amino acids domains, some of which are characteristic of each retro-element group (Xiong and Eickbush, 1990). Degenerate oligonucleotide primers have been designed to amplify these domains by PCR and are used for detection and assessment of their distribution and evolution (Flavell et al., 1992; Voytas et al., 1992; Matsuoka and Tsunewaki, 1999; Friesen et al., 2001). The detailed characterization of different plant taxa with respect to the content, variability, and physical distribution of retro-transposons provides a major contribution to our understanding of host genome organization and evolution (Bennetzen, 2000; Bennetzen, 2002).

Plant retro-transposons are largely inactive, which suggests that silencing is a major mechanism of control for these retro-transposons (Casacuberta and Santiago, 2003). The presence of retro-transposons in databases of expressed sequence tags (ESTs) suggests that they can escape silencing (Echenique et al., 2002). Moreover, they can be induced by various stresses, such as tissue culture and exposure to cell-wall hydrolases (Grandbastien et al., 1997). High copy retro-transposons, which have enhancer and transcriptional readout activity, have a genome-wide effect on the expression of flanking genes (Freschotte et al., 2002). One example is the Wis 2-1A retro-transposon in newly synthesized wheat (*Triticum aestivum* L.) amphiploids (Kashkush et al., 2002a). Both LTRs in Wis 2-1A drive the readout synthesis of new transcripts from adjacent sequences, including the antisense or sense strands (Kashkush et al., 2002b). These data, together with their abundance and ability to be activated, support the theory that retro-transposons are potential controlling elements (Casacuberta and Santiago, 2003).

The genus *Gossypium* is a facile system for investigating the genomic organization and evolution of repetitive DNA sequences that become united in a common nucleus (Wendel, 1989). Ty1/*copia* retro-transposons were identified in allotetraploid cottons, *Gosspyium hirsutum* L. (Vanderwiel et al., 1993), and *G. barbadense* (Abdel Ghany and Zaki, 2003). Fluorescent *in-situ* hybridization was used to study their chromosomal distribution (Hanson et al., 1999). Ty3/*gypsy* elements were recently reported in the upland cottons, *G. hirsutum* (Taliercio and Ulloa, 2003). In this study, the presence of Ty3/*gypsy* elements in Egyptian cotton, *G. barbadense,* was investigated. Partial sequences of the RT gene of Ty3/*gypsy* retro-transposons from the genomes of two of the most widely cultivated *G. barbadense* cottons in Egypt, which are particularly prized for their superior fiber quality, were amplified.

E. A. Zaki, Genetic Engineering and Biotechnology Institute (GEBRI), Post Code 21934, Alexandria, Egypt; A. A. Abdel Ghany, Institute of Efficient Productivity, Zagazig University, El Zagazig, Egypt
*Corresponding author: gossypium@link.net

MATERIALS AND METHODS

Plant growth. Seedlings from *G. barbadense* cultivars, Giza 45 and 84, were grown in the greenhouse at the Genetics Engineering and Biotechnology Research Institute in Alexandria, Egypt. DNA and RNA were extracted from 3-day-old seedlings.

DNA and RNA extraction. Total DNA was extracted using a Qiagen DNeasy kit (Qiagen; Valencia, CA). Total RNA was extracted using a Qiagen RNeasy kit. Total RNA was treated with RNase-free DNaseI (Boehringer; St. Joseph, MO) (2U for 30 min at 37°C), stopped with 0.01 mM Tris/EDTA pH 8.0. RNA was purified through two ethanol precipitation steps, and the pellet suspended in diethyl pyrocarbonate-treated water.

Isolation of the T*y*3/*gypsy* RT gene in *Gossypium*. Total DNA was subject to PCR with two primers specific to the T*y*3/*gypsy* RT gene:

1) 5`-CTAAGCTTTAYCAYCARHTNMGNAT-3`

2) 5`- TCGAATTCTGNCCNARNMDYTYNAC-3`,

where N = A+C+G+T, M = A+C, H = A+C+T, R = A+G, and Y = T+C (Xiong and Eickbush, 1990). DNA amplifications were conducted in an ABI GeneAmp PCR system 9700 cycler (Applied Biosystems; Foster City, CA) with a denaturing step at 95°C for 5 min and with the step cycle program set for 35 cycles (with a cycle consisting of denaturing 94°C for 30 s, annealing at 50°C for 30 s, and an extension step at 72°C for 30 s), followed by a final extension step at 72°C for 10 min. Fragments were cloned in pCR 4 plasmid (Invitrogen; Carlsbad, CA) and plasmid DNAs were sequenced in both directions using BigDye sequencing kit with an ABI 377 DNA sequencer (Applied Biosystems; Foster City, CA). DNA sequences were deposited in the GenBank database with the accession numbers U75247 and U75248.

Reverse transcriptase-PCR. RT-PCR for the detection of the expressed RT domain was conducted on cDNA prepared with the Qiagen OneStep RT-PCR kit (Qiagen; Valencia, CA) according to the manufacturer's instructions using 1µg total RNA extracted from 3-day-old cotton seedlings. The cycle program consisted of 35 cycles of 30 s at 94°C, 30 s at 50°C, and 30 s at 72°C, followed by a final extension step at 72°C for 10 min. Controls for DNA contamination consisted of treating RNA with RNase-free DNaseI, and reactions lacking the reverse transcriptase enzyme in the reverse transcription step.

Alignments and phylogenetic analysis. Pairwise and multiple DNA sequence alignment were carried out using CLUSTALW 1.82 (Thompson et al., 1994). A bootstrap neighbor-joining tree was generated using MEGA 2.1 (Kumar et al., 2001) from CLUSTALW alignments.

DNA & RNA hybridization. PCR-amplified probes were labeled with [α-^{32}P] dCTP using the random primer method (Feinberg and Vogelstein, 1983), and used for DNA, and RNA hybridization as described (Sambrook et al., 1989). Filters were hybridized overnight at 42°C in a solution containing [50% formamide, 5 x sodium saline citrate (SSC), 10 x Denhardt's, and 0.5% sodium dodecyl sulfate (SDS)]. Hybridization wash was conducted at 50°C in 0.1 x SCC containing 0.5% SDS for 1 h. Labeled filters were autoradiographed for 1 to 2 d at room temperature.

RESULTS

Isolation of the T*y*3/*gypsy* RT gene in *G. barbadense*. The search for the T*y*3/*gypsy* RT gene in the *G. barbadense* cultivars, Giza 45 and Giza 84, was conducted using a specific oligonucleotide for the T*y*3/*gypsy* reverse transcriptase gene (Xiong and Eickbush, 1990). A 310 bp fragment was detected in each cultivar, which was designated as G45 and G84 for Giza 45 and Giza 84, respectively. Amino acid identities between G45 and G84, and *gypsy* from Drosophila (Yuki et al., 1986), T*y*3/*gypsy* RT fragments from lilies (*del)* (Smyth et al., 1989), and IFG7 from pine (Kossack and Kinlaw, 1999) range from 56% to 70% (Table 1 and Fig. 1). The high amino acid identities observed among these RT fragments support the interpretation that G45 and G84 sequences generated in this study represent portions of the RT gene of T*y*3/*gypsy* retro-transposons.

Table 1. Amino acid identities between G45 and G84, and the corresponding portions of T*y*3/*gypsy* retro-transposons from *Drosophila*, yeast, *Lilium*, and *Pinus*

Source [z]	Retro-transposon	Amino acid identities	
		G45	G84
Drosophila	Gypsy	57	56
Yeast	Ty3	43	49
Lilium	Del	70	53
Pinus	IFG7	57	47

[z] ***Drosophila*** **from Yuki et al., 1986; yeast from Hansen et al., 1988;** ***Lilium*** **from Smyth et al., 1989; and** ***Pinus*** **from Kossack and Kinlaw, 1999.**

```
          *  *   * * *    **    ******    ** ***   *     * **
G45    RDSDVPKTAFRTRYGHYEFLVMPFGLTNAPAVFMDLMNRIFRQYLDRFVVVF
G84    REGDEWKIAFKTKHSLYEWLVMPFGLTNTSSTFMRLMNHVLRAFIGKFCVVY
Del    RDSDIPKTAFRTRYGHYEFLVMPFGLTNVPTAFMNLMNRVFREYLDKFIVVF
Gypsy RDSDHDRTSFSTNGGKYEFCRLPFGLRNASSIFMRLMNDVLREQLGKFCVVY

          **   **     *  ***  *  **   ***   ** *    * ***  *
G45    IDDILVYSGDETEHAEHLRLVLQILRDKQLYAKFSKCEFWLREVSFLGHVV
G84    FDDILVYSRSLDDHLKHLRAVLDVLRKENLYANLKKCTFCSNQVVFLGFVV
Del    VDDVLIYSRTQKDHEHHLRISLQLLRNNQLYAKLSKCEFWMEKVKFLGHVV
Gypsy VDDVIIYSENESDHVRHLRTVLKCLRDANLYASQEKCRFFKESVEYLGFIV
```

Figure 1. Alignment of the reverse transcriptase domain of G45, G84, Lilium *del,* and Drosophila *gypsy* retrotransposons (Smyth et al., 1989; Yuki et al., 1986). Identical amino acids are identified by asterisks.

```
G45 RDSDVPKTAFRTRYGHYEFLVMPFGLTNAPAVFMDLMNRIFRQYLDRFVVVFIDDILVYS 60
G84 REGDEWKIAFKTKHSLYEWLVMPFGLTNTSSTFMRLMNHVLRAFIGKFCVVYFDDILVYS 60
    *:.*  * **:*::. **:*********:.:.** ***:::* ::.:* **::*******

G45 GDETEHAEHLRLVLQILRDKQLYAKFSKCEFWLREVSFLGHVV 103
G84 RSLDDHLKHLRAVLDVLRKENLYANLKKCTFCSNQVVFLGFVV 103
     .  :* :*** **::**.::***::.** *  .:* ***.**
```

Figure 2. Comparative amino acids sequence analysis of G45 & G84 using CLUSTALW. Residues or nucleotides in that are identical in all sequences in the alignment are designated by '*'; conserved substitutions are designated by ':'; and semi-conserved substitutions are designated by '.'.

The identities of comparative nucleotide and amino acid sequence analysis of G45 and G84 using ClustalW program were 54.1% and 51.4%, respectively (Fig. 2). Gaps were introduced in the nucleotide sequence analysis at positions 1 and 117 for G45 and at 150 and 310 for G84, but they did not affect the inferred amino acids at these corresponding positions or the coding information. The observed sequence heterogeneity between G45 and G84 could be due to the inherent greater heterogeneity of the T*y*3/*gypsy* retro-transposons (Malik et al., 2000) compared with T*y*1/*copia* retro-transposons (Peterson-Burch and Voytas, 2002). The fact that the DNA sequence was determined for only two clones, and taking into consideration that it is most likely that an extremely high copy number of sequences related to G45 and G84 exist in *G. barbadense* (see below), indicates that G45 and G84 represent only part of the diversity of the RT genes in *G. barbadense*.

Phylogenetic analysis of G45 and G84. Elements of G45 and G84 were compared and aligned with other RT genes of plant T*y*3/*gypsy* retro-transposons and with T*y*3 as the outgroup to study the evolutionary relationships of the identified retro-elements in *G. barbadense* (Fig. 3). The neighbour-joining phylogram provided strong bootstrap support for a monophyletic origin of plant T*y*3/*gypsy* retro-transposons, yet showed high diversity within all species. G45 had the strongest affinity with *G. hirsutum* G1 and *Ananas comosus* (L.) Merr. *Dea*1

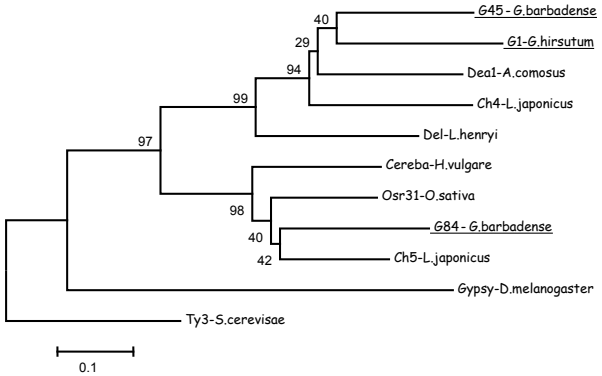

Figure 3. Phylogenetic tree showing relationship between reverse transcriptase nucleotide sequences of *G. barbadense* G45 and G84 and plant T*y*3/*gypsy* retro-transposons. *Gossypium* sequences are underlined. The neighbor-joining method (Saitou and Nei, 1987) was used to construct the tree. The numbers on the branches represent bootstrap support for 1,000 replicates. Names refer to the accession number of the nucleotide sequences that encode the corresponding reverse transcriptase genes. Sequences used and their GenBank accession numbers include: AY181254 *G. hirsutum* G1 T*y*3/*gypsy* LTR-retrotransposon (Taliercio and Ulloa, 2003); ACY12432 *A. comosus Dea*1 T*y*3/*gypsy* LTR-retrotransposon (Thomson et al., 1998); AP004470 *L. japonicus* genomic DNA Chromosome 4 27319-27627 (Sato et al., 2001); X13886 *L. henryi del* T*y*3/*gypsy* LTR-retrotransposon (Smyth et al., 1989); X03734 *D. melanogaster gypsy* LTR-retrotransposon (Yuki et al., 1986); AC005561 *O. sativa osr31/rire7* T*y*3/*gypsy* LTR-retrotransposon (McCarthy et al., 2002); AP004525 *L. japonicus* genomic DNA Chromosome 5 65809-66114 (Sato et al., 2001); AY040832 *H. vulgare cereba* T*y*3/gypsy retrotransposon (Hudakova et al., 2001), and M34549 yeast T*y*3 retrotransposon (Hansen et al., 1988).

Ty3/*gypsy* retro-transposons (Taliercio and Ulloa, 2003; Thomson et al., 1998) with 81% and 74% amino acids identity, respectively. On the other hand, G84 has the strongest affinity with *Lotus japonicus* L. genomic DNA chromosome 5, *Hordeum vulgare* L. *cereba,* and *Oryza sativa* L. *osr31/rire7* Ty3/*gypsy* retro-transposons (Sato et al., 2001; Hudakova et al., 2001; McCarthy et al., 2002) with 75% amino acid identity. It is noteworthy that G45 and G84 did not show close correspondence to previously reported Ty1/*copia* retro-transposons in *G. barbadense* and *G. hirsutum* (Abdel Ghany and Zaki, 2003; Vanderwiel et al., 1993), nor to the major dispersed repeats in the cotton genome (Zhao et al., 1998), which suggests that a very complex population of such elements exists in cotton.

Repetitive nature of Ty3/*gypsy* elements in *G. barbadense*. The detection of Ty3/*gypsy* retro-elements in *G. barbadense* prompted an investigation into the molecular nature of these elements. Total genomic DNA from Giza 45 and 84 were digested with *Hin*dIII and hybridized with their respective related RT domain probes (Fig. 4). The detection of multiple hybridizing bands is indicative of the repetitive nature of these retro-elements in the cotton genome. Moreover, the observation of different band intensities within a single track may be attributed to sequence and/or copy number variation between Ty3/*gypsy* retro-transposon-related sequences in these genomes. The presence of sequences ranging from larger than 23 kb to smaller than 2 kb indicates the presence of a large family of related sequences. The fact that these genomes display a similar set of Ty3/*gypsy* retro-transposon-specific restriction fragments suggests that the genomic locations of these retro-elements are similar within these two closely related genomes. These results indicate that G45 and G84-related sequences are present in relatively high copy number in *G. barbadense*.

Transcription of G45 and G84 in cotton. To determine whether G45 and G84 are transcribed, Northern blot hybridization using ^{32}P-labeled PCR amplified probes were performed (Fig. 5). A transcript of approximately 7 kb in molecular weight was detected in Giza 45 and Giza 84, respectively. Control reactions consisted of bacterial RNA with no detectable products (data not shown), and the use of DNaseI for genomic DNA contamination. The transcriptional activity of G45 and G84 was further investigated by RT-PCR using the same primers for the RT domain (Fig. 6). Product amplification reac-

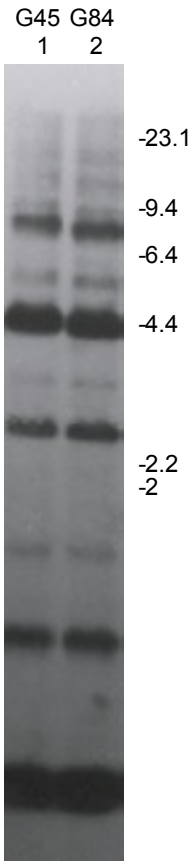

Fig. 4 Molecular distribution of G45- and G84-related sequences from two *G. barbadense* cultivars, Giza 45 and 84. DNA (10 µg) of each cultivar were digested with *Hin*dIII and hybridized with their respective related G45 and G84 ^{32}P-labeled specific probes. Sizes of molecular weight markers in kb are indicated on the right. Lane 1. Cotton cultivar Giza 45 hybridized with G45 probe. Lane 2. Cotton cultivar Giza 84 hybridized with G84 probe.

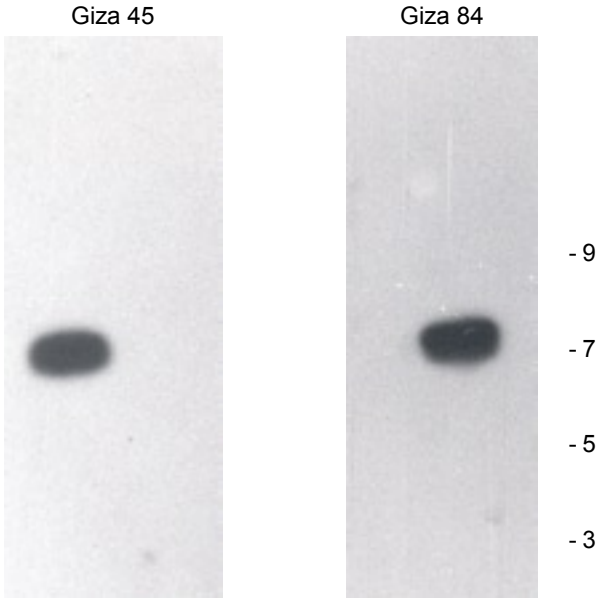

Fig. 5 Northern blot hybridization for the detection of G45- and G84-encoded transcript. Total RNA (10 µg) extracted from young cotton seedlings from Giza 45 and 84, were treated with DNaseI for genomic DNA contamination and hybridized with their respective related G45 and G84 probes. Bacterial RNA was included as a negative control (not shown). Sizes of molecular weight markers in kb are indicated on the right.

tions were identical in length to those from genomic DNA. Control reactions that lacked the addition of the reverse transcriptase enzyme in the reverse transcription step failed to produce products, thus confirming the transcriptional origin of the amplified products. Two amplified cDNAs from G45 and G84 were cloned and sequenced to verify their identity. The predicted proteins from these cDNAs were 84% and 87% to their respective DNA clones (Abdel Ghany and Zaki, unpublished, 2004).

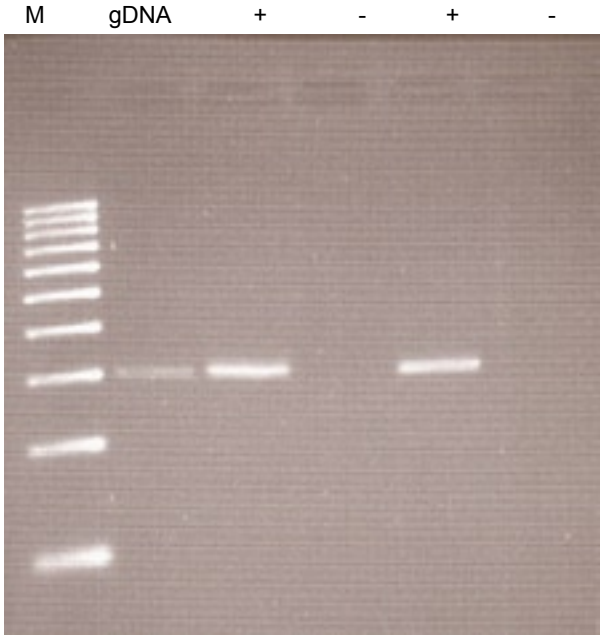

Fig 6. RT-PCR for the detection of the expressed RT domain. Total RNA (1 µg) extracted from young cotton seedlings from Giza 45 and 84, was treated with DNase I prior to the reverse transcription step. Amplification products were then electrophoresed in 2% agarose gels and visualized by ultraviolet illumination after staining with ethidium bromide. As a standard, PCR with 10ng of genomic DNA (gDNA) is shown. The reactions for lanes labeled '+' contained the reverse transcriptase enzyme, and those labeled '–' do not contain the reverse transcriptase enzyme during the reverse transcription step as controls for genomic DNA contamination. M is the 100 bp molecular weight ladder.

DISCUSSION

Ty3/*gypsy* retro-transposons in *G. barbadense*. Degenerate primers used to amplify the RT domain of Ty3/*gypsy* retro-transposons in plants demonstrated their universal nature (Levin, 2002). Using this approach, partial Ty3/*gypsy* RT sequences in two *G. barbadense* cultivars were identified. Comparative nucleotide and amino acid sequences analysis revealed that *G. barbadense* retro-elements, G45 and G84, are heterogeneous and exist in high copy number. The detection of Ty3/*gypsy* retro-transposons in *G. barbadense* and *G. hirsutum* (Taliercio and Ulloa, 2003) supports the fact that Ty3/*gypsy* retro-transposons represent a major component of plant genomes (Freschotte et al., 2002).

Transcriptional activity of G45 and G84. Plant retro-transposons are known to be transcriptionally silent in most plant tissues during development, which suggests transcriptional control is a major mechanism of control for their retro-transposition (Kumar and Bennetzen, 1999). Their expression and transposition are inducible by stresses, such as protoplast isolation and tissue culture (Grandbastien et al., 1997). Detection of their transcripts under ordinary growth conditions has also been reported (Suoniemi et al., 1996; Pearce et al., 1997). In this regard, G45- and G84-encoded transcripts were detected in their related respective young seedlings using RNA slot-blot hybridization. In addition, Northern blot analysis revealed the presence of a transcript of approximately 7 kb. These results suggest that G45 and G84 are transcriptionally active retro-transposons; however, the presence of stop codons in G45- and G84-derived amino acid sequences suggests that these clones represent defective retro-transposons. In addition, a number of the retro-transposons in Egyptian cotton obviously are not functional and must be evolving as pseudogens. Nevertheless, the detection of G45- and G84-encoded transcripts and intermediates in the retro-transposition process (Kumar and Bennetzen, 1999) suggests that a subset of these molecules is competent for retro-transposition.

It is noteworthy that Zhao et al. (1998) reported a highly repetitive DNA clone p067 in cotton, which corresponds to the *Lilium henryi* Baker integrase domain, that is expressed in young cotton seedlings. It is unknown whether G45, or G84, and pX067 represent the same retro-transposon. Further experimental analysis is required to address this question. In addition, we are seeking to determine if the molecular nature of the *G. barbadense* transcript originates from LTRs from transcriptionally active elements, or if it initiates from conventional cellular promoters or adjacent full-length retro-transposons (Kumar and Bennetzen, 1999).

Origin and evolutionary relationships of Ty3/*gypsy* RT of *G. barbadense*. The phylogenetic analysis of the RT domain provides the evolutionary relationships among Ty3/*gypsy* retro-transposons

to be inferred (Malik et al., 2000, Eickbush and Malik, 2002). Our phylogenetic analysis revealed that the closest relatives of G45 and G84 are found in other T*y*3/*gypsy* RT of plants (e.g. *A. comosus* and *H. vulgare*) than to each other. This evolutionary relationship suggest either an ancient origin of plant retro-transposons (vertical transmission), or horizontal transmission in which these retro-transposons have jumped the species-gap (Eickbush and Malik, 2002). Although horizontal transfer has been demonstrated between *Drosophila* species (Jordan et al., 1999, Terzian et al., 2000), data are inconclusive for plant species (Friesen et al., 2001, Stuart-Rogers and Flavell, 2001). The observation that branch lengths separating plant retro-transposons are usually similar indicates a similar evolutionary distance, argues against the horizontal transmission hypothesis, and supports the existence of a diverse group of retro-transposon families in the progenitor of plants. A more comprehensive survey of T*y*3/*gypsy* retro-transposons in *Gossypium* species is required to further clarify their evolutionary relationships.

ACKNOWLEDEGMENT

This research supported by US-Egypt Science & Technology Foundation grant #BIO4-001-004.

REFERENCES

Abdel Ghany, A.A., and E.A. Zaki. 2003. Isolation, characterization, and phylogenetic analysis of T*y*1/*copia*-like retrotransposons in the Egyptian cotton *G. barbadense* and its progenitors. Afr. J. Biotech. 2:165-168.

Bennetzen, J.L. 2000. Transposable element contributions to plant gene and genome evolution. Plant. Mol. Biol. 42:251-269.

Bennetzen, J.L. 2002. Mechanisms and rates of genome expansion and contraction in flowering plants. Genetica 115:29-36.

Casacuberta, J.M., and N. Santiago. 2003. Plant LTR-retrotransposons and MITEs: control of transposition and impact on the evolution of plant genes and genomes. Gene 311:1-11.

Echenique, V., Stamova, B., Wolters, P., Lazo, G., Carolli, V.I., and J. Dubcovsky. 2002. Frequencies of Ty1-copia and Ty3-gypsy retroelements within the Triticeae EST databases. Theor. Appl. Genet. 104:840-844.

Eickbush, T.H., and H.S. Malik. 2002. Origins and evolution of retrotransposons. p 1111-1144. *In* N.L. Craig et al. (ed.) Mobile DNA II. ASM Press, Washington, D.C.

Flavell, A.J., Dunbar, R., Anderson, R., Pearce, S.R., Hartley, R., and A. Kumar. 1992. T*y*1-*copia* group retrotransposons are ubiquitous and heterogeneous in higher plants. Nucleic Acids Res. 20:3639-3644.

Feinberg, A.P., and B. Vogelstein. 1983. A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. Anal. Biochem. 132:6-13.

Feschotte, C., Jiang, N., and S.R. Wessler. 2002. Plant transposable elements: where genetics meets genomics. Nat. Rev. Genet. 3:329-41.

Friesen, N., Brandes, A., and J. Heslop-Harrison. 2001. Diversity, origin and distribution of retrotransposons in conifers. Mol. Biol. Evol. 18:1176-1188.

Grandbastien, M.A., Lucas, H., Morel, J.B., Mhiri, C., Vernhettes, S., and T.M. Casacuberta. 1997. The expression of the tobacco *Tnt1* retrotransposon is linked to plant defense responses. Genetica 100:241-52.

Hansen, L.J., Chalker, D.L. and Sandmeyer, S.B. 1988. *Ty3*, a yeast retrotransposon associated with tRNA genes, has homology to animal retroviruses. Mol. Cell. Biol. 8: 5245-5256.

Hanson, R.E., Islam-Faridi, N., Crane, C.F., Zwick, M.S., Czeschin, D.G., Wendel, J.F., McKnight, T.D., Price, H.J., and D.M. Stelly. 1999. T*y*1-*copia*-retrotransposons behaviour in a polyploid cotton. Chromosome Res. 8:73-76.

Hudakova, S., Michalek, W., Presting, G.G., Hoopen, R., Santos, K., Jasencakova, Z., and I. Schubert. 2001. Sequence organization of barley centromeres. Nucleic Acids Res. 29:5029-5035.

Jordan, I.K., Matyunina, L.V., and J.F. McDonald. 1999. Evidence for the recent horizontal transfer of long terminal repeat retrotransposon. Proc. Natl. Acad. Sci. USA 96:12621-12625.

Kashkush, K., Feidman, M., and A.A. Levy. 2002a. Gene loss, silencing and activation in a newly synthesized wheat allotetraploid. Genetics 160:1651-1659.

Kashkush, K., Feidman, M., and A.A. Levy. 2002b. Transcriptional activation of retrotransposons alters the expression of adjacent genes in wheat. Nat. Genet. 10:1-5.

Kossack, D.S., and C.S. Kinlaw. 1999. *IFG*7, a gypsy-like retrotransposon in *Pinus* (Pinaceae) has extensive history in pines. Plant Mol. Biol. 3:417-426.

Kumar, A., and J.L. Bennetzen. 1999. Plant retrotransposons. Annu. Rev. Genet. 33:479-532.

Kumar, S.K., Tamura, K., Jakobsen, I.B., and M. Nei. 2001. MEGA2: molecular evolutionary genetics analysis software. Bioinformatics 17:1244-1245.

Levin, H.L. 2002. Newly identified retrotransposons of the T*y*3/*gypsy* class in fungi, vertebrates, and plants. p 684-704. *In* N.L. Craig et al. (ed.) Mobile DNA II. ASM Press, Washington, D.C.

Malik, H.S., Henikoff, S., and T.H. Eickbush. 2000. Poised for contagion: evolutionary origins of the infectious abilities of invertebrate retroviruses. Genome Res. 10:1307-1318.

Matsouka, Y., and K. Tsunewaki K. 1999. Evolutionary dynamics of T*y*1-*copia* group retrotransposons in grass shown by reverse transcriptase domain analysis. Mol. Biol. Evol. 13:1384-1392.

McCarthy, E.M., Liu, J., Lizhi, G., and J.F. McDonald. 2002. Long terminal repeat retrotransposons of *Oryza sativa*. Genome Biol. 3:1-11.

Pearce, S.R., Harrison, G., Heslop-Harrison, P., Flavell, A.J., and A. Kumar. 1997. Characterization and genomic organization of T*y*1-*copia* group retrotransposons in rye. Genome 40:617-625.

Peterson-Burch, B.D., and D.F. Voytas. 2002. Genes of the pseudoviridea. Mol. Biol. Evol. 19:1832-1845.

Sato, S., Kaneko, T., Nakamura, Y., Asamizu, E., Kato, T., and S. Tabata. 2001. Structural analysis of a *Lotus japonicus* genome. I. Sequence features and mapping of fifty-six TAC clones which cover the 5.4 Mb regions of the genomes. DNA Res. 8:311-318.

Saitou, N., and M. Nei. 1987. The neighbour-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. 4:406-425.

Sambrook, J., Fritsch, E., and T. Maniatis. 1989. Molecular cloning: A laboratory manual, 2nd ed. Cold Spring Harbour Laboratory Press, Cold Spring Harbour, New York.

Smyth, D.R., Kaltisis, P., Joesph, J., and J.W. Sentry. 1989. Plant retrotransposon from *Lilium henryi* related to T*y*3 of yeast and *gypsy* group of *Drosophila*. Proc. Natl. Acad. Sci. USA 86:5013-5019.

Suoniemi, A.K., Narvanto, A. and A.H. Schulman. 1996. The BARE-1 retrotransposon is transcribed in barely from an LTR promoter active in transient assays. Plant Mol. Biol. 31:295-306.

Stuart-Rogers, C., and A.J. Flavell. 2001. The evolution of T*y*1-*copia* group retrotransposons in gymnosperms. Mol. Biol. Evol. 18:155-163.

Taliercio, E., and M. Ulloa. 2003. The DNA sequence of a gypsy element from *Gossypium hirsutum* L. and characterisation of gypsy elements in three *Gossypium* species. DNA Sequence 14:319-325.

Terzian, C., Ferraz, C., Demaille, J., and A. Bucheton. 2000. Evolution of the *Gypsy* endogenous retrovirus in the *Drosophila melanogaster* subgroup. Mol. Biol. Evol. 17:908-914.

Thompson, J.D., Higgins, D.G., and T.J. Gibson. 1994. CLUSTALW: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 22:4673-6480. (Available on-line at http://www2.ebi.ac.uk/clustalw) (Verified 15 Apr. 2004).

Thompson, K.G., Thomas, J.E., and R.G. Dietzgen. 1998. Retrotransposon-like sequences integrated into the genome of pineapple, *Ananas comosus*. Plant Mol. Biol. 38:461-465.

Vanderwiel, P.S., Voytas, D.F., and J.F. Wendel. 1993. *Copia*-like retrotransposable element evolution in diploid and polyploid cotton (*Gossypium* L.). J. Mol. Evol. 36:429-447.

Voytas, D.F., Cummings, M.P., Konieczny, A., Ausubel, F.M., and S. Rodermel. 1992. *Copia*-like retrotransposons are ubiquitous among plants. Proc. Natl. Acad. Sci. USA 89:7124-7128.

Wendel, J.F. 1989. New world cottons contain old world cytoplasm. Proc. Natl. Acad. Sci. USA 86:4132-4136.

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

Xiong, Y., and T.H. Eickbush. 1990. Origin and evolution of retroelements based upon their reverse transcriptase sequences. EMBO J. 9:3353-3362.

Yuki, S., Ishimaru, S., Inouye, S., and K. Saigo. 1986. Identification of genes for reverse transcriptase-like enzymes in two Drosophila retrotransposons, 412 and gypsy; a rapid detection method of reverse transcriptase genes using YXDD box probes. Nucleic Acids Res. 11:3017-30.

Zhao, X.P., Si, Y., Hanson, R.E., Crane, C.F., Price, J.H., Stelly, D.M., Wendel, J.F., and A.H. Paterson. 1998. Dispersed repetitive DNA spread to new genomes since polyploid formation in cotton. Genome Res. 8:479-492.