

Full Length Research Paper

LTR-retrotransposons-based molecular markers in cultivated Egyptian cottons *G. barbadense* L.

Elsayed E. Hafez^{1Φ}, Abdel Ghany A. Abdel Ghany^{2Φ} and Essam A. Zaki^{1*}

¹Nucleic Acids Research Department, Genetic Engineering and Biotechnology Research Institute, GEBRI, Research Area, Borg El Arab, Post Code 21934, Alexandria.

²Institute of Efficient Productivity, Zagazig University, El Zagazig, Egypt.

Accepted 5 May, 2006

Long terminal repeat (LTR)-retrotransposons are mobile genetic elements that are ubiquitous in plants and constitute a major portion of their nuclear genomes. LTR-retrotransposons possess unique properties that make them appropriate for investigating relationships between closely related species and populations. The aim of the current study was to employ *Ty1-copia* group retrotransposons as molecular markers in cultivated Egyptian cottons, *G. barbadense* L. Restriction site analysis of PCR-amplified *Ty1-copia* RT domain promoted the construction of a restriction map for each Egyptian cultivar. These maps display distinctive patterns of restriction site variation. Furthermore, these patterns are capable of differentiating even between cultivars that appear to have diverged only in the past 50 years. These results demonstrate that retrotransposon-based molecular markers are particularly valuable tools for plant molecular phylogenetic and population genetic studies.

Key words: DNA fingerprinting, *Gossypium*, repetitive DNA, restriction site polymorphisms, sequence diversity, *Ty1-copia* group retrotransposons.

INTRODUCTION

Long terminal repeat (LTR)-retrotransposons are mobile genetic elements that are ubiquitous in plants and constitute a major portion of their nuclear genomes (reviewed in Kumar and Bennetzen, 1999). They are distributed as interspersed repetitive sequences throughout the length of all host chromosomes (Casacuberta and Santiago, 2003). LTR-retrotransposons' high copy number and genome-wide distribution is the result of their inherent replicative mode of transposition within the host genome (Vitte and

Panaud, 2005). They transpose via an RNA intermediate that is reverse transcribed into extrachromosomal DNA and inserted into the genome by the encoded reverse transcriptase (RT), RNaseH and integrase (IN) enzymes (Freschotte et al., 2002)

Molecular marker technology is playing a vital role in plant biology including: DNA fingerprinting, genetic linkage mapping, phylogenetic relationships and molecular breeding (Gebhardt et al., 2005). Several DNA-based markers have been developed to detect polymorphisms by studying subsets of the total amount of DNA in a genome (Varshney et al., 2005). In this context, LTR-retrotransposons possess unique properties that make them appropriate for investigating relationships between closely related species and populations (Kumar and Hirochika, 2001). LTR-retrotransposons appear to evolve at significantly higher rates than conventional nuclear loci (Purugganan and Wessler, 1995). Furthermore, they are: ubiquitous, present in high copy numbers, highly heterogeneous, and show insertional polymorphisms both within and between plant species

*Corresponding author. E-mail: gossypium@link.net. Phone: (+203) 459-3413. Fax: (+203) 459-3423.

ΦThese authors contributed equally to this work.

Abbreviations: IN, integrase; LTR, long terminal repeat; PCR, polymerase chain reaction; RT, reverse transcriptase gene.

Table 1. *G. barbadense* L. cultivars used in this study.

Cultivar	Key
Ashmouni	Ash
Bahtim163	Bah163
Bahtim185	Bah185
Giza31	G31
Giza36	G36
Giza70	G70
Giza71	G71
Giza75	G75
Giza77	G77
Giza83	G83

(Queen et al., 2004).

In this paper, the development of LTR-retrotransposons, Ty1-*cop*ia group as molecular markers in the Egyptian cotton, *Gossypium barbadense* L., is described. We have previously isolated and characterized Ty1-*cop*ia group retrotransposons in *G. barbadense* L. and its progenitors (Abdel Ghany and Zaki, 2003a). They were then further investigated in cultivated Egyptian cottons (Hafez and Zaki, 2005; Zaki, 2005). Egyptian cottons are worldwide renowned for being the highest fiber quality among the world's cottons (Anonymous, 1997). *G. barbadense* L. was brought in Egypt in the late 19th century and this was followed by the Egyptian breeders' efforts to develop several elite *G. barbadense* cultivars, including the Giza cultivars (Anonymous, 1992).

MATERIALS AND METHODS

Plant materials and genomic DNA extraction

G. barbadense cultivars, employed in the current study, are listed in Table 1. Total DNA was extracted using Qiagen DNeasy kit (Qiagen, Germany).

PCR

Total DNA was subject to PCR with specific primers to amplify an approximately 280 bp region of the Ty1-*cop*ia group retrotransposons reverse transcriptase (RT) as described previously (Abdel Ghany and Zaki, 2003a). Briefly, DNA amplifications were carried in an ABI GeneAmp PCR system 9700 cyler with a denaturing step at 95°C for 5 min and the step cycle program set for 45 cycles (with a cycle consisting of denaturing 94°C for 30 s, annealing at 47°C for 1 min and extension step at 72°C for 2 min), followed by a final extension step at 72°C for 10 min.

Restriction site variation analysis

For each cultivar, PCR-amplified fragment was cut separately with each of the following restriction enzymes: *Alu*I, *Alu*II, *Alu*211, *Bfa*I, *Bgl*I, *Cfo*I, *Dpn*I, *Dpn*II, *Hgi*AI, *Hinf*I, *Hind*III, *Hae*III, *Rsa*I, *Sau*3AI,

and *Taq*I. The restricted DNA was then run in 3% NuSieve agarose gel and visualized using ethidium bromide staining.

Restriction maps and phylogenetic analysis

Restriction maps were constructed for the PCR-amplified RT sequence. PAUP was employed to infer the molecular evolutionary relationships between the Egyptian cotton cultivars (Swofford, 1993). Sites were scored as being present, absent or polymorphic as previously described (Purugganan and Wessler, 1995).

RESULTS AND DISCUSSION

Despite the fact that the Egyptian cottons produce the longest, strongest and finest cotton fibers, yet they show significant variation for agriculturally important traits (Abdalla et al., 2001). Thus, an understanding of the genetic relationships of modern cultivars and their progenitors is imperative for further utilization of cotton genetic diversity in the development of superior cultivars (Abdel Ghany and Zaki, 2003b). The aim of the current study, therefore, was to develop LTR-retrotransposons as molecular markers in *G. barbadense* L.

PCR amplification generates the expected 280 bp Ty1-*cop*ia group retrotransposons from all Egyptian cotton cultivars (Abdel Ghany and Zaki, 2003a). This 280 bp region contains the RT domain of Ty1-*cop*ia group retrotransposons. The PCR-amplified band should contain copies from different Ty1-*cop*ia family members residing within each genome (Hafez and Zaki, 2005; Zaki, 2005). Fifteen different restriction enzymes (ten four-cutter, three five-cutter, and two six-cutter enzymes respectively) were employed to analyze the PCR-amplified DNA from different Egyptian cotton cultivars.

The restriction site analysis promoted the construction of a restriction map for each Egyptian cultivar (Figure 1). These maps display distinctive patterns of restriction site variation. Furthermore, these patterns are capable of differentiating even between cultivars that appear to have diverged only in the past 50 years (Giza 70, 71, and 83). According to pedigree records, Ashmouni is the progenitor of the current Egyptian cultivars (Anonymous, 1992). This is substantiated by Ty1-*cop*ia RT restriction maps, which show HindIII restriction site conservation in all cultivars (Figure 1).

Restriction maps of PCR-amplified Ty1-*cop*ia RT domain from different Egyptian cotton cultivars revealed restriction site variations. These variations include: restriction site(s) loss and gain (e.g., *Alu*I in G36 and 70 respectively). The observed restriction site polymorphisms result from the amplification of divergent copies of Ty1-*cop*ia family within a genome (Abdel Ghany and Zaki, 2003a). This high level of polymorphisms was confirmed by sequence analysis of Ty1-*cop*ia RT domains within *G. barbadense* genomes (Hafez and Zaki, 2005; Zaki, 2005). It is noteworthy that the high retrotransposon sequence divergence is also promoted

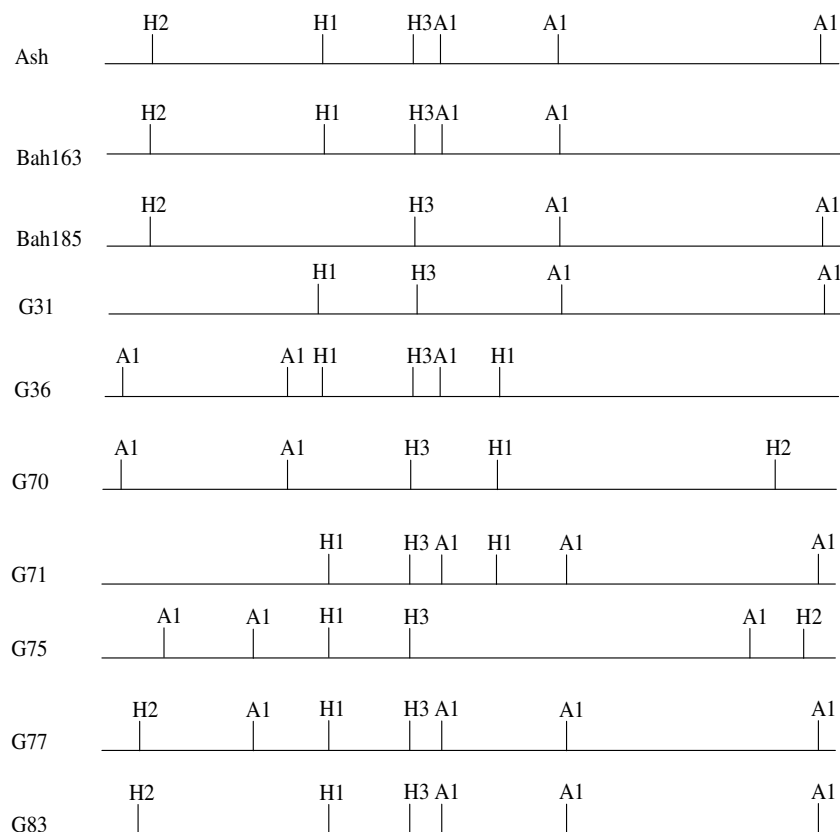


Figure 1. Restriction map of PCR-amplified Ty1-copia RT domain from different Egyptian cotton cultivars. A1: *AluI*; H1: *HgiAI*; H2: *HinfI*, and H3: *HindIII*.

by the RT-mediated retrotransposition process, which is error-prone and results in mutation rates as much as 10^3 times the rate of their counterpart conventional genes (Coffin et al., 1997).

Restriction maps of PCR-amplified Ty1-copia DNA were then employed to deduce the relationships between the different Egyptian cultivars (Figure 2). Sites were scored as being present, absent or polymorphic (Swofford, 1993). The phylogenetic tree, which was inferred using parsimony analysis, clusters all *G. barbadense* L. cultivars together in one group and thus showing their monophyletic origin. Furthermore, it mirrors sequence site variations as evident to the branch lengths which are proportional to the degree of sequence divergence. The inferred tree is concordant with the phylogeny of *Gossypium* as determined from chloroplast DNA restriction site polymorphism data (Wendel, 1989).

Our data indicate that different Ty1-copia group retrotransposons show different levels of polymorphism in the Egyptian cotton cultivars. This is not surprising, as it is known that these elements encompass very heterogeneous families (Hafez and Zaki, 2005; Zaki, 2005). Their encoded transcripts were detected in their

related respective young seedlings using RNA slot-blot hybridization, suggesting their transcriptional activity (Abdel Ghany and Zaki, 2003a). Furthermore, *G. barbadense* L. represents an allopolyploid cotton that appears to have arisen as a consequence of trans-oceanic dispersal of an A-genome taxon to the New World followed by hybridization with an indigenous D-genome diploid (Wendel and Cronn, 2003). As they replicate and insert at multiple sites into the *Gossypium* genome, individual retrotransposons can potentially serve as markers of diversity. Different retrotransposon families, each with its own lineage and structure, have the potential to have been active at distinct phases in the evolution of these cultivars.

The current report demonstrates the utility of a PCR-based method to deduce patterns of restriction site variation between Ty1-copia RT copies within a genome. Ty1-copia RT substantial sequence diversity promotes these sequences to be utilized as molecular markers. LTR-retrotransposons proliferation via transposition, duplication, and interspecific transfer, provides a potentially large pool of useful polymorphic nuclear loci for phylogenetic analysis (Purugganan and Wessler,

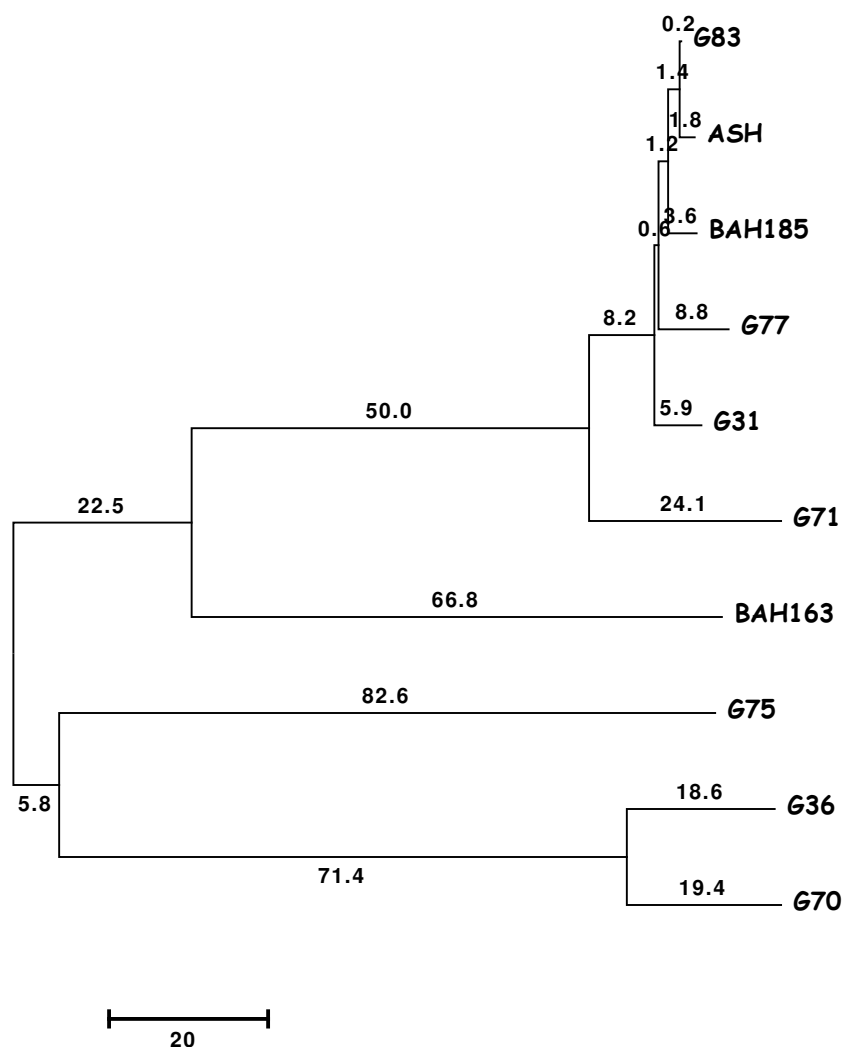


Figure 2. Phylogenetic tree showing evolutionary relationship between Egyptian cotton cultivars. The numbers on the branches represent branch length.

1995). Our results contribute and exemplify the increasingly reports of retrotransposon-based molecular markers in plant genomes (Ellis et al., 1998; Casa et al., 2000; Yu and Wise, 2000; Leigh et al., 2003; Queen et al., 2004). These examples, taken from across the phylogenetic spectrum, illustrate that retrotransposon-based molecular markers are particularly valuable tools for plant molecular phylogenetic and population genetic studies.

LTR-retrotransposons represent a standard component of the *Gossypium* Genome (Zaki and Abdel Ghany, 2003). The analysis of the molecular existence and distribution of ancient and active LTR-retrotransposons, therefore, provides a comprehensive evaluation of the evolutionary history of *Gossypium*. Their differential distributions promote their valuable use for marker-

assisted selection (Abdel Ghany and Zaki, 2003b). It is known that one of major limitations to the application of genomic technology in cotton improvement is the paucity of informative DNA markers. In this context, LTR-retrotransposons map in the *Gossypium* genome is currently being investigated. The availability of such map will indeed create new research opportunities and have many immediate applications.

ACKNOWLEDGEMENTS

This work was supported by a grant from the US-Egypt Science and Technology Foundation to E.A. Zaki. This article is dedicated to the soul of A.M. Zaki.

REFERENCES

- Abdalla AM, Reddy OU, El-Zik KM, Pepper AE (2001). Genetic diversity and relationships of diploid and tetraploid cottons revealed using AFLP. *Theor. Appl. Genet.* 102: 222-229.
- Abdel Ghany AA, Zaki EA (2003a). Isolation, characterization, and phylogenetic analysis of Ty1/*copia*-like retrotransposons in the Egyptian cotton *G. barbadense* and its progenitors. *Afr. J. Biotech.* 2: 165-168.
- Abdel Ghany AA, Zaki, E.A. (2003b). DNA sequences of RAPD fragments in the Egyptian cotton *G. barbadense*. *Afr. J. Biotech.* 2: 129-132.
- Anonymous (1992). Egyptian cotton varieties report. Cotton Research Institute, Agriculture Research Center, Ministry of Agriculture, Giza, Egypt, pp. 2-8.
- Anonymous (1997). In "Agricultural Statistics Board". NASS, USDA. www.usda.gov/nass.
- Casa AM, Brouwer C, Nagel A, Wang L, Zhang Q, Kresovich A, Wessler SR (2000). The MITE family *Heartbreaker (Hbr)*: molecular markers in maize. *Proc. Natl. Acad. Sci. USA.* 97: 10083-10089.
- Casacuberta JM, Santiago N (2003). Plant LTR-retrotransposons and MITEs: control of transposition and impact on the evolution of plant genes and genomes. *Gene* 311: 1-11.
- Coffin JM, Hughes SH, Varmus HE (1997). *Retroviruses*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.
- Ellis THN, Poyser SJ, Knox MR, Vershinin AV, Ambrose MJ (1998). Ty1-*copia* class retrotransposon insertion site polymorphism for linkage and diversity analysis in pea. *Mol. Gen. Genet.* 260: 9-19.
- Feschotte C, Jiang, N, Wessler SR (2002). Plant transposable elements: where genetics meets genomics. *Nat. Rev. Genet.* 3: 329-41.
- Gebhardt C, Schmidt R, Schneider K (2005). Plat genome analysis: the state of the art. *Int. Rev. Cytol.* 247: 223-284.
- Kumar A, Hirochika H (2001). Applications of retrotransposons as genetic tools in plant biology. *Trends Plant Sci.* 6: 127-134.
- Kumar A, Bennetzen JL (1999). Plant retrotransposons. *Annu. Rev. Genet.* 33: 479-532.
- Hafez EE, Zaki EA (2005). Phylogenetic and molecular evolutionary analyses of Ty1-*copia* group retrotransposons in cultivated Egyptian cotton, *Gossypium barbadense* L. *Afr. J. Biotech.* 4: 1275-1280.
- Leigh F, Kalendar R, Lea V, Lee D, Donni P, Schulman AH (2003). Comparison of the utility of barley retrotransposon families for genetic analysis by molecular marker techniques. *Mol. Gen. Genet.* 269: 464-474.
- Purugganan MD, Wessler SR (1995). Transposon signature: species-specific molecular markers that utilize a class of multiple-copy nuclear DAN. *Mol. Ecol.* 4: 265-269.
- Queen RA, Gribbon BM, James C, Jack P, Flavell AJ (2004). Retrotransposon-based molecular markers for linkage and genetic diversity analysis in wheat. *Mol. Gen. Genet.* 271:91-97.
- Swofford D (1993). *PAUP*: Phylogenetic Analysis using Parsimony. Illinois Natural History Survey, Champaign.
- Varshney RK, Graner A, Sorrells ME (2005). Genomic-assisted breeding for crop improvement. *Trends Plant Sci.* 10: 621-630.
- Vitte C, Panaud Q (2005). LTR retrotransposons and flowering plant genome size: emergence of the increase/decrease model. *Cytogenet. Genome Res.* 110: 91-107
- Wendel JF (1989). New world cottons contain old world cytoplasm. *Proc Natl Acad Sci USA.* 86: 4132-4136.
- Wendel JF, Cronn R (2003). Polyploidy and the evolutionary history of cotton. *Advances in Agronomy.* 78: 139-186.
- Yu GX, Wise RP (2000). An anchored AFLP- and retrotransposon-based map of diploid *Avena*. *Genome* 43: 736-749.
- Zaki EA (2005). Ty1-*copia* group retrotransposon families in cultivated cottons *G. barbadense* L. identified by reverse transcriptase domain analysis. *DNA Seq.* 16: 288-294.
- Zaki EA, Abdel Ghany AA (2003). Molecular distribution of *gypsy*-like retrotransposons in cotton *Gossypium* Spp. *Afr. J. Bio.* 2: 124-128.