Isolation and characterization of *Ty1-copia* group of LTRs in genome of three species of *Datura*: *D. innoxia*, *D. stramonium* and *D. metel*

Alka Singh · N. K. Nirala · Alka Narula · Sandip Das · Prem S. Srivastava

Published online: 21 May 2011

© Prof. H.S. Srivastava Foundation for Science and Society 2011

Abstract Retrotransposons (RT) constitute a major fraction of plant genome. They are implicated in evolution and sequence organization. These elements have been proposed to have major role in evolution and variation in genome size. The sequence information of these RT regions in terms of divergence and conservation could be utilized for determining the interrelationship among various copia retrotransposons within the genome. In order to assess the diversity of *Ty1-copia* group of retroelements, reverse transcriptase (RT) sequence was amplified from genomes of three medicinally important *Datura* species: *D. innoxia*, *D. stramonium* and *D. metel* using the primers derived from two conserved domains of RT region. A total of twenty one independent amplicons from RT regions were cloned, sequenced and compared. The intra-family divergence at amino acid level ranged from 4 to 52 %. Though intra-family RT sequences are conserved, no two sequences are identical. Southern blot hybridization suggested that *Ty1-copia*-like retrotransposons are dispersed throughout the *Datura* genome. The results indicate a high degree of heterogeneity among the *Ty1-copia* group of retroelements in *Datura* species.

Keywords *Datura* · Retrotransposon · *Ty1-copia* · Reverse transcriptase

Abbreviations

LTR	Long terminal repeats
RT	Reverse transcriptase
PCR	Polymerase chain reaction
RT-PCR	Reverse transcription-polymerase chain reaction

Introduction

Transposable elements are considered integral constituents of all genomes from prokaryotes to eukaryotes. Of the various transposable elements present in plant genome retrotransposons constitute a large proportion, ranging from 14 in *Arabidopsis* to over 70 % in maize. Depending on the internal placement of RT and endonuclease domains, the LTR-retroelements have been named *Ty1-copia* or *Ty3-gypsy* group of retrotransposons (Xiong and Eickbush 1990). These classes transpose through RNA intermediates which are reverse transcribed in to DNA and inserted in to the genome with the help of encoded reverse transcriptases, RNaseH and integrase. As a source of hypermutagenicity, these elements have been proposed to have major role in evolution and variation in genome size. By comparative sequence analysis of rice and *Arabidopsis*, it has been shown that unequal intrastrand recombination between homologous LTRs retrotransposons terminate to generate solo-LTRs. This illegitimate recombination is associated with the high frequency of loss of genomic DNA by the accumulation of small deletions (Vitte and Bennetzen 2006). Fully sequenced genomes of *Arabidopsis*, rice, black cottonwood and two cultivars of grapevine (*Arabidopsis* genome initiative 2000; Sequencing Project International Rice Genome 2005; Tuskan et al. 2006; Velasco et al. 2007; Cédric et al. 2008) indicate that LTR retrotransposons of *Ty1-copia* and *Ty3-gypsy* are the major components of

A. Singh · N. K. Nirala · A. Narula · S. Das · P. S. Srivastava (✉)
Plant Biotechnology Laboratory, Hamdard University,
New Delhi 110062, India
e-mail: pss410@rediffmail.com

Present Address:

A. Singh · N. K. Nirala
Danforth Plant Science Center,
975 N. Warson Rd.,
St. Louis, MO 63132, USA

intergenic regions. Human genome sequencing has revealed that retrotransposon sequences are located in the coding regions of at least 4 % of the genes and in the promoter region of at least 25 % of the genes. The “C-value-paradox”, i.e. the non-correspondence of the structural complexity in organisms, could largely be explained by the proportion of retroelements in the genome. These elements have also been implicated in genome expansion during evolution (Kalendar et al. 2000; Feschotte et al. 2002).

All the known plant retrotransposons including active ones are largely quiescent during normal development but some get activated transcriptionally as well as transpositionally in response to various biotic and abiotic stresses (Beguiristain et al. 2001). Transcriptional activation of an element indicates its role in stress alleviation of plants (Moreau-Mhiri et al. 1996; Salazar et al. 2007). *Tto1*, a copia-like retrotransposon, active in rice tissues, gets silenced when plants were regenerated (Hirochika et al. 2006). It has also become evident that specific group of retrotransposons, mainly *Ty1-copia* are more likely to be activated by different stresses. Due to their high copy number in highly heterogeneous population and dispersal throughout the genome, retrotransposons are now being utilized as molecular markers in DNA fingerprinting, genetic linkage mapping and phylogenetic analyses (Ellis et al. 1998; Kumar and Hirochika 2001; Moisy et al. 2008; Jia et al. 2009).

Various strategies have been employed to identify retrotransposon sequences in genomes. Prior to the characterization of the complete genome sequence of an organism, full length or partial sequences of *Ty1-copia* like retroelements were identified by a PCR based method of Pearce et al. (1996a, b, Ma et al. 2008). Because *Ty1-copia* retrotransposons make a significant contribution to the size, organization and genetic diversity of genome, it would be worthwhile to characterize such elements in a wide variety of plants. To analyze the heterogeneity among the *copia* family of retroelements in the *Datura* genome, we have isolated reverse transcriptase (RT) sequences by PCR amplification using primers designed from the conserved regions of RT domain and analyzed the heterogeneity in the sequence, phylogenetic relationship, and abundance in the genome.

Material and methods

Plant material and DNA isolation

Three *Datura* species, *D. innoxia*, *D. stramonium*, *D. metel* selected for the present studies were procured from the Herbal Garden at Jamia Hamdard, New Delhi. Genomic DNA for PCR and Southern blotting was extracted from the leaves as described by Doyle and Doyle (1990) with some modifications.

Table 1 Accessions of RT sequences isolated from *Datura innoxia*, *D. metel* and *D. stramonium* deposited in GenBank

Species	Abbreviation used	Accession number
<i>Datura innoxia</i>	Di1	FJ829035
<i>Datura innoxia</i>	Di2	FJ829036
<i>Datura innoxia</i>	Di3	FJ829037
<i>Datura innoxia</i>	Di4	FJ829038
<i>Datura innoxia</i>	Di5	FJ829039
<i>Datura innoxia</i>	Di6	FJ829040
<i>Datura innoxia</i>	Di7	FJ829041
<i>Datura metel</i>	Dm1	FJ829042
<i>Datura metel</i>	Dm2	FJ829043
<i>Datura metel</i>	Dm3	FJ829044
<i>Datura metel</i>	Dm4	FJ829045
<i>Datura metel</i>	Dm5	FJ829046
<i>Datura metel</i>	Dm6	FJ829047
<i>Datura metel</i>	Dm7	FJ829048
<i>Datura stramonium</i>	Ds1	FJ829049
<i>Datura stramonium</i>	Ds2	FJ829050
<i>Datura stramonium</i>	Ds3	FJ829051
<i>Datura stramonium</i>	Ds4	FJ829052
<i>Datura stramonium</i>	Ds5	FJ829053
<i>Datura stramonium</i>	Ds6	FJ829054
<i>Datura stramonium</i>	Ds7	FJ829055

Polymerase chain reaction

Two conserved regions (DVKTAFLL and YVDDMDP) of the reverse transcriptase domain were used to design forward and reverse primers for PCR amplification (Voytas and Ausubel 1988; Flavell et al. 1992). PCR was performed in a 25 µl reaction mixture, with 100 ng of the genomic DNA, 50 µm each of dNTPs, 25 pmol of each primer (5'-GGGATCCAYRTCRTCNACRTANARNA-3' and 5'-ATTCGAYGTNAARCANGCNTTYT-3'), and 1U of

Fig. 1 Amplification of *Ty1-copia* retrotransposon-like sequences from three *Datura* species (Di: *D. innoxia*, Dm: *D. metel*, Ds: *D. stramonium*). Arrow indicates the expected ~280 bp band. Lane M contains 100 bp DNA ladder

Fig. 2 Genomic organization of the *Ty1-copia*-like retroelements in *Datura* species (Di: *D. innoxia*, Dm: *D. metel*, Ds: *D. stramonium*). Leaf DNA digested with *EcoRI* or *HindIII* were hybridized with the clone Di7 probe. Lane M contains *HindIII* digested lambda DNA as a marker

Fig. 3 Alignment of the deduced amino acid sequences corresponding to the reverse transcriptase domains (RT) of the *Ty1-copia* group retrotransposons in the three species of *Datura*. Dashes (-) show gaps. Numerals on the right are the number of amino acid residues in the sequences

Di2	WYKRF-DSFMTDY-GYDRTNSDNCVYFKKFSEGDFFIILLLYVDDMDP---	99
Di7	WYKRF-DSFMTDY-GYDRTNSDNCVYFKKFSEGDFFIILLLYVDDMDP---	95
Di1	WYKRF-DSFMTDH-GYDRTNSDHYVYFKKFSEGDFFIILLLYVGRGQSLS	95
Ds1	WYKRF-DSFIISL-GYNRYSSDTCVCYKRFCGDGDFIILLLFYVDDMDP---	92
Ds2	WYKRF-DSFIISL-GYNRYSSDTCVCYKRFCGDGDFIILLLFYVDDMDP---	92
Di3	WYKRF-DSFIISL-GYNRHSSDPCVYKRFIDGDF-IILLYVDDMDP---	92
Di5	WYKRF-DSFIISL-GYNRHSSYPCVYKRFIDGDF-IILLYVDDVDP---	92
Dm7	WYKRF-DFFIISL-GYNRHSSNPCVYKRFIDGDFIILLLYVDDMDP---	93
Ds6	WYKRF-DFFIISL-GYNRHSSNPCVYKRFIDGDFIILLLYVDDMDPNH-	97
Ds3	WYKRF-DSFIISL-RYNRHSSDPCVYKRFIDGDFIILLLYVDDMDP---	94
Di6	WYKRF-DKFMLGQ-MYTRSKYDHCVYLCKLQDGSFIYLLLYVDDVDP---	93
Ds7	WYKRF-DAFMTTH-KFSRSFSDSCVYHKKMSGNSMIYLLLYVDDVDPNH-	97
Dm1	-YIKF-YNLVSSF-EFTENIVDQCIYLT-ISGNKYIFLVLYVDDVDP---	90
Dm2	-YIKF-YNLVSSF-EFTENIVDQCIYLT-ISGNKYIFLVLYVDDVDP---	90
Dm3	-YIKF-YNLVSSF-EFTENIVDQCIYLT-ISGNKYIFLVLYVDDVDP---	90
Dm5	-YIKF-YNLVSSF-EFTENIVDQCIYLT-ISGNKYIFLVLYVDDVDP---	90
Dm6	WYKRF-ESVMGE--GYKKTSSDNCVYVQKFSDEDFIILLLYVDDMDP---	92
Ds5	WYKRF-ESFMGEQ-GYKKTSSDHCVFVQKFSKDFIILLLYVDDMDPNH-	97
Dm4	WYKRF-DSFIISL-GYNRHSSDPSVSYKRFIDGDFIILLLYVDDMDP---	92
Di4	GMQI--PDYVVK--WFQD--V--MSLH-EYSKSNH-HCLLYVDDMDP---	80
Ds4	LFQPINGSFQLTQHVLIAIFHETL-LYHIYSLIKTSIRNAVLTSSNN----	94

Taq DNA polymerase (Promega) using 0.75 mM MgCl₂. PCR was performed in a thermal cycler (Eppendorf) with the following parameters: 94 °C for 5 min, followed by 35 cycles of 94 °C for 1 min, 47 °C for 1 min and 72 °C for 1.5 min, followed by a final extension step at 72 °C for 10 min. The resultant products were separated by electrophoresis on 1.5 % agarose gels in TBE buffer and visualized under UV light after staining with ethidium bromide.

Cloning

The expected amplified product was eluted from the agarose gel, purified and cloned into pGEM-T Easy vector (Promega) and then transferred to *E.coli* (DH5α strain). The recombinant plasmid was isolated, and the presence of insert was confirmed by digesting it with *EcoRI*. Subsequently, 20 clones were randomly selected (seven for each *D. innoxia*, *D. metel*, *D. stramonium*) and sequenced. DNA sequencing was performed on ABI PRISM 3730 version 3.0 with T7 and SP6 primers, at DNA sequencing facility (UDSC), New Delhi.

Fig. 4 Phylogenetic analysis of the deduced amino acid sequences representing fragments of the RT domain of tree *Datura* genome *Ty1-copia* group retrotransposons using the Neighbour-Joining-method

Sequencing and database search

DNA sequences were translated to reading frames using EBI Translation tool online (<http://www.expasy.ch/tools/dna.html>). Sequence homology queries were performed by submitting DNA sequences to the database searching algorithms at the TIGR Plant Repeat Database (<http://tigrblast.tigr.org/euk-blast/data/blastn/>).

Multiple sequence alignment was done using program CLUSTAL X version 1.83. (<http://www.ebi.ac.uk/clustalx>). Sequence alignments were annotated using Gene Doc (version 2.6.0.2). The bootstrap neighbour-joining tree with 1,000 replications was generated by MEGA version 3.1. Genomic DNA sequences were deposited in the GenBank databases under accession numbers FJ829035-FJ829055 (Table 1).

Southern blotting

Genomic DNA of all the three *Datura* species was digested using 1 µg/unit of restriction enzymes (EcoRI or HindIII) in 1× appropriate buffer at 37 °C for 8 h. Digested genomic DNA was electrophoresed on 0.8 % agarose gel in 1× TAE for about 14 h to facilitate the separation which was then blotted onto Hybond N+ nylon membrane (Amersham). The hybridization was performed under conditions of high stringency using the clone Di7 as a probe and following the methods of Sambrook et al. (1989).

Result

Isolation and characterization of *Ty1-copia* RT sequence

The *Ty1-copia* RT specific degenerate primer set produced an amplicon of ~280 bp (Fig. 1), that was eluted from the gel, purified, cloned in to pGEM-T easy vector and rechecked by restriction digestion. A total of 21 clones (seven for each species) were randomly selected for sequence analysis.

Southern hybridization of digested three species of *Datura* genome with the clone Di7 sequence probe showed

Table 2 Accessions of RT sequences isolated from *Datura innoxia*, *D. metel* and *D. stramonium* deposited in GenBank

Species name	Abbreviation used	Family	Accession number
<i>Lycopersicum esculentum</i>	lyco	<i>Ty1-copia</i>	AF072655
<i>Solanum tuberosum</i>	sol	<i>Ty1-copia</i>	AJ228810
<i>Nicotiana tabacum</i>	nic	<i>Ty1-copia</i>	D83003
<i>Helianthus annus</i>	hel	<i>Ty1-copia</i>	M94494
<i>Ipomea batatas</i>	ipo	<i>Ty1-copia</i>	AF231939
<i>Ephedra distachya</i>	eph	<i>Ty1-copia</i>	DQ054416
<i>Triticum aestivum</i>	triti	<i>Ty1-copia</i>	AB061329
<i>Pinus taeda</i>	pin	<i>Ty3-gypsy</i>	AJ290632
<i>Musa acuminata</i>	mus	<i>Ty3-gypsy</i>	AJ971814
<i>Brassica</i>	bras	<i>Ty3-gypsy</i>	AJ415653
<i>Arabidopsis</i>	ara	<i>Ty3-gypsy</i>	AJ295139
HIV-1 virus	hiv	Retrovirus	AY877248
Lentivirus	len	Retrovirus	AY454232
Human adenovirus	humad	Retrovirus	U34348
Humanart	humar	Retrovirus	M25768

the existence of a large population of *Ty1-copia*-like retrotransposons in all three species of *Datura* (Fig. 2). The presence of hybridizing bands across entire lanes indicates the dispersal of the *Ty1-copia*-like sequences throughout the *Datura* genome. The observation of different band intensities within a single track could be attributed to the copy number of copia like RT within the genome.

Sequence analysis

The alignment and the *in silico* analysis of all the RT sequences showed strong homology to the conserved RT domains of *Ty1-copia* family of retrotransposons, though no two sequences were identical (Fig. 3). The intra-family and intra-subfamily degree of divergence at the amino acid level ranged from 4 to 52 % (Fig. 4). The amino acid motifs DVKTAFLHG at their 5'end and, LYVDDMDLP at 3'end are conserved and also used for designing primers. LLLYVDDMD present at the 3'end of the isolated RT sequences in *Datura* is another conserved motif. The RT

sequences from *D. metel* show more homology within species. For comparative analysis, sequences of *Ty1-copia*, *Ty3-gypsy* and retroviruses identified from GenBank database (Table 2) were used for alignment and the phylogenetic tree was generated (Fig. 5). These sequences included those of *Lycopersicum esculentum*, *Solanum tuberosum*, *Ipoemaea batatas*, *Nicotiana tabacum*, *Ephedra*, *Triticum aestivum*, *Pinus*, *Musa*, *Brassica*, *Arabidopsis*, *HIV-1 virus*, humanadenovirus and Human RT. The analysis of phylogenetic tree suggested that the sequences of *Datura* species did not fall in one group. These infact were dispersed throughout. Within the family, the sequences that show branching from a common nearest point in the phylogenetic tree have higher amino acid homology than the others. Extreme heterogeneity was detected among all sequences that exceed the heterogeneity seen in *D. melanogaster*, which has more *Ty1-copia* group elements per genome than does *Arabidopsis*. It is, therefore assumed that in plants *Ty1-copia* group of retrotransposons are inherently more susceptible to sequence variation.

Fig. 5 Phylogenetic analysis based on RT sequences isolated from *Datura* species (our studies) and comparison with sequence from *Lycopersicum*, *Solanum*, *Nicotiana*, *Helianthus*, *Ipoemaea*, *Ephedra*, *Triticum*, *Pinus*, *Musa*, *Brassica*, *Arabidopsis*, *HIV-1*, Lentivirus, Human adenovirus, Human RT

Discussion

Structurally the *Ty1-copia* group of retrotransposons contains 5'- and 3'-LTRs flanking an internal *gag-pol* region. The *pol* gene contains many conserved domains but the reverse transcriptase domain of the *pol* gene is highly conserved in various plant species. We have identified *Ty1-copia* retrotransposon in *Datura* sps., using primers for the reverse transcriptase regions. Comparison at the nucleotide as well as the amino acid level revealed a significant heterogeneity among the isolated conserved RT domains of the *Ty1-copia* retrotransposon in *Datura* sps. The divergence of predicted amino acid sequences of RT domain was from 4 to 52 %. This is in agreement with the results from potato, where 31 subcloned fragments of the reverse transcriptase region of *Ty1-copia*, were found to be different, with amino acid similarities varying from 5 to 75 % (Flavell et al. 1992). Extreme heterogeneity has been observed in several plant species like *Arabidopsis* (Konieczny et al. 1991), rice (Noma et al. 1997), *Vicia* (Pearce et al. 1996a), *Allium* (Pearce et al. 1996b), grass (Matsuoka and Tsunewaki 1999). In *Ipoemaea batatas* (L), the sequence divergence of *Ty1-copia* retrotransposon was from 2 to 73 % (Villordon et al. 2000). Similar results have been reported in tomato, conifers, and sorghum (Su and Brown 1997; Friesen et al. 2001; Muthukumar and Bennetzen 2004).

Southern blot analysis in this study shows the existence of dispersed copies of *Ty1-copia* element in *Datura* genome as well. The fact that these genome display a similar set of copia-like retrotransposon-specific restriction fragments suggests that the genomic location could be similar within these different genomes. This result may possibly indicate that this element is an ancient component of the *Datura* genome, introduced before the divergence of the species and conserved during evolution. The uniform pattern of Southern blotting could explain that these elements might possess uniform localization and favourable habitat of insertion in the host genome. This result agreed with previous report in *Citrus* (De Felice et al. 2009). According to Kumar and Bennetzen (1999) there could also be an existence of a mixed population of active and defective retrotransposons which accumulate mutations over time, and eventually give rise to highly heterogeneous population. All transposons are integrated in to chromosomal DNA, so the heterogeneity of retrotransposons is maintained between generations through vertical transmission.

It has been observed that in RT genomic sequence from three species of *Datura* contain one or more chain terminating codons in all the reading frame. However, the presence of multiple chain terminating codons in a single ORF reflected the non functionality and long insertion time of the other two elements. Analysis of all the known

retrotransposons show that, the older the element the greater the number of termination codons. In other words, the age of an element is directly proportional to the number of termination codons incorporated in it since the termination codons are not accumulated in an active element due to its continuous retrotransposition (Gao et al. 2003).

The isolation and characterization of hereogeneity study of *Ty1-copia* RT genomic sequence from *D. innoxia*, *D. stramonium*, and *D. metel* represents the first known report. The study will contribute an important background for further understanding of retrotransposons in *Datura* genome, as it symbolizes an important medicinal herb in the family Solanaceae.

Acknowledgements The study was supported by financial assistance by Department of Science and Technology, Govt. of India to PSS, and Junior/Senior Research Fellowship from University Grants Commission (UGC) to AS.

References

- Arabidopsis Genome Initiative (2000) Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. Nature 408:796–815
- Beguistain T, Grandbastien MA, Puigdomenech P, Casacuberta JM (2001) Three *Tnt1* subfamilies show different stress-associated patterns of expression in tobacco. Consequence of retrotransposon control and evolution in plants. Plant Physiol 127:212–221
- Cédric M, Keith EG, Carole PM, Frédérique (2008) Characterization of ten novel *Ty1-copia*-like retrotransposon families of the grapevine genome. BMC Genomics 9:1–37
- De Felice B, Wilson RR, Argenziano C, Kafantaris I, Conicella C (2009) A transcriptionally active copia-like retroelement in *Citrus limon*. Cell Mol Biol Lett 14:289–304
- Doyle JJ, Doyle JL (1990) Isolation of plant DNA from fresh tissue. Focus 12:13–15
- Ellis TH, Poyser SJ, Knox MR, Vershinin AV, Ambrose MJ (1998) Polymorphism of insertion sites of *Ty1-copia* class retrotransposons and its use 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 3:329–341
- Flavell AJ, Smith DB, Kumar A (1992) Extreme heterogeneity of *Ty1-copia* group retrotransposons in plants. Mol Gen Genet 231:233–242
- Friesen N, Brandes A, Heslop-Harrison (2001) Diversity, origin and distribution of retero-transposons (gypsy and copia) in conifers. Mol Biol Evol 18:1176–1188
- Gao X, Havecker ER, Baranov PV, Atkins JF, Voytas DF (2003) Translational recoding signals between *gag* and *pol* in diverse LTR retrotransposons. RNA 9:1422–1430
- Hirochika H, Otsuki H, Yoshikawa M, Otsuki Y, Sugimoto K, Takeda S (2006) Autonomous transposition of the tobacco retrotransposon *Tto1* in rice. Plant Cell 8:725–734
- Jia J, Yang G, Li G, Liu Ch, Lei M, Zhang T, Zhou J, Ren Z (2009) Isolation and distribution of a novel *Ty1-copia*-like sequence from *Secale*, which enable identification of wheat-*Secale africanum* introgression line. J Appl Genet 50:25–28

- Kalendar R, Tanskanen J, Immonen S, Nevo E, Schulman AH (2000) Genome evolution of wild barley (*Hordeum spontaneum*) by BARE-1 retrotransposon dynamics in response to sharp microclimatic divergence. *Proc Natl Acad Sci USA* 97:6603–6607
- Konieczny A, Voytas DF, Cummings MP, Ausubel FM (1991) A superfamily of *Arabidopsis thaliana* retrotransposons. *Genetics* 127:801–809
- Kumar A, Bennetzen JL (1999) Plant retrotransposon. *Annu Rev Genet* 33:479–532
- Kumar A, Hirochika R (2001) Applications of retrotransposons as genetic tools in plant biology. *Trends Plant Sci* 6:127–134
- Ma Y, Sun HZG, Dai H, Gao X, Li H, Zhang Z (2008) Isolation and characterization of genomic retrotransposon sequences from octaploid strawberry (*Fragaria X ananassa* Duch.). *Plant Cell Rep* 27:499–507
- Matsuoka Y, Tsunewaki K (1999) Evolutionary dynamics of *Ty1*-copia group retrotransposons in grass shown by reverse transcriptase domain analysis. *Mol Biol Evol* 16:208–217
- Moisy C, Garrison KE, Meredith CP, Pelsy F (2008) Characterization of ten novel *Ty1/copia*-like retrotransposon families of the grapevine genome. *BMC Genomics* 9:469–483
- Moreau-Mhiri C, Morel JB, Audeon C, Ferault M, Grandbastien MA, Lucas H (1996) Regulation of expression of the tobacco *Tnt1* retrotransposon in heterologous species following pathogen related stresses. *Plant J* 9:409–419
- Muthukumar B, Bennetzen JL (2004) Isolation and characterization of genomic and transcribed retrotransposon sequences from sorghum. *Mol Genet Genomics* 271:308–316
- Noma K, Nakajima R, Ohtsubo H, Ohtsubo E (1997) RIRE1, a retrotransposon from wild rice (*Oryza australiensis*). *Genes Genet Syst* 72:131–140
- Pearse SR, Harrison G, Li D, Heslop-Harrison JS, Kumar A, Flavell AJ (1996a) The *Ty1*-copia group retrotransposons in *Vicia* species: copy number, sequence heterogeneity and chromosomal localization. *Mol Gen Genet* 250:305–315
- Pearse SR, Pich U, Harrison G, Flavell AJ, Heslop-Harrison JS (1996b) The *Ty1*-copia group retrotransposons of *Allium cepa* are distributed throughout the chromosomes but are enriched in the terminal heterochromatin. *Chromosome Res* 4:357–364
- Salazar M, González E, Casaretto JA, Casacuberta JM, Ruiz-Lara S (2007) The promoter of the TLC1.1 retrotransposon from *Solanum chilense* is activated by multiple stress-related signaling molecules. *Plant Cell Report* 26:1861–1868
- Sambrook J, Fritsch EF, Maniatis T (1989) *Molecular cloning: A laboratory manual*, 2nd edn. Cold Spring Harbor Lab. Press, Cold spring Harbor
- Sequencing Project International Rice Genome: The map-based sequence of the rice genome (2005) *Nature* 436:793–800
- Su PY, Brown TA (1997) Ty3-gypsy like retrotransposon sequences in tomato. *Plasmid* 38:148–157
- Tuskan GA, DiFazio S, Jansson S, Bohlmann J, Grigoriev I, Hellsten U, Putnam N, Ralph S, Rombauts S, Salamov A (2006) The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray). *Science* 313:1596–1604
- Velasco R, Zharkikh A, Troglio M, Cartwright DA, Cestaro A, Pruss D, Pindo M, FitzGerald LM, Vezzulli S, Reid J (2007) A High quality of draft consensus sequence of the genome of a heterozygous grapevine variety. *PLoS ONE* 2:e1326
- Villordon AQ, Jarret RL, LaBonte DR (2000) Detection of *Ty1*-copia like reverse transcriptase sequences in *Ipoemaea batatas* (L) Poir. *Plant Cell Report* 19:1219–1225
- Vitte C, Bennetzen JL (2006) Eukaryotic transposable elements and genome evolution special feature: Analysis of retrotransposon structural diversity uncover properties and propensities in angiosperm genome evolution. *Proc Natl Acad Sci USA* 103:17638–17643
- Voytas DF, Ausubel FM (1988) A copia like transposable element family in *Arabidopsis thaliana*. *Nature* 336:242–244
- Xiong Y, Eickbush TH (1990) Origin and evolution of retroelements based upon their reverse transcriptase sequences. *EMBO J* 9:3353–3360