Cloning and analysis of reverse transcriptases from Ty1-\textit{copia} retrotransposons in \textit{Camellia sinensis}

Jin Yao\textsuperscript{a}, Li Xiaoyu\textsuperscript{a,b}, Pan Cheng\textsuperscript{a}, Li Yeyun\textsuperscript{a}, Jiang Jiayue\textsuperscript{a} and Jiang Changjun\textsuperscript{a}

\textsuperscript{a}State Key Laboratory of Tea Plant Biology and Utilization, College of Tea and Food Technology, Anhui Agricultural University, Hefei, PR China; \textsuperscript{b}Department of Biological Resources, School of Life Science and Environmental Science, Huangshan University, Huangshan, PR China

\textbf{ABSTRACT}

As mobile genetic elements, the diversity and activity of the Ty1-\textit{copia} retrotransposons are key contributors to genome organisation and evolution, which have been investigated in many plants but little in the tea plant, \textit{Camellia sinensis}. We selected a total of 12 varieties of tea plant distributed across a large geographical area and sequenced the reverse transcriptase (RT) sequences of Ty1-\textit{copia} retrotransposons using degenerate primers. The sequences that were widespread among tea varieties, were approximately 260 bp in length and exhibited high heterogeneity among tea plants. The RT sequences from tea plants were similar to other known Ty1-\textit{copia} retrotransposon RT sequences. Phylogenetic analysis showed that tea RT sequences were closely related to those from woody plants, such as pear, poplar and apple. In contrast, they were more distantly related to RT sequences from herbaceous plants, such as tomato and rice. Activity assay revealed that Ty1-\textit{copia} retrotransposons are transcriptionally active during the normal development of tea plants. These results will support further research on Ty1-\textit{copia} retrotransposons in \textit{C. sinensis}.

\textbf{Introduction}

Retrotransposons are mobile genetic elements that are widespread among eukaryotes [1]. During transposition, transposable DNA elements are transcribed into RNA intermediates, which are then reverse transcribed back into DNA with reverse transcriptase (RT). Following reverse transcription, the newly produced DNA copies are inserted into new chromosomal locations. Among the family of retrotransposons, there are long terminal repeat (LTR) and non-long terminal repeat (non-LTR) elements. LTRs can be divided into two subclasses: Ty1-\textit{copia} and Ty3-gypsy, which are characterized by the properties of universality, high copy number and single insertion sites in plant genomes [2]. The cloning and analysis of these sequences can be of great interest when analysing plant genome evolution, expression and transcriptional regulation. These sequences have also represented an important tool for studying biological diversity [3,4] and phylogenetic evolution [5–7], as well as for genetic linkage mapping [8,9], gene tagging [10] and gene function analysis.

Ty1-\textit{copia} retrotransposons are the most widespread and well-studied LTR retrotransposons. Thus far, a number of research projects have been devoted to studying their sequence characteristics, system evolution and transpositional activity in various plants [11], such as: pear trees [12,13], \textit{Populus × Canadensis} [14], \textit{Eucalyptus} [15], \textit{Gossypium} [16], \textit{Fragaria × ananassa} [17], \textit{Diospyros kaki} [18], \textit{Hylocereus undatus} [19], \textit{Paeonia suffruticosa} (Andrews) [20], etc. Meanwhile, it has been shown that retrotransposons can be transferred not only via vertical transfer, but also via horizontal transfer by viruses and bacteria [21]. However, few investigations have focused on Ty1-\textit{copia} retrotransposons in the tea plant, \textit{Camellia sinensis} (L.) O. Kuntze. To examine the possible influence of Ty1-\textit{copia} retrotransposons on the evolution and transcriptional regulation of the tea plant genome, it is first necessary to determine whether these elements exist in the tea plant genome, and, if so, to then characterize these elements. Because RT plays an important role in the transposition of Ty1-\textit{copia} retrotransposons, it can be used to shed light on the elements themselves. Therefore, the aim of this study was to analyse sequence characteristics, phylogenetic relationships and transcriptional activity of the Ty1-\textit{copia} RT genes in the hopes of providing new information about \textit{C. sinensis} genome evolution and genome regulation research.
Materials and methods

Plant materials

A total of 12 tea varieties distributed across different geographical regions and including large and small trees and shrubs were used in this study. These varieties included *C. sinensis* cv. *Shuchazao* from the North Yangtze river area; *C. sinensis* var. *assamica* cv. Yinhong 1 and *C. sinensis* var. *assamica* cv. Yunkang 10 from south China; and *C. sinensis* cv. Jinfeng, *C. sinensis* cv. Zhenong 113, *C. sinensis* cv. Longjing-changye, *C. sinensis* cv. Anhui 3, *C. sinensis* cv. Anhui 1, *C. sinensis* cv. Tieguyanyin, *C. sinensis* cv. Baihaozao and *C. sinensis* cv. Wannong 95 from the Jiangnan region. *C. sinensis* var. *assamica* cv. Yinhong 1 and *C. sinensis* var. *assamica* cv. Yunkang 10 were acquired from the Key Laboratory of Anhui Agricultural University greenhouse, and the others were acquired from the tea garden of Anhui Agricultural University. These young leaves were harvested (October 2015) and immediately frozen in liquid nitrogen and stored at −80 °C. Three independent biological replicates of each variety were acquired. Each replicate was collected from randomly selected tea plants.

Polymerase chain reaction (PCR) amplification of Ty1-copia RTs

DNA was isolated by a modified hexadecyl trimethyl ammonium bromide (CTAB) method. PCR amplification was conducted with degenerate primers designed by Kumar [22]: RTF was used as the forward primer (5’-ACNGCNYTYYCAYGG-3’) and RTS was used as the reverse primer (5’-ARCATRTCRTNACRTA -3’). PCR was performed in 25 μL reactions including 10 μL *Taq* buffer (2.5 μL), 25 mmol/L MgCl2 (1.5 μL), 2.5 mmol/L deoxynucleoside triphosphates (dNTPs; 1.0 μL), 10 μmol/L primers (0.8 μL each RTS and RTF), 5 U/μL *Taq* enzyme (0.2 μL) and DNA template (100 ng). The PCR conditions used were as follows: 94 °C for 2 min; 35 cycles of 94 °C for 1 min, 48–56 °C for 50 s, and 72 °C for 50 s; and a final extension of 72 °C for 5 min (S1000™ Thermal Cycler, Bio-RAD Co., California, USA). Products were visualized with 1.2% agarose gel electrophoresis in 0.5 × TBE buffer.

Cloning of Ty1-copia RTs

An AxyPrep™ DNA gel extraction kit by Axygen Co. (California, USA) was used to purify PCR products, which were then directly cloned into the PMD19-T vector by TaKaRa Co. (Tokyo, Japan). After incubation for 8–10 h, vectors were transformed into trans1-T1 *Escherichia coli* competent cells by TransGen Co. (Beijing, China), which were cultured in solid Luria–Bertani (LB) medium containing ampicillin by Sangon Biotech Co. (Shanghai, China) at 37 °C for 10–12 h. Individual white colonies were then isolated for culture in 100 mg/L liquid LB medium containing ampicillin overnight, and colony PCR was performed.

Sequencing and analysis of Ty1-copia RTs

Positive clones were selected by a PCR procedure the same as the PCR amplification procedure of Ty1-copia RT and were sequenced by Sangon Biotech Co. (Shanghai, China). The sequencing procedure was as follows: sequencing reactions were carried out in both directions using M13 universal forward and reverse primers and BigDye version 3.1 kit by Applied Biosystems Co. (California, USA) on an ABI 3730XL DNA Sequencer by Applied Biosystems Co. Sequences were trimmed according to chromatogram quality criteria, using the base-calling program Phred 50 at a score >20, and vector-derived sequences were removed using NCBI (National Center for Biotechnology Information) VecScreen (http://blast.ncbi.nlm.nih.gov/Vectorscreen, html) and BioEdit software. Sequences were analysed through comparison with homologous sequences from NCBI detected by The Basic Local Alignment Search Tool (BLAST) (http://blast.ncbi.nlm.nih.gov/Blast.cgi) and were analysed by DNASTar and DNAMAN software. The phylogenetic analysis of these sequences was performed using the neighbour-joining algorithm method in MEGA 6.0 software.

Detection of Ty1-copia retrotransposon activity

RNA was extracted from cultivars selected from the 12 tea plant varieties using the RNA Prep Pure Plant Kit (QIAGEN Co., Dusseldorf, Germany). Genomic DNA was removed with Recombinant *DNasel* (TaKaRa Co.) RNA was reverse-transcribed into cDNA using the PrimeScript™ 1st Strand cDNA Synthesis Kit (TaKaRa Co.). The PCR procedures were as follows: 37 °C for 15 min, 85 °C for 5 s. The products were diluted to 100 ng/μL and used as templates to amplify RT sequences. The primers, PCR amplification procedures and conditions were the same as those for PCR amplification of Ty1-copia RTs.

Results and discussion

PCR amplification and sequencing of Ty1-copia RTs

Genomic DNA from 12 tea plant varieties was amplified by PCR. RT PCR products ranging from 260 to 270 bp
were found in all tea varieties (Figure 1). This demonstrates that Ty1-copia RT genes are widespread among tea varieties from different geographical regions, which is consistent with studies on these elements in other plants [23–27].

The amplified RT products were then cloned and sequenced. This resulted in 114 Ty1-copia-like RT sequences (see the Online Supplemental Data) after removing duplicate and short sequences using DNAstar. C. sinensis sequences from the NCBI database (https://www.ncbi.nlm.nih.gov/). We identified 14 unique RT sequences from C. sinensis cv. Shuchaozao, 16 ones from C. sinensis cv. Tieguanyin, 12 from C. sinensis cv. Baihaozao, 8 from C. sinensis cv. Jinfeng, 24 from C. sinensis var. assamica cv. Yinghong 1, 11 from C. sinensis cv. Longjing-changye, 7 from C. sinensis cv. Anhui 1, 8 from C. sinensis cv. Anhui 3, 1 from C. sinensis cv. Fuyun 6, 4 from C. sinensis var. assamica cv. Yunkang 10, 5 from C. sinensis cv. Zhenong 113 and 4 from C. sinensis cv. Wannong 95. Homology-based BLAST searches against the NCBI database revealed that these putative RT clones exhibited nucleotide similarity to the RT domains of other known plant Ty1-copia retrotransposons. These results further verify the authenticity of the tea retrotransposon RT gene fragments.

**Sequence analysis and heterogeneity of Ty1-copia RTs**

For a more in-depth understanding of the characteristics of the RT sequences, we performed a detailed comparison of 16 RT sequences from three of the 12 tea varieties from two different regions: C. sinensis var. assamica cv. Yunkang 10, a large-leaved arboreal variety from south China (sequences YK1–4), C. sinensis cv. Jinfeng, a medium-leaved, small, arboreal species from Jiangnan (sequences JF1–8), and C. sinensis cv. Wannong 95, a medium-leaved shrub species from Jiangnan (sequences WN1–4).
the 5'-TAFF(L)HG sequence, the 3'-YVDDM sequence and the central LYGLKQ sequence (Figure 3). Most of the 16 RT sequences exhibited strong homology to the conserved RT domains of Ty1-copia retrotransposons from various plant species [22], confirming that the cloned sequences belong to Ty1-copia RTs.

However, among the 16 sequences, frameshift mutations were identified in seven sequences: WN2 (at amino acid positions 23 and 74), YK1 (positions 23 and 74), YK2 (position 74), YK4 (position 74), JF1 (position 74), JF2 (position 74) and JF6 (position 74). In addition, mutations that introduce an early termination codon were identified in five sequences: YK1 (position 62), YK3 (position 44), JF3 (position 44), JF4 (position 44), and JF5 (position 16). Therefore, it appears that frameshift and termination codon mutations are important contributors to heterogeneity among tea plant Ty1-copia retrotransposons.

The analysis of nucleotide and amino-acid sequences revealed several differences among the 16 sequences from the three varieties. Presumably, these are due to the fact that retrotransposons lack an automatic repair
mechanism during the transposition process, resulting in mutations introduced during reverse transcription, such as base mismatches, base losses and code-shifting mutations; the resulting sequences thus demonstrate a relatively high degree of heterogeneity [28]. This research is consistent with studies on apple [25], capsicum [26], cucumber [29] and other plants, which shows that Ty1-copia retrotransposons exhibit high heterogeneity within individual species.

**Phylogenetic analysis of RT amino-acid sequences**

For comparative purposes, we included in the phylogenetic analysis of deduced Ty1-copia RT amino-acid sequences those from other species that were deposited in the GenBank database of NCBI (Figure 4). The phylogenetic relationships among the Ty1-copia retrotransposon RT sequences were determined by constructing a neighbour-joining tree based on P-distances and supported with 1000 bootstrap replicates. The results revealed that RT sequences from tea plants were more closely related to those from Malus × domestica (DQ410750.1), Populus ciliata (AY624351.1) and Pyrus × bretschneideri (JX083313.1) and less closely related to those from herbaceous plants like Oryza sativa (AB017997.1) and Lycopersicon esculentum (AF072656.1).

The high heterogeneity and widespread distribution of retrotransposons among plants indicate that an ancient transposable element existed in early plants. As a component of the genome, these elements not only are able to be passed on from generation to generation [30], but also can transfer horizontally from one species to another. Comparing the RT amino-acid sequences from tea to those of other plants, it is evident that tea Ty1-copia RTs are closely related to those from woody plants and are less related to those from herbaceous plants. This indicates that tea and other woody plants may share a recent common ancestor or similar
retrotransposons. However, this could also be an indication of horizontal gene transfer between species.

**Detection of Ty1-copia retrotransposon activity**

cDNA used as a template for the PCR amplification of RT sequences showed there are PCR products (Figure 5), implying that Ty1-copia reverse retrotransposons have transcriptional activity in the normal growth of tea plants, thus stably existing in the genome.

The Ty1-copia retrotransposon activity plays an important role during transposition. Future research should investigate the relationship between the tea stress response mechanism and Ty1-copia retrotransposons activity in order to provide new information on the genetic variability of the tea plant.

**Conclusions**

In this study, we isolated 114 Ty1-copia RT sequences from 12 tea varieties, then performed sequence and phylogenetic analysis. The results suggested that RT genes are widespread among tea varieties and transfer not only horizontally, but also vertically. We also determined that Ty1-copia retrotransposons have transcriptional activity in the normal growth of tea plant. The results suggested that Ty1-copia retrotransposons stably exist in the genome and are transcriptionally active in tea plants. These results will support further research on Ty1-copia retrotransposons in *C. sinensis*.

**Acknowledgement**

The authors are very grateful to the editors and reviewers for their constructive comments on this paper.

**Disclosure statement**

No potential conflict of interest was reported by the authors.

**Funding**

This work was supported by the National Natural Science Foundation of China [grant number 31270729].

**References**

[1] Kumar A, Bennetzen JL. Plant retrotransposons. Annu Rev Genet. 1999;33:479–532.
[2] Schulman AH, Flavell AJ, Paux E, et al. The application of LTR retrotransposons as molecular markers in plants. Methods Mol Biol. 2012;859:145–173.
[3] Kalendar R, Flavell AJ, Ellis THN, et al. Analysis of plant diversity with retrotransposon-based molecular markers. Heredity. 2011;106:520–530.
[4] Biswas MK, Baig MNR, Cheng YJ, et al. Retro-transposon based genetic similarity within the genus *Citrus* and its relatives. Genet Resor Crop Evol. 2010;57:963–972.
[5] Slotkin RK. The epigenetic control of the Athila family of retrotransposons in Arabidopsis. Epigenetics. 2010;5:483–490.
[6] Elena B, Lucia N, Tommaso G, et al. LTR retrotransposon dynamics in the evolution of the olive (*Olea europaea*) genome. DNA Res. 2015;22:91–100.
[7] Andrea Z, Douglas GS, Emanuele DP, et al. The Ty1-copia LTR retroelement family PARTC is highly conserved in conifers over 200 MY of evolution. Gene. 2015;568:89–99.
[8] Sung-II, Jong HK, Kyong CP, et al. LTR-retrotransposons and inter-retrotransposon amplified polymorphism (IRAP) analysis in *Lilium* species. Genetica. 2015;143:343–352.
[9] Liang YH, Ryan R. L, Dai WH. Development of retrotransposon-based molecular markers and their application in genetic mapping in chokecherry (*Prunus virginiana* L.). Mol Breeding. 2016 [cited 2016 Nov 29];36:109. DOI:10.1007/s11032-016-0535-2.
[10] Sivalingam A, Abhilash N, Dipali SK, et al. Retrotransposon based TRAP marker displays diversity among onion (*Allium cepa* L.) genotypes. Sci Hortic-Amsterdam. 2015;190:123–127.
[11] Elaine SD, Clémence H, Serge H, et al. Large distribution and high sequence identity of a Copia-type retrotransposon in angiosperm families. Plant Mol Biol. 2015;89:83–97.
[12] Jiang S, Luo J, Sun YW, et al. Prediction of ORF for GAG of Ty1-copia retrotransposons in pear and its expression analysis in response to light. Acta Hortic Sin. 2016;43:1577–1584.
[13] Zhou P, Chou ZH, Zhai R, et al. Cloning and comparative analysis of reverse transcriptase of Ty1-copia retrotranspon in ‘Zaosu’ (Pyrus bretschneideri Rehd) pear and its red mutant. J Northwest Sci-Tech Univ Agric Forest (Nat Sci Ed). 2014;42:162–182.

[14] Tommaso G, Rosa MC, Flavia M, et al. Genome-wide analysis of LTR-retrotransposon expression in leaves of Populus × canadensis water-deprived plants. Tree Genet Genomes. 2016;12:1–14.

[15] Marcon HS, Domingues DS, Costa Silva J, et al. Transcriptionally active LTR retrotransposons in Eucalyptus genus are differentially expressed and insertionally polymorphic. BMC Plant Biol. 2015 [cited 2016 Nov 29];15:198. DOI:10.1186/s12870-015-0550-1

[16] Cao YF, Jiang YR, Ding MQ, et al. Molecular characterization of a transcriptionally active Ty1/copia-like retrotransposon in Gossypium. Plant Cell Rep. 2015;34:1037–1047.

[17] Ma Y, He P, Sun HY, et al. Isolation and characterization of transcriptionally active Ty1-copia retrotransposons in Fragaria X ananassa. Agr Sci China. 2010;9:337–345.

[18] Du XY, Zhang QL, Luo ZR. Isolation and characterization of RNaseH-LTR sequence of Ty1-copia retrotransposons in Oriental Persimmon (Diospyros kaki Thunb. Loutian-tian-shi). Acta Hortic Sin. 2008;35:501–508.

[19] Fan FH, Qiao G, Zheng SC, et al. Cloning and analysis of reverse transcriptase of Ty1-copia retrotransposons in Hyllocereus undatus. Acta Hortic Sin. 2012;39:265–272.

[20] Jia TT, Guo DL, Hou XG, et al. Isolation and characterization of Ty1-copia retrotransposons sequence from suffruticosus Andrews. Acta Agric Boreali-Sin. 2012;27:30–33.

[21] Kumar A. The evolution of plant retroviruses: moving to green pasture. Trends Plant Sci. 1998;3:371–374.

[22] Kumar A, Pearce SR, Mclean K, et al. The Ty1-copia group of retrotransposons in plants: genomic organisation, evolution, and use as molecular markers. Genetica. 1997;100:205–217.

[23] Li C, Si JP, Gao YH, et al. Cloning and analysis of reverse transcriptase (RT) of Ty1-copia retrotransposon in D endrobiurn officinale. China J Chin Mater Me d. 2014;39:209–215.

[24] Li WT, Yang T, Zhang XY, et al. Cloning and analysis of Ty1-copia-like reverse transcriptase in modern rose. Mol Plant Breeding. 2014;12:1216–1221.

[25] Sun J, Fang JG, Gao B, et al. Ty1-copia-like retrotransposon in Gala apple. J Fruit Sci. 2005;22:193–197.

[26] Diao WP, Tu SB, Liu JB, et al. Cloning and sequence analysis of reverse transcriptase of Ty1-copia-like retrotransposons in Capsicum annuum. Mol Plant Breeding. 2012;10:55–61.

[27] Zhang JG, Hu HJ, Tian R, et al. Cloning and analysis of reverse transcriptase of Ty1-copia-like retrotransposons in Pyrus. Hubei Agric Sci. 2012;51:4388–4390.

[28] Vershinin AVL, Ellis THN. Heterogeneity of the internal structure of PDRI, a family of Ty1-copia like retrotransposons in pea. Mol Genet Genomics. 1999;262:703–713.

[29] Jing B, Lou QF, Diao WP, et al. Cloning and analysis of reverse transcriptase of Ty1-copia-like retrotransposons in Cucumis. Acta Hortic Sin. 2008;35:1147–1154.

[30] Doolittle RF, Feng DF, Johnson MS, et al. Origins and evolutionary relationships of retroviruses. Q Rev Biol. 1989;64:1–30.