Scientia Horticulturae 127 (2010) 46–53

Contents lists available at ScienceDirect

Scientia Horticulturae

journal homepage: www.elsevier.com/locate/scihorti

Genetic diversity of Ty1-*copia* retrotransposons in a wild species of *Cucumis* (*C. hystrix*)

Biao Jiang[a], Zhi-Ming Wu[a,b], Qun-Feng Lou[a], Dong Wang[a], Wan-Ping Zhang[c], Jin-Feng Chen[a,*]

[a] *State Key Laboratory of Crop Genetics and Germplasm Enhancement, Key Laboratory of Southern Vegetable Crop Genetic Improvement, College of Horticulture, Nanjing Agricultural University, Nanjing 210095, China*
[b] *Institute of Economic Crop Research, Hebei Academy of Agricultural and Forestry Sciences, Shijiazhuang 050051, China*
[c] *College of Agriculture, Guizhou University, Guiyang 550025, China*

ARTICLE INFO

Article history:
Received 15 April 2010
Received in revised form 12 August 2010
Accepted 11 September 2010

Keywords:
Cucumis hystrix
Retrotransposons
Reverse transcriptase
Heterogeneity
Transcriptional activity

ABSTRACT

Ty1-*copia* retrotransposons are ubiquitous in the plant kingdom and have an important effect on genome structure and evolution. *Cucumis hystrix* ($2n = 2x = 24$) is an exotic wild species of *Cucumis* with a number of valuable traits that are not available in cultivated cucumber. However, lack of the basic genome structure and evolutionary information about *C. hystrix* has hindered its utilization in cucumber breeding and evolutionary research. In the present study, reverse transcriptase (RT) sequences of Ty1-*copia* retrotransposons were isolated from genome of *C. hystrix* by PCR amplification using degenerate oligonucleotide primers. The 24 clones obtained were different from each other and showed high heterogeneity. Of the 24 sequences, 13 are intact whereas the remaining sequences have a stop codon, a frameshift, or both. Southern hybridization revealed that Ty1-*copia* retrotransposons distributed widely in *C. hystrix* by multiple copies, and the total number was estimated to be about 5460 as calculated by dot-blot hybridization. The retrotransposons of this kind comprise approximately 8.5% of the *C. hystrix* genome. No transcription of Ty1-*copia* retrotransposons was detected by reverse transcriptase PCR (RT-PCR) in leaves of chilling or heat treated plants. Moreover, phylogenetic analysis showed that the 24 RT sequences were divided into five families and had homology with other species, indicating that both vertical transmission and horizontal transmission are the source of Ty1-*copia* retrotransposons in *C. hystrix*. FISH with Ty1-*copia* retrotransposons in *C. hystrix* revealed the retrotransposons of this kind were widely dispersed over all the chromosomes with clustering in terminal heterochromatin regions. These results contribute to our understanding about the organization and evolution of *C. hystrix* genome and will provide valuable information for the utilization of retrotransposons in *C. hystrix*.

1. Introduction

The mobile genetic elements called retrotransposon are the largest class of transposable elements (TEs). They are ubiquitous in the plant kingdom, and present in high copy numbers in most plants, making them major constituents of plant genomes (Kumar and Bennetzen, 1999). They move to new genome locations via an RNA intermediate, which are reversely transcribed into DNA by the encoded reverse transcriptase/RNaseH enzymes prior to reinsertion into the genome (Grandbastien, 1992). Retrotransposons can be divided into two major groups: LTR (long terminal repeat) and non-LTR retrotransposons according to the presence or absence of LTR (Kumar and Bennetzen, 1999; Bennetzen, 2000). LTR retrotransposons are further subclassified into the Ty1-*copia* and the Ty3-*gypsy* groups that differ from each other by sequence similarity and the order of their catalytic enzymes, encoded gene products (Kumar and Bennetzen, 1999).

Ty1-*copia* retrotransposons are present throughout the plant kingdom, ranging from algae to bryophytes, gymnosperms, and angiosperms (Voytas et al., 1992). However, due to the presence of stop codons, frameshifts, and deletions, most retrotransposons appear to be nonfunctional (Bennetzen, 2000), but might be activated by various stresses (Grandbastien, 1998). Only a handful of them have been reported to be active, including Tnt1 (Grandbastien et al., 1989) and Tto1 (Hirochika, 1993) from tobacco, the Tos elements from rice (Hirochika et al., 1996), BARE-1 from barley (Manninen and Schulman, 1993), CIRE1 from *Citrus sinensis* (Rico-Cabanas and Martínez-Izquierdo, 2007), and Reme1 from melon (Ramallo et al., 2008). Retrotransposons are present in high copy numbers in most plants. They typically account for more than half of the total DNA (Kumar and Bennetzen, 1999), and in some

Abbreviations: LTR, long terminal repeat; ORF, open reading frame; PCR, polymerase chain reaction; RT, reverse transcriptase; RT-PCR, reverse transcription PCR; TE, transposable element.

* Corresponding author. Tel.: +86 25 84396279; fax: +86 25 84396279.
E-mail address: jfchen@njau.edu.cn (J.-F. Chen).

0304-4238/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.scienta.2010.09.013

species such as bread wheat they even accounts for 90% of the nuclear genome (Flavell, 1986). It is proved that retrotransposons have contributed to increasing genome size in the plant kingdom, and play important roles in the evolution of genome (Kumar and Bennetzen, 1999; Feschotte et al., 2002). Furthermore, retrotransposons can also generate mutations by inserting within or near genes (Kashkush et al., 2002, 2003), and cause chromosomal rearrangements by means of ectopic recombination (Bennetzen, 2002).

Cucumis hystrix ($2n = 2x = 24$) is an exotic wild species of *Cucumis* with a number of valuable traits that are not available in cultivated cucumber (Qian et al., 2002; Zhuang et al., 2002; Chen et al., 1994). It is the closest wild species of cucumber and the only one that is cross-compatible with cucumber, which makes it possible to transfer the unique disease resistance genes from *C. hystrix* into cucumber (Chen et al., 1997, 2004). However, lack of the basic genetic and evolutionary information about *C. hystrix* has hindered its further utilization in cucumber breeding and evolutionary research.

In *Cucumis*, numerous sequences of LTR retrotransposons were isolated and characterized from *Cucumis melon* genome, and the transcription of a *copia* retrotransposon Reme1 could be induced by UV light (Ramallo et al., 2008). In our former research, we isolated Ty1-*copia* retrotransposons from *Cucumis sativus* (Jiang et al., 2008). However, little is known about Ty1-*copia* retrotransposons and their genetic distribution in *C. hystrix*. The genome characterization with reference to the content, heterogeneity, and overall distribution of retrotransposons will contribute to our thorough understanding about *C. hystrix* genome organization and its evolution. Therefore, in the present study, we isolated the reverse transcriptase sequences of Ty1-*copia* retrotransposons in *C. hystrix* genome, and investigated their sequence heterogeneity, phylogenetic relationships, genetic distribution, copy numbers, and transcriptionally activity.

2. Materials and methods

2.1. Plant material and isolation of nucleic acids

Seedlings of *C. hystrix* Chakr. ($2n = 2x = 24$) were grown in the research experiment field of Nanjing Agricultural University. Total DNA was extracted from the fresh leaves using the CTAB method described by Murray and Thompson (1980) with some modifications.

Seedlings with two fully expanded leaves were used as experimental material for cold and heat treatment respectively. Some seedlings were transferred to a growth chamber and the temperature was set at 15 °C (day)/8 °C (night) lasted 1 day for chilling treatment. Meanwhile, others were laid in another chamber and treated at 38 °C for 2 h. Total RNA was isolated from fresh leaves of cold and heat treated plants using the Trizol Kit (Promega, USA) according to the manufacturer's instructions. RNA quality was checked by running samples on a formaldehyde agarose gel. The residual DNA was removed by DNase I (RNase-free) (Takara, Japan), and RNA concentration was estimated using a spectrophotometer.

2.2. Polymerase chain reaction (PCR) and cloning of PCR products

Using total genomic DNA as a template, the RT domains of Ty1-*copia* group retrotransposons was amplified by PCR with the degenerate primes Co1 5′-CAN GCN TTY YTN CAY GG-3′ encoding TAFLHG and Co2 5′-ARC ATR TCR TCN ACR TA-3′ encoding YVDDML (Kumar et al., 1997). PCR amplifications were performed in 20 μl reaction mixtures with 50 ng gDNA, 50 pmol of each primer of Co1 and Co2, 0.2 mmol/l of dNTP, 2.5 mmol/l of $MgCl_2$ and 1 U of Taq polymerase (Takara, Japan). Reactions were denatured at 94 °C for 1 min, followed by 35 cycles of 1 min at 94 °C, 1 min at 42 °C and 1 min at 72 °C, with a final elongation step of 7 min at 72 °C.

The PCR products were separated by electrophoresis on 1% agarose gels in 1 × TAE buffer and visualized under UV light after staining with ethidium bromide.

2.3. Cloning and sequencing

The PCR products of expected size were recovered and purified by DNA Gel Extraction Kit (Karroten, China). The purified fragments were cloned into the vector pMD19-T (Takara, Japan) and then transformed into *E. coli* strain DH5α. Recombinant clones were used directly as templates for PCR to identify the insert sizes in clones, and the clones having expected size were sequenced, which was done by Invitrogen Bio-Technology Co., Ltd., Shanghai, China.

2.4. Sequence analysis

Searching for homologous sequences of our clones were performed with the BLASTN, BLASTX and TBLASTX searching tools as implemented on the National Center for Biotechnology Information (NCBI, http://www.ncbi.nlm.nih.gov/) non-redundant databases. Other RT sequences were used to create a comparative phylogenetic dendrogram. *C. sativus*: SC1 (EU162112), SC2 (EU162114), SC3 (EU162121), and SC4 (EU162124); *C. melon*: CM1 (AM182641), CM2 (AM182630), CM3 (AM182625), and CM4 (CAJ76068); *C. sinensis* (AAT72463); *Oryza sativa* (CAD40121); *Malus* × *domestica* (AAX56355); *Beta vulgaris* (CAA93148); *Nicotiana tabacum* (X13777); *Triticum aestivum* (AAA34262); *Camellia sinensis* (AAY98501); *Populus ciliate* (AAT73707); *Prunus mume* (ABF57057); and *Solanum lycopersicum* (AAC34611). Multiple sequence alignments were performed using Clustal W (Thompson et al., 1994). The construction of phylogenetic trees and bootstrap tests with 1000 replications were carried out using MEGA 4.0 software by NJ method.

2.5. Southern hybridization and dot-blot analysis

To detect the distribution of the LTR retrotransposons in *C. hystrix*, genomic DNA was digested with two restriction enzymes (*Eco*RV and *Hind*III), and then separated on 1% agarose gels and transferred to positively an Immobilon-Ny$^+$ membrane (Millipore, Boston, MA, USA). The 260 bp PCR product was used as a probe, which were labeled using digoxigenin-dUTP by random prime-labeling. Hybridization was carried out at 65 °C for 18 h, with washing conditions as recommended by the manufacturer (DIG HIGH Prime DNA Labeling and Detection Starter Kit II; Roche, Mannheim, Germany). Hybridization was performed at 42 °C overnight, while membranes were washed twice with 2 × SSC/0.1% SDS at 65 °C.

The copy number of the Ty1-*copia* retrotransposons was estimated using Southern dotblotting hybridization. The copy numbers was calculated using the method described by MacRae (1998) (copy number = the size of haploid genome × average proportion of nuclear genomic DNA hybridizing to the probe/the size of probe). Serial dilutions of *C. hystrix* genomic DNA (300 ng, 200 ng, 100 ng, 80 ng, 50 ng) and the RT gene fragments of Ty1-*copia* retrotransposons (1000 pg, 500 pg, 400 pg, 250 pg, 100 pg) were spotted on an Immobilon-Ny$^+$ membrane (Millipore, Boston, MA, USA). The PCR product of RT fragments was used as a probe. Hybridization and washing conditions were the same as described for Southern hybridization. The hybridization signals were subsequently detected with the Ultraviolet Transilluminator Bioimaging Systems equipped with Gel Base-Gel software (UVP, USA), and analyzed on a Vision Works LS Image Acquisition and Analysis Software Ver5.5.5. Then two linear regression equations reflected hybridiza-

tion signals of genomic DNA and RT genes were calculated and the corresponding were used to estimate the copy number of the sequenced probe in the sample of genomic DNA.

2.6. Chromosome preparation and FISH

Fresh roots of *C. hystrix* were harvested and fixed in freshly prepared 3/1 mixture of 100% methanol/glacial acetic acid for at least 1 day. The roots were then rinsed for 5 min in distilled water for three times. Subsequently, cut off the root-tips and transferred into enzyme mixtures containing 4% cellulose and 2% pectinase at 28 °C for 2 h. The enzyme solutions were replaced using distilled water and left on ice for 30 min before transferring the root-tips into the fixative solution (100% methanol/glacial acetic acid, 3/1). Completely macerated the root tip on a slide and flame dry the slide. Then check the quality of chromosomes under a phase contrast microscope and stored in −70 °C freezer until use.

FISH was conducted essentially according to published procedures (Jiang et al., 1995). The total RT genes of Ty1-*copia* retrotransposons were labeled with biotin-16-dUTP (Roche) and detected with antibiotin antibody (Roche). Chromosomes were counterstained with DAPI in an anti-fade solution VectorShield (Vector Laboratories, Burlingame, CA), and images were captured using a SenSys CCD camera attached to an Olympus BX60 microscope. The CCD camera was controlled using IPLab Spectrumv3.1 software (Signal Analytics, Vienna, VA) on a Macintosh computer.

2.7. Reverse transcription polymerase chain reaction (RT-PCR)

First-strand cDNA was synthesized from total RNA using ReverTra Ace Kit (M-MLV) following the instruction (Toyobo, Japan). The reverse transcription reaction mixture contained 1–2 μg of total RNA, 1 μl 10 mmol/l each dNTP, 0.5 μl random primer (50 μmol/l) and 0.5 μl Oligo d(T)18 primer (50 μmol/l) and RNase-free water up to 10 μl. The mixture was incubated at 65 °C for 5 min, and subsequently kept in ice water immediately for 2 min. Then 4 μl 5 × first-strand synthesis buffer, 1 μl RNase inhibitor, 1 μl reverse transcriptase, and 4 μl RNase-free water were added. The reaction mixture were incubated at 25 °C for 10 min, then at 42 °C for 60 min, and finally kept in ice water for 2 min. PCR was carried out using the same degenerate primers and cycling conditions as above, except that 1 μl of first-strand cDNA was used as the template. Meanwhile, water was used as a negative control and gDNA as a positive control.

3. Results

3.1. Ty1-copia RT sequences and alignments

Degenerate primers Co1 and Co2 based on conserved amino acid sequence domains of Ty1-*copia* RT yielded an expected 260 bp fragment. When recovered and cloned, a total of 35 clones were randomly selected for sequencing. When blast and compared with each other, 24 unique clones (named as CORT1 to CORT24) were obtained. Sequences ranged in length from 263 bp to 270 bp. These sequences were AT-rich, with AT content ranging from 59.2% and 66.5% (Table 1). Homology between these nucleotide sequences ranged from 51.0% to 98.5%. The conceptual open reading frames (ORFs) of the sequences corresponded to the conserved motifs expected by Ty1-*copia* retrotransposons. These sequences have been submitted to GenBank with the accession numbers of GU569968–GU569991.

The putative Ty1-*copia* group RT sequences were translated into amino acids and where necessary, were edited for frameshift, based on published residue landmarks of plant retrotransposon RTs. Among the 24 sequences, 11 (46%) contain stop codons and/or frameshift mutations, while the remaining 13 (54%) sequences are intact, which means there is neither a stop codon nor frameshift within the coding regions when translated (Fig. 1). The homology matrix between individual sequences ranged from 38.8% to 100%.

Table 1
The nucleotide composition of reverse transcriptase domains of Ty1-*copia* retrotransposons in *C. hystrix*.

Sequence no.	Size (bp)	A + T content (%)	Accession number
CORT1	263	64.3	GU569968
CORT2	266	64.7	GU569969
CORT3	270	65.2	GU569970
CORT4	266	66.2	GU569971
CORT5	266	63.5	GU569972
CORT6	269	60.6	GU569973
CORT7	266	63.5	GU569974
CORT8	264	63.3	GU569975
CORT9	263	64.6	GU569976
CORT10	270	65.6	GU569977
CORT11	270	65.9	GU569978
CORT12	266	60.5	GU569979
CORT13	266	66.2	GU569980
CORT14	266	65.0	GU569981
CORT15	266	66.5	GU569982
CORT16	264	62.5	GU569983
CORT17	263	65.4	GU569984
CORT18	266	64.3	GU569985
CORT19	265	63.4	GU569986
CORT20	264	62.5	GU569987
CORT21	265	59.2	GU569988
CORT22	263	61.6	GU569989
CORT23	266	65.8	GU569990
CORT24	263	65.8	GU569991

3.2. Phylogenetic analysis of Ty1-copia RT sequences from C. hystrix

To study the relationships among the obtained RT sequences representing the Ty1-*copia* retrotransposons in *C. hystrix*, a neighbor-joining tree was constructed by MEGA 4.0 software. These RT sequences were divided into five groups (Fig. 2). Group 1 is the largest one, which contained 17 sequences, and its members were very similar to each other, indicating that they were recently duplicated. Two groups (2 and 4) each containing one single sequence were separated as independent groups because of their long-distance to all the other sequences.

For comparative purpose, sequences of Ty1-*copia* group RT sequences from other species identified from GenBank database were included in the phylogenetic analysis (Fig. 3). The comparative tree sustained the five groups identified in Fig. 2. The first group in Fig. 3 seemed to be associated with EU162124 isolated from *C. sativus*, and EU162121 also from *C. sativus* was associated with group 5. Group 4 appeared to be associated with EU162112 from *C. sativus* as well as ABF57057 isolated from *Prunus mume*. Besides, the RT sequences we cloned were divided into different group, which showed the high diversity of retrotransposon within *C. hystrix*. From these results, we concluded that horizontal transmission had happened during the evolution of these retrotransposons in *C. hystrix*.

3.3. Southern hybridization and dot-blot analysis

The genomic organization of Ty1-*copia* retrotransposons was revealed by Southern hybridization. Total genomic DNA from *C. hystrix* was digested with the *Eco*RV and *Hind*III restriction enzyme, and then probed with the total *C. hystrix* RT-PCR products. A strong hybridization signal was obtained from genomic DNA, indicating the presence of high copy numbers of homologous sequences in *C. hystrix* (Fig. 4). A smear was present over the whole track, indicating

```
Group 1
CORT9     TAFLHGY-LDETIYMVQPKGFEVQGKEDL- -YCLLKKSIYRLKQSPRCWYRRFD-DFIASLGFQRSSYDMCVYINS-TCKDKVYLLLYVDDM
CORT17    TAFLHGY-LDETIYMVQPKGFEVQGKEDL- -YCLLKKSIYRLKQSPRCWYRRFD-DFIASLGFQRSSYDMCVYINS-TCKDKVYLLLYVDDM
CORT24    TAFLHGY-LDETIYMVQPKGFEVQGKEDL- -YCLLKKSIYRLKQSPRCWYRRFD-DFIASLGFQRSSYDMCVYINS-TCKDKVYLLLYVDDM
CORT5     TAFLHGY-LEETIYMV*PKGFEVQGKENL- -YCLLKRSIYGLKQSPRCWYRRFD-DFIASLGFQRSSYDMCVYINSTTYKDKVYLLLYVDDM
CORT14    TAFLHGY-LEETIYMV*PKGFEVQGKENL- -YCLLKRSIYGLKQSPRCWYRRFD-DFIASLGFQRSSYDMCVYINSTTYKDKVYLLLYVDDM
CORT7     TAFLHGY-LEDTMYMVQPKGFEVQGKEDL- -YCLLKKSIYGLKQSPRCWYRRFD-DFIASLGFQRSSYDMCVYINSTTYKDKVYLLLYVDDM
CORT3     TAFLHGYLLDETIYMVQPKGFEVQGKEDL- -YYLLKKSISGLKQSPRCWYRRFD-DFIASLSFQISSYDMCVYINPTTYKDKVYLLLYVDDM
CORT10    TAFLHGYLLDETIYMVQPKGFEVQGKEDL- -YYLLKKSISGLKQSPRCWYRRFD-DFIASLSFQISSYDMCVYINPTTYKDKVYLLLYVDDM
CORT11    TAFLHGYLLDETIYMVQPKGFEVQGKEDL- -YYLLKKSISGLKQSPRCWYRRFD-DFIASLSFQISSYDMCVYINPTTYKDKVYLLLYVDDM
CORT18    TAFLHGY-LDQTIYMVQPKGFEVQGKEDL- -YCLLKKSICGLKQSPRCWDRRFD-DFIASLSFQRSSYDMCGYINSTTYKEKVYLLLYVDDM
CORT4     TAFLHGY-LDQTIYMVQPKGFEVQGKEDL- -YYLLKKSIYGLKQSPRCWCRRFD-DFIASLGFQRSSYDMCVYINSTTYKDKVYLLLYVDDM
CORT2     TAFLHGY_LDEIIYMVQPKGFEVQGKEDL- -YYLLKKLIYGLKQLPRCWYRRFD-DFIASLGFQRSSYDMCVYINSTTYKGKVYLLLYVDDM
CORT8     TAFLHGY-LDETIYMEQPKGFEVQGKEDL- S -LLKKSIYGLKQSPRCW*RRFN-DFIASLGFQRSSYDTCVYINSTSYKDKVYLLLYVDDM
CORT16    TAFLHGY-LDETIYMEQPKGFEVQGKEDL- S -LLKKSIYGLKQSPRCW*RRFN-DFIASLGFQRSSYDTCVYINSTSYKDKVYLLLYVDDM
CORT20    TAFLHGY-LDETIYMEQPKGFEVQGKEDL- S -LLKKSIYGLKQSPRCW*RRFN-DFIASLGFQRSSYDTCVYINSTSYKDKVYLLLYVDDM
CORT19    TAFLHGY-LVETIYMVQPQSFEAQGKADL- -YCLLKKSIYRLKQSPRCWYKRFD-DFIASLGFQRSSYDMCGYINSTTYKDKVYLLLYVDDM
Group 2
CORT13    TAFLHGD-LTEDIYMNQPQGYIEKGKEDQ- -VCYLKKSIYGLKQSPRCWYKRFD-DFISKLGFNKSSYDSCAFINSNSYSSKVYLLLYVDDM
CORT23    TAFLHGD-LTEDIYMNQPQGYIEKGKEDQ- -ICYLKKSIYGLKQSPRCWYKRFD-DFISKLGFNKSSYDSCAFINSNSYSSTVYLLLYVDDM
CORT15    TAFLHGD-LAEDIYMNQPQGYIEKGKEDQ- -VRYLKKSIYGLKQSPRCWYKRFD-DFISKLGFNKSSYDSCAFINSTSYSSKVYLLLYVDDM
Group 3
CORT21    TAFLHGD-LQKEVYMEQPPGFVAQGESDK- -VCRLRKSLYSLKQSPRA*FGKFSSQALVCFGMKKSTSDHSVFYCR-SDNGIVLLVVYVDDM
CORT22    TAFLHGD-LQEKVYMEQSPRFVA*RKSDK- -VCRLRKSLYSLKQSPRAWFGKFS-QALVCFSMKTSTSDHSVFYRR-SDNGIVLLVVYVDDM
ungrouped
CORT12    TAFLHGK-LEERIYMAQLEGHLKKRSEDM- -VCLLKKSLCGLK*SPRCWYKRFD-EFISKLKFKRSSHDSGVYIRECEEKGEV*LLMYVDDM
CORT6     TAFLHGI-LTEEVYMEQPVGFMAESISQ-KPLVCKLKKALYGLKQAPRAWFDRLKT-FLLSQG-ITNSKADCLLFMKLTNGSSCYILIYVDDM
```

Fig. 1. Alignment of the conceptual translations of the nucleotide sequences representing the 24 RT sequences of Ty1-*copia* retrotransposons isolated from *C. hystrix*. Gaps and stop codons are indicated as (–) and (*), respectively. Frameshift mutations are underlined.

that the Ty1-*copia* retrotransposons existed widely in the genome of *C. hystrix*.

To estimate its copy number, Southern dot-blot hybridization was carried out. Serial dilutions of total genomic DNA from *C. hystrix* and the RT gene fragments of Ty1-*copia* retrotransposons were used to compare signal intensities (Fig. 5). The 260 bp fragments of RT were used as a probe. As *C. hystrix* has the same number of chromosomes with *C. melon*, we assumed the haploid genome size of *C. hystrix* to be 450 Mb (Arumuganathan and Earle, 1991). According to the signal hybridization intensity, then we calculated that the total number of Ty1-*copia* retrotransposons in *C. hystrix* was approximately 5460. Assuming that all the RT copies detected in the present study were representing full-length retrotransposons and the average size of them was about 7 kb (Hill et al., 2005), Ty1-*copia* retrotransposons accounted for approximately 8.5% of the *C. hystrix* genome.

3.4. Chromosomal localization of Ty1-copia retrotransposons in C. hystrix

To study the physical distribution of Ty1-*copia* retrotransposons on interphase nuclei and metaphase chromosomes of *C. hystrix*, fluorescent *in situ* hybridization was used. Total RT genes representing the entire populations of Ty1-*copia* retrotransposons were used as probe to hybridize to the chromosomes of *C. hystrix*. Sites of hybridization of probe to chromosomes were detected with anti-digoxigenin-FITC antibodies (green) and chromosomes were counterstained with DAPI (blue). From the interphase nuclei it is evident that the Ty1-*copia* retrotransposons are widely distributed on the chromatin with clustering in some areas (Fig. 6A–C). The signal of Ty1-*copia* retrotransposons is more remarkable on metaphase chromosomes which indicated that retroelements of this kind are clustered in all terminal heterochromatin (DAPI-blight regions) regions of *C. hystrix* (Fig. 6D–F). On most metaphase chromosomes the probe of retrotransposons shows a dispersed organization with clustering in only one terminal region, however, in some chromosomes, they are clustered on both terminal regions (Fig. 6F, arrows).

3.5. Transcriptional analysis

The expression of retrotransposons and their transposition frequency in the host genome is regulated (Hirochika, 1993; Hirochika et al., 1996). Most retrotransposons are transcriptionally silent in most plant tissues during development. However, some retrotransposons are transcriptionally activated by various biotic and abiotic stress factors (Grandbastien, 1998). These abiotic stresses included protoplast isolation, cell culture, wounding, methyl jasmonate, $CuCl_2$, salicylic acid and UV light. However, there is little research about the effects of abiotic stresses such as cold and heat treatment on the activation of retrotransposons. So we carried out RT-PCR to investigate whether the cold and heat treatment can induce transcription of Ty1-*copia* retrotransposons in *C. hystrix*. Water was used as a negative control and gDNA as a pos-

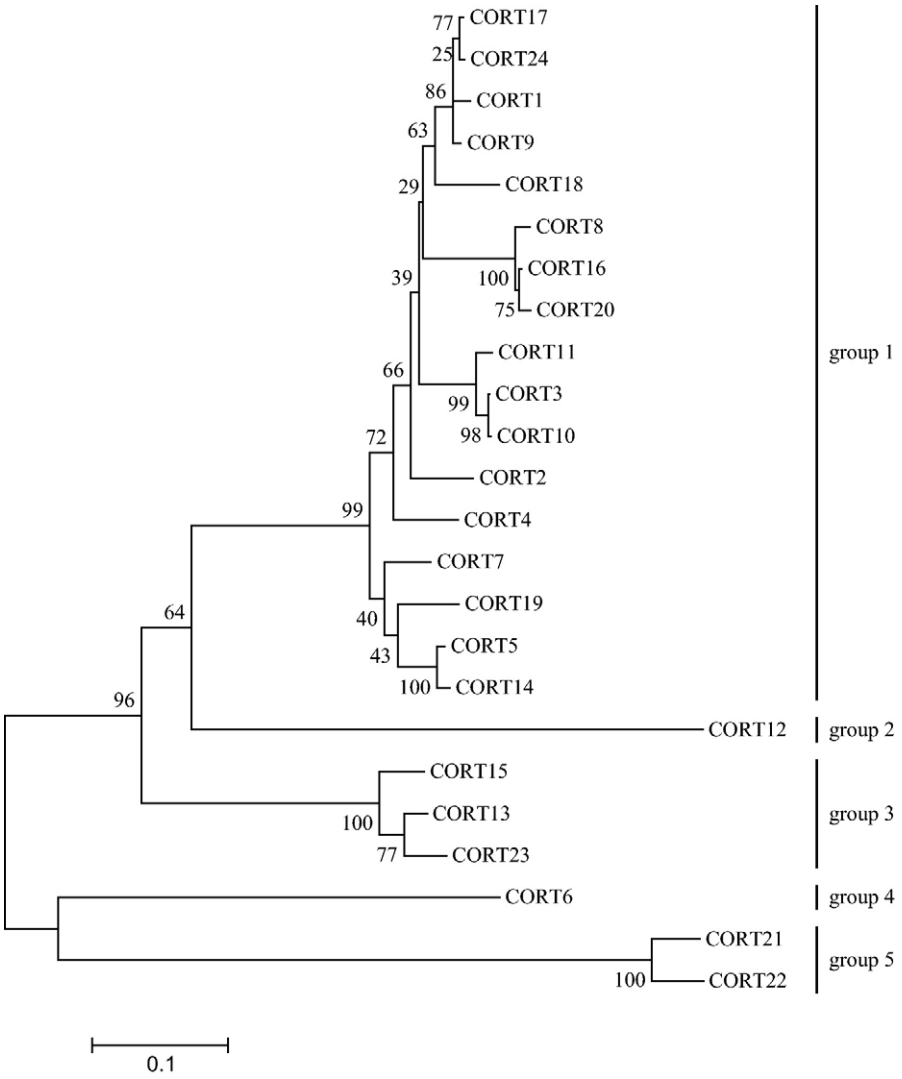

Fig. 2. Phylogenetic analysis of conceptually translated nucleotide sequences representing the Ty1-*copia* RT sequences from *C. hystrix*. The tree was generated using the MEGA program. Numerals at the branch nodes indicate the bootstrap support out of 1000 replications. The branch lengths are proportional to the genetic distances as estimated by the neighbor-joining method (Saitou and Nei, 1987).

itive control. However, no product corresponding to the expected size of Ty1-*copia* retrotransposons was amplified from cDNAs of treated seedlings as well as water, while the expected product was amplified from gDNA. The result indicated that Ty1-*copia* retrotransposons in *C. hystrix* were transcriptionally inactive under these conditions.

4. Discussion

Retrotransposons comprise a significant proportion of plant genomes, which make them one of the most important components affecting the structural and evolution of genomes. To identify Ty1-*copia* retrotransposons in *C. hystrix*, we used a PCR strategy to clone the fragments of reverse transcriptase, by using the degenerate RT primers correspond to two conserved amino acid domains of the Ty1-*copia* retrotransposons, TAFLHG (amino-terminal) and YVDDML (carboxy-terminal) of reverse transcriptase (Kumar et al., 1997). Twenty four RT sequences identified were highly heterogeneous. All the 24 sequences obtained were different from each other. The high diversity of RT sequence of retrotransposon had also been reported in other plants, such as tomato (Rogers and Pauls, 2000), strawberry (Ma et al., 2008), and melon (Ramallo et al., 2008). There could be several reasons for the high sequences heterogeneity. One reason is the high mutation rates associated with retrotransposition, which is the way of retrotransposons proliferation. However, in the process of retrotransposition, there is no proof-reading function of reverse transcriptase, which leads to mutation increases with each replication cycle (Steinhauer and Holland, 1986). Another reason is homologous recombination and frequent mutation of methylated cytosine to thymine in the RT sequences (Heslop-Harrison et al., 1997). As the RT sequences of AT-rich could increase DNA flexibility, the increase of thymine in RT sequence would result in high heterogeneity. In addition, both horizontal and vertical transmissions are also thought to be contributed to the high heterogeneity (Kumar and Bennetzen, 1999). The retrotransposons from *C. hystrix* made clusters, which proved the vertical transmission of retrotransposons. In the meantime, the result of phylogenetic analysis showed that the sequences from *C. hystrix* had high homology with that from other species, such as *C. sativus* and *P. mume* (Fig. 3), which supports the horizontal transmission. Furthermore, the presence of distantly related retrotransposons within *C. hystrix* also supported the horizontal transmission.

Previous studies indicated that the copy number of retrotransposons might have association with the size of plant genomes (Konieczny et al., 1991). Therefore, the study about copy number

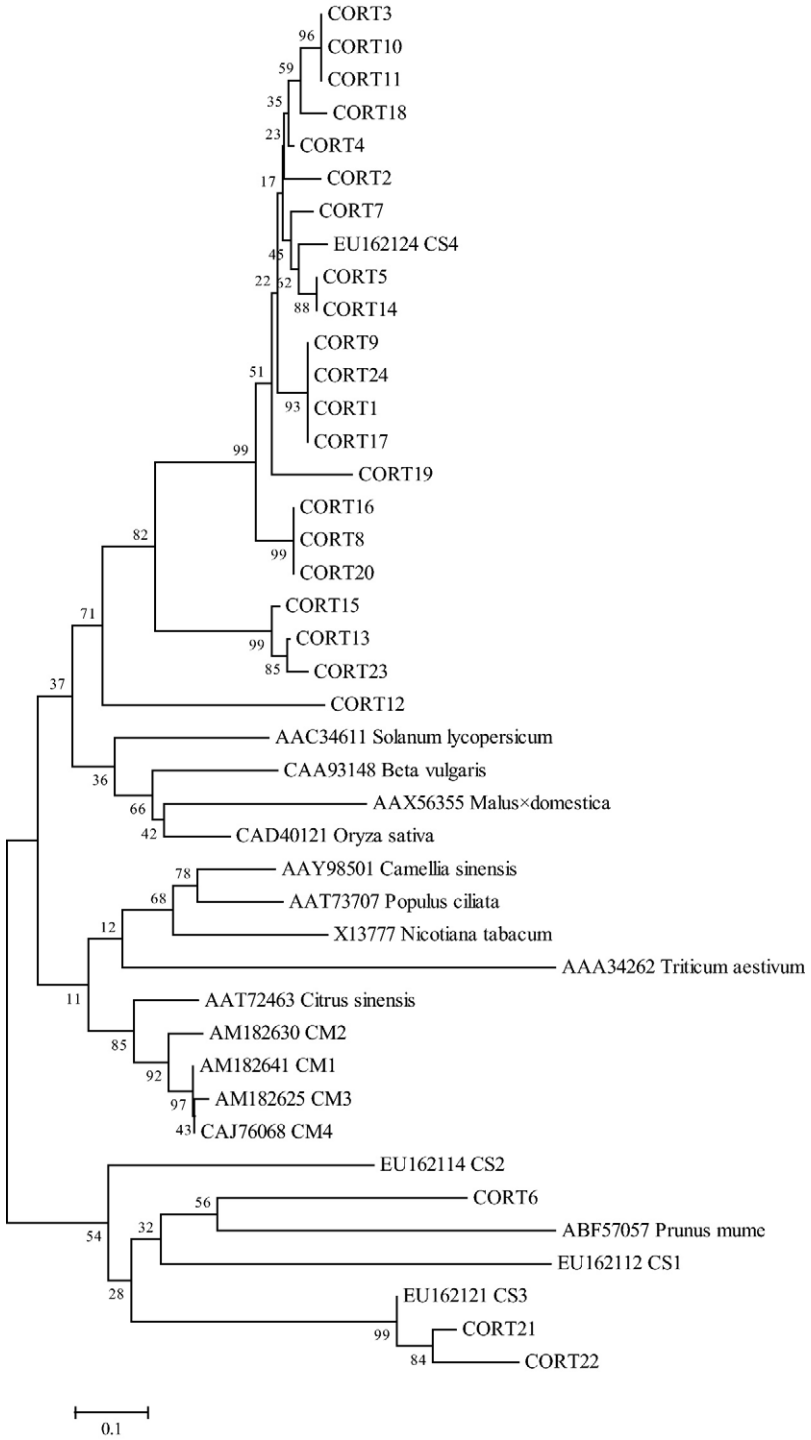

Fig. 3. Phylogenetic relationships among the predicted amino acids for the RT genes of *C. hystrix* and RT sequences of other plants obtained from GenBank using the MEGA 4.0 program. The branch lengths are proportional to the genetic distances as estimated by the neighbor-joining method (Saitou and Nei, 1987). This study employed 1000 bootstrap replication.

of retrotransposons has important reference value to the research of plant genome evolution. The copy number of retrotransposons found in many plants varies from low (1–100 copies) to very high (1000–1,000,000 copies), depending on the host plant species and the types of retrotransposons. In tobacco genome, retrotransposon Tnt1 shows about 100 copies as reported by Grandbastien et al. (1989), while barley genome contains 10,000 copies of BARE-1 (Kalendar et al., 2000). In our present study, when the RT sequences of Ty1-*copia* retrotransposons were used as a probe in Southern hybridization analysis, the hybridization exhibited a smear bands in *C. hystrix* genome, which indicated that Ty1-*copia* retrotransposons displayed multiple copies in *C. hystrix* genome (Fig. 4). The total number of Ty1-*copia* retrotransposons is approximately 5460 and comprise about 8.5% of the *C. hystrix* genome. The high copy number might indicate the transposition activity of retrotransposons during at least part of their evolutionary process. These results also indicated that Ty1-*copia* retrotransposons might play a significant role in *C. hystrix* genome evolution.

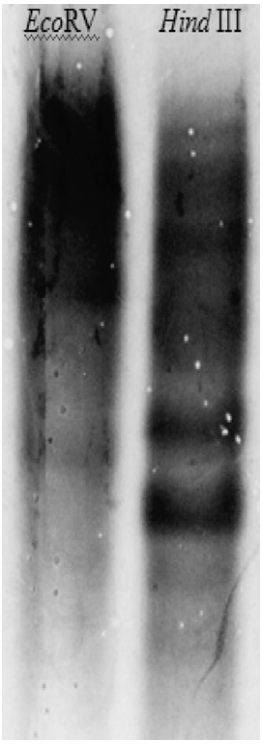

Fig. 4. Southern hybridization analysis of *C. hystrix* genomic DNA, digested with *Eco*RV and *Hind*III and probed with the total RT-PCR product of Ty1-*copia* retrotransposons from *C. hystrix*.

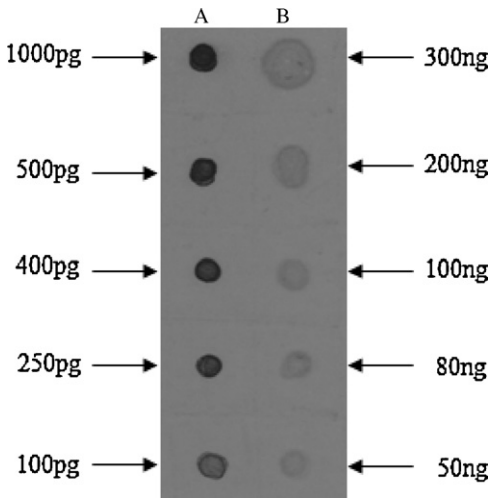

Fig. 5. Dotblots for estimating the copy number of Ty1-*copia* retrotransposons in *C. hystrix*. Different amounts of Ty1-*copia* PCR products were dotblotted on row A, while row B contains serial dilutions of genomic DNA of *C. hystrix*. The PCR product of RT sequences was used as a probe.

Chromosomal distribution of LTR retrotransposons has been analyzed in many different plant species (Katsiotis et al., 1995; Brandes et al., 1997; Friesen et al., 2001; Ruas et al., 2008). The Ty1-*copia* elements in most plants were present along the entire length of chromosomes with possible exceptions of some regions, such as NORs, centromeres, telomeres and heterochromatic regions (Kumar and Bennetzen, 1999). However, in some species it was reported that Ty1-copia elements clustered in certain chromosomal regions (Belyayev et al., 2001; Friesen et al., 2001). This is also the case Ty1-*copia* elements in *C. hystrix*, where the remarkable enrichment of retrotransposons is observed in the terminal heterochromatic (DAPI-blight) regions. This indicates the important role of retrotransposons of this kind in terminal heterochromatin formation of this species.

Because of the presence of stop codons and frameshits mutations in the coding regions, most of them are inactive during plant

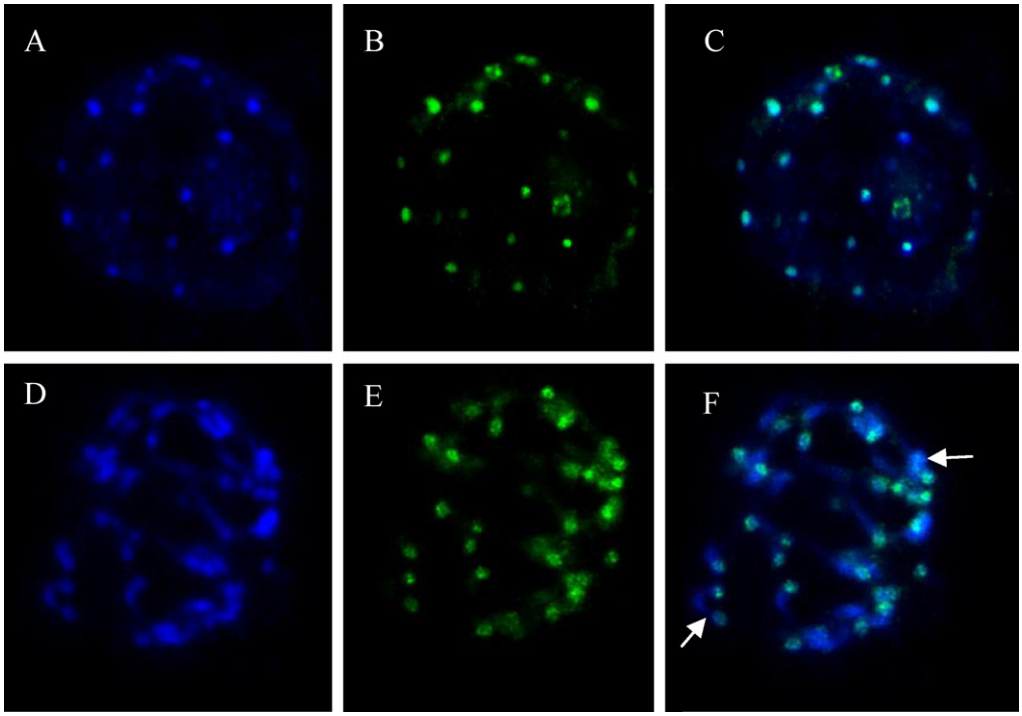

Fig. 6. Localization of Ty1-*copia*-like retrotransposons sequences on interphase nuclei and metaphase chromosomes of *C. hystrix* by fluorescence *in situ* hybridization. Nucleus and metaphase chromosomes were stained with DAPI (A and D) (green). Total RT-PCR product from *C. hystrix* probed on interphase nucleus (B) and metaphase chromosomes (E) (blue). C and F showed the overlaid images. Arrows indicated chromosomes showing signal accumulation on both terminal heterochromatins. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

development, even that their genomes contain high copies of retrotransposons. In strawberry, 5 of 19 RT gene fragments of Ty1-*copia* retrotransposons contained stop codons and/or frameshifts (Ma et al., 2008), while 51% of the Ty1-*copia* RT sequences presented stop codons and/or frameshifts in persimmon (Nakatsuka et al., 2002). In this study, about 46% of Ty1-*copia* RT sequences contained stop codons and/or frameshifts when translated. Furthermore, no transcriptional fragment corresponding to the RT domains of Ty1-*copia* retrotransposons was detected from leaves of chilling and heat treated seedlings. The result indicated that chilling and heat stress treatments is insufficient to induce activation of Ty1-*copia* retrotransposons in *C. hystrix* leaves.

In conclusion, we have showed that Ty1-*copia* retrotransposons in *C. hystrix* genome were highly heterogeneous and examined their phylogenetic relationships, genetic distribution, copy numbers and transcriptional activity. These results provide basic genetic and evolutionary information of *C. hystrix* genome, and will provide valuable information for the utilization of retrotransposons in *C. hystrix*.

Acknowledgements

This research was partially supported by the Key Program (30830079) and the General Program (31071801, 30972007) from the National Natural Science Foundation of China; National Basic Research Program of China (973 Program) (2009CB119000); the '863' Program (2008AA10Z150); the National Supporting Programs (2008BADB105) from the Ministry of Science and Technology of China; Ph.D. funding (20070307034, 20090097110024) from the Ministry of Education of People's Republic of China.

References

Arumuganathan, K., Earle, E.D., 1991. Nuclear DNA content of some important plant species. Plant Mol. Biol. Rep. 9, 208–219.
Belyayev, A., Raskina, O., Nevo, E., 2001. Chromosomal distribution of reverse transcriptase-containing retroelements in two *Triticeae* species. Chromosome Res. 9, 129–136.
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.
Brandes, A., Heslop-Harrison, J.S., Kamm, A., Kubis, S., Doudrick, R.L., Schmidt, T., 1997. Comparative analysis of the chromosomal and genomic organization of Ty1-*copia*-like retrotransposons in pteridophytes, gymnosperms and angiosperms. Plant Mol. Biol. 33, 11–21.
Chen, J.F., Isshiki, S., Tashiro, Y., Miyazaki, S., 1997. Biochemical affinities between *Cucumis hystrix* Charkr and two cultivated *Cucumis* species (*C. sativus* L. and *C melo* L.) based on isozyme analysis. Euphytica 97, 139–141.
Chen, J.F., Moriarty, G., Jahn, M., 2004. Some disease resistance tests in *Cucumis hystrix* and its progenies from interspecific hybridization with cucumber. In: Proceeding of 8th RUCARPIA Meeting on *Cucubit* Genetics and Breeding, pp. 189–196.
Chen, J.F., Zhang, S.L., Zhang, X.G., 1994. The Xishuangbanna gourd (*Cucumis sativus* var. *xishuangbanna* Qi et Yuan), a traditionally cultivated plant of the Hanai people, Xishuangbanna, Yunan, China. Cucurbit Genet. Coop. Rpt. 17, 18–20.
Feschotte, C., Jiang, N., Wessler, S.R., 2002. Plant transposable elements: where genetics meets genomics. Nat. Rev. Genet. 3, 329–341.
Flavell, R.B., 1986. Repetitive DNA and chromosome evolution in plants. Philos. Trans. R. Soc. Loond. B. Biol. Sci. 312, 227–242.
Friesen, N., Brandes, A., Heslop-Harrison, J.S., 2001. Diversity, origin, and distribution of retrotransposons (*gypsy* and *copia*) in conifers. Mol. Biol. Evol. 18, 1176–1188.
Grandbastien, M.A., 1992. Retroelements in higher plants. Trends Genet. 8, 103–108.
Grandbastien, M.A., 1998. Activation of plant retrotransposon under stress conditions. Trends Plant Sci. 3, 181–187.
Grandbastien, M.A., Spielmann, A., Caboche, M., 1989. Tnt1, a mobile retroviral-like transposable element of tobacco isolated by plant cell genetics. Nature 337, 376–380.
Heslop-Harrison, J.S., Brandes, A., Taketa, S., Schmidt, T., Vershinin, A.V., Alkimova, E.G., Kamm, A., Doudrick, R.L., Schwarzacher, T., Katsiotis, A., Kubis, S., Kumar, A., Pearch, S.R., Flavell, A.J., Harrison, G.E., 1997. The chromosomal distributions of Ty1-*copia* group retrotransposable elements in higher plants and their implication for genome evolution. Genetica 100, 197–204.
Hill, P., Burford, D., Martin, D.M.A., Flavell, A.J., 2005. Retrotransposon populations of *Vicia* species with varying genome size. Mol. Genet. Genomics 273, 209–216.
Hirochika, H., 1993. Activation of tobacco retrotransposons during tissue culture. EMBO J. 12, 2521–2528.
Hirochika, H., Otsuki, H., Yoshikawa, M., Otsuki, Y., Sugimoto, K., Takeda, S., 1996. Autonomous transposition of the tobacco retrotransposon Tto1 in rice. Plant Cell 8, 725–734.
Jiang, B., Lou, Q.F., Diao, W.P., Chen, L.Z., Zhang, W.P., Chen, J.F., 2008. The cloning and analysis of reverse transcriptase of Ty1-*copia*-like retrotransposons in *Cucumis*. Acta Horticult. Sinica 35, 1147–1154.
Jiang, J.M., Gill, B.S., Wang, G.L., Ronald, P.C., Ward, D.C., 1995. Metaphase and interphase fluorescence *in situ* hybridization mapping of the rice genome with bacterial artificial chromosomes. Proc. Natl. Acad. Sci. U. S. A. 92, 4487–4491.
Kalendar, R., Tanskanen, J., Immonen, S., Nevo, E., Schulman, A.H., 2000. Genome evolution of wild barley (*Hordeum spontaneum*) by BARE-1 retrotransposon dynamics in response to sharp microclimatic divergence. Proc. Natl. Acad. Sci. U.S.A. 97, 6603–6607.
Kashkush, K., Feldman, M., Levy, A.A., 2002. Gene loss, silencing and activation in a newly synthesized wheat allotetraploid. Genetics 160, 1651–1659.
Kashkush, K., Feldman, M., Levy, A.A., 2003. Transcriptional activation of retrotransposons alters the expression of adjacent genes in wheat. Nat. Genet. 33, 102–106.
Katsiotis, A., Schmidt, T., Heslop-Harrison, J.S., 1995. Chromosomal and genomic organization of Ty1-*copia*-like retrotransposon sequences in the genus *Avena*. Genome 39, 410–417.
Konieczny, A., Voytas, D.F., Cumings, M.P., Ausubel, F.M., 1991. A super family of *Arabidopsis thaliana* retrotransposon. Genetics 127, 801–809.
Kumar, A., Bennetzen, J.L., 1999. Plant retrotransposons. Annu. Genet. 33, 479–532.
Kumar, A., Pearce, S.R., McLean, K., Harrison, G., Heslop-Harrison, J.S., Waugh, R., Flavell, A.J., 1997. The Ty1-*copia* group of retrotransposons in plants: genomic organization, evolution and use as molecular markers. Genetica 100, 205–217.
Ma, Y., Sun, H.Y., Zhao, G.L., Dai, H.Y., Gao, X.Y., Li, H., Zhang, Z.H., 2008. Isolation and characterization of genomic retrotransposon sequences from octoploid strawberry (*Fragaria ananassa* Duch.). Plant Cell Rep. 27, 499–507.
MacRae, A.F., 1998. A pentamer-repeat-containing DNA sequence in Texas bluebonnet (*Lupinus texensis* Hook.). Genome 41, 553–559.
Murray, H.G., Thompson, W.F., 1980. Rapid isolation of higher weight DNA. Nucl. Acids Res. 8, 4321–4326.
Manninen, I., Schulman, A.H., 1993. BARE-1, a *copia*-like retroelement in barley (*Hordeum vulgare* L.). Plant Mol. Biol. 22, 829–846.
Nakatsuka, A., Iwami, N., Matsumoto, S., Itamura, H., Yamagishi, M., 2002. Ty1-*copia* group retrotransposons I persimmon (*Diospyros kaki* Thumb.). Genes Genet. Syst. 77, 131–136.
Qian, C.T., Chen, J.F., Zhuang, F.Y., Xu, Y.B., Li, S.J., 2002. Several photosynthetic characters of the synthetic species *Cucumis hytivus* Chen and Kirkbride under weak light condition. Plant Physiol. Commun. 38, 336–338.
Ramallo, E., Kalendar, R., Schulman, A.H., Martıínez-Izquierdo, J.A., 2008. Reme1, a *Copia* retrotransposon in melon, is transcriptionally induced by UV light. Plant Mol. Biol. 66, 137–150.
Rico-Cabanas, L., Martínez-Izquierdo, J.A., 2007. CIRE1, a novel transcriptionally active Ty1-*copia* retrotransposon from *Citrus sinensis*. Mol. Genet. Genomics 277, 365–377.
Rogers, S.A., Pauls, K.P., 2000. Ty1-*copia*-like retrotransposons of tomato (*Lycopersicon esculentum* Mill.). Genome 43, 887–894.
Saitou, N., Nei, M., 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. 4, 406–425.
Steinhauer, D.A., Holland, J.J., 1986. Direct method for quantitation of extreme polymerase error frequencies at selected single base sites in viral RNA. J. Virol. 57, 219–228.
Ruas, C.F., Weiss-Schneeweiss, H., Stuessy, T.F., Samuel, M.R., Pedrosa-Harand, A., Tremetsberger, K., Ruas, P.M., Schlüter, P.M., Herrera, M.A.O., König, C., Matzenbacher, N.I., 2008. Characterization, genomic organization and chromosomal distribution of Ty1-*copia* retrotransposons in species of *Hypochaeris* (*Asteraceae*). Gene 412, 39–49.
Thompson, J.D., Higgins, D.G., Gibson, T.J., 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, population-specific gap penalties and weight matrix choice. Nucleic Acids Res. 22, 4673–4680.
Voytas, D.F., Cummings, M.P., Konieczny, A., Ausubel, F.M., Rodermel, S.R., 1992. *Copia*-like retrotransposons are ubiquitous among plants. Proc. Natl. Acad. Sci. U.S.A 89, 7124–7128.
Zhuang, F.Y., Chen, J.F., Qian, C.T., Li, S.J., Ren, G., Wang, Z.J., 2002. Responses of seedlings of *Cucumis hytivus* and progenies to low temperature. J. Nanjing Agric. Univ. 25, 24–30.