LTR retrotransposons cause expression changes of adjacent genes in early generations of the newly formed allotetraploid *Cucumis hytivus*

Li Jia [a,1], Qunfeng Lou [a,1], Biao Jiang [b], Dong Wang [a], Jinfeng Chen [a,*]

[a] *State Key Laboratory of Crop Genetics and Germplasm Enhancement, Nanjing Agricultural University, Nanjing 210095, China*
[b] *Vegetable Research Institute, Guangdong Academy of Agriculture Science, Guangzhou 510640, China*

ARTICLE INFO

ABSTRACT

Cucumis hytivus is a newly synthesized allotetraploid in which retrotransposons was strongly activated owing to allopolyploidization. In present study, the chromosomal distribution of LTR retrotransposons and their effects on expression of adjacent genes were observed by investigating the first four generations of *C. hytivus*. Fluorescent *in situ* hybridization (FISH) analysis revealed that LTR retrotransposons are distributed throughout all the chromosomes. The clusters on terminal regions indicated the high copy number of retrotransposons in *C. hytivus*. The extensive epigenetic changes (gene silencing and gene activation) were detected by cDNA-SSAP analysis. Totally, twenty transcripts subjected to gene expression alterations were sequenced, including 12 gene silencing and 8 gene activation. However, the silenced/activated transcripts consisted of known genes or putative proteins as well as new sequences that had no similarity to any known genes. Both gene silencing and activation mainly occurred in the early generations of allotetraploid. The interlaced distribution of retrotransposons and genes might lead to the transcriptional interference of retrotransposons on expression of their adjacent genes, which may contribute to the rapid genetic stabilization of newly formed allotetraploid.

1. Introduction

LTR retrotransposons are mobile genetic elements which are ubiquitous throughout the plant kingdom and present in high copy numbers in plants, which make them major constituents of plant genomes (Kumar and Bennetzen, 1999). Retrotransposons are normally inactive during plant development, but their expression and mobility might be activated by specific stimuli, such as various biotic and abiotic stresses (Bradshaw and McEntee, 1989; Grandbastien, 1998; Grandbastien et al., 2005; Hirochika, 1993; Strand and McDonald, 1985; Takeda et al., 1998; Yun and Davis, 1989). Since retrotransposons constitute a major fraction of plant genomes, they can be basic resources for illegitimate recombination, which can induce various mutations to plant genomes, such as deletions, insertions, translocations and duplications (Lonning and Saedler, 2002). Consequently, retrotransposons play an important role in genome structure through transposition and recombination (Vitte and Panaud, 2005). Given that the expression of retrotransposons can be induced by different stresses, their insertions may participate indirectly in further regulation of gene expression (Slotkin and Martienssen, 2007).

Allopolyploidy has played an important role in the process of specification in plants. The most common form of polyploidy is allopolyploidy which happens through wide hybridization and polyploidy. The newly formed species is subjected to two kinds of 'genomic shock': hybridization and polyploidy (McClintock, 1984). Allopolyploidization often induced a surprisingly large array of genetic and epigenetic changes. Many researches have focused on the genetic and epigenetic of allopolyploidy, but the mechanism still remains to be figured out, since its complex and elusive (Michaud et al., 1994; Kashkush et al., 2003; Adams and Wendel, 2004; Parisod et al., 2009; Petit et al., 2010). McClintock (1984) postulated that the activation of transposable elements could play an important role in genomic changes and might occur as a response to the 'genomic shock'. This hypothesis is supported by experimental evidence from different allopolyploid plants. Kashkush et al. (2003) found that transcriptional activation of *Wis2-1*A retrotransposon in synthetic allopolyploid from the cross of *Aegilops sharonensis* and *Triticum monococcum* alter the expression of adjacent genes. Petit et al. (2010) studied the mobilization of retrotransposons in synthetic allotetraploid tobacco, and found that major changes (such as retrotransposon amplification and molecular restructuring in or around insertion sites) occurred rapidly in response to allopolyploidy. TE insertions was also found to induce significantly

more frequent methylation changes around the insertions in the young allopolyploid *Spartina anglica* (Parisod et al., 2009). Given the transposable elements (TEs) are highly mutagenic and silenced by overlapping epigenetic mechanisms, they are likely to play a pivotal role in fuelling genome reorganization and functional changes following allopolyploidization (Parisod et al., 2010). Further researches from different species gave inconsistent results (Lim et al., 2007; Parisod et al., 2009; Petit et al., 2010). Therefore, investigations using more allopolyploid species will be helpful to the elucidating of genome reorganization and functional changes during the stabilization of allopolyploid. The newly synthesized allopolyploid should be a very good model system for these studies, as it is young and the genetic information of its parents can be obtained exactly.

Cucumis hytivus ($2n = 4x = 38$) is a newly synthesized allotetraploid where two different genomes, *C. sativus* L. ($2n = 2x = 14$) and *C. hystrix* Char. ($2n = 2x = 24$, a wild species of *Cucumis* originated from Asia) joined into the same nucleus through wide cross and subsequently chromosome doubling (Kirkbride, 1993; Chen et al., 1997; Chen and Kirkbride, 2000). It is a new model system to investigate genome evolution in the early generation of an allopolyploid. Our previous studies demonstrated that various morphological changes occurred (Chen et al., 2002, 2003, 2004), and that extensive genomic and epigenetic changes were triggered in the early generations of this newly synthesized allopolyploid (Chen et al., 2007; Chen and Chen, 2008; Zhuang and Chen, 2009). Furthermore, our recent study found that allopolyploidization could induce the activation of Ty1-*copia* retrotransposons in *C. hytivus* (Jiang et al., 2011). The investigation of the effect of LTR retrotransposons activations on the expression of adjacent genes in the *C. hytivus* could be expected to provide more valuable clue for understanding the formation of the new allopolyploid.

In the present study, to further investigate potential epigenetic changes associated with the the newly synthesized *Cucumis* allotetraploid, we focused on the effects of retrotransposons on gene expression of the first four generations of *C. hytivus* compared with its two diploid parents, following the analysis of chromosome distribution of LTR retrotransposons on this allotetraploid. We employed a cDNA-SSAP strategy for genome-wide display of transcripts flanking LTR retrotransposons to show a close link between LTR retrotransposons and the expression changes of their adjacent genes. Our research may help to understand the phenotypic variation of *C. hytivus* and also accelerate the utilization of allopolyploid for further relative epigenetic studies.

2. Materials and methods

2.1. Plant materials

Plant materials used for the present study consisted of the first four generations of the newly synthesized allotetraploid *C. hytivus* (genome HHCC, $2n = 4x = 38$) and its diploid parents, a common cucumber *C. sativus* cv. Beijingjietou (genome CC, $2n = 2x = 14$) and a wild *Cucumis* species *C. hystrix* (genome HH, $2n = 2x = 24$). Both of the two diploid parents are highly inbred lines. The primary synthetic allotetraploid (S_0) was previously obtained from wide hybridization, and followed by embryo rescue and chromosome doubling by somaclonal variation (Chen and Kirkbride, 2000). The S_0 were maintained by tissue culture and somatic chromosome numbers were determined as 38. Then the allotetraploid had restored fertility and set fruits with viable seeds. Inbred lines of the allotetraploid were subsequently advanced by self-pollination to obtain the first four generations as S_1, S_2, S_3 and S_4 (Chen et al., 2007; Chen and Chen, 2008). Five seedlings selected randomly from each generation along with the two parents were grown in greenhouse

at Jiangpu Horticulture Experiment Station of Nanjing Agricultural University, Nanjing, Jiangsu, China (N 31.2°, E 118.4°).

2.2. Chromosome preparation and FISH analysis

Fresh roots of *C. hytivus* 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 with distilled water for three times. The root tips about 1 mm length were cut and transferred into enzyme mixtures containing 4% cellulose and 2% pectinase at 28 °C for 2 h. The enzyme solution were replaced with distilled water and left on ice for 30 min before transferring into the fixative solution as above. Completely macerated root tips were put on a slide and then flame dried. The quality of chromosomes were checked under an Olympus BX41 contrast microscope and stored in −70 °C freezer until use.

Fluorescence *in situ* hybridization was conducted essentially according to published procedures (Jiang et al., 1995). The total RT genes of both LTR (Ty1-*copia* and Ty3-*gyspy*) retrotransposons were used as probes for *in situ* hybridization (Jiang et al., 2010; unpublished). RT genes of Ty1-*copia* retrotransposons were labeled with biotin-16-dUTP (Roche) and detected with fluorescein isothiocyanate-conjugated anti-biotin antibody (Roche). And the RT genes of Ty3-*gyspy* retrotransposons were labeled with digoxigenin-dUTP and detected with rhodamine-conjugated anti-digoxigenin antibody (Roche). Chromosomes were counterstained with 4, 6-diamidino-2-phenylindole (DAPI) in an anti-fade solution VectorShield (Vector Laboratories, Burlingame, CA), and images were captured using a Photometrics (Tuc-son, AZ) SenSys CCD camera attached to an Olympus BX51 microscope. The CCD camera was controlled using Applied Spectral Imaging FISH view 5.5 software (Applied Spectral Imaging, Inc., USA) on a computer. The optimal image was performed using Adobe Photoshop (Adobe Systems).

2.3. SSAP analysis

Total RNA was isolated from young fresh leaves of five individual seedlings with two fully expanded leaves using the Trizol kit (Promega, USA). First and second strand cDNA synthesis was carried out according to manufacturer's instructions (TaKaRa, Japan). Five microliters of the double-stranded cDNA was checked on 1% agarose gel, and when the expected smear bands ranging between 100 bp and 3000 bp were observed, the rest cDNA was used for SSAP procedure according to Waugh et al. (1997). The LTR primer and adaptor sequences employed in this study were listed in Table 1. L1, L2 and L3 were designed based on the retrotransposon sequences of *Tcs1* (EF122141), *Tcs2* (EF122142) and *Tcs3* (EF122143) in *C. sativus*, respectively (Lou and Chen, 2007). L4 was designed based on the retrotransposon of *Reme*1 in *C. melon* (AM1174993, Ramallo et al., 2008).

2.4. Sequence isolation and analysis

The target fragments were excised from the SSAP gel and dissolved in 100 μl ddH$_2$O, incubated at 100 °C for 15 min and deposited slowly at room temperature. Then 3 μl of the sample was directly used as a DNA template using the same primers and protocol as pre-amplification. The PCR products were visualized and quantified on 1% agarose gels in 1× TAE buffer and visualized under UV light after staining with ethidium bromide. The expected fragments were recovered by DNA Gel Extraction Kit (Karroten, China) and then ligated into the pMD19-T vector (TaKaRa, Japan) and transformed into *Escherichia coli* strain DH5α. Recombinant clones were directly used as templates for PCR and clones of the expected size were sequenced by Invitrogen Bio-Technology Co., Ltd., Shanghai, China.

Table 1
Pre-amplification and selective amplification primer sequences used in the study.

Type	Primer/adaptor name	Sequence
SSAP adaptors	*Eco*R I-adaptor 1	5′-CTCGTAGACTGCGTACC-3′
	*Eco*R I-adaptor 2	3′-CATCTGACGCATGGTTAA-5′
	Mse I-adaptor 1	5′-GACGATGAGTCCTGA-3′
	Mse I-adaptor 2	3′-TACTC-AGGACTCAT-5′
Primers of pre-amplification	E00	5′-GACTGCGTACCAATTC-3′
	M00	5′-GATGAGTCCTGAGTAA-3′
Primers for SSAP	L1	5′-GTTTACGTCGTCATTCGACCAC-3′
	L2	5′-AATCAAGTTTATATTTCAACAC-3′
	L3	5′-ACAATTCTCAATATATACTCCAG-3′
	L4	5′-ATTCTTCTATCCTTGCCCGTGG-3′
	MCAA	5′-GATGAGTCCTGAGTAACAA-3′
	MCAC	5′-GATGAGTCCTGAGTAACAC-3′
	MCAG	5′-GATGAGTCCTGAGTAACAG-3′
	MCAT	5′-GATGAGTCCTGAGTAACAT-3′
	MCTA	5′-GATGAGTCCTGAGTAACTA-3′
	MCTC	5′-GATGAGTCCTGAGTAACTC-3′
	MCTG	5′-GATGAGTCCTGAGTAACTG-3′
	MCTT	5′-GATGAGTCCTGAGTAACTT-3′
	MCGA	5′-GATGAGTCCTGAGTAACGA-3′
	MCGC	5′-GATGAGTCCTGAGTAACGC-3′
	MCGG	5′-GATGAGTCCTGAGTAACGG-3′
	MCGT	5′-GATGAGTCCTGAGTAACGT-3′

The gene sequences obtained were analyzed for similarity in public database using the BLAST and BLASTX on the National Center Biotechnology Information Server (http://www.ncbi.nlm.nih.gov/) and cucurbit genomic database (http://www.icugi.org/).

2.5. Reverse transcription (RT)-PCR and genomic PCR (gPCR) confirmation

Specific primers were designed according to the representative sequence obtained, using the software Primer 5.0. Specific primers are the following: 5′-ATACGACCAGGTGGAGTG-3′ and 5′-CTCTTCTAATTCGTCGATACG-3′ for IF5; 5′-TGCAGCTCTGTA-CAATTC-3′ and 5′-ATATCATGAAGGAGATAGGG-3′ for IF14 (IF is the abbreviation of 'isolated fragment'). Then the primers were synthesized by Invitrogen Biotech. Total RNA (5 μg) was treated with DNase I (RNase-free) (TaKaRa, Japan) and first cDNA strand was synthesized according to instructions of the manufacturer (TaKaRa, Japan). One microliter of cDNA or 50 ng genomic DNA was used as templates with the following profile: 5 min at 94 °C; 30 s at 94 °C, 30 s at 55 °C, and 30 s at 72 °C, followed by 35 cycles; and a final extension step of 7 min at 72 °C. *Actin* gene was used as controls for expression analysis. The amplified products were analyzed in 2% agarose gel.

3. Results

3.1. Chromosomal distribution of LTR retrotransposons in C. hytivus

The results of FISH showed that the two types of LTR retrotransposons were distributed throughout all the chromosomes of *C. hytivus*, with clusters on the terminal regions (Fig. 1). Most chromosomes showed clusters only on one terminal region, however, some had clusters on both terminal regions (indicated by arrows Fig. 1). Furthermore, among the chromosomes with two clusters, most of them exhibited two intensive signals on both terminal regions, and the rest of chromosomes showed two clusters with one intensive signal on one side and one weak signal on the other side.

3.2. Gene expression alterations in the newly synthesized allotetraploid induced by retrotransposons

To detect the changes of gene expression which is induced by the insertion of retrotransposon in the synthetic allotetraploid *Cucumis hytivus* compared to its two diploid parents *C. hystrix* (HH) and *C. sativus* (CC), cDNA-SSAP technique was employed. One primer was anchored in retrotransposons and another was random primer. The
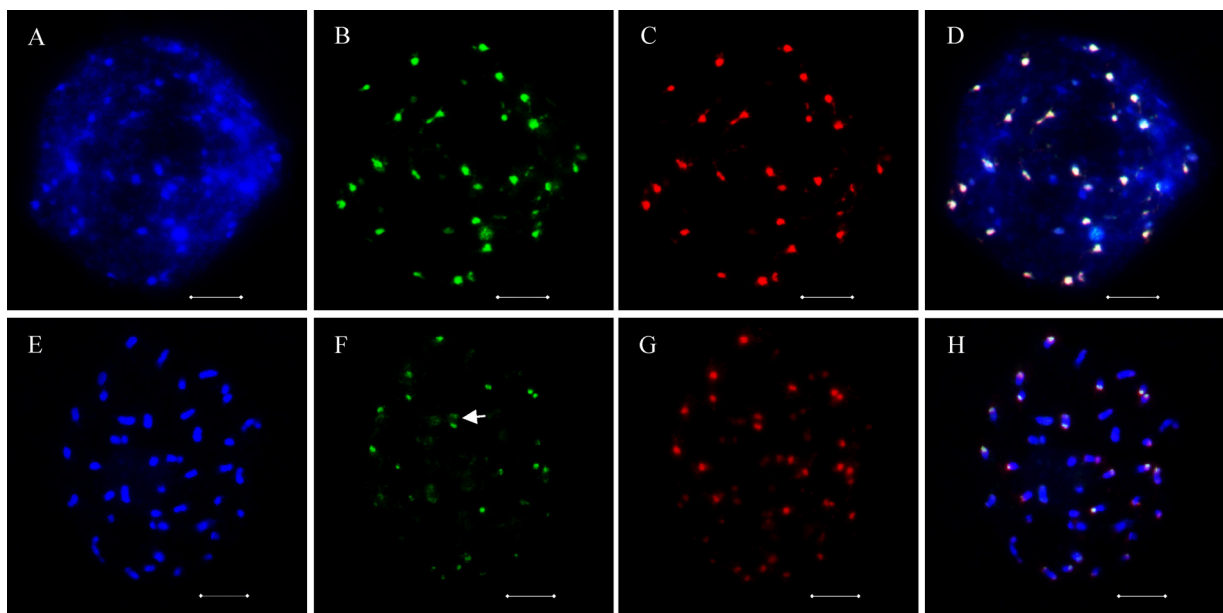

Fig. 1. Distribution of two types of LTR retrotransposons on the somatic mitotic chromosome of *Cucumis hytivus*. The reverse transcriptase genes of Ty1-*copia* and Ty3-*gypsy* retrotransposons were used as probes. The chromosomes were counterstained with DAPI (blue). (A, E) Interphase and metaphase chromosomes were counterstained with DAPI (blue). (B, F) Reverse transcriptase genes of Ty1-*copia* retrotransposons probed on interphase and metaphase chromosomes (green). (C, G) RT genes of Ty3-*gypsy* retrotransposons probed on interphase and metaphase chromosomes (red). (D, H) Interphase and metaphase chromosomes probed with RT genes of both LTR retrotransposons. Arrows pointed at the chromosomes with clusters in both terminal regions. Bars = 5 μm. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 2
cDNA-SSAP analysis in the first four generations of the synthetic allotetraploid compared to their diploid parent.

Transcripts in diploid parents			Observed transcripts in allotetraploid			
Polymorphic	Monomorphic	Total transcripts[a]				
HH	CC			S_1	S_2	S_3	S_4
125	81	912	1118	1091	1086	1087	1087

[a] Each cDNA-SSAP band was considered as a different transcript.

Table 3
Type and number of transcripts that subjected to gene expression changes in the allotetraploid compared to their diploid progenitors.

Transcript absent in allotetraploid				New transcript in allotetraploid			
S_1	S_2	S_3	S_4	S_1	S_2	S_3	S_4
35	42	45	45	8	10	14	14

S_1–S_4 represent the first four generations of the synthetic allotetraploid *C. hytivus*, respectively. We defined that as gene expression alteration only when a new transcript appeared in the allotetraploid that is absent in the diploid parents, or conversely, when a specific transcript present in one of the parents or both disappeared in the allotetraploid. At the same time, only the qualitative differences were investigated and we ignored the difference of strong and weak bands showed by the gel.

Of the 48 pairs of selective primers detected herein, 42 produced good amplification results. In total, 1118 transcripts were amplified using 42 pairs of selective primers in the two diploid progenitors (Table 2). Out of these, 912 transcripts were monomorphic and the polymorphic transcripts in *C. hystrix* and *C. sativus* were 125 and 81, respectively. The total transcripts observed in the first four generations (from S_1 to S_4) of allotetraploid were 1091, 1086, 1087, and 1087, respectively (Table 2). The gene expression patterns were compared across the first four generations of the synthetic allotetraploid and two diploid parents. The number of transcripts that disappeared in the S_1–S_4 generations was 35, 42, 45 and 45, respectively. Meanwhile, the number of new transcripts in S_1–S_4 generations was 8, 10, 4, and 14, respectively (Table 3).

Temporal differences on the type of gene expression changes were observed in the first four generations of allotetraploid. The result revealed that the silencing transcripts existed in the maternal, paternal or both parents (Fig. 2). Moreover, some transcripts were silent in S_1–S_4 generations, whereas others were expressed in one parent and retained in S_1 or S_2 generation, but became silent in subsequent generations (Fig. 2; Table 4).

Among the transcripts that were only expressed in allotetraploid, some of the activated transcripts started from S_1 generation of allotetraploid (Fig. 2D, upper arrow), some from the S_2 generation (Fig. 2E, arrow) and some from S_3 generation (Fig. 2D, lower arrow), and then could be inherited in the later generations. Furthermore, there was one transcript only expressed in the S_2 generation, but silent in both parental lines and other generations of allotetraploid (Table 4). These results indicated that the changes of gene expression mainly occurred in the S_1 and S_2 generation, and subsequently stabilized in the S_3 and S_4 generations.

3.3. Characterization of the altered transcripts detected by cDNA-SSAP technique

The transcripts that exhibited gene expression alteration were recovered and clear sequences were obtained for 20 fragments (see Supplement data). Table 4 summarizes the expression patterns and molecular characterization of the loci with gene expression alterations. Twelve transcripts (IF1-IF12) expressed in either one or both parental lines were silent in the allotetraploid, while eight loci (IF13-IF20) were expressed only in the allotetraploid. Among the silent transcripts, seven (IF1-IF7) were expressed in either one or both parents and could not be detected in the allotetraploid, and five transcripts (IF8-12) were expressed in one parent and inherited in the S_1 or S_2 generation of the allotetraploid, but disappeared in the subsequent generations (Table 4). Among the 8 activated transcripts, one (IF13) was only expressed in S_2 generation of *C. hytivus*, and the expression of the remaining 7 transcripts started from the S_1 generation (IF14-16), the S_2 generation (IF17), or the S_3 generation (IF18-20), and then stably inherited in the subsequent generations (Table 4).
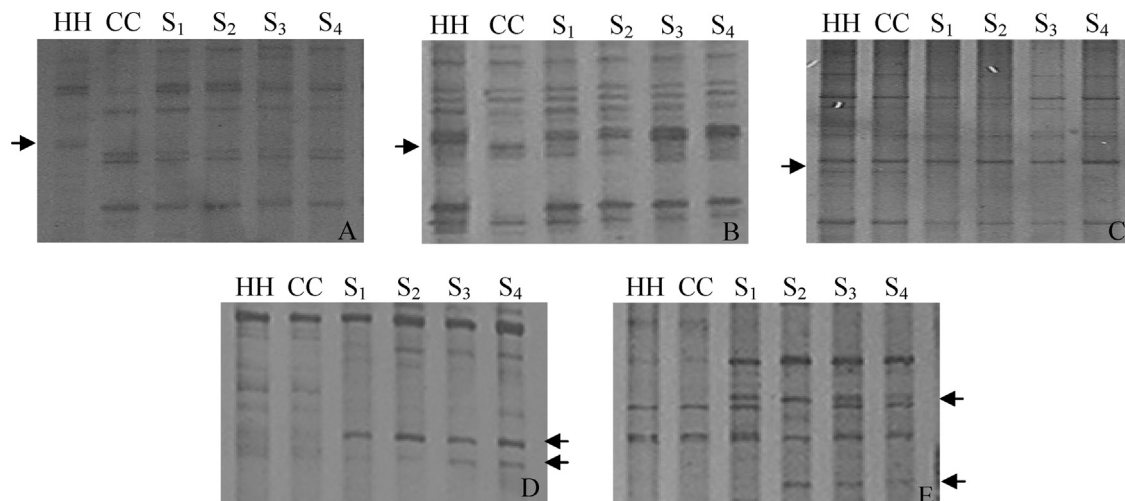

Fig. 2. cDNA-SSAP patterns detected in the two diploid parents, *C. hystrix* (HH), and *C. sativus* (CC), and the first four generations of the synthetic allotetraploid *C. hytivus* (S_1–S_4). (A) Primer L1/MCAC: female parent transcript absent in allotetraploid. (B) Primer L4/MCTA: male parent transcript absent in allotetraploid. (C) Primer L4/MCGC: transcript existed in both parents absent in allotetraploid. (D) Primer L3/MCTG: the allotetraploid gained a new transcript from the S_1 generation and maintained in the S_2–S_4 generations (upper arrow). The allotetraploid gained a new transcript from the S_3 generation and maintained in the S_4 generations (lower arrow). (E) Primer L2/MCGT: the allotetraploid gained a new transcript from the S_1 generation and maintained in the S_2–S_4 generations (upper arrow). The allotetraploid gained a new fragment from the S_2 generation and maintained in the S_3 and S_4 generations (lower arrow).

Table 4
Molecular characterization of cDNA clones obtained by cDNA-SSAP.

Sequence code[a]	Size	The pattterns of transcript[b]						Similarity[c]
		HH	CC	S_1	S_2	S_3	S_4	
IF1	122 bp	+	–	–	–	–	–	Cucumber sterol 3-beta-glucosyltransferase
IF2	93 bp	+	–	–	–	–	–	Cucumber polyadenylate-binding protein 2
IF3	240 bp	+	–	–	–	–	–	Vitis hybrid cultivar transposon Tvv1
IF4	240 bp	+	–	–	–	–	–	Cucumber unigene CU53230
IF5	139 bp	–	+	–	–	–	–	Silene vulgaris NADH dehydrogenase subunit 7 mRNA
IF6	178 bp	+	+	–	–	–	–	Ricinus communis cytokinin dehydrogenase
IF7	120 bp	+	+	–	–	–	–	None
IF8	213 bp	–	+	+	–	–	–	Vitis vinifera hypothetical protein
IF9	256 bp	–	+	+	+	–	–	Cucumber ribosomal RNA processing protein
IF10	155 bp	–	+	+	+	–	–	Putative protein
IF11	147 bp	–	+	+	+	–	–	Putative protein
IF12	248 bp	–	+	+	+	–	–	Arabidopsis thaliana ferritin2
IF13	138 bp	–	–	–	+	–	–	Arabidopsis thaliana short-chain dehydrogenase/reductase
IF14	130 bp	–	–	+	+	+	+	Cucumber transmembrane protein
IF15	164 bp	–	–	+	+	+	+	None
IF16	164 bp	–	–	+	+	+	+	Cucumber calcium-transporting ATPase 4
IF17	147 bp	–	–	–	+	+	+	Cucumber unigene CU15537
IF18	107 bp	–	–	–	–	+	+	None
IF19	93 bp	–	–	–	–	+	+	None
IF20	85 bp	–	–	–	–	+	+	None

[a] IF is the abbreviation of 'isolated fragment'.
[b] +, transcript present; –, transcript absent.
[c] The search was done using the BLAST package and similarity was considered as significant for E-values < e^{-10}.

Sequence analysis revealed that among the 12 silenced genes (Table 4), one transcript (IF7) had no similarity with any known gene. Three (IF8, IF10, and IF11) showed high similarity to the putative proteins. One transcript (IF1) showed high similarity to cucumber sterol 3-beta-glucosyltransferase gene; one (IF2) corresponded to cucumber polyadenylate-binding protein 2 gene; one (IF3) was similar to *Vitis* hybrid cultivar transposon Tvv1; one (IF4) showed high similarity to cucumber unigene CU53230; one (IF5) had a high similarity to *Silene vulgaris* NADH dehydrogenase subunit 7 mRNA. IF6 showed high similarity to *Ricinus communis* cytokinin dehydrogenase. IF9 was similar to cucumber ribosomal RNA processing protein, and IF12 was similar to *Arabidopsis thaliana* ferritin2.

Among the eight transcripts (IF13-IF20) that were expressed only in the allotetraploid, four transcripts (IF15 and IF18-IF20) had no similarity to any known genes. One transcript (IF13) was similar to *Arabidopsis thaliana* short-chain dehydrogenase/reductase; one (IF14) showed high similarity to cucumber transmembrane protein; one (IF16) had high homology to cucumber

calcium-transporting ATPase 4; and the remaining one (IF17) showed high similarity to cucumber unigene CU15537.

3.4. RT-PCR and genomic PCR (gPCR) validation of the expression altered genes

To verify the hypothetical expression changes of corresponding genes, we performed RT-PCR analysis. Two pairs of specific primers were designed based on two representative isolated fragments IF5 (*Silene communis* dehydrogenase subunit 7 mRNA) and IF 14 (cucumber transmembrane protein). In agreement with the results of cDNA-SSAP, the result RT-PCR confirmed that IF5 showed expression activity only in one parental lines *C. sativus*, while IF14 showed expression in all the four generations of *C. hytivus*, but silent in the diploid parents (Fig. 3).

In order to ensure that the observed differentiation of gene expression was caused by gene regulation rather than DNA loss, genomic PCR (gPCR) was carried out using the same primers as RT-PCR analysis. The result showed that both IF5 and IF14 existed in all
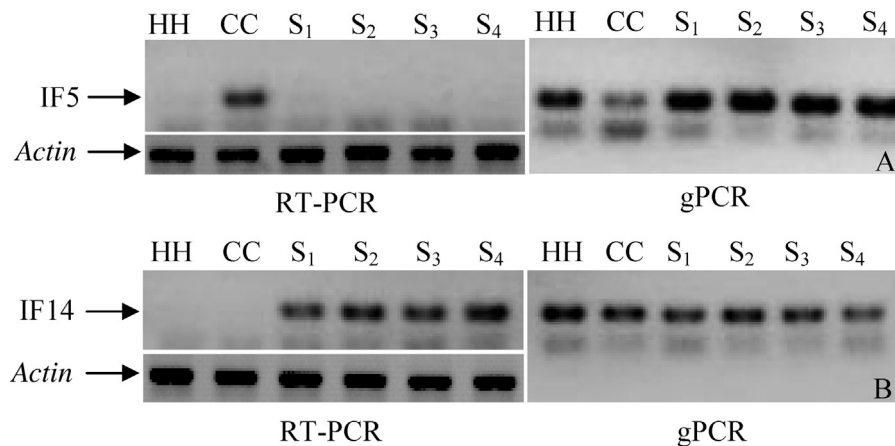

Fig. 3. RT-PCR and genomic PCR analysis of representative transcripts that showed changed expression in the first four generations of *C. hytivus versus* the diploid parents. The IF5 transcript similar to *Silene vulgaris* NADH dehydrogenase (A) and IF14 which had high homology to cucumber transmembrane protein (B) were chosen for validation. *Actin* gene was used as a positive control. HH, *C. hystrix*; CC: *C. sativus*; S_1–S_4: the first four generations of *C. hytivus*.

the materials (the first four generations of *C. hytivus* and diploid parents, *C. sativus* and *C. hystrix*) (Fig. 3), which removed the possibility of DNA loss.

4. Discussion

Retrotransposons are mobile genetic elements that are distributed as interspersed repetitive sequences throughout plant chromosomes (Kumar and Bennetzen, 1999). LTR-retrotransposons are well used as molecular markers in *Cucumis*, since they are ubiquitous and present in high copy numbers as well as their chromosomal distribution (Ramallo et al., 2008; Jiang et al., 2010). In *Cucumis hytivus*, RT genes of LTR retrotransposons are distributed throughout all the chromosomes, with clusters on the terminal regions (Fig. 1), which indicate high copy numbers of retrotransposons in this species. The high copies of LTR retrotransposons may have a genome-wide effect on the expression of flanking genes (Kashkush et al., 2003). In addition, a part of chromosomes with clusters only on one terminal region showed a weak hybridization signal (short arrows in Fig. 1), indicating that the copy number of LTR retrotransposons in these chromosomes is low. The non-uniform distribution of retrotransposons, such as the different hybrid signal intensity and different clusters in terminal regions, might suggests the stronger adaptability of species for the insertion of retrotransposon in heterochromatins regions. Retrotransposons transpose *via* an RNA intermediate, which indicates that they can alter the expression of adjacent genes by insertion interference.

Allopolyploidization was a huge 'genome shock' to host genome, and consequently could induce activation of retrotransposons (McClintock, 1984). In wheat, allopolyploidization activated the transcription of retrotransposons *Wis* 2-1A (Kashkush et al., 2002). In Tobacco, *Tnt1* A retrotransposons were mobilized in response to genome shock of allopolyploidy (Petit et al., 2010). Moreover, transcriptional active of retrotransposons have important influence on the expression of adjacent genes. For example in wheat, Kashkush et al. (2003) reported that the activated retrotransposons *Wis 2-1*A altered the expression of their adjacent genes. Our recent studies revealed that allopolyploidization induced the activation of Ty1-*copia* retrotransposons in *Cucumis* allotetraploid, *Cucumis hytivus* (Jiang et al., 2011). In the present study, we focused on the effect of LTR retrotransposons to the expression of their adjacent genes in the allotetraploid and its diploid parents. Sequence specific amplification polymorphism (SSAP) is the most useful retrotransposon-based marker which had already been developed in *Cucumis* (Lou and Chen, 2007). We employed a cDNA-SSAP strategy for genome-wide display of transcripts flanking LTR retrotransposons to show a close link between LTR retrotransposons and the expression changes of their adjacent genes.

Extensive expression alterations such as silencing and activation of adjacent genes were observed in *Cucumis* allotetraploid when compared with its diploid parents. The number of gene silencing was about three or four fold more than gene activation (Table 4), which indicated that gene silencing was a more remarkable event during the formation of allopolyploid. Most of the silencing genes were present in either one or both parents and the events started from the first two generations of the allotetraploid *C. hytivus*. Subsequently, we carried out RT-PCR and gPCR analysis for validation of the expression alteration of genes detected by cDNA-SSAP, which proved the authenticity of epigenetic changes (Fig. 3). Similarly, activation genes also mainly occurred in the first two generations and subsequently could be stably inherited. These results indicated that the epigenetic changes were mainly present in the early stage of allopolyploid, and it would be stably inherited in the subsequent generation. Our findings are in agreement with previously reported

in other allopolyploids, such as *Arabidopsis* (Comai et al., 2000), wheat (Kashkush et al., 2002, 2003) and cotton (Adams and Wendel, 2004).

All the sequences obtained corresponded to nested transcripts which overlapped sequences of the retrotransposons and the adjacent gene. Among the obtained transcripts, 12 were silent while the remaining 8 were new transcripts only expressed in the newly synthesized allotetraploid. Most of them had high similarity to known genes or putative protein. Similarly, in wheat newly synthesized allotetraploid *Aegilops sharonensis* × *Triticum monococcum*, Kashkush et al. (2003) recovered 22 transcripts that were subjected to expression alterations flanking the transcriptional active retrotransposons *Wis 2-1*A, and found that many of the transcripts showed high similarity to known genes. The interlaced distribution of retrotransposons and genes might lead to transcriptional interference of normal gene expression, or generate new patterns of gene expression, which in turn may contribute to the stabilization of host genome. This point of view is in accordance with the view proposed by McClintock (1984) that transposons act as controlling elements. The genes flanking LTR retrotransposons with altered expression may be beneficial in the stabilization of *Cucumis* allotetraploid, *C. hytivus*.

Acknowledgements

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.scienta.2014.05.022.

References

Adams, K.L., Wendel, J.F., 2004. Exploring the genomic mysteries of polyploidy in cotton. Biol. J. Linn. Soc. Lond. 82, 573–581.

Bradshaw, V.A., McEntee, K., 1989. DNA damage activates transcription and transposition of yeast Ty retrotransposons. Mol. Gen. Genet. 218, 465–474.

Chen, J.F., Kirkbride, J.J., 2000. A new synthetic species of *Cucumis* (Cucurbitaceae) from interspecific hybridization and chromosome doubling. Brittonia 52, 315–319.

Chen, J.F., Luo, X.D., Qian, C.T., Molly, M.J., Staub, J.E., Zhuang, F.Y., Lou, Q.F., Ren, G., 2004. *Cucumis* monosomic alien addition lines: morphological, cytological, and genotypic analyses. Theor. Appl. Genet. 108, 1343–1348.

Chen, J.F., Staub, J.E., Adelberg, J., Lewis, S., Kunlle, B., 2002. Synthesis and preliminary characterization of a new species (Amphidiploid) in *Cucumis*. Euphytica 123, 315–322.

Chen, J.F., Staub, J.E., Qian, C.T., Luo, X.D., Zhuang, F.Y., 2003. Reproduction and cytogenetic characterization of interspecific hybrids from *Cucumis hystrix* Chakr × *C. sativus* L. Theor. Appl. Genet. 106, 688–695.

Chen, J.F., Staub, J.E., Tashiro, Y., Isshiki, S., Miyazaki, S., 1997. Successful interspecific hybridization between *Cucumis sativus* L. and *Cucumis hystrix* Chakr. Euphytica 96, 413–419.

Chen, L.Z., Chen, J.F., 2008. Changes of cytosine methylation induced by wide hybridization and allopolyploidy in *Cucumis*. Genome 51, 789–799.

Chen, L.Z., Lou, Q.F., Zhuang, Y., Chen, J.F., Zhang, X.Q., Wolukau, J.N., 2007. Cytological diploidization and rapid genome changes of the newly synthesized allotetraploids *Cucumis* × *hytivus*. Planta 225, 603–614.

Comai, L., Tyagi, A.P., Winter, K., Holmes-Davis, R., Reynolds, S.H., Stevens, Y., Byers, B., 2000. Phenotypic instability and rapid gene silencing in newly formed *Arabidopsis* allotetraploids. Plant Cell 12, 1551–1568.

Grandbastien, M.A., 1998. Activation of plant retrotransposons under stress conditions. Trends Plant Sci. 3, 181–187.

Grandbastien, M.A., Audeno, C., Bonnivard, E., Casacuberta, J.M., Chalhoub, B., Costa, A.P.P., Le, Q.H., Melayah, D., Petit, M., Poncet, C., Tam, S.M., van Sluys, M.A., Mhiri, C., 2005. Stress activation and genomic impact of *Tnt1* retrotransposons in *Solanaceae*. Cytogenet. Genome Res. 110, 229–241.

Hirochika, H., 1993. Activation of tobacco retrotransposons during tissue culture. EMBO J. 12, 2521–2528.

Jiang, B., Lou, Q.F., Wang, D., Wu, Z.M., Zhang, W.P., Chen, J.F., 2011. Allopolyploidization induced the activation of Ty1-*copia* retrotransposons in *Cucumis hytivus*, a newly formed *Cucumis* allotetraploid. Bot. Stud. 52, 145–1552.

Jiang, B., Wu, Z.M., Lou, Q.F., Wang, D., Zhang, W.P., Chen, J.F., 2010. Genetic diversity of Ty1-*copia* retrotransposons in a wild species of *Cucumis* (*C. hystrix*). Sci. Hortic. 127, 46–53.

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.

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.

Kirkbride, J.J., 1993. Biosystematic Monograph of the Genus Cucumis (*Cucurbitaceae*). Parkway Publishers, Boone, pp. 84–88.

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

Lim, K.Y., Kovarik, A., Matyasek, R., Chase, M.W., Clarkson, J.J., Grand bastien, M.A., Leitch, A.R., 2007. Sequence of events leading to near-complete genome turnover in allopolyploid *Nicotiana* within five million years. New Phytol. 175, 756–763.

Lonning, W.E., Saedler, H., 2002. Chromosome rearrangements and transposable elements. Annu. Rev. Genet. 36, 389–410.

Lou, Q.F., Chen, J.F., 2007. Ty1-*copia* retrotransposon-based SSAP marker development and its potential in the genetic study of cucurbits. Genome 50, 802–810.

McClintock, B., 1984. The significance of responses of the genome to change. Science 226, 792–801.

Michaud, E.J., Van Vugt, M.J., Bultman, S.J., Sweent, H.O., Davisson, M.T., Woychik, R.P., 1994. Differential expression of a new dominant agouti allele (A^{iapy}) is correlated with methylation state and is influenced by parental lineage. Gene Dev. 8, 1463–1472.

Parisod, C., Alix, K., Just, J., Petit, M., Sarilar, V., Mhiri, C., Ainouche, M., Chalhoub, B., Grandbastien, M., 2010. Impact of transposable elements on the organization and function of allopolyploid genomes. New Phytol. 186, 37–45.

Parisod, C., Salmon, A., Zerjal, T., Tenaillon, M., Grandbastien, M.A., Ainouche, M.L., 2009. Rapid structural and epigenetic reorganization near transposable elements in hybrid and allopolyploid genomes in *Spartina*. New Phytol. 184, 1003–1015.

Petit, M., Guidat, C., Daniel, J., Denis, E., Montoriol, E., Bui, Q.T., Lim, K.Y., Kovarik, A., Leitch, A.R., Grandbastien, M.A., Mhiri, C., 2010. Mobilization of retrotransposons in synthetic allotetraploid tobacco. New Phytol. 186, 135–147.

Ramallo, E., Kalendar, R., Schulman, A.H., Martínez-Izquierdo, J.A., 2008. Reme1, a *Copia* retrotransposon in melon, is transcriptionally induced by UV light. Plant Mol. Biol. 66, 137–150.

Slotkin, R.K., Martienssen, R., 2007. Transposable elements and the epigenetic regulation of the genome. Nat. Rev. Genet. 8, 272–285.

Strand, D.J., McDonald, J.F., 1985. Copia is transcriptionally responsive to environmental stress. Nucl. Acids Res. 13, 4401–4410.

Takeda, S., Sugimoto, K., Otsuki, H., Hirochika, H., 1998. Transcriptional activation of the tobacco retrotransposon *Tto1* by wounding and methyl jasmonate. Plant Mol. Biol. 36, 365–376.

Vitte, C., Panaud, O., 2005. LTR retrotransposons and flowering plant genome size: emergence of the increase/decrease model. Cytogenet. Genome Res. 110, 91–107.

Waugh, R., McLean, K., Flavell, A.J., Pearce, S.R., Kumar, A., Thomas, B.B.T., Powell, W., 1997. Genetic distribution of *Bare-1*-like retrotransposable elements in the barley genome revealed by sequence-specific amplification polymorphisms (S-SAP). Mol. Gen. Genet. 253, 687–694.

Yun, Y., Davis, R.L., 1989. Copia RNA levels are elevated in dunce mutants and modulated by cAMP. Nucl. Acids Res 17, 8313–8326.

Zhuang, Y., Chen, J.F., 2009. Changes of gene expression in early generations of the synthetic allotetraploid *Cucumis × hytivus* Chen et Kirkbride. Genet. Resour. Crop. Evol. 56, 1071–1076.