Arch Biol Sci. 2017;69(2):315-322 https://doi.org/10.2298/ABS160326103J

Identification and characterization of *REC66*, a *Ty1-copia*-like retrotransposon in the genome of red flower of *Mirabilis jalapa* L.

Jiang Shunri [1], Liang G. Joshua[2], Feng Haiyan[1], Luo Dan[1], Blake Shester[4], Yang Liang[1], Lu Wen[3], Zhang Suzhi[3], Yang Yi[1] and Liang Peng[1,2,4,*]

[1] *Department of Biochemistry and Molecular Biology, School of Life Sciences, Sichuan University*, Chengdu, Sichuan, P. R. China

[2] *Clover Biopharmaceuticals*, Chengdu, P. R. China

[3] *Maize Research Institute, Sichuan Agricultural University*, Wenjiang, Chengdu, Sichuan, P. R. China

[4] *GenHunter Corporation*, 624 Grassmere Park, Nashville, TN37211, USA

*__Corresponding author__: liangpeng@scu.edu.cn

Received: March 26, 2016; **Revised**: May 30, 2016; **Accepted**: July 11, 2016; **Published online**: November 4, 2016

Abstract: *Mirabilis jalapa* L is the most commonly grown ornamental species of *Mirabilis* and is available in a range of brilliant colors. However, genetic research on *Mirabilis jalapa* L is limited. Using fluorescent differential display (FDD) screening, we report the identification of a novel *Ty1-copia*-like retrotransposon in the genome of the red flower of *Mirabilis jalapa* L, and we named it *REC66* based on its sequence homology to the GAG protein from *Ty1-copia* retrotransposon. Using degenerate primers based on the DNA sequence of *REC66*, a total of fourteen different variants in reverse transcriptase (RT) sequence were recovered from the genomic DNA. These RT sequences show a high degree of heterogeneity characterized mainly by deletion mutation; they can be divided into three subfamilies, of which the majority encode defective RT. This is the first report of a *Ty1-copia* retrotransposon in *Mirabilis jalapa* L. The finding could be helpful for the development of new molecular markers for genetic studies, particularly on the origin and evolutionary relationships of *M. jalapa* L, and the study of *Ty1-copia* retrotransposons and plant genome evolution in the genus *Mirabilis* or family Nyctaginaceae.

Key words: *Mirabilis jalapa* L.; *Ty1-copia* like retrotransposons; *REC66*; SEFA-PCR; reverse transcriptase; fluorescent differential display

INTRODUCTION

Transposable elements (TEs) are classified into Class I transposable elements and Class II DNA transposons based on their transposition mechanisms. Class I transposable elements that mainly consist of retrotransposons is the largest class of transposable elements and the most abundant and widespread components of plant genomes [1]. Based on their mechanistic features, organization and reverse transcriptase (RT) phylogeny, Class I retrotransposons can be divided into five groups: LTR (long terminal repeat) retrotransposons, DIRS (Dictyostelium intermediate repeat sequence)-like elements, PLEs (*Penelope*-like elements), LINEs (long interspersed repetitive elements) and SINEs (short interspersed repetitive elements) [2]. LTR retrotransposons are further classified into *Ty1-copia* and *Ty3-gypsy* groups, which differ in the order of RT and integrase (INT) in the POL locus [2]. The LTRs start with 5'-TG-3' and end with 5'-CA-3', with sizes ranging from a few hundred base pairs to over 5 kb, and they do not encode any known proteins except promoter and terminator cis-acting elements for the transcription of LTR retrotransposons. The LTR retrotransposons encode GAG and POL proteins, the latter of which encodes an aspartic proteinase (AP), reverse transcriptase (RT), RNase H (RH) and integrase (INT). Together, these proteins are involved in the life cycle of LTR retrotransposons [2]. *Ty1-copia* and *Ty3-gypsy* group retrotransposons are ubiquitous in many higher plants genomes and play a crucial role in plant evolution [3-6].

Mirabilis jalapa L is a herbaceous plant that belongs to the Nyctaginaceae family in the order Caryophyllales, and it is one of the oldest flowering plants that has been experimented on. Native to Peru, it was intro-

duced to Europe by the Spaniards in 1596. It has many different color types: pure white, red and yellow; bicolored variegates of white/red, white/yellow or yellow/red and a tricolored variegate of white/yellow/red [7].

Gregor Mendel [8] was the first to perform genetic research on this plant. In a letter to Carl Naegell in 1870, Mendel uncovered the dominant and recessive genes in determining the genetic inheritance of the flower colors red (crimson), yellow and white in *M. jalapa*. However, molecular genetic studies on *M. jalapa* has been rare since then. It was not until recently that the "R" gene responsible for the formation of the red color flowers was identified through both biochemical and molecular biologic analyses [9]. Shortly thereafter, a *dTmj1* transposon interrupting the R gene was discovered to be the cause for the generation of red/yellow mosaic flowers [10]. Here, using fluorescent differential display (FDD) screening for flower-specific transcripts, we report the discovery of a full-length *Ty1-copia*-like retrotransposon from the red flower of *M. jalapa*. The phylogenetic relationship of this type of retrotransposons that reside in the genome of the flower based on the heterogeneity of its conserved RT sequences were also investigated.

MATERIALS AND METHODS

Plant material

The seeds of red-flower *M. jalapa* were scattered and planted in the same bed of soil in early spring. The plants flowered during mid-summer and their flowers were harvested in late summer. A total of 10-12 flowers was taken from the plants, and the styles and anthers with pollen were removed, leaving only the funnel-shaped petals for FDD screening and total genome DNA isolation.

Isolation of nucleic acids

Total genomic DNA was isolated from the fresh red flowers of *M. jalapa* by using the cetyltrimethylammonium bromide (CTAB) method [11]. Total RNA was purified from the red flowers and leaves of *M. jalapa* using RNApure Reagent (GenHunter Corp.) according to the manufacturer's instructions.

Fluorescent differential display (FDD) screening of the red flower of *M. jalapa*

DNase I treatment of RNA prior to FDD was carried out using the MessageClean kit (GenHunter Corp.). FDD-PCR was carried out for the comparison of mRNA expression profiles between the flower and leaves of the red *M. jalapa.* Combinations of one-base anchored oligo (dT) primers and arbitrary 13-mers with fluorescent labels from the RNA spectra kits (GenHunter Corp.) were used for the FDD screening [12]. cDNAs that were amplified from the 3' terminal of the mRNAs were separated on a 6% denaturing polyacrylamide gel, and visualized via fluorescent imaging on a Typhoon 9410 imager (GE biosciences). Bands representing cDNAs of interest and amplified with the anchor primers H-T_{11}C combined with random HAP-66 primer were excised from the gel, cloned into the PCR-TRAP cloning vector (GenHunter Corp.) and then analyzed by DNA sequencing.

Self-formed adaptor polymerase chain reaction (SEFA-PCR) genome walking

Primers used for SEFA-PCR were developed based on the 320 bp C66 cDNA sequence obtained by FDD. SEFA-PCR amplification of the 5'-flanking region was performed by using three gene-specific primers (C66-5-SP1, C66-5-SP2 and C66-5-Hemi-SP3), while the amplification of the 3'-flanking region was performed using nine gene-specific primers (C66-3-1-SP1, C66-3-1-SP2, C66-3-1-Hemi-SP2, C66-3-2-SP1, C66-3-2-SP2, C66-3-2-Hemi-SP3, C66-3-3-SP1, C66-3-3-SP2 and C66-3-3-Hemi-SP3) (Table 1) [13]. The PCR products were purified, cloned into a pKILLIN vector [14] and sequenced. In order to verify SEFA-PCR sequences, we used LA-PCR (TaKaRa) with REC66-F-P and REC66-R-P primers (Table 1). The strategy to isolate the complete *Ty1-copia*-like retrotransposon sequence is summarized in the schematic Fig. 2A.

Polymerase chain reaction (PCR) amplification of *Ty1-copia*-like RT fragments

The degenerate primers RT-F-P and RT-R-P (Table 1) were used to amplify the RT domains of *Ty1-copia* group retrotransposons (corresponding to motif I – TAFLHG and motif III – YVDDML) from the

Table 1. Primers used in this study.

Name	Sequence
H-T11C	5'-AAGCTTTTTTTTTTTC-3'
AP66	5'-AAGCTTGCCTTTA-3'
C66-5'-SP1	5'- GGCGGCAGTATCTTGATCAATTAATAGTTC -3'
C66-5'-SP2	5'- TGGAGAAGCGGGAGTAATAGTGCCATCAAG -3'
C66-5' -Hemi -SP3	5'-ACTTGAACATGACGCCNNNNNNNNNTAACGG-3'
C66-3'-1-SP1	5'- CCGTTACCAATATCCGGCGTCATGTTCAAGT-3'
C66-3'-1-SP2	5'- CACTGGTTATCTTGATGGCACTATTGTTC -3'
C66-3'-1-Hemi -SP3	5'-GAACTATTAATTGATNNNNNNNNNGCCGCC-3'
C66-3'-2-SP1	5'- GACATCACCCGTGTCGAGTTGTTCTGGCTAC-3'
C66-3'-2-SP2	5'- CCCATAAAAATTCTACAATGTGACAACGGTCG-3'
C66-3'-2-Hemi -SP3	5'- GGACCATAAGAACCANNNNNNNNNCTCTTC-3'
C66-3'-3-SP1	5'- TCTCATTCCTGGCAGGTTCATCAGCTTG-3'
C66-3'-3-SP2	5'- GGCTTTATATGGCCTTAAGCAAGCACCTC-3'
C66-3'-3-Hemi -SP3	5'-TGTATCGACAGGGCTNNNNNNNNNCTTATC-3'
REC66-F-P	5'-TTGAAAGACAAATAAGATAAAAGATGAAC-3'
REC66-R-P	5'-GAAGAGTGATGCCGAATGTGTTTGTATT-3'
RT-F-P	5'-ACNGCNTTPyPyTNCAPyGG-3'
RT-R-P	5'-APuCATPuTCPuTCNACPuTA-3'

DNA letter Codes: N=A+T+C+G, Pu=A+G, Py=T+C.

genomic DNA of the red flowers [15]. Reaction mixtures in a final volume of 20 μL contained 10-50 ng of the genomic DNA, 0.2 mmol/L of dNTPs, 50 pmol of each of the primers RT-F-P and RT-R-P, 2.5 mmol/L of $MgCl_2$ and 1 U of rTaq polymerase (TaKaRa). The PCR amplification conditions were: 94°C denatured for 3 min, followed by 35 cycles at 94°C for 30 s, 55°C for 30 s and 72°C for 1 min, with a final elongation step at 72°C for 5 min. The PCR products were purified, cloned and sequenced as above.

Sequencing and phylogenetic analysis

DNA sequencing analysis was performed by INVITROGEN (Shanghai, China). Sequences of *C66*, *REC66*, and *MYRT1* to *MYRT14* were submitted to National Center for Biotechnology Information (NCBI) using Sequin software.

Homology searches were carried out using Blastn and Blastx against the sequences in the GenBank database (http://www.ncbi.nlm. nih.gov). Five typical sequences of retrotransposons used for comparison were obtained from the GenBank database: Tnt1-94 (GenBank: X13777) [16], TLC1.1 (GenBank: AF279585) [17], Sto-4 (GenBank:AF082133) [18], Hopscotch (GenBank: U12626) [19] and *Copia* (GenBank: X02599) [20]. Multiple alignments of the sequences were performed by the ClustalX version1.83 [21]. The nucleotide sequences were translated into amino acid

Fig 1. The 320-bp C66 fragment encodes a GAG peptide from *Ty1-copia* retrotransposon. **A** – Nucleotide sequence of C66 cDNA fragment (320 bp) from the FDD screening. Highlighted in red are the corresponding flanking FDD PCR primers used. **B** – The open reading-frames (ORF) from the 320-bp fragment matching domains UBN2_3 (Accession: pfam14244), GAG polypeptide of LTR *copia* is as indicated.

sequences and the sequence alignments were annotated with DNAMAN (Woffelman C. DNAMAN for Windows. Version; 2004.). The phylogenetic tree of RT sequences was constructed through the maximum likelihood method of the MEGA 5.1 software [22].

RESULTS

FDD screening involving a combination of 250 primer combinations

A 320-bp red flower-specific cDNA fragment that was absent in the leaf was obtained (Fig. 1A) and designated as *C66* (GenBank: KU720516). Database search by BLAST with NCBI showed that the nucleic acid sequence of *C66* had moderate homology (37%) to a putative plant retroelement from *Oryza sativa* (GenBank accession numbers: AAN34944.1 and AAM74347.1). In addition, the conserved domain database (CDD) identified the open reading frames (ORF) of the 320-bp cDNA with a sequence homology to the GAG protein of the *Ty1-copia* type retrotransposon (Fig. 1B). These results reveal that the 320-bp nucleic acid sequence from the FDD is part of the *Ty1-copia* retrotransposon GAG peptide-related gene in the red flower of *M. jalapa*.

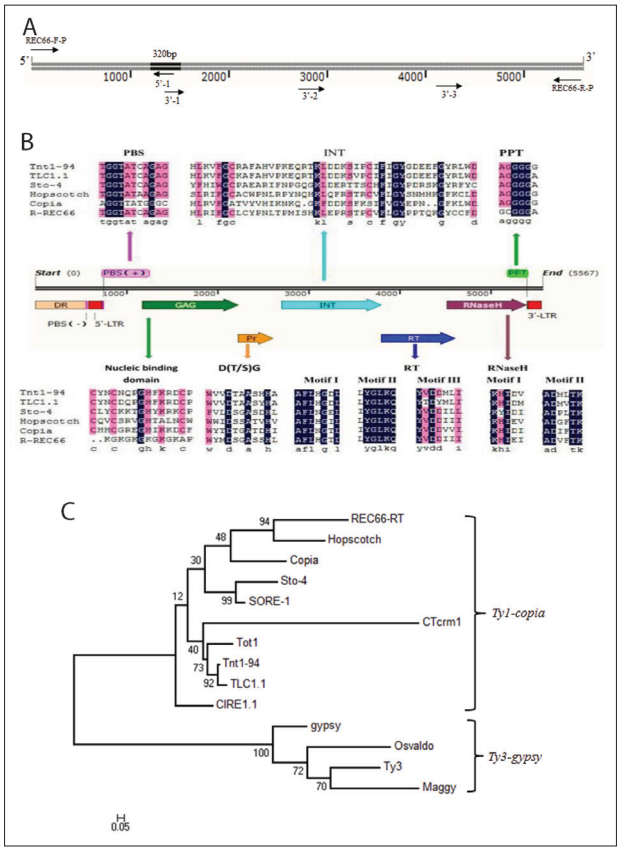

Fig 2. Schematic illustration of *REC66*, a novel *Ty1-copia* retrotransposon from *M. jalapa*. **A** – Schematic diagram of SEFA-PCR used for chromosome walking based on the 320-bp C66 cDNA sequence from the FDD screening. The solid line represents the known 320-bp C66 cDNA sequence, while the dotted line represents the unknown 5' and 3' flanking sequences. 5'-1 represents the first amplification from the SEFA-PCR walking for 5'-flanking sequences; 3'-1, 3'-2 and 3'-4 represent the first, second and third round of amplifications of the SEFA-PCR walking for 3'-flanking sequences, respectively. **B** – Conserved motifs in *REC66* from the red flower of *M. jalapa*. **C** – Maximum likelihood tree based on amino acid sequence of the conserved RT domains from *REC66* and other retrotransposons. The tree was obtained by the maximum likelihood method using MEGA 5.1 software. The numbers at the branch nodes showed bootstrap values with 1000 replications.

After one round of 5' and three rounds of 3' chromosome walking, LA-PCR was performed in order to identify the contiguous longer genomic fragments (Fig. 2A). As a result, a 5567-bp full-length sequence of a *Ty1-copia*-like retrotransposon designated as *REC66* was recovered from the red flower of *M. jalapa* (GenBank: KU720517). Based on the PR-INT-RT-RNaseH gene order and the POL gene-encoding protein sequence homology, the *REC66* retrotransposon belongs to typical *Ty1-copia* retrotransposons [2]. The structure of the *REC66* sequence contains the characteristics of the primer binding site (PBS) sequence TGGT, polypurine tract (PPT) sequence AGGGG, and LTR boundary motif 5'-TG-CA-3'. It also has conserved motifs, including: C-C-H-C motif in GAG; TSG sites in PR; G-K-F-GY-G motif in INT; motif I (AFLGL), motif II (YGLKQ), motif III (YVDDI) in RT; motif I (KHI) and motif II (ADTK) in RNaseH (Fig. 2B). In addition, sequence homology analysis and phylogeny relationship study between *REC66* and other *Ty1-copia* retrotransposons in RT indicated that the REC66 retrotransposon indeed belongs to the *Ty1-copia* type of retrotransposons (Fig. 2C).

Table 2. Base composition of reverse transcriptase (RT) of *Ty1-copia* retrotransposon in red flower of *M. jalapa*.

Sequence NO	Base composition				Size (bp)	AT/GC
	A (%)	C (%)	G (%)	T (%)		
MYRT1	30.0	15.6	24.3	30.0	263	1.5
MYRT2	28.5	18.5	24.2	28.8	260	1.34
MYRT3	19.8	20.2	21.3	38.8	263	1.41
MYRT4	28.2	16.6	24.3	30.9	259	1.44
MYRT5	27.7	18.8	20.0	33.5	260	1.58
MYRT6	31.6	15.6	23.2	29.7	263	1.58
MYRT7	31.2	16.2	21.5	31.2	260	1.66
MYRT8	28.5	19.2	23.8	28.5	260	1.33
MYRT9	28.8	18.1	23.5	29.6	260	1.4
MYRT10	29.7	18.1	22.0	30.1	259	1.49
MYRT11	27.4	19.3	20.1	33.2	259	1.54
MYRT12	32.3	17.1	22.4	28.1	263	1.53
MYRT13	28.5	18.1	23.8	29.6	260	1.39
MYRT14	28.5	18.1	23.8	29.6	260	1.39

To analyze the heterogeneity and phylogenetic relationships of *Ty1-copia*-type retrotransposons within the *M. jalapa* genome, degenerated genomic PCR targeting the RT loci was carried out, and the expected size of putative RT sequences (around 260 bp) were recovered, purified and cloned. A total of 30 clones were randomly selected for sequencing, of which 14 were unique sequences and were designated as MYRT1 to MYRT14 (GenBank numbers of KU720518-KU720531), respectively. These RT sequences were AT rich, with an AT;GC ratio ranging from 1.21 to 1.66 (Table 2), which is similar to *Ty1-copia*-type retrotransposons found in other species [23]. The length of these RT sequences from *M. jalapa* ranged from 256-263 bp, which was slightly shorter than the 273 bp reported for the RT sequences from other *Ty1-copia*

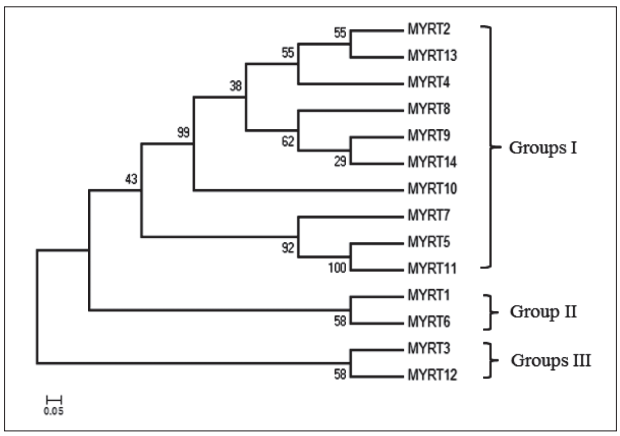

Fig 3. Phylogenetic analysis of the reverse transcriptase (RT) sequences of *Ty1-copia* retrotransposons in red flowers of *M. jalapa*. The tree was obtained by the maximum likelihood method using MEGA 5.1 software. The numbers at the branch nodes showed bootstrap values with 1000 replications.

types of retrotransposons [24]. This finding suggested that REC66 from *M. jalapa* represents a new variant of the *Ty1-copia* type of retrotransposons.

Phylogenetic analysis of these RT sequences divided them into three groups (Fig. 3). Group I (with 10 sequences, 71.4%), which includes the RT from *REC66*, was the largest one, followed by Group II (with 2, 14.3%) and group III (with 2, 14.3%). Group III has a 58 bootstrap value with Groups I and II, and Group I a 43 bootstrap value with Group II. The bootstrap value of the sequences was high in Group I, and bootstrap values were relatively low in Groups II and III. The results suggested that the RT sequences of the Group I family, which also includes REC66, were the major type of *Ty1-copia*-like retrotransposons in the red flower of *M. jalapa*.

The RT sequences of *Ty1-copia* type retrotransposons were translated into amino acids, and the similarity matrix between the predicted proteins ranged from 80.7% to 100%. The majority of the RT coding regions contained the three conserved motifs TAFLHG, SLYGLKQ and YVDDM (Fig. 4). For the conserved motif of TAFLHG, it was found that "F" was found to be replaced by "L" in the sequences of MYRT2, MYRT13, MYRT4 and MYRT12 isolates; while "A" was substituted for "S" for the conserved motif of SLYGLKQ in the sequences of MYRT12, MYRT3 and MYRT6; the "L" was replaced by "F" in the sequences of MYRT3 and the "G" replaced by "D" in the sequences of MYRT7; while all 14 RT amino acid sequences contained the conserved motif of YVDDM. Among the 14 sequences, 9 (64%) contained stop codons and/or frame-shift mutations, while the remaining 5 (36%) RT sequences, including that of *REC66* were intact and without disrupted ORF (Fig. 4).

DISCUSSION

In an attempt to identify flower-specific transcripts from *M. jalapa*, the most commonly grown ornamental species of *Mirabilis*, we conducted FDD screening to compare the mRNA expression between the flower and leaves of the red plant. DNA sequence analysis of the cDNA fragments of interest recovered from FDD revealed that one of the flower-specific transcripts, 320 bp in length (designated *C66* cDNA), encoded a GAG-like peptide with homology to *Ty1-copia* retrotransposons. Then SEFA-PCR chromosome walking was employed to obtain the full-length retrotranspo-

Fig 4. Amino acid sequence alignment of reverse transcriptase (RT) sequences from the *Ty1-copia* like the retrotransposons in red flowers of *M. jalapa.* The four shade levels indicate degree of residue conservation: deep blue (100% conserved), red (75% or greater conserved), light green (50% or greater conserved) and no shading (<50% conserved). The right hand numbers represent the total numbers of amino acid residues of each RT sequence.

son, 5567 bp in length, which exhibited a high degree of sequence homology to *Ty1-copia* retrotransposons. Thus, we have designated the new transposon as *REC66* as it was discovered from the red flower of *M. jalapa*. Based on the PR-INT-RT-RNaseH gene order and the POL gene-encoding protein sequence homology, *REC66* is a first retrotransposon belonging to the *Ty1-copia* retrotransposon family ever discovered from *M. jalapa* [2]. The structure of the *REC66* sequence contains characteristics of the PBS sequence TGGT, polypurine tract (PPT) sequence AGGGG, and LTR boundary motif 5'-TG-CA-3'. It also has conserved motifs, including C-C-H-C motif in GAG; TSG sites in PR; G-K-F-GY-G motif in INT; motif I (AFLGL), motif II (YGLKQ), motif III (YVDDI) in RT, motif I (KHI) and motif II (ADTK) in RNaseH [25-27]. Finally, sequence homology analysis and phylogeny relationship study between *REC66* and other *Ty1-copia* retrotransposons in RT indicated that the *REC66* retrotransposon indeed belongs to the *Ty1-copia* type of retrotransposons.

In the past, the most common strategy to obtained the full-length and complete sequences of *Ty1-copia*-type retrotransposons often involved two steps: amplification of RT domains based on RT conserved motifs I (AFLGL) and III (YVDDI) [15], and using chromosome walking strategies to recover the 3'- and 5'-flanking regions of the RT fragments until the full-length and complete sequences was obtained. Many novel and complete sequences of *Ty1-copia* retrotransposons have been successfully isolated and identified with this approach [28-33]. In our study, we discovered and characterized a novel *Ty1-copia*-type retrotransposon, *REC66*, not based on the RT sequence homology but rather on a 320-bp expressed cDNA sequence from FDD, which showed unmistakable amino acid-sequence homology to the GAG protein from Ty1-retrotransposons. Then SEFA-PCR chromosome walking was employed to obtain the full-length retrotransposon. Thus, FDD can be used as a new approach to obtain retrotransposons, especially when their transpositions may be involved in tissue-specific gene expression and the resulting phenotypes, such as variations in floral colors.

Our investigation of the sequence heterogeneity and phylogenetic relationship of the RT sequences from the *Ty1-copia*-like retrotransposons in *M. jalapa* revealed that of the 14 unique RT sequences discovered, 64% of them have defective RT with either stop codons and/or frame-shift mutations, while the remaining 36%, including that of *REC66*, were intact and without disrupted ORF. This finding suggests that *REC66* might have been active or functional in the past [34,35]. Of the 14 RT sequences recovered from *M. jalapa*, all of them were slightly shorter (around 10 bp) than *Ty1-copia*-type retrotransposons previously reported. Phylogenetic analysis divided them into three groups, with Group I being the most abundant. Similar results were also reported for *Ty1-copia*-type retrotransposons in wax gourd [36], as well as for other species [32,37-39].

In conclusion, this is the first report of *Ty1-copia* retrotransposons in *M. jalapa*. The finding is of importance, not only in light of *M. jalapa* phylogenetic studies, but also for the studies on *Ty1-copia*-type retrotransposons and plant genome evolution as a whole. Furthermore, the discovery of *Ty1-copia* retrotransposons could be helpful for the development of new molecular markers for future genetic studies, such as floral color variations, on *M. jalapa* and other closely related plant species.

Acknowledgments: We thank Yu, Ruoxuan and Zhao, Rongchuan for their technical assistance and Jiang, Pu for logistic support. We also thank scientists from Clover Biopharmaceuticals for the assistance with the fluorescent differential display (FDD) screening.

Authors' contribution: Shunri Jiang and Peng Liang conceived and designed the experiments; Shunri Jiang performed the experiment, analyzed data and wrote the manuscript. Joshua G. Liang and Blake Shester helped with FDD screening. Haiyan Feng, Dan Luo, Liang Yang, Wen Lu, Suzhi Zhang and Yi Yang helped in data analysis, in the manuscript discussion and editing. Peng Liang supervised the research project and revised the manuscript. All authors have contributed their efforts to this work, read and approved the final manuscript.

Conflict of interest disclosure: The authors declare that they have no conflict of interest.

REFERENCES

1. Kumar A, Bennetzen JL. Plant retrotransposons. Ann Rev Genet. 1999;33(1):479-532.
2. Wicker T, Sabot F, Hua-Van A, Bennetzen JL, Capy P, Chalhoub B, Flavell A, Leroy P, Morgante M, Panaud O. A unified classification system for eukaryotic transposable elements. Nature Reviews Genetics. 2007;8(12):973-82.

3. Wessler SR, Bureau TE, White SE. LTR-retrotransposons and MITEs: important players in the evolution of plant genomes. Curr Opin Genet Devel. 1995;5(6):814-21.
4. Kumar A, Bennetzen JL. Retrotransposons: central players in the structure, evolution and function of plant genomes. Trends Plant Sci. 2000;5(12):509-10.
5. Havecker ER, Gao X, Voytas DF. The diversity of LTR retrotransposons. Genome Biol. 2004;5(6):1.
6. Zhao M, Ma J. Co-evolution of plant LTR-retrotransposons and their host genomes. Prot Cell. 2013;4(7):493-501.
7. Engels J, Van Kester W, Spitters C, Vosselman L, Zeven A. Investigations of the inheritance of flower variegation in *Mirabilis jalapa* L. 1. General introduction and 2. Inheritance of colour in uniformly coloured flowers. Euphytica. 1975;24(1):1-5.
8. Mendel G. Gregor Mendel's letters to Carl Nägeli, 1866-1873. Genetics. 1950;35(5 2):1-29.
9. Hatlestad GJ, Sunnadeniya RM, Akhavan NA, Gonzalez A, Goldman IL, McGrath JM, Lloyd AM. The beet R locus encodes a new cytochrome P450 required for red betalain production. Nature Genet. 2012;44(7):816-20.
10. Suzuki M, Miyahara T, Tokumoto H, Hakamatsuka T, Goda Y, Ozeki Y, Sasaki N. Transposon-mediated mutation of CYP76AD3 affects betalain synthesis and produces variegated flowers in four o'clock (*Mirabilis jalapa*). J Plant Physiol. 2014;171(17):1586-90.
11. Murray M, Thompson WF. Rapid isolation of high molecular weight plant DNA. Nucl Acids Res. 1980;8(19):4321-6.
12. Liang P, Meade JD, Pardee AB. A protocol for differential display of mRNA expression using either fluorescent or radioactive labeling. Nature Prot. 2007;2(3):457-70.
13. Wang S, He J, Cui Z, Li S. Self-formed adaptor PCR: a simple and efficient method for chromosome walking. Appl Environ Microbiol. 2007;73(15):5048-51.
14. Ma Z, Luo D, Huang A, Xu Y, Wang Y, Wei Y, Liang P. pKILLIN: a versatile positive-selection cloning vector based on the toxicity of Killin in *Escherichia coli*. Gene. 2014;544(2):228-35.
15. Kumar A, Pearce SR, McLean K, Harrison G, Heslop-Harrison J, Waugh R, Flavell AJ. The Ty1-*copia* group of retrotransposons in plants: genomic organisation, evolution, and use as molecular markers. In: Capy P, editor. Evolution and Impact of Transposable Elements. Amsterdam: Springer; 1997. p. 205-17.
16. Grandbastien M A, Spielmann A, Caboche M. Tnt1, a mobile retroviral-like transposable element of tobacco isolated by plant cell genetics. Nature. 1989;337(6205):376-80.
17. Tapia G, Verdugo I, Yañez M, Ahumada I, Theoduloz C, Cordero C, Poblete F, González E, Ruiz-Lara S. Involvement of ethylene in stress-induced expression of the TLC1. 1 retrotransposon from *Lycopersicon chilense* Dun. Plant Physiol. 2005;138(4):2075-86.
18. Marillonnet S, Wessler SR. Extreme structural heterogeneity among the members of a maize retrotransposon family. Genetics. 1998;150(3):1245-56.
19. White SE, Habera LF, Wessler SR. Retrotransposons in the flanking regions of normal plant genes: a role for *copia*-like elements in the evolution of gene structure and expression. Proc Natl Acad Sci USA. 1994;91(25):11792-6.
20. Emori Y, Shiba T, Kanaya S, Inouye S, Yuki S, Saigo K. The nucleotide sequences of *copia* and *copia*-related RNA in Drosophila virus-like particles. Nature. 1984;315(6022):773-6.
21. Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG. The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucl Acids Res. 1997;25(24):4876-82.
22. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Molec Biol Evol. 2011,28(10):2731-9.
23. Stergiou G, Katsiotis A, Hagidimitriou M, Loukas M. Genomic and chromosomal organization of Ty1-*copia*-like sequences in *Olea europaea* and evolutionary relationships of Olea retroelements. Theor Appl Genet. 2002;104(6-7):926-33.
24. Voytas DF, Cummings MP, Koniczny A, Ausubel FM, Rodermel SR. *Copia*-like retrotransposons are ubiquitous among plants. Proc Natl Acad Sci USA. 1992;89(15):7124-8.
25. Malik HS, Eickbush TH. Phylogenetic analysis of ribonuclease H domains suggests a late, chimeric origin of LTR retrotransposable elements and retroviruses. Genome Res. 2001;11(7):1187-97.
26. Peterson-Burch BD, Voytas DF. Genes of the Pseudoviridae (Ty1/*copia* retrotransposons). Molec Biol Evol. 2002;19(11):1832-45.
27. Xiong Y, Eickbush TH. Origin and evolution of retroelements based upon their reverse transcriptase sequences. EMBO J. 1990;9(10):3353.
28. Zhao G, Dai H, Chang L, Ma Y, Sun H, He P, Zhang Z.Isolation of two novel complete Ty1-*copia* retrotransposons from apple and demonstration of use of derived S-SAP markers for distinguishing bud sports of *Malus domestica* cv. Fuji. Tree Genet Gen. 2010;6(1):149-59.
29. Pearce SR, Harrison G, Heslop-Harrison P, Flavell AJ, Kumar A. Characterization and genomic organization of Ty1-*copia* group retrotransposons in rye *(Secale cereale*). Genome. 1997;40(5):617-25.
30. Rogers SA, Pauls KP. Ty1-*copia*-like retrotransposons of tomato (*Lycopersicon esculentum* Mill.). Genome. 2000;43(5):887-94.
31. Kimura Y, Tosa Y, Shimada S, Sogo R, Kusaba M, Sunaga T, Betsuyaku S, Eto Y, Nakayashiki H, Mayama S. OARE-1, a Ty1-*copia* retrotransposon in oat activated by abiotic and biotic stresses. Plant Cell Physiol. 2001;42(12):1345-54.
32. Ma Y, Sun H, Zhao G, Dai H, Gao X, Li H, Zhang Z. Isolation and characterization of genomic retrotransposon sequences from octoploid strawberry (*Fragaria× ananassa* Duch.). Plant Cell Rep. 2008;27(3):499-507.
33. Rico-Cabanas L, Martínez-Izquierdo JA. CIRE1, a novel transcriptionally active Ty1-*copia* retrotransposon from *Citrus sinensis*. Molec Genet Genome. 2007;277(4):365-77.
34. Dixit A. Flavell AJ, Smith DB, Kumar A. Extreme heterogeneity of Ty1-*copia* group retrotransposons in plants. Molec Gen Genet. 1992;231(2):233-42.
35. Flavell AJ, Dunbar E, Anderson R, Pearce SR, Hartley R, Kumar A. Ty1-*copia* group retrotransposons are ubiquitous and heterogeneous in higher plants. Nucl Acids Res. 1992;20(14):3639-44.

36. Jiang B, Liu W, Peng Q, He X, Xie D. Characterization and chromosomal organization of Ty1-*copia* retrotransposons in wax gourd. Gene. 2014;551(1):26-32.
37. Dixit A, Ma K H, Yu J W, Cho E-G, Park Y-J. Reverse transcriptase domain sequences from Mungbean (*Vigna radiata*) LTR retrotransposons: sequence characterization and phylogenetic analysis. Plant Cell Rep. 2006;25(2):100-11.
38. Jiang B, Wu Z M, Lou Q F, Wang D, Zhang W-P, Chen J-F. Genetic diversity of Ty1-*copia* retrotransposons in a wild species of Cucumis (C. hystrix). Scientia Horticult. 2010;127(1):46-53.
39. Kolano B, Bednara E, Weiss-Schneeweiss H. Isolation and characterization of reverse transcriptase fragments of LTR retrotransposons from the genome of Chenopodium quinoa (Amaranthaceae). Plant Cell Rep. 2013;32(10):1575-88.