

LTR-retrotransposon activation, IRAP marker development and its potential in genetic diversity assessment of masson pine (*Pinus massoniana*)

Fuhua Fan · Bowen Cui · Ting Zhang · Guijie Ding · Xiaopeng Wen

Received: 1 July 2013 / Revised: 9 October 2013 / Accepted: 17 October 2013 / Published online: 2 November 2013
© Springer-Verlag Berlin Heidelberg 2013

Abstract Activation of retrotransposon is the pivotal factor for the genesis of genetic polymorphism. Retrotransposon-based molecular markers are excellent tools for detecting genetic diversity and genomic changes associated with retrotransposon activity. The scarcity of retrotransposon long terminal repeat (LTR) sequence limits the application of retrotransposon-based molecular marker systems. In this study, retrotransposon 3'-LTR segments were firstly isolated and characterized from the masson pine genome using a genome walking strategy, subsequently inter-retrotransposon amplified polymorphism (IRAP) markers were explored for genetic diversity assessment. These were shown to have clearly distinguishable amplification bands and high levels of polymorphism for masson pine. Dendrogram of our IRAP data highlighted that IRAP markers for individual elements are distinguishable and will shed light on the usage in genetic diversity studies of the masson pine. Transcription activation of Ty1-*copia* and Ty3-*gypsy* group retrotransposons in masson pine was investigated with exposure to various abiotic stresses. The insertional polymorphism and the transposition activation were detected by IRAP. Results revealed that none of the analyzed materials in the course of 2-month trials displayed fingerprint changes.

Keywords LTR retrotransposons · Activity · Genetic diversity · Masson pine · IRAP

Communicated by D. Neale

F. Fan · B. Cui · T. Zhang · X. Wen (✉)
Guizhou Key Laboratory of Agricultural Bioengineering, Guizhou University, Guiyang 550025, Guizhou Province, China
e-mail: xpwense@hotmail.com

F. Fan · G. Ding
School of Forestry Science, Guizhou University,
Guiyang 550025, Guizhou Province, China

Introduction

Retrotransposons are mobile genetic elements, and transpose via an RNA intermediate. They are classified as LTR or non-LTR retrotransposons dependent on the presence or absence of long terminal repeats (LTRs), which are direct repeats flanking the internal coding regions (Staton et al. 2009). LTR retrotransposons can be further divided into two groups, the Ty1-*copia* group and the Ty3-*gypsy* group, based on the order of their coding domains, as well as their sequence similarity to the prototype elements in *Drosophila* and yeast (Boeke and Corces 1989; Zhao et al. 2010). Both Ty1-*copia* and Ty3-*gypsy* retrotransposons exist in plant genomes, usually as high-copy-number dispersed sequences, especially in gymnosperms (Nystedt et al. 2013). As the mechanism of retrotransposition is replicative, the parental copy is preserved with transpositional activity generating new insertions of the element at other sites of the genome, resulting in a high degree of heterogeneity and insertional polymorphism (Kumar and Bennetzen 1999; Carrier et al. 2012). In the cases examined, the increase of the retrotransposon copy number appeared to have been a major factor in the genome size growth of plants (Park et al. 2011; Zedek et al. 2010). Conversely, with a few exceptions, loss of retrotransposons through the recombinational production of solo LTRs and accumulation of deletions helps to keep genome size stable (Wicher and Keller 2007; Fedoroff 2013). Transcriptional and transpositional activations of transposable elements are mainly induced by abiotic and biotic stresses (Grandbastien 1998; Huang et al. 2012), and the activation has been regarded as the mechanism of genotypic remodeling. These characteristics provide an excellent basis for the development of molecular marker systems (Schulman et al. 2012; Kalendar et al. 2011).

Several different retrotransposon-based molecular marker systems have been developed to visualize the genetic diversity

(Kalendar et al. 2011; Smýkal et al. 2011; Jing et al. 2012; Abdollahi Mandoulakani et al. 2012). Most of the retrotransposon marker developments take advantage of two basic properties, namely the large insertions through transpositional activity and the conserved domains from which polymerase chain reaction (PCR) primers can be designed. The inter-retrotransposon amplified polymorphism (IRAP) method displays insertional polymorphisms by amplifying the segments of DNA between two retrotransposons. One significant virtue of IRAP is its experimental simplicity, which needs simple PCR followed by electrophoresis to resolve the PCR products. IRAP can be carried out with a single primer matching the conserved motifs, or with two primers. Adjacent TEs may be found in different orientations in the genome (head-to-head, tail-to-tail or head-to-tail), increasing availability to detect polymorphism resting with the method and primer combinations (Kalendar et al. 2011). To date, IRAP markers have been widely used to elucidate the genetic diversity in many species (Kalendar and Schulman 2006; Vukich et al. 2009; Campbell et al. 2011; Abdollahi Mandoulakani et al. 2012).

Masson pine (*Pinus massoniana*), a gymnosperm genus conifer native to Southern China, has been one of the most economically important forest trees, because it can be widely used for timber, pulp and resin production. Synecological studies had revealed that masson pine germplasm was prone to become short (Ding and Song 1998), chiefly due to the loss of genetic diversity. Molecular markers are extremely useful in the aspects of assessing genetic diversity and identifying potential novel genotypes among the masson pine germplasm. Several types of marker systems, including random amplified polymorphic DNA (RAPD), inter-simple sequence repeat (ISSR), and simple sequence repeat (SSR), have been used to analyze masson pine germplasm (Peng et al. 2003; Li et al. 2009; Cai and Ji 2009). Taken together, these studies show that masson pine has low genetic diversity, possibly resulting from a domestication bottleneck, or aforementioned markers produce a less even distribution across the genome compared to markers based on retrotransposons (Abdollahi Mandoulakani et al. 2011; Biswas et al. 2010; Yuan et al. 2012).

Previously, we demonstrated the occurrence of Ty1-*cop*ia and Ty3-*gypsy* group retrotransposons with a high heterogenous, and the total number of Ty1-*cop*ia group retrotransposons is approximately 89,577 molecules, and Ty3-*gypsy* group retrotransposons is about 29,310 molecules per genome (Fan et al. 2013). The objectives of the present study were (1) to detect the activation of masson pine retrotransposon under several abiotic stresses; (2) to isolate Ty1-*cop*ia and Ty3-*gypsy* group retrotransposons 3'-LTR segments; (3) to develop the IRAP marker based on the sequences of retrotransposons; and (4) to assess the genetic diversity of masson pine germplasm of China.

Materials and methods

Plant materials

Thirty-four masson pine cultivated lines were obtained from the National Masson Pine Germplasm Collection of Duyun Forestry Station (Guizhou Province, China) for genetic diversity analysis.

Seeds of masson pine from a same line obtained from the Duyun Forestry Station were germinated on moist filter paper in the dark and transferred to pots in the Plant Chamber Facility of Guizhou University. Seedlings irrigated with Hoagland and Arnon trace elements nutrient solution regularly were grown for about 2 months. Parts of these materials were used for genome walking, and others used for various stresses, such as heat (42°C), cold (−2°C), salicylic acid (2 mM), gibberellic acid (2 mM), 2, 4-D (50 mM) and UV (wavelength: 315–400 nm, irradiance: 0.68–0.8 W/m²) for 24 h. These needle materials were stored at −80°C until RNA isolation.

DNA extraction

Total DNA was extracted from fresh needles using Tiangen DNAsure Plant Kit (Tiangen, Beijing, China) according to the manual.

RNA isolation and RT-PCR

Total RNA of normal and stressed materials was isolated using Invitrogen Plant RNA Purification Reagent (USA) according to the manual. All RNA samples were treated with DNase I (TaKaRa, Dalian, China) at 37°C for 2 h before RNA precipitation. First-strand cDNA synthesis was carried out using RNA LA PCR™ KIT (TaKaRa), following the manufacturer's instructions, and choosing the oligo dT-adaptor primer. Control reactions without RT were routinely included in the PCR amplification of reverse-transcribed products. Relative RT-PCR was performed to detect transcriptional activation of LTR-retrotransposons. The degenerate primers Rtp1 5'-ACNGCNTTYTNCAYGG-3' and Rtp2 5'-ARCATRTCRTCACRTA-3' were used to amplify RT domains of Ty1-*cop*ia group retrotransposons. The degenerate primers Gyrt1 5'-AGMGRTATGTGYGT SGAYTAT-3' and Gyrt2 5'-CAMCCMRAAMW CACAMTT-3' were employed to amplify RT domains of Ty3-*gypsy* group retrotransposons.

Determination of LTR sequences by genome walking

The principal protocol of GenomeWalker™ Universal Kit (Clontech, USA) was adopted to isolate 3'-LTR sequences of retrotransposons. Four independent pools of high-quality

masson pine DNA (2.5 µg) were digested by the blunt end restriction enzymes *Dra*I, *Eco*RV, *Sst*I, and *Pvu*II. Then each batch of digested genomic DNA fragments was ligated separately to the GenomeWalker adaptor, which consisting of a 48-mer (5'-GTAATACGACTCACTATAGGGCACGC GTGGTCGACGGCCCCGGGCTGGT-3') and an 8-mer with the 3' end capped by an amino residue (5'-PO₄-ACCAGCCC-N₂H-3'). A walker step consisted primary and secondary PCR amplifications. The primary amplified mixture between the outer adaptor primer (5'-GTAATACGACTCACTATAGG GC-3') and a primary gene-specific primer was used as the template for a secondary amplification with a nested gene-specific primer and the inner adaptor primer, which overlapped the primary primer (5'-ACTATAGGGCACGC GTGGT-3'). The gene-specific primers (Table 1) were designed from known sequences using the Primer Premier 5.0 software and were synthesized by Generay Biotechnology (Shanghai, China). Twenty-five-microliter PCR reaction mixtures included 20 ng of template DNA approximately, 2.5 µl of 10× Advantage 2 PCR buffer (Clontech, USA), 0.5 µl of dNTP (10 mM each), 0.5 µl of each primer (10 µM) and 0.5 µl of 50× Advantage 2 Polymerase Mix (Clontech, USA). The primary PCR program used the following two-step cycle parameters: seven cycles of 94°C for 25 s and 72°C for 3 min, followed by 32 cycles of 94°C for 25 s and 67°C for 3 min, then 67°C for an additional 7 min after the final cycle. The secondary PCR program was the same as described above, except that 5 cycles was used instead of 7 cycles and 20 cycles instead of 32 cycles. Aliquot parts of the second PCR products were separated, recovered, cloned and sequenced as described previously (Fan et al. 2013). Three or four independent subclones were sequenced to correct errors generated by amplification and sequencing.

Sequence analysis

Sequence assembly was done using DNAMAN software. Homology queries were carried out by submitting our sequences to the BLASTn and BLASTx non-redundant database searching algorithms at the National Center for Biotechnology Information (<http://www.ncbi.nlm.nih.gov/>). Multiple sequence alignments were made under the aid of ClustalX (version 2.1) software. Sequence alignments were annotated with GeneDoc. The sequences used for comparison were obtained from the GenBank database, such as Tnt1 (CAA32025) of *Nicotiana tabacum*, Copia (CAA26444) of *Drosophila melanogaster*, CIRE1 (CAJ09951) of *Citrus sinensis* and TLC1.1 (AAK29467) of *Solanum chilense* for Ty1-copia group comparison, and AFK13856 of *Beta vulgaris*, AAG51046 of *Arabidopsis thaliana* and ABA97923 of *Oryza sativa Japonica* for Ty3-gypsy group comparison. Both of the masson pine retrotransposons obtained here have been submitted to GeneBank with the

accession numbers of KC355438 and KC355439. The conceptual translations were implemented corresponded to the conserved motifs expected by LTR retrotransposons.

IRAP amplification and data analysis

The IRAP primers (Table 2) were designed to match the LTRs and RT conserved regions with the aid of Primer Premier 5.0 software. PCR amplification performed in 10 µl reaction mixtures with 10–50 ng genomic DNA, 1 µl of primer (10 µM), 5 µl Mix (Tiangen, China. Containing 0.1 U µl⁻¹ *Taq* polymerase, 500 µM dNTP, 20 mM Tris-HCl (pH 8.3), 100 mM KCl, 3 mM MgCl₂ and other stabilizer and intensifier) and appropriate ddH₂O. The cycling program consisted of initial denaturation at 94°C for 5 min, followed by 35 cycles at 94°C for 45 s, 55°C primer annealing for 45 s and 72°C extension for 1 min, with a final elongation step at 72°C for 7 min. The PCR products were separated either by electrophoresis on 1.8 % agarose gels in 1× TBE buffer and visualized under UV light after staining with ethidium bromide, or on 8 % PAGE gels followed by silver staining for better resolution.

Each IRAP band was treated as a single locus. The presence or absence of a given length fragment in each sample was recorded manually in binary code. DNA marker data was processed by NTSYS-pc version 2.10 e using the SIMQUAL module with the Jaccard genetic similarity coefficient (GSJ), and the similarity data was used to perform an unweighted pair group method with arithmetic mean (UPGMA) cluster analysis using the SHAN module.

Results

Detection of transcriptional activation

To determine the transcriptional activation function of retrotransposons after exposure to abiotic stresses, reverse transcription (RT-PCR) was carried out in masson pine subjected to heat, cold, salicylic acid, gibberellic acid, 2,4-D and UV stress. Genomic DNA was used as a positive control and water as a negative control. The results revealed that salicylic acid, gibberellic acid, 2,4-D and UV light induced transcriptional activation of both Ty1- and Ty3-type retrotransposons, with 2,4-D as the strongest inducer on both. However, heat and cold treatment did not lead to transcription of LTR-retrotransposons (Fig. 1).

Isolation and characterization of retrotransposon 3'-LTRs

In order to isolate the 3'-LTRs of retrotransposon from the masson pine genome, two sequences capable of generating the subsequent steps were used as original to carry out genome

Table 1 The gene-specific primers employed in the genome walk

| Original fragment | Primer name | Primer sequences (5'→3') | Primer location in retrotransposon (bp) | Step generated | Retrotransposon generated | |
|-------------------|-------------|---------------------------------|---|----------------|---------------------------|------|
| PMRT16 REPM6 | 1 | GGCTAAAAACAGGCTCCAAGGTAGTGG | 111–138 | D 1 | PmRT | |
| | 2 | GTCTGAACAACGTTATACTAGATGTCACCTC | 161–191 | | | |
| | 3 | TAAGGTGGTGTGTCCCAAGACTCAGG | 503–529 | D 2 | | |
| | 4 | GGCTTGATATTGCACATGCAGTGGGAG | 604–631 | | | |
| | 5 | ACCAGGGAGAAATTGAGAATCTGCTCA | 1,161–1,188 | D 3 | | |
| | 6 | GAAGCAGCAGTGGAGTTGCTATAATCAG | 1,264–1,292 | | | |
| | 7 | ACGCCCCAGCTACATTCCAGTCTTGC | 235–261 | D' 1 | | REPM |
| | 8 | TCCAGCACATAAAGACAGTGTTCGCGCAT | 352–380 | | | |
| | 9 | ATCACAAAGAGCCTCAGGTTTCGTTTTGG | 936–963 | D' 2 | | |
| | 10 | ATGAGCTTAGACGCAGACTGGAGGGC | 1,100–1,126 | | | |
| | 11 | TCACCCACAGAGGGACGGGTAGACAG | 1,762–1,788 | D' 3 | | |
| | 12 | CTGGAGGACTACTTGCCTAACTATG | 1,805–1,830 | | | |
| | 13 | GGAGTGATAGTTTGAGGTAGGGGAC | 2,098–2,123 | D' 4 | | |
| | 14 | AGGGTGATACGCAGGATTGGACAGGTTG | 2,213–2,241 | | | |

walking. One RT sequence of 266 bp named as PMRT16 (GenBank accession no. JQ975194) was isolated with degenerate primers corresponding to the conserved reverse transcriptase domains of Ty1-*copia* retrotransposons. The other Ty3-*gypsy* RT sequence of 432 bp, named as REPM6 (GenBank accession no. JQ975242), was also obtained by the similar method. The unknown fragments adjacent to both known sequences were isolated by GenomeWalker™ Universal Kit (Clontech, USA) with reference to the instruction. The largest and distinguishable PCR fragments from the secondary amplification were selected for sequence analysis. Identity of overlap sequences between the subsequently obtained fragment and the previous sequence implied that both fragments derived from a unique retrotransposon. Three successive steps downstream from PMRT16 and four successive steps downstream from REPM6

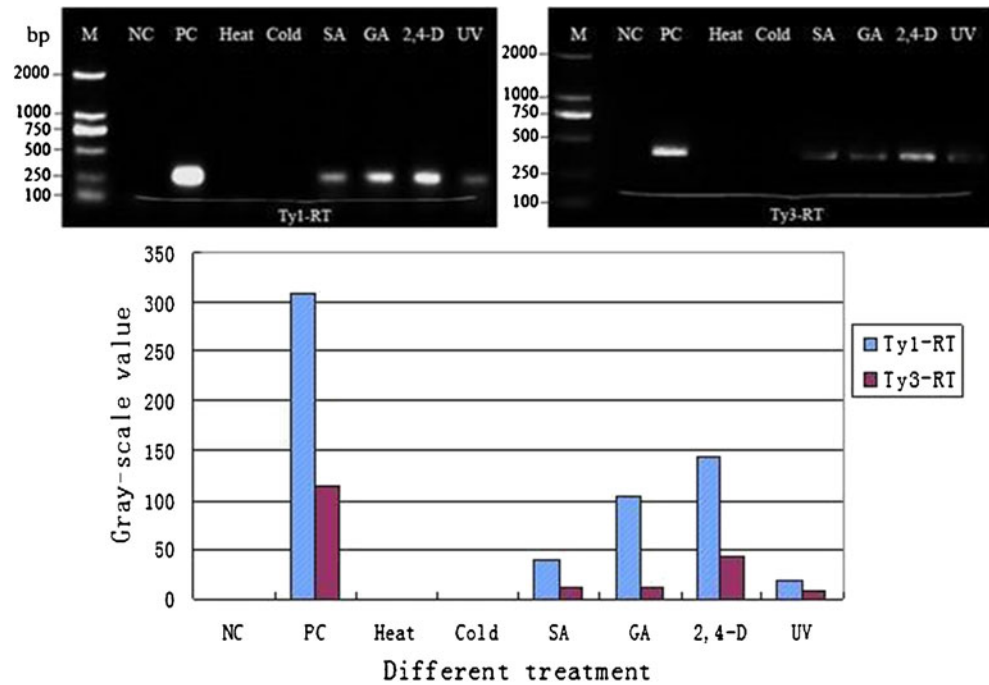
generated fragments of different sizes. Fragment assembly results revealed that both of the contig sequences contained almost exact LTRs that start with TG and end with CA. Therefore, an RT–RNaseH–LTR sequence of Ty1-*copia* retrotransposon, named as PmRT (GenBank accession no. KC355438) and an RT–RNaseH–INT–LTR sequence of Ty3-*gypsy* retrotransposon, named as REPM (GenBank accession no. KC355439), were characterized from the masson pine genome. The lengths are 1,766 and 2,533 bp for PmRT and REPM, respectively.

The sequences of different domains of PmRT and REPM retrotransposons were compared with the LTR retrotransposons of other reported organisms (tobacco, fruit fly, sweet orange, tomato, sugar beet, arabidopsis, japonica rice), and the results indicated that both retrotransposons contain conserved motifs (Fig. 2), such as a TG/CA inverted

Table 2 IRAP primers with sequence and polymorphism level

| Primer number | Sequence (5'→3') | GenBank accession number | Position and orientation | Amplified bands per sample (mean) | Percentage polymorphism |
|---------------|------------------------------|--------------------------|--------------------------|-----------------------------------|-------------------------|
| SAF2 | GTGGAGCCTGATTATCTTCTAT | KC355438 | 1,316→1,338 | 24 | 50 % |
| PPT2 | ACCAGGGAGAAATTGAGAATCT | KC355438 | 1,161→1,183 | 26 | 77.8 % |
| SAR1 | CAACTCATAACGAACCTGTCCAATCCTG | KC355439 | 2,224←2,252 | 17 | 41.2 % |
| P3 | AGGGTAAGAAAGAAGACTGGTATG | JQ975223 | 68→89 | 13 | 84.6 % |
| P12 | AACAGGCTTCAAGACAGTGGTA | JQ975192 | 119→140 | 15 | 40 % |
| R3 | GCCCTGAACAAGAAGACACTG | JQ975245 | 25→45 | 16 | 43.8 % |
| R13 | AGAGGTATGTGCGTGGATTA | JQ975250 | 1→20 | 22 | 50 % |
| R16 | AGCGGTATGTGCGTGGACTA | JQ975244 | 1→20 | 28 | 67.9 % |
| R17 | AGACCTCAAAGAATCGCTACCC | JQ975244 | 38→59 | 12 | 41.7 % |

Fig. 1 RT-PCR detection of LTR retrotransposon subjected to abiotic treatments. *NC* negative control, *PC* positive control. Heat, cold, SA, GA, 2,4-D and UV stands for heat, cold, salicylic acid, gibberellic acid, 2,4-D and UV treatment, respectively



repeat in the LTRs. In addition, PmRT had three conserved amino acid motifs (TAF-HG, YGLKQ, and YVDDML) in RT sequence and two characteristic motifs (KHID and DMLTK) in the RNaseH region followed by a polypurine tract as an RNaseH–LTR junction (Pearce et al. 1999). REpm contained a zinc-finger domain (H-6aa–H-29aa–C-2aa–C) in the integrase followed by a polypurine tract besides the conserved coding region motifs of both the RT and RNaseH (Suoniemi et al. 1998).

Development of IRAP makers for masson pine germplasm genotyping

The IRAP primers were designed to match the LTRs of PmRT and REpm, as well as the RT conserved regions of other reverse transcriptase sequences (Table 2). In order to test the suitability of the IRAP markers for genotyping, a set of six genotypes randomly selected from cultivated masson pine lines was used to carry out PCR amplification (Fig. 3). The scoring criteria consisted of the sharpness, number, and evenness in intensity of the PCR products following electrophoresis detection, as well as the degree of polymorphism among the genotypes. Some primers, which yielded either poor amplifications or few products, or produced primarily monomorphic products, were discarded. A set of nine primers might yield polymorphic and evenly distributed fragments (Table 2; Fig. 3), and those generating more than 40 % of polymorphism, was retained for further work.

Retrotransposons stability in the genome detected by IRAP markers

In order to see if we might detect retrotransposon mobilization, we elected the materials which stressed by 2,4-D (50 mM) for 24 h, and tested the stability of IRAP fingerprints of the materials from three sampled time. None of the analyzed seedlings in course of 2 months' trials displayed fingerprint changes (Fig. 4).

Genetic diversity within masson pine germplasm

A total of 34 masson pine lines were scored by IRAP with the most informative nine primers, yielding 153 discernible reproducible fragments, among which 82 were polymorphic with a polymorphic rate as 53.6 %. The number of scorable bands per primer ranged from 13 (R17; 41.7 % polymorphic) to 28 (R16; 67.9 % polymorphic). The genetic relationships of the tested genotypes were unraveled using UPGMA method based on Jaccard similarity coefficients (ranged from 0.680 to 0.863 and overall mean: 0.77) computed with IRAP markers, and seven groups were obtained from dendrogram taking 0.73 as a threshold (Fig. 5). In the dendrogram, Group I included most of the accessions (15 genotypes), followed by group VII having nine masson pine lines. And remarkably, only one genotype was subclassed into group V, which indicated that this genotype had further genetic distance with other masson pine lines.

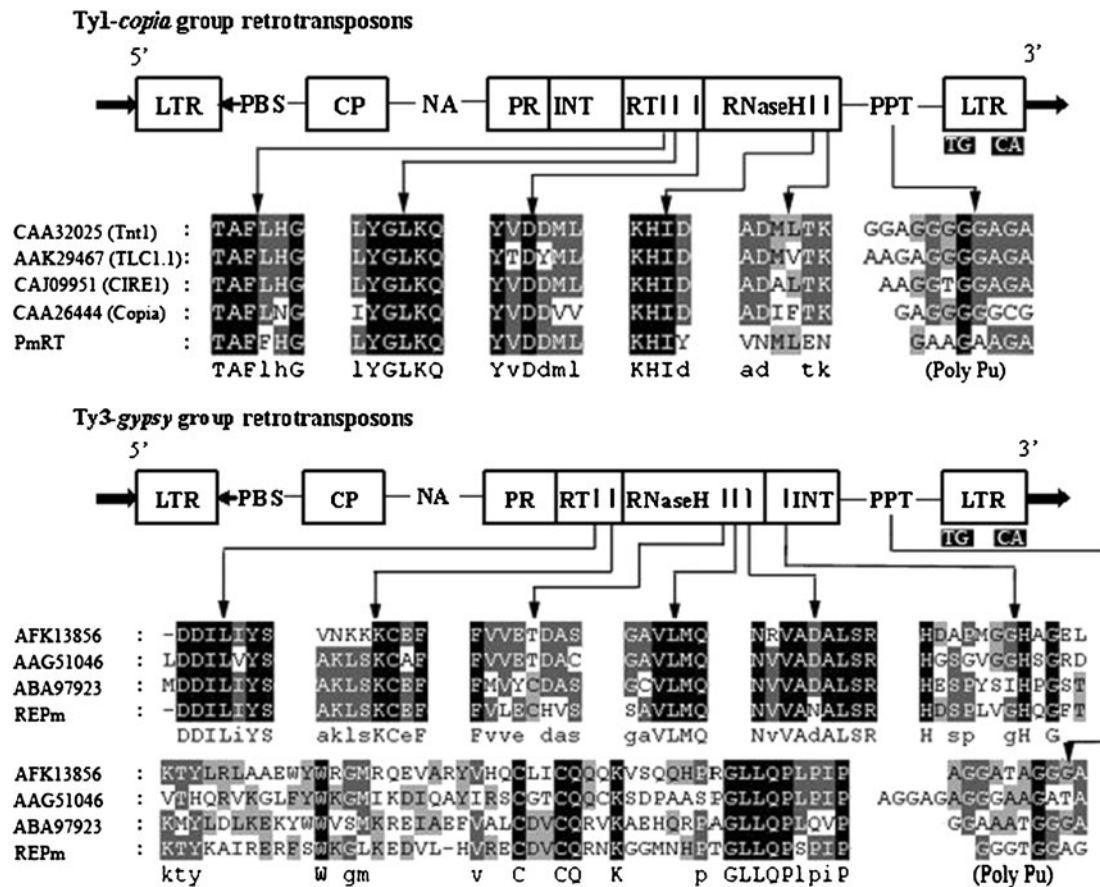


Fig. 2 The structure and conserved motifs of Ty1-copia group and Ty3-gypsy group retrotransposons. Sequences used for the alignment were obtained from the GenBank. The black vertical bars in the given regions represent sites of characteristic motifs. Completely conserved or nearly

invariant amino acids among retrotransposons are indicated by black shade, and partially conserved signatures among retrotransposons are marked by gray shade

Discussion

Previous evidences had showed that RT sequences of LTR retrotransposons in the masson pine genome were highly

heterogeneous and in high copy number (Fan et al. 2013). In the present study, we had taken advantage of the ubiquity and abundance of LTR retrotransposons in plant genomes and their role in genomic diversification to develop IRAP marker,

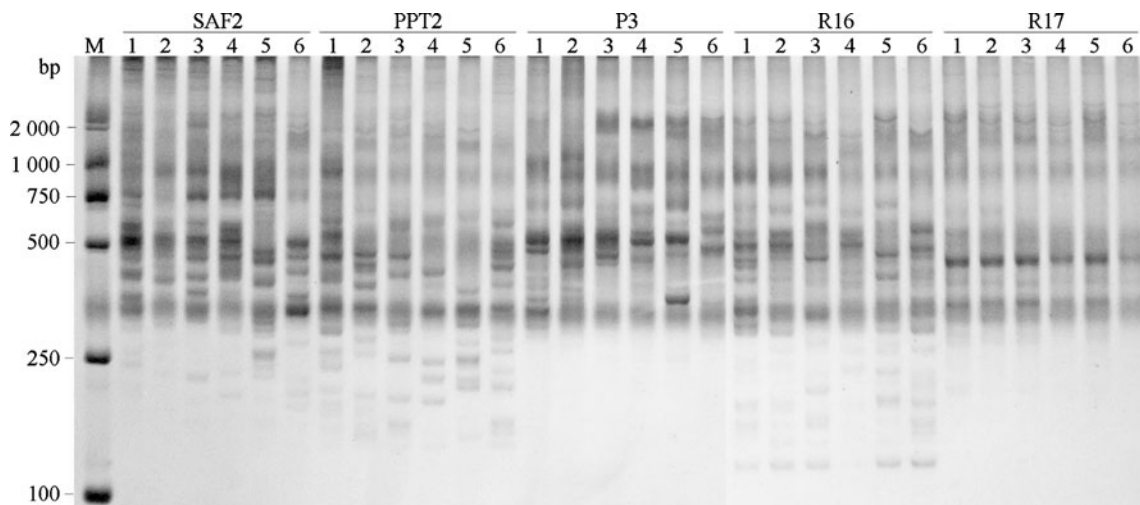


Fig. 3 IRAP analysis for a set of six masson pine lines. The primer codes used are shown above each set of lanes

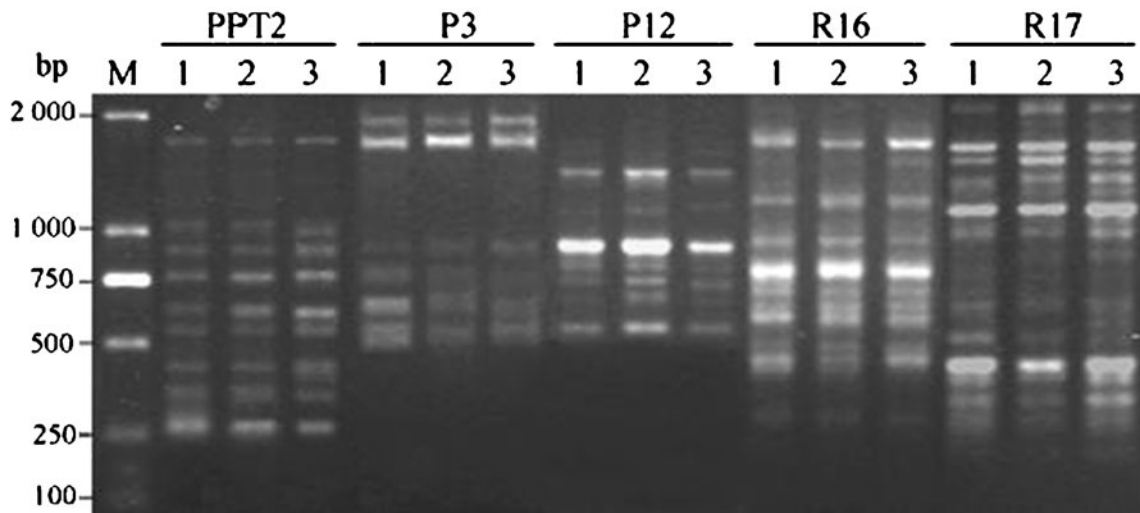


Fig. 4 IRAP fingerprints detection. "1", "2" and "3" stand for the day before stress treatment, the fifth day and the 20th day after stress treatment, respectively. The primer codes used are shown *above* each set of lanes

subsequently, the retrotransposon-based marker was used to genotype the masson pine lines, as well as to elucidate their genetic relationship. Two retrotransposon segments were isolated through genome walking based on adaptor-mediated nested PCR in this work. It showed that both of the contig sequences contained almost exact LTRs that start with TG and end with CA (Jin and Bennetzen 1989), which facilitate the design of LTR-specific primers. Out of tested primers, nine might yield discernible and polymorphic banding pattern, were applied to detect the transpositional activation of retrotransposon, and to illuminate the genetic diversity among 34 masson pine

genotypes. Several cases have demonstrated that primers designed based on LTR sequences of retrotransposon families can be readily used across species lines (Lou and Chen 2007; Kalendar et al. 2011; Abdollahi Mandoulakani et al. 2012). In our study, some single IRAP primers based on retrotransposons of masson pine might produce polymorphic banding patterns in *Hylocereus undatus* and *Prunus pseudocerasus* (data not showed), indicating the presence of the homologous retroelements among them. This could be expected since horizontal transmission had happened during the evolution of retrotransposons (Stuart-Rogers and Flavell 2001).

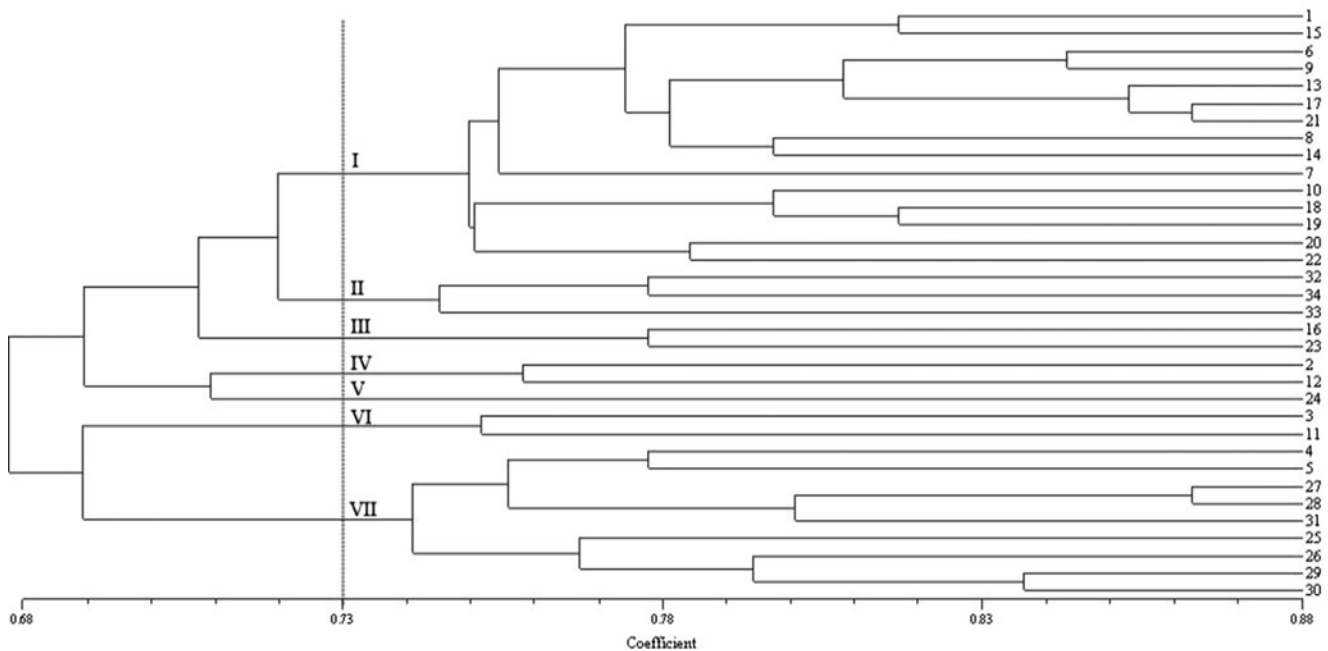


Fig. 5 Dendrogram tree of 34 masson pine genotypes obtained from UPGMA cluster analysis using IRAP loci based on Jaccard similarity coefficients

Retrotransposons transcription and transposition activity

It is a common feature for some retrotransposons being transcriptionally activated by various biotic and abiotic stress factors (Grandbastien 1998; Ungerer and Kawakami 2013). We investigated the transcriptional activation of LTR-retrotransposons in masson pine by RT-PCR. The results revealed that abiotic stresses such as salicylic acid, gibberellic acid, 2,4-D and UV light induced transcriptional activation with the exception of heat and cold treatment, which is consistent with previous reports in oat (Kimura et al. 2001) and in cucumis (Jiang et al. 2010). The expression of retrotransposons and their transposition frequency in the host genome are regulated (Hirochika et al. 1996). In general, plant genomes remain stable in the face of retrotransposon replication through loss of the inserted copies overtime (International Brachypodium Initiative 2010). Practically, LTR-retrotransposons recombinational loss has less effect on display pattern than do new insertions (International Brachypodium Initiative 2010). Consequently, retrotransposon markers are sensitive enough to detect rapid genome changes (Belyayev et al. 2010). Genome instability was reported in most of subgenus *Pinus* (Grotkopp et al. 2004), many life-history patterns were indirectly but consistently associated with genome size. Previous reports demonstrated that transposable elements insertional dynamics could promote morphological and karyotypical changes, some of which might be potentially important for the process of microevolution, allowing species with plastic genomes to survive as new forms or even as new species in times of rapid climatic change (Belyayev et al. 2010). Morse et al. (2009) concluded that most of the enormous genome complexity of pines could be explained by divergence of retrotransposons. Therefore, the stability of IRAP fingerprints of stressed materials was conducted to see whether we might detect retrotransposon mobilization associated with genotrophs. None of the analyzed materials in course of 2-month trials displayed fingerprint changes. It suggested that LTR retrotransposons in conifers might be less frequently removed by unequal recombination than in other plant genomes, which was consistent with the conclusions from the conifer genome evolution as documented by Nystedt et al. (2013). Also, some small deletions possibly occurred through unequal recombination same as that in *Arabidopsis* (Devos et al. 2002), which could not be identified herein by agarose gel electrophoresis. It is also likely that the IRAP experiments did not display all the retrotransposons in the genome and would not detect the nested insertions located within the retrotransposon "behind" (5' to) the PCR priming sites (Smýkal et al. 2011). Hence, we concluded that stress might activate retrotransposons transcription, but cannot exclude its subsequent transposition in masson pine.

IRAP markers for genetic diversity in masson pine

Genetic diversity is a commonly thought to be narrow as a consequence of plant breeding (Tanksley and McCouch 1997). As to masson pine, the elucidation of genetic diversity is highly important for genetic improvement, whereas, genetic uniformity was seen throughout South China caused probably by large-scale artificial afforestation during variety testing (Li and Peng 2001; Peng et al. 2003; Li et al. 2009). Up to date, it has been the first report of IRAP-based assessment of genetic diversity in masson pine. The dendrogram demonstrated that 7 groups were obtained taking 0.73 as a threshold (Fig. 4). Populations from different groups could be introduced as parents with enough genetic distance to produce hybrids on masson pine. Results from IRAP maker data presented here revealed no significant differentiation among tested masson pine germplasms, which was consistent with previous studies (Peng et al. 2003; Li et al. 2009). Numerous lines of evidence indicated that genetic diversity across the cultivars was lower compared with the wild lines (Mandel et al. 2011; Aranzana et al. 2012; Liu et al. 2012). Therefore, it was essential to enforce the protection of natural populations and to expand the scope of protection of natural populations on masson pine. In our study, a set of retrotransposon primers were developed, but only nine single polymorphic primers were used to test the genetic diversity of masson pine, while a series of primers can be derived in the future if they successively pairs with each other in IRAP analysis, with various anchored SSR primers for REMAP analysis or with different kinds of restriction endonuclease adapter primers for SSAP and MSAP analysis (Du et al. 2009).

In conclusion, we isolated the 3'-LTR sequences of Ty1-*copia* and Ty3-*gypsy* retrotransposon sequence separately in the masson pine genome. This study demonstrated the utility of high-resolution IRAP markers based on the LTRs and RTs of retrotransposons for distinguishing masson pine cultivated lines, and confirmed that those markers are highly informative for genetic diversity and phylogenetic studies. It also showed that primers designed based on masson pine retrotransposon families can be readily used across species lines. We did not aim here to show primer combinations that are more informative than single IRAP — this task awaits further large-scale genetic diversity analysis of masson pine.

Acknowledgments The project is supported by grants from Special Core Program of Guizhou Province, P. R. China (20126011–1), the 863 Program of China (2011AA10020301), and the National Natural Science Foundation of China (31260183). We gratefully acknowledge Lianfeng Wu, Harvard Medical School, for critical reading of the manuscript.

Data archiving statement In this study, there are two obtained sequences have been submitted to the National Center for Biotechnology Information (<http://www.ncbi.nlm.nih.gov/>), as listed below:

PmRT (GenBank accession no. KC355438)

REPM (GenBank accession no. KC355439)

References

- Abdollahi Mandoulakani B, Piri Y, Darvishzadeh R, Beroosi I, Jafari M (2011) Diversity in *Medicago sativa* populations revealed by IRAP and REMAP markers. *Plant Mol Biol Report*. doi:10.1007/s11105-011-0338-x
- Abdollahi Mandoulakani B, Piri Y, Darvishzadeh R, Beroosi I, Jafari M (2012) Retroelement insertional polymorphism and genetic diversity in *Medicago sativa* populations revealed by IRAP and REMAP markers. *Plant Mol Biol Report* 30(2):286–296
- Aranzana MJ, Barreneche T, Arús P (2012) Diversity analysis. In: Kole C, Abbott AG (eds) *Genetics, genomics and breeding of stone fruits*, 55. CRC Press, Boca Raton, USA
- Belyayev A, Kalendar R, Brodsky L, Nevo E, Schulman AH, Raskina O (2010) Transposable elements in a marginal plant population: temporal fluctuations provide new insights into genome evolution of wild diploid wheat. *Mob DNA* 1:6
- Biswas MK, Xu Q, Deng XX (2010) Utility of RAPD, ISSR, IRAP and REMAP markers for the genetic analysis of *Citrus* spp. *Sci Hortic* 124(2):254–261
- Boeke JD, Corces VG (1989) Transcription and reverse transcription of retrotransposons. *Annu Rev Microbiol* 43:403–434
- Cai JJ, Ji KS (2009) Diversity of SSR markers and strategy of genetic linkage map making with masson's pine megagametophytes. *Mol Plant Breed* 7(5):934–940 (in Chinese with English abstract)
- Campbell BC, LeMare S, Piperidis G, Godwin ID (2011) IRAP, a retrotransposon-based marker system for the detection of somaclonal variation in barley. *Mol Breed* 27:193–206
- Carrier G, Cunff LL, Dereeper A, Legrand D, Sabot F, Bouchez O, Audeguin L, Boursiquot JM, This P (2012) Transposable elements are a major cause of somatic polymorphism in *Vitis vinifera* L. *PLoS One* 7(3):e32973
- Devos KM, Brown JKM, Bennetzen JL (2002) Genome size reduction through illegitimate recombination counteracts genome expansion in *Arabidopsis*. *Genome Res* 12:1075–1079
- Ding SY, Song YC (1998) Declining causes of *Pinus massoniana* in the processes of succession of evergreen broad leaved forest. *Acta Bot Sin* 40:755–760 (in Chinese with English abstract)
- Du XY, Zhang QL, Luo ZR (2009) Development of retrotransposon primers and their utilization for germplasm identification in *Diospyros* spp. (Ebenaceae). *Tree Genet Genomes* 5:235–245
- Fan FH, Wen XP, Ding GJ, Cui BW (2013) Isolation, identification, and characterization of genomic LTR retrotransposon sequences from masson pine (*Pinus massoniana*). *Tree Genet Genomes*. doi:10.1007/s11295-013-0631-y
- Fedoroff NV (2013) *Plant transposons and genome dynamics in evolution*. Wiley-Blackwell, Oxford, UK. doi: 10.1002/9781118500156.finatter
- Grandbastien MA (1998) Activation of plant retrotransposons under stress conditions. *Trends Plant Sci* 3(5):181–187
- Grotkopp E, Rejmánek M, Sanderson MJ, Rost TL (2004) Evolution of genome size in pines (*Pinus*) and its life-history correlates: supertree analyses. *Evolution* 58(8):1705–1729
- 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
- Huang CRL, Burns KH, Boeke JD (2012) Active transposition in genomes. *Genetics* 46(1):651
- International Brachypodium Initiative (2010) Genome sequencing and analysis of the model grass *Brachypodium distachyon*. *Nature* 463:763–768
- Jiang B, Wu ZM, Lou QF, Wang D, Zhang WP, Chen JF (2010) Genetic diversity of Ty1-copia retrotransposons in a wild species of *Cucumis* (*C. hystrix*). *Sci Hortic* 127:46–53
- Jin YK, Bennetzen JL (1989) Structure and coding properties of Bs1, a maize retrovirus-like transposon (transposable element/retrotransposon/reverse transcription/translation frameshift). *Proc Natl Acad Sci* 86:6235–6239
- Jing R, Ambrose MA, Knox MR, Smykal P, Hybl M, Ramos A, Ellis TH (2012) Genetic diversity in European *Pisum* germplasm collections. *Theor Appl Genet* 125:367–380
- Kalendar R, Schulman AH (2006) IRAP and REMAP for retrotransposon-based genotyping and fingerprinting. *Nat Protoc* 1:2478–2484
- Kalendar R, Flavell AJ, Ellis THN, Sjakste T, Moisy C, Schulman AH (2011) Analysis of plant diversity with retrotransposon-based molecular markers. *Heredity* 106:520–530
- Kimura Y, Tosa Y, Shimada S, Sogo R, Kusaba M, Sunaga T, Betsuyaku S, Eto Y, Nakayashiki H, Mayama S (2001) *OARE-1*, a Ty1-copia retrotransposon in oat activated by abiotic and biotic stresses. *Plant Cell Physiol* 42(12):1345–1354
- Kumar A, Bennetzen JL (1999) Plant retrotransposons. *Annu Rev Genet* 30:479–532
- Li D, Peng SL (2001) Genetic diversity in three *Pinus massoniana* populations in different elevations and its relationship with ecological factors. *Acta Ecol Sin* 21:415–421 (in Chinese with English abstract)
- Li ZH, Chen Y, Zhang DL, Yang MH, Jiang Y, Ding GJ, Tan XF (2009) ISSR analysis of genetic diversity of *Pinus massoniana* on Gupeng and Langshui nature populations in Guangxi. *Chin Agric Sci Bull* 16:116–119 (in Chinese with English abstract)
- Liu FL, Yao JT, Wang XL, Repnikova A, Galanin DA, Duan DL (2012) Genetic diversity and structure within and between wild and cultivated *Saccharina japonica* (Laminariales, Phaeophyta) revealed by SSR markers. *Aquaculture* 358–359:139–145
- Lou Q, Chen J (2007) Ty1-copia retrotransposon-based SSAP marker development and its potential in the genetic study of cucurbits. *Genome* 50:802–810
- Mandel JR, Dechaine JM, Marek LF, Burke JM (2011) Genetic diversity and population structure in cultivated sunflower and a comparison to its wild progenitor, *Helianthus annuus* L. *Theor Appl Genet* 123(5):693–704
- Morse AM, Peterson DG, Islam-Faridi MN, Smith KE, Magbanua Z, Garcia SA, Kubisiak TL, Amerson HV, Carlson JE, Nelson CD, Davis JM (2009) Evolution of genome size and complexity in *Pinus*. *PLoS One* 4(2):e4332
- Nystedt B, Street NR, Wetterbom A, Zuccolo A, Lin YC, Scofield DG, Vezzi F, Delhomme N, Giacomello S, Alexeyenko A, Vicedomini R, Sahlin K, Sherwood E, Elfstrand M, Gramzow L, Holmberg K, Hällman J, Keech O, Klasson L, Koriabine M, Kucukoglu M, Käller M, Luthman J, Lysholm F, Niittylä T, Olson A, Rilakovic N, Ritland C, Rosselló JA, Sena J, Svensson T, Talavera-López C, Theißen G, Tuominen H, Vanneste K, Wu ZQ, Zhang B, Zerbe P, Arvestad L, Bhalariao R, Bohlmann J, Bousquet J, Gil RG, Hvidsten TR, de Jong P, MacKay J, Morgante M, Ritland K, Sundberg B, Thompson SL, de Peer YV, Andersson B, Nilsson O, Ingvarsson PK, Lundeberg J, Jansson S (2013) The Norway spruce genome sequence and conifer genome evolution. *Nature*. doi: 10.1038/nature12211
- Park M, Jo S, Kwon JK, Park J, Ahn JH, Kim S, Lee YH, Yang TJ, Hur CG, Kang BC, Kim BD, Choi D (2011) Comparative analysis of pepper and tomato reveals euchromatin expansion of pepper genome caused by differential accumulation of Ty3/Gypsy-like elements. *BMC Genomics* 12:85
- Pearce SR, Stuart-Rogers C, Kumar A, Flavell AJ (1999) Rapid isolation of plant Ty1-copia group retrotransposons LTR sequences for molecular marker studies. *Plant J* 19:1–7
- Peng SL, Li QF, Li D, Wang ZF, Wang DP (2003) Genetic diversity of *Pinus massoniana* revealed by RAPD markers. *Silvae Genet* 52(2):60–63
- Schulman AH, Flavell AJ, Paux E, Ellis TH (2012) The application of LTR retrotransposons as molecular markers in plants. In *Mobile Genetic Elements* 859:115–153. Humana Press, USA

- Smýkal P, Bačová-Kertessová N, Kalendar R, Corander J, Schulman AH, Pavelek M (2011) Genetic diversity of cultivated flax (*Linum usitatissimum* L.) germplasm assessed by retrotransposon-based markers. *Theor Appl Genet* 122:1385–1397
- Staton SE, Ungerer MC, Moore RC (2009) The genomic organization of Ty3/gypsy-like retrotransposons in *Helianthus* (Asteraceae) homoploid hybrid species. *Am J Bot* 96(9):1646–1655
- Stuart-Rogers C, Flavell AJ (2001) The evolution of Ty1-*copia* group retrotransposons in Gymnosperms. *Mol Biol Evol* 18(2):155–163
- Suoniemi A, Tanskanen J, Schulman AH (1998) Gypsy-like retrotransposons are widespread in the plant kingdom. *Plant J* 13(5):699–705
- Tanksley SD, McCouch SR (1997) Seed banks and molecular maps: unlocking genetic potential from the wild. *Science* 277(5329):1063–1066
- Ungerer MC, Kawakami T (2013) Transcriptional dynamics of LTR retrotransposons in early generation and ancient sunflower hybrids. *Genome Biol Evol* 5(2):329–337
- Vukich M, Schulman AH, Giordani T, Natali L, Kalendar R, Cavallini A (2009) Genetic variability in sunflower (*Helianthus annuus* L.) and in the *Helianthus* genus as assessed by retrotransposon-based molecular markers. *Theor Appl Genet* 119:1027–1038
- Wicher T, Keller B (2007) Genome-wide comparative analysis of *copia* retrotransposons in Triticeae, rice, and *Arabidopsis* reveals conserved ancient evolutionary lineages and distinct dynamics of individual *copia* families. *Genome Res* 17:1072–1081
- Yuan LX, Zhang QL, Guo DY, Luo ZR (2012) Genetic differences among ‘Luotian-tianshi’ (*Diospyros kaki* Thunb.) genotypes native to China revealed by ISSR and IRAP markers. *Sci Hortic* 137(1):75–80
- Zedek F, Šmerda J, Šmarda P, Bureš P (2010) Correlated evolution of LTR retrotransposons and genome size in the genus *Eleocharis*. *BMC Plant Biol* 10:265
- Zhao GL, Dai HY, Chang LL, Ma Y, Sun HY, He P, Zhang ZZ (2010) 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 Genomes* 6:149–159