# Terminal-Repeat Retrotransposon In Miniature (TRIM) elements and their potential utility as DNA markers in *Brassica* relatives

**Tae-Jin Yang[1], Soo-Jin Kwon[2], Jung Sun Kim[2], Mina Jin[2], Beom-Soon Choi[1], Yong Pyo Lim[3], Beom-Seok Park[2]**

[1]*Department of Plant Science, College of Agriculture and Life Sciences, Seoul National University, San 56-1, Sillim-dong, Gwanak-gu, Seoul, 151-921 Korea Email:tjyang@snu.ac.kr*

[2]*Brassica Genomics Team, National Institute of Agricultural Biotechnology (NIAB), RDA, Suwon, 441-707, Korea*

[3]*Department of Horticulture, Chungnam National University, Daejeon, 305-764 Korea*

**Abstract**

Based on genome wide search for identification of Terminal-Repeat Retrotransposon In Miniature (TRIM) elements, we characterized eight TRIM families; four from *Brassica* and four from *Arabidopsis* (including three families which were identified before) in the family *Brassicaceae*. The elements contain 100~257 bp terminal repeats (TR) and 100~800 bp internal sequences. Studies of insertion site polymorphisms of each element across taxa infer the taxonomic lineage in the family *Brassicaceae* and the insertion time of each element. TRIM display experiments revealed abundant polymorphic bands among taxa in *Brassicaceae* and among commercial F1 varieties of Chinese cabbage and radish suggesting the potential utility of TRIM elements as DNA markers for the study of genome evolution, genome mapping, and variety protection for any *Brassica* crops. Other TRIM features that are advantageous for such studies include their miniature size, insertion preference in euchromatin regions, and ubiquitous distribution as evident in the *Arabidopsis* genome.

**Key words:** Brassica, Arabidopsis, TRIM, Retrotransposon, DNA Markers

## Introduction

The genus *Brassica* is one of the core genera including many important vegetables. Six *Brassica* species are cultivated worldwide: three diploids plant species such as *B. rapa* (syn. *campestris*, A genome, n=10), *B. nigra*(B genome, n=8) *and B. oleracea* (C genome, n=9); and three amphidiploids species such as *B. juncea* (AB genome, n=18), *B. napus* (AC genome, n=19), and *B. carinata* (BC genome, n=17) (U. 1935).

Transposons comprise a significant fraction of the genomes in most organisms, particularly in plants. They play an important role in maintaining chromosome structure, especially for the centromere and telomere structure and euchromatin components (Kumar and Bennetzen 1999; Kumekawa et al. 2001; Yang et al. 2005b; Yang et al. 2005c; Lim et al. 2006). In plant genome, LTR retrotransposons are the major components consisting the large blocks of heterochromatic regions. Recently, a small non-autonomous LTR-type retrotransposons are identified and coined as terminal-repeat retrotransposons in miniature (TRIM) (Witte et al. 2001; Yang et al. 2005c). TRIM elements have terminal repeat (TR) sequences ranging from 100 bp to 250 bp in length that encompass an internal domain of ~300 bp and create 5 bp target site duplications (TSD) by their insertion. Even though TRIM elements lack the coding domains required for mobility, they can transpose, probably by trans-mobilization with help of other retrotransposons, and play roles in restructuring plant genomes by affecting the promoter, coding region, and intron-exon structure of genes (Witte et al. 2001; Yang et al. 2006b). TRIM generally show widespread chromosome dispersion, in many cases at gene-rich euchromatin regions, as several hundreds of members in most *Brassica* relatives.

DNA markers used as powerful tools for practical application in plant breeding and for studying plant diversity. Several molecular marker systems such as RFLP, RAPD, SSR, AFLP, and SNP were used for developing agronomically important tools. Recently, transposon display have been devised, which reveal insertion polymorphism generated by the transposition of these elements. The ubiquitous distribution, high copy number and widespread chromosomal dispersion of TRIM elements will provide excellent potential for developing a DNA marker system. Here, we report characters of eight distinct families of TRIM and a successful application of insertion polymorphism of TRIM elements in various crops in the *Brassicaceae* family.

## Material and Methods

Plant materials of various species in the tribe *Brassiceae* are showin in the previous paper (Yang et al. 2006b). Data analysis, various PCR condition, and TRIM display are described in Yang et al (2006b) and Kwon et al. (in preparation).

## Results

Characterization of TRIM elements Br1, Br2, Br3, and Br4 from Brassica species and At1, At2, At3, and At4 from Arabidopsis and their insertion polymorphism as clade markers

TRIM elements were identified by pairwise comparison of *B. rapa* BAC clones and their homologous sequence in *Arabidopsis* and subsequent BLAST and data mining (Yang et al. 2006a, b). A total of 146 TRIM elements derived from the Brassicaceae are grouped as eight independent lineages: At1, At2, At3, and At4 derived from *Arabidopsis*; Br1, Br2, Br3, and

Br4 derived from the genus *Brassica* (Yang et al. 2006b). All the members have distinct feature of TRIM elements of which have common sequence and structure shown in Fig. 1. Among all the members, Br1 and Br2 were the most abundant members and similar to At1 of *Arabidopsis* as shown in Fig. 2.

The TRIM elements derived from *Arabidopsis* and *Brassica* are belonged to independent lineages by phylogenetic analysis. In contrast, elements derived from *B. rapa* (red dots), *B. oleracea* (gray dots), and *B. napus* (green dots) are intermingled in each Br lineage (Fig. 2). This suggests that TRIM elements were activated after the divergence of *Brassica* and *Arabidopsis*, that is supported by PCR analysis using flanking sequences of the insertion sites against 17 taxa in the family Brassicaceae (Yang et al. 2006b). The distribution patterns of Br element across taxa show their evolutionary lineage and permit inference of the time periods in which the insertion events occurred because it move "copy and paste" mechanism and fix in the host genome (Kumar and Bennetzen 1999). Inspection of various insertion sites of Br elements can be used as clade markers for the species in Brassicaceae such as the previous study using insertion sites of SINE element (Lenoir et al. 1997; Tatout et al. 1999). We have inspected insertion polymorphism against several Br1 and Br2 membes. Among them, one member, Br2-6P20, revealed that it is inserted into the ancient *rapa-oleracea* (A-C genome) clade after divergence from *B. nigra* (B genome) approximately 7.9~4.0 MY ago. Another member, Br2-80A08, is *B. rapa* (A genome)-unique and likely to have been inserted into the *B. rapa* lineage within 3.7~0.01 MYA, after the *rapa-oleracea* split, but before the allopolyploidization of the AB and AC genomes. Another member, Br1-52O08, is identified in only one of four inbred lines of *B. rapa* and one of allopolyploids (AC), indicating it was activated in a certain *B. rapa* lineage during 0.8-0.01 MYA. And one member, Br1-12I15 is unique in only one of four inbred lines of *B. rapa*, indicating it was activated after allopolyploidization events fewer than 10,000 years ago (Yang et al. 2006b).

*Copy numbers and distribution of the TRIM elements in* B. rapa *and* B. oleracea

We have estimated copy numbers of the TRIM elements based on the occurrence of TRIM elements in 96.5 Mb of *B. rapa* BAC end sequences (133,644 BAC end sequences with an average length of 723 bp) and 434.3 Mb of *B. oleracea* shotgun sequences (538,418 whole genome shotgun sequences with an average length of 807 bp), respectively. The numbers are roughly estimated to be present in 660 and 530 copies in *B. oleracea* and *B. rapa* genomes, respectively (Yang et al. 2006b). Distribution of these elements are unknown yet in *B. rapa*. However, five TRIM elements were mapped on gene-rich regions of *B. rapa* chromosomes based on BAC sequences. In silico mapping of the insertion sites in *Arabidopsis* showed that the Katydid-At elements were evenly distributed throughout the genome and some were directly related to the structure of the expressed genes in *Arabidopsis* (Witte et al. 2001; Yang et al. 2006b). Based on evidence of distribution of TRIM elements in the *Arabidopsis* euchromatin regions and occurrence of chimeric feature of TRIM elements in lots of ESTs, the abundant TRIM elements seem to play an important role for restructuring the host genome in *B. rapa* as do MITE elements in the rice genome (Jiang and Wessler 2001; Jiang et al. 2003; IRGSP 2005). Moreover TRIM elements have promoter and terminator sequence in their small terminal repeat sequences that can provide active modification of gene feature by internal insertion.

*TRIM display*

To investigate distribution of the TRIM elements across taxa in the family Brassicaceae, we performed transposon display (TRIM display) (Fig. 3). All the taxa (including *Arabidopsis thaliana*, At) produced over than 30 bands (30-60 bands). Most bands are polymorphic among different genera indicating that Br1- or Br2- like elements are ubiquitous in the genome of family Brassicaceae. Some of polymorphic bands among basic *Brassica* species (dotted line) are appeared in their allopolyploids indicating that they were inserted into the subgenome before allopolyploidization. Bands polymorphic among inbred lines of *B. rapa* (arrows showing polymorphism among lanes A1~A4, Fig. 3) can be mapped because these four lines are parental lines of two broadly known mapping population of *B. rapa* (www.brasssica-rapa.org). To further investigate the potential utility of the TRIM elements as DNA markers, we have inspected polymorphism among commercial F1 cultivars of Chinese cabbage and radish in Korea. Many polymorphic bands were detected among F1 cultivars of Chinese cabbage and Radish (data not shown). Some of bands are common in all cultivars of *B. rapa* and *R. sativus* indicating this marker system can be broadly applied for variety identification in the genus *Brassica*.

*Conclusion: Potential utility of TRIM elements as DNA markers*

We have identified many of expressed sequence tags (ESTs) that include a part of Br elements as chimeric forms. Sequence-level analysis of the triplicated *B. rapa* genome sequence revealed that about 12% and 44% of triplicate genes remained as triplet and doublet paralogues, (Yang et al. 2005a; 2006a). These duplicated genes can attain a selective advantage by gaining new function or silencing in each duplicate in a complementary manner (Hurles 2004). We found that one of the triplicated genes is dramatically rearranged at its 5′ coding and promoter region. Furthermore, TRIM display against cDNA gave more than 11 clear bands indicating many expressed genes are modified by TRIM insertion (Kwon et al. In preparation). Collective data suggest that TRIM elements play an active role for gaining new function by rearrangement of the duplicate genes in the highly replicated *Brassica* genome.

We applied TRIM display against 19 commercialized F1 cultivars in Korea. Data demonstrate most of commercialized F1 cultivars can be distinguished from each other by TRIM display using one or at least two primer pairs. Combination of PCR markers using flanking sequence of the insertion site can be applied for F1 purity test, variety identification, or classification of the commercialized cultivars. The ubiquitous TRIM elements should be applied as a DNA marker system for the studies of genome evolution, genome mapping, and variety identification to protect breeder's right. Their insertional

preference to euchromatin regions and miniature size of the elements (less than 350 bp) providing clear InDel (actually insertion) polymorphism may be another big advantage for their utility as DNA markers.

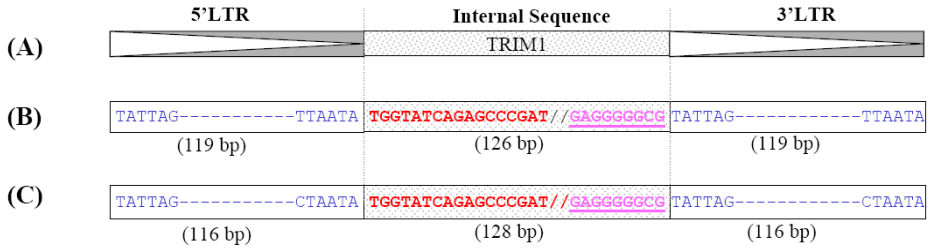


Fig. 1. The general structure of TRIM-Br1 and TRIM-At1. Shematic representation of TRIM (A), signature sequence of Br1 (B) and At1 (C). Overall length of Br1 and At1 is 364 bp and 360 bp, respectively. The LTR sequence (blue letters) contain 5 or 6 bp invert repeat at 5′ and 3′ terminals and the internal sequence begins with PBS of tRNA methionine (red letters) and ends with PPT (Pink letters).

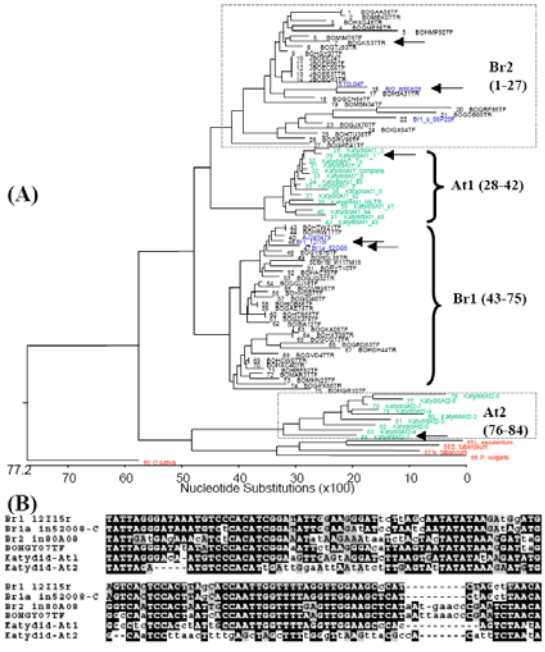


Fig. 2 (Left). Phylogeny of TRIM elements, Br1 and Br2, in Brassicaceae and other plant family (A) and multiple alignment of 5′ terminal repeat (TR) sequences of selected elements (B). The alignment of the selected LTR, arrows in the A panel (A), was created with CLUSTALW (default parameter) and the program boxshade was used for shading. The origins of each element are as follows based on the TIGR database accession numbers (black letters), BES names (blue letters), GenBank accession numbers (red letters), and *Arabidopsis* elements represented in Table 2 (green letters). The real positions of all elements are represented in supplemental Tables 2 and 3.

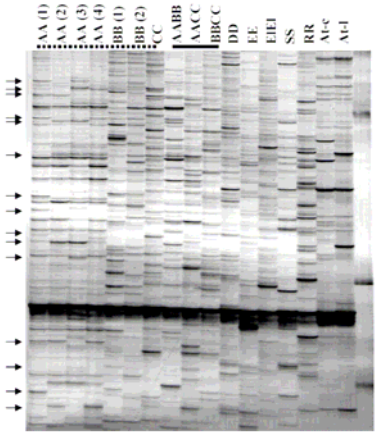


Fig. 3 (Right). TRIM display. Insertion sites for Br1 and Br2 family elements are displayed using the primer set Br1n2DP-L/BfaI-G across 17 species (listed in Fig. 2B) in the family Brassicaceae. Arrows denote polymorphic markers in four inbred lines of *B. rapa*. Bold lines indicate allopolyploids of *Brassica*. Letters on each lane indicate genome; AA for *Brassica rapa* (A genome) , BB for *B. nigra* (B), CC for *B. oleracea* (C), AABB for *B. juncea* (AB), AACC for *B. napus* (AC), BBCC for *B. carinata* (BC), DD for *Diplotaxis erucoides*, EE for *Eruca stenocarpa*, ElEl for *Erucastrum laevigatum*, SS for *Sinapis alba*, RR for *Raphanus sativus*, At-c for *Arabidopsis thaliana* cv. Columbus, At-l for *A. thaliana* cv. Landsberg. Last lane is DNA marker.