Plant Breed. Biotech. 2014 (December) 2(4):322~333
http://dx.doi.org/10.9787/PBB.2014.2.4.322
Online ISSN: 2287-9366
Print ISSN: 2287-9358

REVIEW ARTICLE

Miniature Inverted-repeat Transposable Elements (MITEs) as Valuable Genomic Resources for the Evolution and Breeding of *Brassica* Crops

Perumal Sampath[1], Tae-Jin Yang[1]*

[1]Deptartment of Plant Science, Plant Genomics and Breeding Institute, and Research Institute of Agriculture and Life Sciences, College of Agriculture and Life Sciences, Seoul National University, Seoul 151-921, Republic of Korea

ABSTRACT Transposable elements (TEs) play important roles in structural and functional diversification, genome enlargement, and speciation in plant genome. Their derivatives or small non-autonomous TEs play important roles in the alteration of homologous genes by epigenetic control or structural modification. The miniature inverted-repeat transposable element (MITE) is one of the representative non-autonomous class II TEs. MITEs include high copy members that are widely distributed and in close association with genic regions, which make MITEs useful targets and resources for in-depth understanding of genome evolution, as well as practical applications in molecular breeding. Here, we discuss the important features of MITEs, such as the identification tools of a novel MITE family, structural characterization, distribution pattern analysis, and impact on evolution in highly duplicated *Brassica* genome. We show the characteristics, copy numbers, and distribution patterns of 20 novel MITE families, and represent their putative roles in the evolution of the triplicated *Brassica* genome. We also introduce our MITE database, and discuss the utility of MITEs for developing MITE-derived markers that are useful for molecular breeding of *Brassica* crops.

Keywords *Brassica*, Transposable elements, Miniature inverted-repeat transposable element (MITE), Molecular markers, Breeding

INTRODUCTION

Transposable elements, also known as "mobile genetic elements", are DNA sequence fragments that move, or are copied from one location to another in the genome, either directly, by a cut-and-paste mechanism (class II DNA transposons), or indirectly, by a copy-and-paste mechanism through an RNA intermediate (class I retrotransposons; (Fig. 1) (Feschotte *et al.* 2002). Transposition of both classes of elements may result in a heritable increase in copy number within the genome; hence, individual TE types are found in multiple copies (often referred to as a TE family), and constitute the majority of the repetitive fraction of eukaryotic genomes (Wicker *et al.* 2007). Large-scale sequencing of eukaryotic genomes has revealed that TEs are the most abundant component of most eukaryotic genomes, are ubiquitously present, and occupy large fractions of genomes: TEs account for 40% of *Oryza sativa* (rice) (Paterson *et al.* 2009), 50% of *Glycine max* (soybean) (Schmutz *et al.* 2010), and >80% of *Zea mays* (maize), *Triticum aestivum* (wheat), and *Hordeum vulgare* (barley) (Bennett and Smith 1976; Paterson *et al.* 2009; Wicker *et al.* 2009) genomes.

Whole-genome analyses estimated that 39.5% of the *B. rapa* ($2n = 2x = 529$ Mb), 38.8% of the *B. oleracea* ($2n = 2x =$ 630 Mb), and 34.8% of its allopolyploid *B. napus* ($2n = 4x =$ 1130 Mb) genomes are occupied by transposon-related sequences (Chalhoub *et al.* 2014; Liu *et al.* 2014; Wang *et al.* 2011). High proportions of TEs are intact in *B. rapa* and *B. oleracea* genomes (68% and 98%, respectively), although TEs have been continuously amplified in both genomes since at least 4.6 million years ago (MYA) (Liu *et al.* 2014). Compared to *B. rapa*, *B. oleracea* has many younger TEs, which are responsible for its increased genome size. *B.*

Received December 26, 2014; Revised December 27, 2014; Accepted December 27, 2014; Published December 31, 2014
***Corresponding author** Tae-Jin Yang, tjyang@snu.ac.kr, Fax: +82-2881-4547

napus genomes contain less than their progenitor, which suggests that a small fraction of the TE has proliferated since the *B. napus* was generated by allotetraploidization (< 0.01 MYA) (Chalhoub *et al.* 2014). Amplification of TEs in the genome can not only cause an increase in genome size, but also help to drive the evolution of genes and genomes (Arkhipova *et al.* 2012), although most TEs are inactive, and are mainly controlled by epigenetic mechanisms (e.g., DNA and histone methylation) (Alzohairy *et al.* 2013; Bire and Rouleux-Bonnin 2012; Fedoroff 2012; Feschotte 2008; Hollister and Gaut 2009; Lisch 2009).

TEs containing their own functional genes for transposition are referred to as autonomous transposable elements (aTEs); whereas, TEs that lack coding genes, and therefore cannot

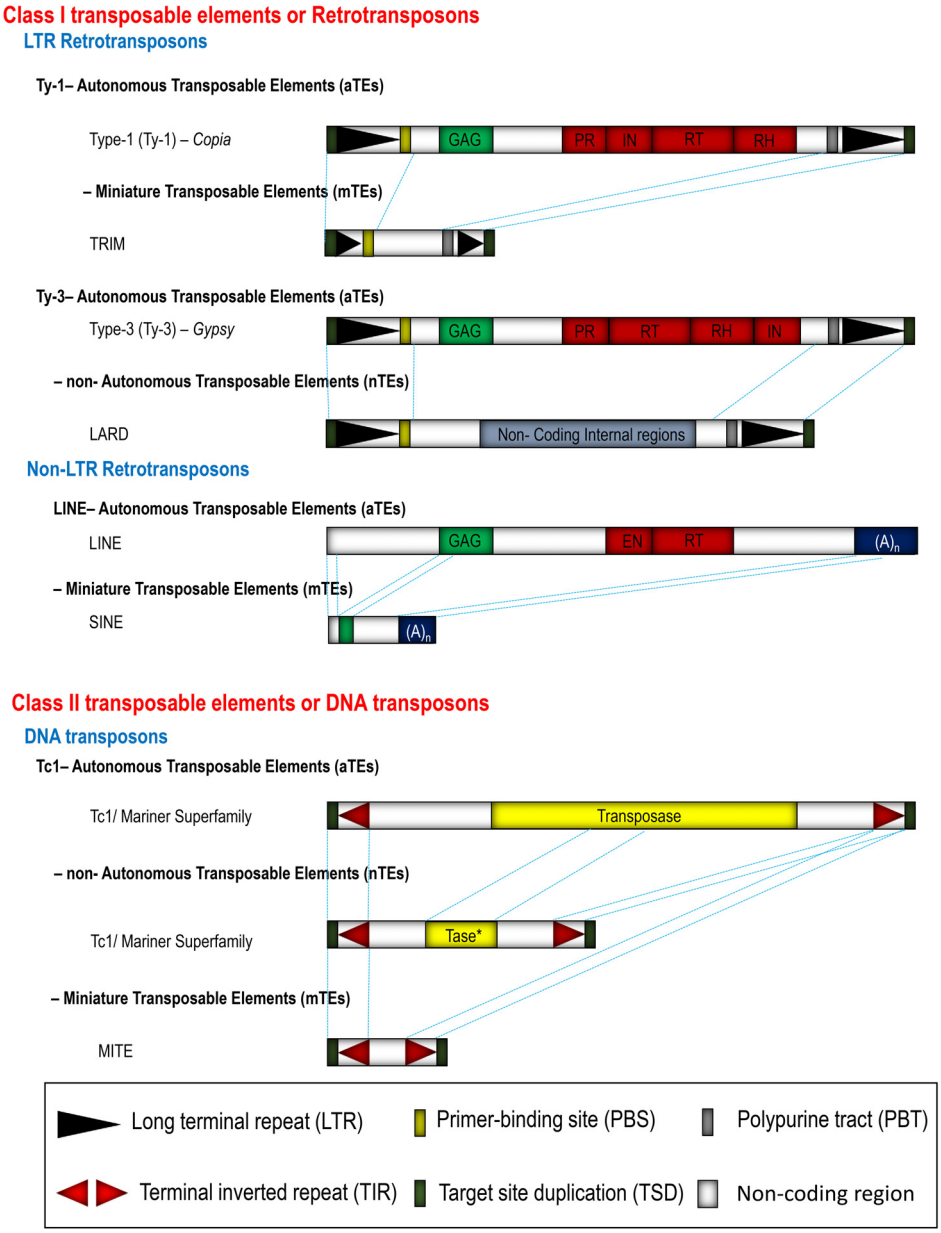

Fig. 1. Classification and structural characteristics of TEs and mTEs. LARD, large retrotransposon derivative; TRIM, terminal-repeat retrotransposons in miniature; LINE, long interspersed nuclear element; SINE, short interspersed nuclear element; GAG, a structural protein for virus-like particles; PR, protease; IN, integrase; RT, reverse transcriptase; RH, RNAse H; EN, endonuclease.

produce their own transposase or reverse transcriptase, are termed non-autonomous or noncoding transposable elements (nTEs). The nTEs, such as large retrotransposon derivatives (LARDs), terminal repeat retrotransposons in miniature (TRIMs), short interspersed elements (SINEs), and miniature inverted-repeat transposable elements (MITEs), are generally deletion derivatives of aTEs, and require a *trans*-acting transposase from their corresponding autonomous partner elements for transposition. TRIMs, SINEs, and MITEs are examples of miniature transposable elements (mTEs; Fig. 1) (Casacuberta and Santiago 2003; Feschotte and Pritham 2007); and families belonging to each type of mTE have had significant influences on gene and genome evolution (Wessler 2006). Here, we summarize the characteristics, copy numbers, and comparative distribution of 20 MITE families in *B. rapa, B. oleracea* and *B. napus*. We also discuss the utility of MITEs for genomics-assisted breeding and evolutionary studies.

Characteristics and distribution of MITEs

MITEs are class II non-autonomous TEs that are characterized by relatively small size (< 800 bp), AT-rich sequences, and flanking terminal inverted repeats (TIRs) ranging from 10 to 200 bp. Insertion of a MITE produces TSD ranging from 2 to 11 bp, depending on the MITE superfamily involved (Fig. 1). The TIRs are more conserved than their respective internal sequences, and act as a recognition site for endonucleases for the integration of TEs via transposition (Casacuberta and Santiago 2003; Lu *et al.* 2012). The TIRs are complementary to each other, leading to the formation of a secondary loop structure, which can be a source of small RNA, and may act in gene regulation. The internal sequences of MITEs have sequence diversity, due to the influence of unrelated autonomous TEs during transposition (Sampath *et al.* 2013; Yaakov and Kashkush 2012). Unlike TRIMs and SINEs, transposition of MITEs occurs by cut-and-paste mechanisms; and MITEs can be amplified in the genome by abortive gap repair, bursts of amplification, or as yet unknown mechanisms under stress (Casacuberta 2013; Fattash *et al.* 2013).

Different MITE families are classified based upon TSD length, structure, and sequence similarity to the putative transposase of the corresponding DNA transposon. MITEs were first identified in the maize genome, and later in various other plant and animal genomes (Bureau and Wessler 1992, 1994; Feschotte *et al.* 2002). So far, seven MITE superfamilies have been identified in plants, although 15 superfamilies of DNA transposons have been reported (Fattash *et al.* 2013). MITEs comprise two major families, namely *Stowaway*-like (with TA as the TSD), and *Tourist*-like (with TAA as the TSD), as well as several other minor families, including *hAT*-like (with 5, 6, or 8 bp TSDs), *MULE* (with 9-10 bp TSDs), and *En/Spm* (3-bp TSDs) MITEs (Oki *et al.* 2008).

MITEs include the mTE members with the most copies, distributed throughout the genome. MITE family members occupy different proportions in plant and animal genomes, reaching up to 10% in rice, 8% in *Medicago*, 4% in *B. rapa,* 0.71% in *A. thaliana,* and 16% in *Aedes aegypti* (yellow fever mosquito) (Chen *et al.* 2013). In silico analysis reveals 174 families with more than 45,821 members, including *Tourist* (56), *Stowaway* (16), *hAT* (90), *Mutator* (11), and *CACTA* (1) in the *B. rapa* genome. Furthermore, we intensively analyzed for 20 MITE families, and identified their roles and utility for genomics and breeding in the *Brassica* genus (Chen *et al.* 2013; Sampath *et al.* 2014).

Identification of a novel MITE family

There are various bioinformatics tools available for the mining of MITEs within genomes, each with its own advantages and drawbacks (Janicki *et al.* 2011). Sequence similarity-based analysis tools require a known MITE family sequence for searches; whereas, structure-based MITE mining tools promote the identification of novel families based on structural characteristics without known sequence, even though a substantial portion of false positive MITEs are identified (Han and Wessler 2010). Currently, a recently developed database, BrassicaTED (http://im-crop.snu.ac.kr/BrassicaTED/index.php) lists 20 different MITE families and their member distribution in *B. rapa* and *B. oleracea*. The BrassicaTED provides tools for mining and characterization of mTEs and TEs (Murukarthick *et al.* 2014).

Mining of MITEs on the genome scale can be performed using various genomics tools. For instance, FINDMITE (Tu 2001), MUST (Chen *et al.* 2009), MITE Hunter (Han

and Wessler 2010), RSBP (Lu *et al.* 2012), and MITE digger (Yang 2013) are available online to identify MITEs, based on signature structures, such as the TIR and TSD. Repbase, Repeatmasker, Inverted Repeat Finder, REPuter, RECON, Micropeats, and STAN can also be used to mine the MITEs, based on sequence similarity. A recently developed database for plant MITEs (P-MITE) contains MITEs from 40 different species, including *Brassica* (Chen *et al.* 2013).

Influence of MITEs on the evolution of the triplicated *Brassica* genome

MITEs can play an important role in the regulation of gene expression and rearrangement of gene structure (Benjak *et al.* 2009). Transposition of MITEs into genes has been found to modify gene structure and function by deletion, point mutation, and by affecting the transcriptional activity (Castelletti *et al.* 2014; Mo *et al.* 2012; Naito *et al.* 2006; Naito *et al.* 2009; Shirasawa *et al.* 2012).

The Brassiceae tribe diverged by the recent hexaploidization approximately 15 MYA, after divergence with Arabidopsis approximately 18 MYA (Yang *et al*. 2006, Liu *et al*. 2014). The *Brassica* genus has overall triplicated chromosome segments, and approximately 12, 44, and 44% of the triplicated genes remain as 3, 2, and 1 copies, respectively, in the current *Brassica* genomes (Yang *et al.* 2006). We have identified that MITE insertion played a role for the modification or sub-functionalization of the duplicated genes in the *Brassica* genome. Specifically, MITE transposition into introns in triplicated *B. rapa* genes appears to underlie their differential expression patterns (Sampath *et al.* 2013). Although most MITEs are associated with genic regions, they are generally not found in exons. An exception is a *tourist* family of MITEs from *B. rapa*, *BraTo-9*, which is preferentially present in the exons of triplicated *B. rapa* genes. *BraTo-9* has provided new exons for functional genes of *B. rapa* (Sampath *et al.* 2014). When *BraTo-9* insertion occurred in triplicated or duplicated genes of *B. rapa*, the element was always found in only one of the duplicated or triplicated genes, suggesting that the *BraTo-9* members were actively amplified in *B. rapa,* after divergence with *B. oleracea* 4.6 MYA.

MITE excision has caused gene knockout or silencing, and up- or down-regulation of gene expression by gene rearrangement, *trans* duplication, and footprint mutation (Naito *et al.* 2009; Shirasawa *et al.* 2012). In addition, MITEs are sources of small interfering RNA, and can control genes in their vicinity (Kuang *et al.* 2009; Piriyapongsa *et al.* 2007). *Trans* duplication in MITEs (i.e., MITEs with host gene sequence captured during excision) increases the likelihood of generating siRNA, which can influence gene regulation (Benjak *et al.* 2009; van Leeuwen *et al.* 2003). For example, a MITE-based siRNA represses the expression of nearby genes, by acting as a functional regulator triggering DNA methylation, and thereby affects agronomic traits, such as leaf angle, plant height, and inflorescence morphology. MITEs also have the ability to escape from silencing more efficiently than other TEs (Parisod *et al.* 2010).

Contribution of MITEs to the evolution of *Brassica* species

Recent genome projects revealed a complete genome sequence of three *Brassica* species, *B. rapa* (AA genome), *B. oleracea* (CC genome), and *B. napus* (AACC genomes). The chromosomal level synteny remained between AA and CC genomes, even though there are some rearrangements between eight and seven chromosomes of AA and CC chromosomes, respectively, after divergence about 4.6 MYA (Wang *et al.* 2000, Liu *et al.* 2014). Both sub genomes remained as intact chromosome in AACC genomes, since allotetraploidization at approximately 8,000 years ago (Chalhoub *et al.* 2014).

We identified the intact copy numbers for each MITE family for each MITE in the genome sequence assembly of *B. rapa*, *B. oleracea,* and *B. napus,* based on homology search using the criteria of 80% sequence similarity with 80% coverage, using representative members of 20 MITE families (Table 1) (Sampath *et al.* 2014). We identified 1,600~3,000 intact members belonging to the 20 MITE families in the three species. MITE members from *B. rapa* and *B. oleracea* were randomly distributed throughout the genome, and resided in various genic regions, intergenic spaces, and near genic regions (Sampath *et al.* 2014). In addition, some MITE families were more abundant in one of the two basal *Brassica* species, *B. rapa* and *B. oleracea,* which suggests that mTE members were greatly amplified after the *B. rapa* and *B. oleracea* diversification about 4.6

MYA (Sampath *et al.* 2014).

MITEs also show significant divergence in copy number between *Brassica* species (Table 1, Fig. 2). *B. napus* has a few less copies compared to its diploid ancestors, *B. oleracea* and *B. rapa,* which suggests that MITEs proliferated in each species, after allotetraploidization of *B. napus* around 8,000 years ago. Meanwhile, BraTo-1 and BraHAT1-1 show higher copy numbers in *B. napus* than its progenitor genomes, which suggests that these two families were selectively amplified in the *B. napus* genome (Table 1). Three species showed micro level synteny between the corresponding segments. Comparison of micro-synteny revealed the insertion time of each MITE member, by comparison of species-specific insertions among the three species. If MITE is common in the counterparts of the three genomes, we can estimate that the MITE member was inserted into the region before 4.6 MYA (Fig. 3A). If the MITE is common in AACC genome and one of the progenitor genomes (AA or CC), but not in the other progenitor genome, the insertion happened during 4.6 ~ 0.01 MYA (Fig. 3B). If the MITE is unique in one of the three species, the insertion happened later than 0.01 MYA (Sampath *et al.* 2013). It is estimated that many of the MITE members were recently amplified by species-unique manner, which indicates that the MITEs are one of the contributors for genome evolution after speciation. We also detect a lot of copy number difference between accessions in the same species, which suggests that MITEs

Table 1. Members of 20 MITE families in the *B. rapa* and *B. oleracea* pseudo-chromosome sequences.

MITE No.	MITE ID[z)]	unit size (bp)	Copies in genome assembly[y)]		
			B. rapa	*B. oleracea*	*B. napus*
1	BraSto-1	267	16	50	131
2	BraSto-2	260	401	210	612
3	BraSto-3	242	6	2	3
4	BraSto-4	558	97	336	374
5	BraTo-1	212	8	127	191
6	BraTo-2	366	61	60	99
7	BraTo-3	252	245	116	309
8	BraTo-4	160	287	36	257
9	BraTo-5	286	118	37	133
10	BraTo-6	257	60	76	100
11	BraTo-7	366	54	199	245
12	BraTo-8	348	29	26	32
13	BraTo-9	264	20	32	81
14	BraTo-10	255	35	50	43
15	BraTo-11	305	4	5	8
16	BraTo-12	273	66	67	80
17	BraTo-13	268	74	85	146
18	BraHAT-1	439	24	55	72
19	BraHAT-2	248	16	19	63
20	BraMu-1	271	24	16	31
		Total	1645	1604	3010

[z)]Conserved MITE sequences were used based on previous study (Sampath *et al.* 2014).

[y)]MITE copies were identified from the available 283, 385, and 850 Mb whole-genome pseudo-chromosome sequences of *B. rapa*, *B. oleracea* and *B. napus*, respectively, with 80% sequence similarity.

are an important target for molecular breeding purposes and evolutionary studies.

Utility of MITEs as molecular markers

DNA markers are used for a wide range of genomic applications, such as the construction of genetic linkage maps, genome-wide association studies, and evolutionary studies (Casa *et al.* 2000; Kalendar and Schulman 2014; Yaakov *et al.* 2012). TEs have been used to develop molecular markers, such as those for inter-retrotransposon amplified polymorphism (IRAP), REtrotransposn-microsatellite amplified polymorphism (REMAP), sequence-specific amplification polymorphism (S-SAP), retrotransposon-based insertion polymorphism (RBIP), inter-MITE polymorphism (IMP), and transposon display (TD) (Agarwal *et al.* 2008). TE-based markers have been successfully utilized for various genomics purposes, such as the analysis of genetic diversity, inspection of clonal variation, and breeding. TE

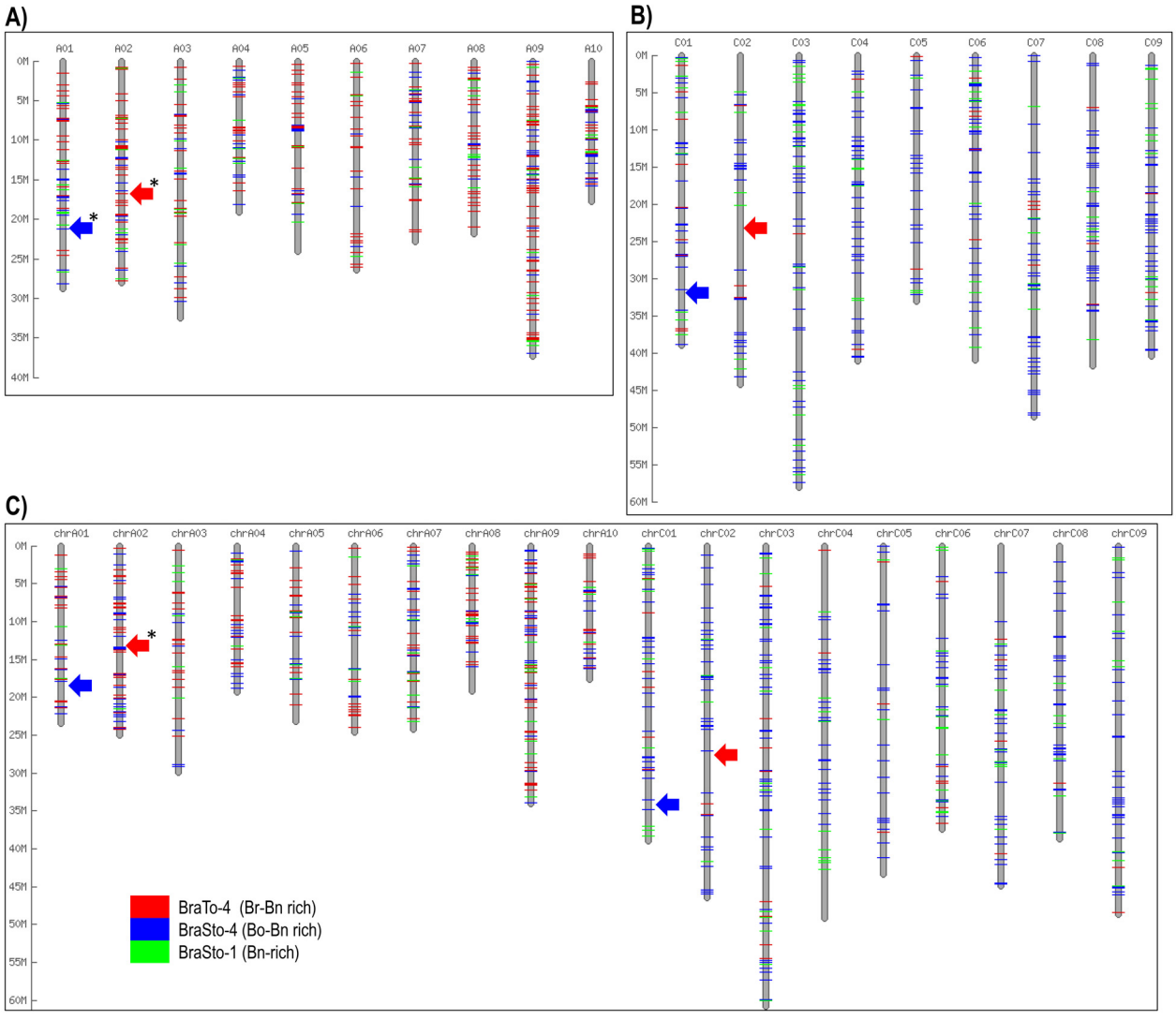

Fig. 2. Differential distribution of MITE family members in *B. rapa*, *B. oleracea*, and *B. napus*. MITE families with intact members were used for *in silico* map construction on the 256-Mb *B. rapa* (A), 385-Mb *B. oleracea* (B), and the 645-Mb *B. napus* pseudo-chromosome sequences, based on physical positions. The red and blue arrows indicate the syntenic regions corresponding to Figs. 3A and 3B, respectively. The arrows with and without star indicate the positions of genes that have MITE insertion and non-insertion, respectively, according to Fig. 3. The physical position information for the MITE families can be found in BrassicTED (Murukarthick *et al.* 2014).

markers are also useful to identify unambiguous gene flow between closely related species (Bire and Rouleux-Bonnin 2012; Carrier *et al.* 2012). The principle characteristics of mTEs, namely their abundance, small size, stability, and distribution in genic regions, are advantageous for DNA marker development in both plants and animals. Thus, so-called mTE markers have been developed from mTEs, such as TRIMs, SINEs, and MITEs (Casa *et al.* 2000; Deragon and Zhang 2006; Kwon *et al.* 2007; Witte *et al.* 2001).

Insertion polymorphism of MITEs

The presence (inserted site) or absence (empty site) of a

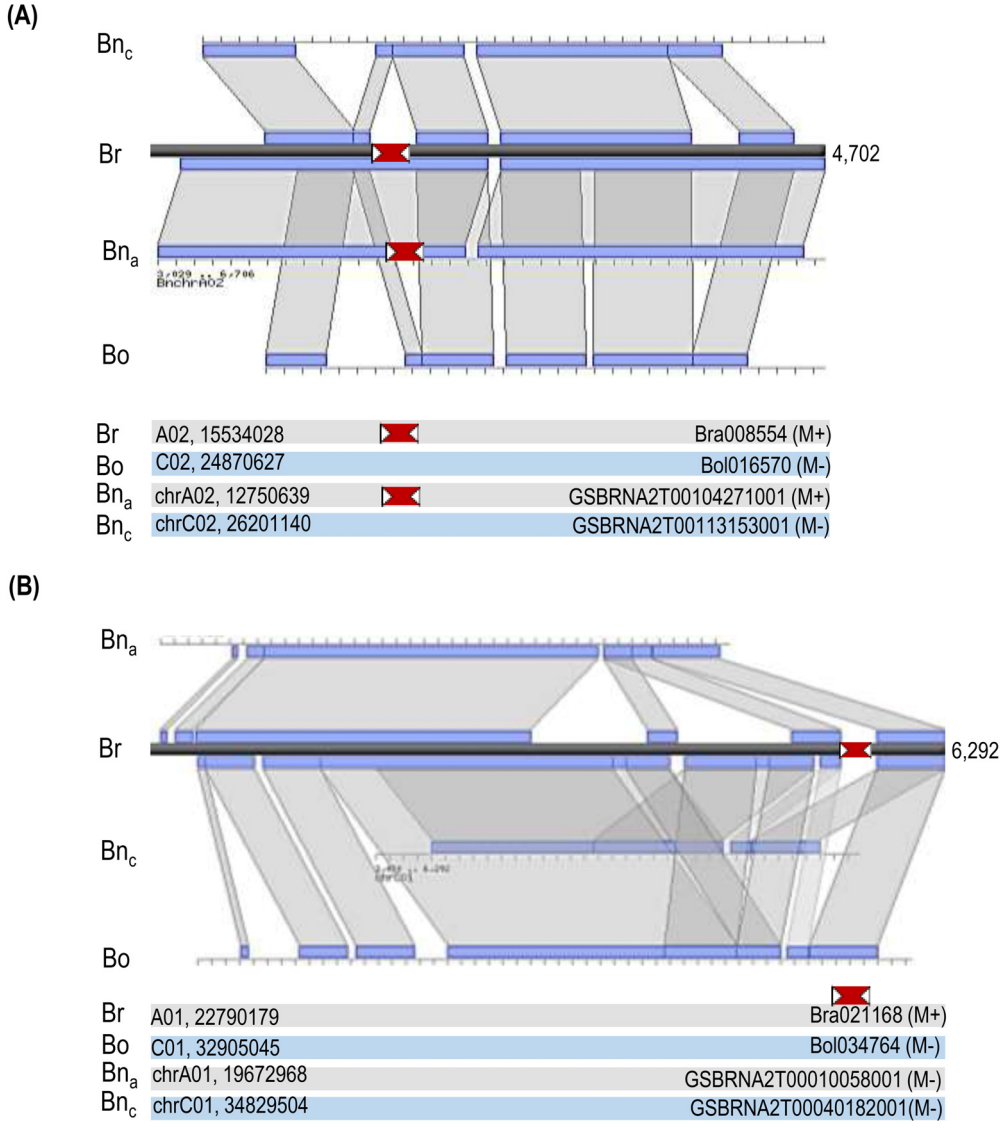

Fig. 3. Micro-synteny comparison of *B. rapa* genomic regions containing MITE (*BraSto-2*) with their non-inserted orthologs (NIOs) in *B. oleracea* and *B. napus*. (A) Micro-synteny between the genomic region, showing the shared insertion of *BraSto-2* in genes of *B. rapa* (Bra008554) and *B. napus* (GSBRNA2T00104271001), compared with those of its NIOs of *B. oleracea* (Bol016570) and *B. napus* (GSBRNA2T00113153001). (B) Micro-synteny between the genomic regions, showing unique insertion of *BraSto-2* in *B. rapa* (Bra021168) gene compared with those of its NIOs of *B. oleracea* (Bol034764) and *B. napus* (GSBRNA2T00010058001 and GSBRNA2T00040182001). MITE element insertions are shown as red bars, and +/- indicate genes with MITE (M) insertion and non-insertion, respectively. The gray bars connecting boxes on genome sequences indicate syntenic blocks present in both sequences.

MITE at a particular locus can be different among accessions; and this insertion polymorphism (IP) can be surveyed by PCR analysis using primers designed from the MITE flanking region (Yaakov *et al.* 2012) (Fig. 4A). The MITE markers have an advantage over other types of markers, because the stability and high copy numbers of MITEs allow the development of abundant markers. In addition, the close association of MITE members with genic regions is beneficial for the development of markers at functional genomic regions and for tightly linked genes, making MITEs good targets for genomics studies (Monden *et al.* 2009). IP markers represent co-dominant alleles at a single locus, and can be used for applications such as the identification of genome duplication or allopolyploidization events, genetic diversity analysis among related accessions or species, and development of markers and mapping using segregating populations between parental lines (Figs. 4B, C, D) (Sampath *et al.* 2013).

IP markers can be produced from sequences harboring MITE elements near the genes of interest. MITE insertion polymorphism (MIP) markers have been extensively studied in rice using a *Tourist* family MITE, *mPing*, the first active MITE identified in eukaryotes (Monden *et al.* 2009). MIP markers based on three MITEs (*Hbr, zmv1,*

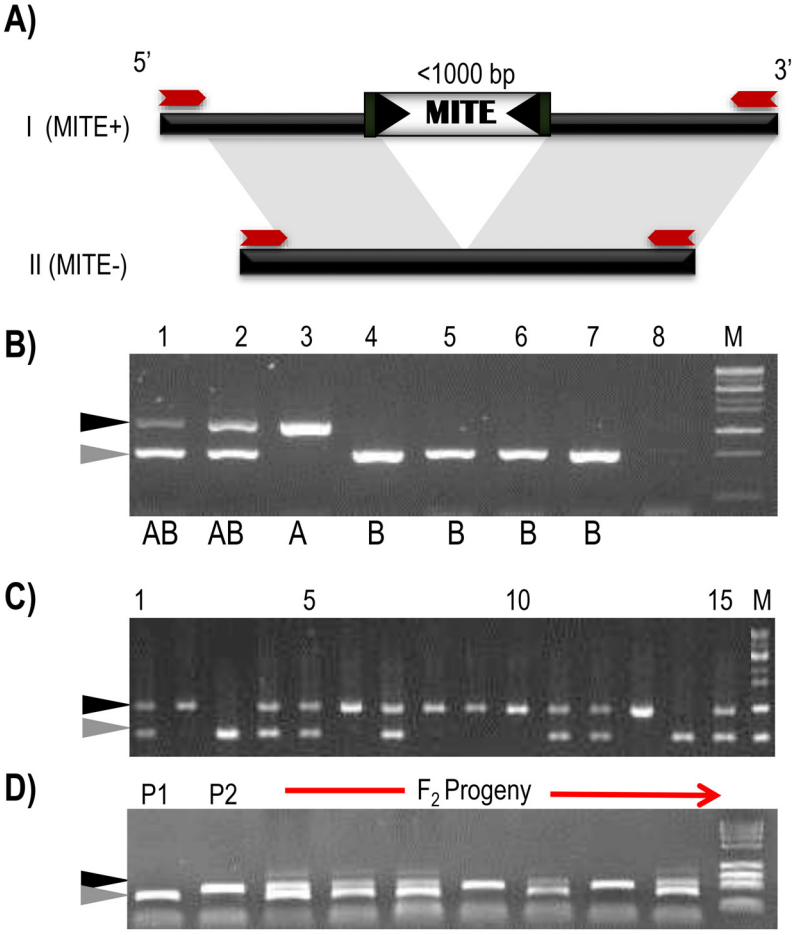

Fig. 4. Utility of MITEs as molecular markers. A) MITE insertion polymorphism analysis using flanking primers. Comparison of DNA fragments showing the presence or absence of MITE insertion. MITE-flanking primer positions are indicated as red arrowheads. B) Polymorphism profile by MIP analysis of 7 *Brassica* accessions based on *BraMi-1*, a *Brassica* MITE. AB, insertion and non-insertion (Heterozygous insertion); A, Insertion (Homozygous insertion); B, non-insertion (Homozygous non-insertion) (Sampath *et al*. 2013). C) Diversity analysis using different *B. oleracea* commercial cultivars. D) Genotyping analysis of 94 *B. oleracea* F_2 plants from a cross between parental lines C1234 (P1) and C1184 (P2).

Ins2) were successfully used to study genetic diversity, and identify a new candidate gene for flowering time variation in maize (Casa *et al.* 2000). MIP markers have also been used for high-resolution genetic diversity analysis, and to elucidate the evolutionary history of *Triticum* (Yaakov and Kashkush 2012). A MIP survey of three different *Brassica* accessions revealed high levels of inter- and intra-species polymorphism, at 52% (150 markers) and 23% (66 markers), respectively. Transposition of MITEs and evolutionary dynamics were also evaluated in *Brassica* species using an MIP approach (Sampath *et al.* 2013). Thus, MITEs can be valuable target genomic resources to produce high numbers of co-dominant markers. It is important to note that high IP ratios are dependent on recent activation and high copy numbers for the target MITEs.

Transposon display (TD) for MITEs

TD is a modification of the AFLP method to target TEs, and amplify most of the insertion sites of TEs. TD is an efficient approach for rapid marker development, because multiple insertion sites can be simultaneously amplified, using conserved sequences of target MITEs that are distributed throughout the genome. TD was first developed and used for the maize *heartbreaker* MITE family (Casa *et al.* 2004). TD can be performed with primers targeting conserved regions of MITEs, such terminal inverted repeats (TIRs) for MITEs. TD-based markers have been effectively utilized for examining the genetic diversity, phylogenetic analysis, genetic mapping, and identification of activation time of TEs, based on divergence time and evolutionary studies (Monden *et al.* 2009; Naito *et al.* 2006). MITE-based display, termed MITE-TD, has been applied for genome-wide detection of insertion sites that are polymorphic between or within species, such as rice, maize, *Brassica*, *Vitis vinifera* (grapevine), and mosquito (Naito *et al.* 2006; Naito *et al.* 2009; Zhang *et al.* 2000). The MITE-TD approach has advantages over AFLP, because MITEs are more widely distributed in genome, especially in euchromatin regions. In addition, MITEs are closely associated with genic regions, which may help to develop markers related to agronomically important traits. Reports have also suggested that MITE-TD identifies a higher proportion of polymorphisms than does AFLP.

Next-generation sequencing (NGS) technology produces numerous short DNA reads at relatively low cost, and in a short period of time. NGS has a wide range of applications, and has revolutionized the use of genomic data for crop improvement (Wei *et al.* 2013). The combination of TD with NGS technology allows the use of different high copy MITE families to detect insertion polymorphism among accessions. This approach will be a powerful tool for molecular breeding and evolutionary analysis.

CONCLUSION

Although MITEs cannot transpose by themselves, due to their lack of protein-coding genes, they have played important roles in plant genome evolution. Understanding the characteristics and member distribution of MITEs will promote their effective utilization to analyze genome evolution, dynamics, and plasticity, as well as to identify the relevant genetic components of germplasm with agronomically important traits in the *Brassica* genome. The MITE-based markers are valuable resources for high-density genetic mapping, diversity analysis, and evolution studies. Furthermore, insertion polymorphism surveys and NGS combined with TD are potential tools for marker systems that are aimed at high throughput marker development, with minimum time and cost.

ACKNOWLEDGMENTS

This work was supported by the Golden Seed Project (Centre for Horticultural Seed Development, No. 213003-04-1-SB430), of the Ministry of Agriculture, Food and Rural Affairs (MAFRA).

REFERENCES

Agarwal M, Shrivastava N, Padh H. 2008. Advances in molecular marker techniques and their applications in plant sciences. Plant cell reports 27: 617-631.

Alzohairy AM, Gyulai G, Jansen RK, *et al.* 2013. Transposable elements domesticated and neofunctionalized by eukaryotic

genomes. Plasmid 69: 1-15.

Arkhipova IR, Batzer MA, Brosius J, *et al.* 2012. Genomic impact of eukaryotic transposable elements. Mobile DNA 3: 19.

Benjak A, Boue S, Forneck A, *et al.* 2009. Recent amplification and impact of MITEs on the genome of grapevine (Vitis vinifera L.). Genome biology and evolution 1: 75-84.

Bennett MD, Smith J. 1976. Nuclear DNA amounts in angiosperms. Philosophical Transactions of the Royal Society of London B, Biological Sciences 274: 227-274.

Bire S, Rouleux-Bonnin F. 2012. Transposable elements as tools for reshaping the genome: it is a huge world after all! Methods Mol Biol 859: 1-28.

Bureau TE, Wessler SR. 1992. Tourist: a large family of small inverted repeat elements frequently associated with maize genes. The Plant cell 4: 1283-1294.

Bureau TE, Wessler SR. 1994. Stowaway: a new family of inverted repeat elements associated with the genes of both monocotyledonous and dicotyledonous plants. The Plant cell 6: 907-916.

Carrier G, Le Cunff L, Dereeper A, *et al.* 2012. Transposable elements are a major cause of somatic polymorphism in Vitis vinifera L. PloS one 7: e32973.

Casa AM, Brouwer C, Nagel A, *et al.* 2000. The MITE family heartbreaker (Hbr): molecular markers in maize. Proceedings of the National Academy of Sciences of the United States of America 97: 10083-10089.

Casa AM, Nagel A, Wessler SR. 2004. MITE display. Methods Mol Biol 260: 175-188.

Casacuberta JM. 2013. MITEs, Miniature Elements with a Major Role in Plant Genome Evolution. Plant Transposable Elements: Impact on Genome Structure and Function 24: 113.

Casacuberta JM, Santiago N. 2003. Plant LTR-retrotransposons and MITEs: control of transposition and impact on the evolution of plant genes and genomes. Gene 311: 1-11.

Castelletti S, Tuberosa R, Pindo M, *et al.* 2014. A MITE transposon insertion is associated with differential methylation at the maize flowering time QTL Vgt1. G3 (Bethesda) 4: 805-812.

Chalhoub B, Denoeud F, Liu S, *et al.* 2014. Plant genetics. Early allopolyploid evolution in the post-Neolithic Brassica napus oilseed genome. Science 345: 950-953.

Chen J, Hu Q, Zhang Y, *et al.* 2013. P-MITE: a database for plant miniature inverted-repeat transposable elements. Nucleic acids research.

Chen Y, Zhou F, Li G, *et al.* 2009. MUST: a system for identification of miniature inverted-repeat transposable elements and applications to Anabaena variabilis and Haloquadratum walsbyi. Gene 436: 1-7.

Deragon J-M, Zhang X. 2006. Short interspersed elements (SINEs) in plants: origin, classification, and use as phylogenetic markers. Systematic biology 55: 949-956.

Fattash I, Rooke R, Wong A, *et al.* 2013. Miniature inverted-repeat transposable elements: discovery, distribution, and activity1. Genome / National Research Council Canada 56: 475-486.

Fedoroff NV. 2012. Transposable elements, epigenetics, and genome evolution. Science 338: 758-767.

Feschotte C. 2008. Transposable elements and the evolution of regulatory networks. Nature reviews Genetics 9: 397-405.

Feschotte C, Jiang N, Wessler SR. 2002. Plant transposable elements: where genetics meets genomics. Nature reviews Genetics 3: 329-341.

Feschotte C, Pritham EJ. 2007. DNA transposons and the evolution of eukaryotic genomes. Annual review of genetics 41: 331.

Han Y, Wessler SR. 2010. MITE-Hunter: a program for discovering miniature inverted-repeat transposable elements from genomic sequences. Nucleic acids research 38: e199.

Hollister JD, Gaut BS. 2009. Epigenetic silencing of transposable elements: a trade-off between reduced transposition and deleterious effects on neighboring gene expression. Genome research 19: 1419-1428.

Janicki M, Rooke R, Yang G. 2011. Bioinformatics and genomic analysis of transposable elements in eukaryotic genomes. Chromosome research : an international journal on the molecular, supramolecular and evolutionary aspects of chromosome biology 19: 787-808.

Kalendar R, Schulman AH. 2014. Transposon-based tagging: IRAP, REMAP, and iPBS. Methods Mol Biol 1115: 233-255.

Kuang H, Padmanabhan C, Li F, *et al.* 2009. Identification of miniature inverted-repeat transposable elements (MITEs) and biogenesis of their siRNAs in the Solanaceae: new functional implications for MITEs. Genome research 19: 42-56.

Kwon SJ, Kim DH, Lim MH, *et al.* 2007. Terminal repeat retrotransposon in miniature (TRIM) as DNA markers in Brassica relatives. Molecular genetics and genomics :

MGG 278: 361-370.

Lisch D. 2009. Epigenetic regulation of transposable elements in plants. Annual review of plant biology 60: 43-66.

Liu S, Liu Y, Yang X, *et al.* 2014. The Brassica oleracea genome reveals the asymmetrical evolution of polyploid genomes. Nature communications 5: 3930.

Lu C, Chen J, Zhang Y, *et al.* 2012. Miniature inverted-repeat transposable elements (MITEs) have been accumulated through amplification bursts and play important roles in gene expression and species diversity in Oryza sativa. Molecular biology and evolution 29: 1005-1017.

Mo Y-J, Kim K-Y, Shin W-C, *et al.* 2012. Characterization of Imcrop, a Mutator-like MITE family in the rice genome. Genes & Genomics 34: 189-198.

Monden Y, Naito K, Okumoto Y, *et al.* 2009. High potential of a transposon mPing as a marker system in japonica x japonica cross in rice. DNA research : an international journal for rapid publication of reports on genes and genomes 16: 131-140.

Murukarthick J, Sampath P, Lee SC, *et al.* 2014. BrassicaTED - a public database for utilization of miniature transposable elements in Brassica species. BMC research notes 7: 379.

Naito K, Cho E, Yang G, *et al.* 2006. Dramatic amplification of a rice transposable element during recent domestication. Proceedings of the National Academy of Sciences of the .United States of America 103: 17620-17625.

Naito K, Zhang F, Tsukiyama T, *et al.* 2009. Unexpected consequences of a sudden and massive transposon amplification on rice gene expression. Nature 461: 1130-1134.

Oki N, Yano K, Okumoto Y, *et al.* 2008. A genome-wide view of miniature inverted-repeat transposable elements (MITEs) in rice, Oryza sativa ssp. japonica. Genes & genetic systems 83: 321-329.

Parisod C, Alix K, Just J, *et al.* 2010. Impact of transposable elements on the organization and function of allopolyploid genomes. New Phytologist 186: 37-45.

Paterson AH, Bowers JE, Bruggmann R, *et al.* 2009. The Sorghum bicolor genome and the diversification of grasses. Nature 457: 551-556.

Piriyapongsa J, Marino-Ramirez L, Jordan IK. 2007. Origin and evolution of human microRNAs from transposable elements. Genetics 176: 1323-1337.

Sampath P, Lee S-C, Lee J, *et al.* 2013. Characterization of a new high copy Stowaway family MITE, BRAMI-1 in Brassica genome. BMC plant biology 13: 56.

Sampath P, Murukarthick J, Izzah NK, *et al.* 2014. Genome-Wide Comparative Analysis of 20 Miniature Inverted-Repeat Transposable Element Families in *Brassica rapa* and *B. oleracea*. PloS one 9: e94499.

Schmutz J, Cannon SB, Schlueter J, *et al.* 2010. Genome sequence of the palaeopolyploid soybean. Nature 463: 178-183.

Shirasawa K, Hirakawa H, Tabata S, *et al.* 2012. Characterization of active miniature inverted-repeat transposable elements in the peanut genome. Theoretical and Applied Genetics 124: 1429-1438.

Tu Z. 2001. Eight novel families of miniature inverted repeat transposable elements in the African malaria mosquito, Anopheles gambiae. Proceedings of the National Academy of Sciences 98: 1699-1704.

van Leeuwen H, Monfort A, Zhang H-B, *et al.* 2003. Identification and characterisation of a melon genomic region containing a resistance gene cluster from a constructed BAC library. Microcolinearity between Cucumis melo and Arabidopsis thaliana. Plant molecular biology 51: 703-718.

Wang X, Wang H, Wang J, *et al.* 2011. The genome of the mesopolyploid crop species Brassica rapa. Nature genetics 43: 1035-1039.

Wei L, Xiao M, Hayward A, *et al.* 2013. Applications and challenges of next-generation sequencing in Brassica species. Planta.

Wessler SR. 2006. Transposable elements and the evolution of eukaryotic genomes. Proceedings of the National Academy of Sciences of the United States of America 103: 17600-17601.

Wicker T, Sabot F, Hua-Van A, *et al.* 2007. A unified classification system for eukaryotic transposable elements. Nature reviews Genetics 8: 973-982.

Wicker T, Taudien S, Houben A, *et al.* 2009. A whole-genome snapshot of 454 sequences exposes the composition of the barley genome and provides evidence for parallel evolution of genome size in wheat and barley. The Plant Journal 59: 712-722.

Witte C-P, Le QH, Bureau T, *et al.* 2001. Terminal-repeat retrotransposons in miniature (TRIM) are involved in restructuring plant genomes. Proceedings of the National Academy of Sciences 98: 13778-13783.

Yaakov B, Ceylan E, Domb K, *et al.* 2012. Marker utility of miniature inverted-repeat transposable elements for wheat

biodiversity and evolution. TAG Theoretical and applied genetics Theoretische und angewandte Genetik 124: 1365-1373.

Yaakov B, Kashkush K. 2012. Mobilization of Stowaway-like MITEs in newly formed allohexaploid wheat species. Plant molecular biology 80: 419-427.

Yang G. 2013. MITE Digger, an efficient and accurate algorithm for genome wide discovery of miniature inverted repeat transposable elements. BMC bioinformatics 14: 186.

Yang TJ, Kim JS, Kwon SJ, *et al.* 2006. Sequence-level analysis of the diploidization process in the triplicated FLOWERING LOCUS C region of Brassica rapa. The Plant cell 18: 1339-1347.

Zhang Q, Arbuckle J, Wessler SR. 2000. Recent, extensive, and preferential insertion of members of the miniature inverted-repeat transposable element family Heartbreaker into genic regions of maize. Proceedings of the National Academy of Sciences of the United States of America 97: 1160-1165.