

## Genomic scale sorting and characterization of MITEs in the entangled mobilome

Guojun Yang

### Abstract

Miniature inverted repeat transposable elements (MITEs) are interesting transposable elements (TEs) because of their high copy numbers and mysterious identifies. Despite their low DNA content percentage in a genome, their numerous copies can disturb genomic stability and cause important genetic variations. Historically, MITE families were often discovered individually, a practice cannot keep pace with the large scale genome sequencing. Automated genome wide discovery of MITE families and their characterization are desirable. We developed a whole pipeline from MITE discovery to detailed analyses of MITEs at genomic scales. Using the pipeline, we performed in depth analyses of the MITE families in multiple recently release crop genomes. These analyses revealed the diversity of MITE families and their evolution in the host genome. We have also predicted the transposases that may be responsible for the mobilization of some MITE families. These MITE families can potentially be used as genetic markers for the improvement of these crop species.

Genome size differences among crop plants are largely due to unequal accumulation of repetitive DNA sequences, mainly transposable elements (TEs). Over the past decade, many families of miniature inverted-repeat transposable elements (MITEs) have been identified and characterized in a variety of organisms including animals and plants. MITEs are characterized by short terminal inverted repeats (TIRs) (10-15 bp), small size (approx 100 to 500 bp), high-copy-number (approx 1000 to 15,000 per haploid genome), and a preference for insertion into 2-bp to 3-bp targets that are rich in A and T residues. In this chapter, we present a modified transposon display procedure based on the maize MITE family Heartbreaker (Hbr). This technique is similar to AFLP in which AFLP adaptors are ligated to compatible ends of digested genomic DNA. Subsets of Hbr-containing fragments are then amplified using one AFLP primer and another primer complementary to an internal sequence of the Hbr element. Like AFLP, the Hbr display method permits the simultaneous analysis of numerous DNA fragments. Given the plethora of available marker systems, the major advantage of Hbr markers, and perhaps most MITE-based markers, is a preference for insertion in or near transcriptionally active genomic regions. This feature may be especially valuable in the large genomes of agriculturally important plants like maize, wheat, and barley where gene-rich islands are thought to exist in a sea of retrotransposons. Having a class of markers that are enriched

in genic regions, coupled with the ease of isolating MITE markers, could expedite chromosome walks and map-based cloning protocols in these organisms.

Hbr markers are based on location presence/absence of the Heartbreaker family of miniature inverted repeat transposable elements (MITEs). Together with the cost-effective technique~Hbr display were developed in 2000. We chose 15 populations from the Slovenian maize germplasm bank and described ten individual samples per each population by 268 Hbr markers and 35 morphological traits (IPGRI descriptors). Samples from the same population had highly similar DNA fingerprints, while the between populations differences were very high. Therefore, only a minor part of the total genetic variance existed within populations (23.3%), and the major part among populations (76.7%). Beli zob (the only dent type population) and stajerski dvanajsterec were the most divergent populations, others were closely related. They shared the majority of bands in the way that each band was shared by different set of populations. This is suggesting the origin from the common gene pool and the high extent of migrations.

Plant molecular breeding is expected to give significant gains in cultivar development through development and utilization of suitable molecular marker systems for genetic diversity analysis, rapid DNA fingerprinting, identification of true hybrids, trait mapping and marker-assisted selection. Transposable elements (TEs) are the most abundant component in a genome and being used as genetic markers in the plant molecular breeding. Here, we review on the high copious transposable element belonging to class-II DNA TEs called "miniature inverted-repeat transposable elements" (MITEs). MITEs are ubiquitous, short and non-autonomous DNA transposable elements which have a tendency to insert into genes and genic regions have paved a way for the development of functional DNA marker systems in plant genomes. This review summarises the characteristics of MITEs, principles and methodologies for development of MITEs based DNA markers, bioinformatics tools and resources for plant MITE discovery and their utilization in crop improvement.

Hbr markers are based on location presence/absence of the Heartbreaker family of miniature inverted repeat transposable elements (MITEs). Together with the cost-effective technique~Hbr display were developed in 2000. We chose 15 populations from the Slovenian maize germplasm bank and described ten

individual samples per each population by 268 Hbr markers and 35 morphological traits (IPGRI descriptors). Samples from the same population had highly similar DNA fingerprints, while the between populations differences were very high. Therefore, only a minor part of the total genetic variance existed within populations (23.3%), and the major part among populations (76.7%).

Beli zob (the only dent type population) and stajerski dvanajsterec were the most divergent populations, others were closely related. They shared the majority of bands in the way that each band was shared by different set of populations. This is suggesting the origin from the common gene pool and the high extent of migrations.

This work is partly presented at 2nd International Conference on Big Data Analysis and Data Mining November 30-December 01, 2015 San Antonio, USA

---

Guojun Yang  
University of Toronto, Canada, E-mail: gage.yang@utoronto.ca