AMPLIFICATION AND TRANSPOSITIONAL ACTIVITY OF *mPing* IN RICE

by

EUNYOUNG CHO

(Under the Direction of Susan R. Wessler)

ABSTRACT

Transposable elements are the most abundant type of DNAs in eukaryotic genomes. The first active MITE, *mPing* has been activated by long-term tissue culture, anther culture, gamma-ray-irradiation. However, in this study, *mPing* was shown for the first time to be activated under normal conditions in the rice cultivar Gimbozu and Nipponbare. In Gimbozu landraces with high-copy number of *mPing*, *mPing* was highly activated, generating ~50 new insertions per generation. This suggests that *mPing* is still accumulating its copies in these lines. In contrast, very low level of *mPing* activity was detected in one Gimbozu and Nipponbare strain for which low copy number of *mPing* is present. Comparison of *de novo* insertions with previously isolated new Gimbozu and older Nipponbare *mPing* elements implied that the negative selection occurred in stages: rapid elimination of insertions into exons and introns, and then slow and gradual removal of insertions near genes.

INDEX WORD: Transposable elements, MITE, Transposon display, Activity, Rice, Evolution

AMPLIFICATION AND TRANSPOSITIONAL ACTIVITY OF *mPing* IN RICE

by

EUNYOUNG CHO

B.S., The Seoul National University, South Korea, 1998

M.S., The Seoul National University, South Korea, 2000

A Thesis Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements of the Degree

MASTER OF SCIENCE

ATHENS, GEORGIA

2007

AMPLIFICATION AND TRANSPOSITIONAL ACTIVITY OF *mPing* IN RICE

by

EUNYOUNG CHO

Major Professor: Susan R. Wessler

Committee: Russell Malmberg
Kelly R. Dawe

Electronic Version Approved:

Maureen Grasso
Dean of the Graduate School
The University of Georgia
August 2007

ACKNOWLEDGEMENTS

I would like to thank you my major professor, Susan Wessler, for her advice and support throughout this work. I thank my committee members, Kelly Dawe and Russell Malmberg, for their advice and interest in this work. I am very grateful to all past and present members of the Wessler lab for helpful discussions and suggestions. Thanks to the greenhouse staff and the Plant Biology Department staff. Most of all, I would like to thank my family, always supporting my husband, Soo Hyung and my angel, Darby.

TABLE OF CONTENTS

LIST OF TABLES

Page

LIST OF FIGURES

Page

CHAPTER 1

INTRODUCTION

Transposable elements (TEs) are DNA sequences capable of moving throughout the genome and often making duplicate copies of themselves during the transposition process (Craig et al. 2002; Feschotte et al. 2002a). First discovered by Barbara McClintock in maize (McClintock 1951), TEs have since been found in all eukaryotic genomes investigated as the single largest component (Feschotte et al. 2002a), accounting for ~ 50% of the human genome (Lander et al. 2001) and 50-80% of some grass species (SanMiguel and Bennetzen 1998; Vicient et al. 1999; Meyers et al. 2001).

Eukaryotic TEs are divided into two classes based on whether their transposition intermediate is RNA or DNA (Capy et al. 1998). Class 1 (RNA) elements transpose via an RNA intermediate which is reverse-transcribed into cDNA that can integrate in a new genomic locus (Capy et al. 1998). In this process, a single copy of an element can potentially generate many new insertions, often contributing to genome expansion. For example, for maize and barley, ~60% and ~70% of their genomes are composed of class 1 elements, respectively (SanMiguel and Bennetzen 1998; Vicient et al. 1999; Meyers et al. 2001). Class 1 elements include long terminal repeats (LTRs) and non-LTR retrotransposons. LTR retrotransposons have LTRs in direct orientation and are classified as either *Ty1 copia*-like or *Ty3 gypsy*-like elements based on the order of their encoded proteins (Doolittle et al. 1989; Xiong and Eickbush 1990; Kumar and Bennetzen 1999). Non-LTR retrotransposons terminate at one end with a poly(A) tract and consist of long interspersed nuclear elements (LINEs) and short interspersed nuclear elements

(SINEs)(Feschotte et al. 2002a). In contrast, class 2 (DNA) elements, which will be the focus of this study, have short terminal inverted repeats (TIRs) and transpose via a DNA intermediate using an element-encoded transposase. In the process of transposition, TIRs are recognized by the element transposase, which catalyzes the excision and insertion of the element to a new genomic locus (Craig et al. 2002). Based on class 2 elements transposition mechanism, the increase in copy number by transposition occurs either during DNA replication, or when the excision site is repaired by the host gap-repair system (Labrador and Corces 1997). Both class 1 and class 2 elements contain autonomous and nonautonomous elements. Unlike autonomous elements that encode the required enzymes for their own transposition, nonautonomous elements lack coding capacity. However, nonautonomous elements can also transpose and often attain very high copy numbers because they retain the *cis*-sequences necessary for transposition (Capy et al. 1998; Feschotte et al. 2002a).

MITEs are a group of nonautonomous class 2 elements that are predominantly associated with plant genes (Zhang et al. 2000; Feschotte et al. 2002a; Feschotte et al. 2002b). Discovered in maize, numerous MITEs families have been later identified in diverse plant and animal genomes mostly through computer-assisted database searches (Feschotte et al. 2002a). Because MITEs lack coding capacity and interfamily sequence similarity, most MITEs have been classified on the basis of their terminal inverted repeats (TIRs) and target site duplications (TSDs)(Jiang et al. 2004). Most MITEs in plants have been grouped into two superfamilies, *Tourist*-like or *Stowaway*-like MITEs (Feschotte et al. 2002a; Jiang et al. 2004). *Tourist*-like MITEs, originally isolated in maize, insert preferentially into TAA (or TTA) and the first seven bases of their TIRs are usually GGGCCTG/GGGCATC. *Stowaway*-like MITEs, first discovered in sorghum, insert preferentially into TA and the first ten bases of their TIRs are usually

CTCCCTCCGT. MITEs are nonautonomous, therefore, they require a source of transposase for mobility which is often provided by the autonomous partner. *Tourist*-like and *Stowaway*-like MITEs have been associated with the *PIF*/*Pong* and the *mariner*/Tc1 DNA transposon superfamily for their mobilization, respectively (Bureau and Wessler 1992, 1994; Feschotte et al. 2003; Zhang et al. 2004).

Despite their structural similarity, the high copy number of MITEs distinguishes them from other class 2 elements (Feschotte et al. 2002a). Most nonautonomous class 2 element families are < 1 kb in length and can amplify to not more than 50 copies (Feschotte et al. 2002a; Naito et al. 2006). In contrast, MITEs are 100 - 500 bp long and MITE families can amplify to very high copy numbers (up to several thousand) in the genome, while each MITE family shows intrafamily structural homogeneity (Wessler et al. 1995; Jiang et al. 2004). Phylogenetic analysis of MITE families implies that they had rapidly and extensively amplified from a small number of copies and then experienced inactivity and drift after that (Feschotte et al. 2002a; Santiago et al. 2002). Because of their preferential transposition into single-copy regions of the plant genome, MITEs could play an important role in genome evolution (Feschotte et al. 2002a). However, the mechanism underlying the ability of MITEs to attain very high copy number is still poorly understood.

Several decades ago, it was a common idea that TEs were selfish and junk DNA. However, it has been revealed that the role of TEs in the genome evolution is as a main source of new genetic variation (Kidwell and Lisch 2001; Kidwell and Lisch 2002). Specifically, TEs have been shown to be directly or indirectly involved in genome restructuring in many different ways such as ectopic recombination, chromosome translocations and inversions, and epigenetic gene regulation (Kidwell and Lisch 2001; Wessler 2006). However, even with this potential

beneficial role on genome evolution, a high level of TE activity can be deleterious to individuals or populations due to mutations they cause in the genome (Kidwell et al. 1977; Bennetzen 2000). Thus, it is not surprising that, although TEs comprise a significant portion of some eukaryotic genomes, most TEs are usually inactive. Part of the reason for this inactivity is because most TEs have become defective with mutations, internal deletions, rearrangements, and/or replacements (Jin and Bennetzen 1994; Bennetzen 2000). However, even structurally complete TEs lack activity mainly because their activity was repressed by host defense system (Feschotte et al. 2002a). Host genomes have successfully regulated the activity of TEs, so that only a few TEs are transpositionally active. For example, only 30-60 *L1* elements of 500,000 LINEs are considered to be currently active in human (Sassaman et al. 1997; Lander et al. 2001). In plants, active retrotransposons have been very rarely identified (Kalendar et al. 2000; Feschotte et al. 2002a) and only a few active DNA elements have been isolated in plants largely by mutant analysis, such as *Ac*/*Ds*, *Spm*/*dspm* (*En*/*I*) and *Mutator* system of maize and the *Tam* elements of snapdragon (Flavell et al. 1994; Kunze et al. 1997).

However, silenced TEs have been activated by various biotic and abiotic stresses (Wessler 1996; Capy et al. 2000; Madlung and Comai 2004). For example, tissue culture has been shown to induce mobilization of the tobacco retrotransposons *Tto1*, *Tto2*, and *Tnt1* (Hirochika 1993), and the rice retrotransposon *Tos17* (Hirochika et al. 1996) and *Karma* (Komatsu et al. 2003). *Tnt1* and *Tto1* have been also reported to be activated by wounding, oxidative stress, pathogen infection and microbial elicitors (Wessler 1996; Madlung and Comai 2004). The *Tam* elements of snapdragon are temperature-sensitive DNA elements for which activity was significantly dependent on lower than normal temperatures (Coen et al. 1986).

As the world's largest crop, providing more than one fifth of the calories consumed by human worldwide (Smith 1998), rice (the genus *Oryza*), has two domesticated species, Asian rice (*O. sativa*) and African rice (*O. glaberrima*), native plants to the tropical and subtropical climates of southern and southeastern Asia and Africa, respectively. The species *O. sativa* was domesticated from wild rice, *O. rufipogon* near the Himalayas, with the subspecies *indica* being cultivated by the Indians and the subspecies *japonica* being cultivated by Chinese and Japanese (Londo et al. 2006). The importance of rice as a main source of human nutrition has made it a major focus of researchers. For these reasons, a significant amount of genetic and genomic resources are available including the complete genomic sequences of *indica* and *japonica*, and thousands of diverse cultivars. With these resources, rice has been validated as a model grass for the study of DNA transposons because its genome contains virtually all the major element families found in other grasses (Jiang et al. 2004). Rice has the smallest genome (430 Mb) among the agriculturally important cereal crops (e.g, 3,000 Mb, maize and 16,000 Mb, wheat)(Burr 2002; Jiang et al. 2004). Approximately 26% of the rice genome sequence is derived from TEs; of this, 15% and 11% are derived from class 1 and class 2 elements, respectively. In terms of element copy number, MITEs are the largest group (with around 90,000 elements divided into hundreds of families), comprising about 6% of the genome (Jiang et al. 2004).

CHAPTER 2

AMPLIFICATION AND TRANSPOSITIONAL ACTIVITY OF *mPing* IN RICE

BACKGROUND AND RATIONALE

mPing is a typical *Tourist*-like MITE. Structurally, *mPing* is small (430 bp), with terminal inverted repeats (GGCCAGTCACAATGG) and target site duplications of TAA (or its complementary sequence, TTA)(Jiang et al. 2003). Discovered by database analyses of rice genomic sequences (Jiang et al. 2003; Kikuchi et al. 2003a) and mutant analyses (Nakazaki et al. 2003). Transpositional activity of *mPing* has been demonstrated in long-term tissue culture derived from the rice subspecies *indica* cultivar C5924 (Jiang et al. 2003), in newly derived anther culture from the rice subspecies *japonica* cultivar Nipponbare (Kikuchi et al. 2003a), and in a gamma-ray-irradiated mutant line from the *japonica* cultivar Gimbozu *in planta* (Nakazaki et al. 2003). Despite its activity, *mPing* was initially found to be present at relatively low copies (for a MITE), with < 10 copies in the subspecies *indica* and < 50 copies in the subspecies *japonica* (Jiang et al. 2003). However, subsequent study of the cultivar Gimbozu mutant line (IM294), where *mPing* was active *in planta*, and its progenitor (EG4) showed that *mPing* was present at very high copy number in these lines (Naito et al. 2006). The cultivar Gimbozu is a Japanese rice popularly cultivated in the first half of the 20th century with many distinct but closely related strains and landraces (Naito et al. 2006). Additional analysis of related Gimbozu landraces revealed that *mPing* had amplified to over thousand copies in a few Gimbozu

landraces, while its copy number remained below fifty in the rest of the landraces (Fig. 1)(Naito et al. 2006).

The cultivar Gimbozu and, the cultivar Nipponbare, whose genome sequence is available, are very closely related (Naito et al. 2006). The conserved insertion sites of *mPing* between Nipponbare and Gimbozu low-copy-number *mPing* (hereafter, low-copy Gimbozu) landraces indicated that *mPing* has not been activated in these lines for about a century. In contrast, *mPing* distribution among the Gimbozu high-copy-number *mPing* (hereafter, high-copy Gimbozu) landraces revealed the independent amplification of over 1,000 copies in each of these landraces, suggesting that a burst of *mPing* was a recent event (Naito et al. 2006). Most *mPing* insertions in high-copy Gimbozu landraces were assumed to be relatively newer than insertions in most low-copy plants, such as Nipponbare (Naito et al. 2006). Previously, extensive analysis of the *mPing* insertion sites was performed for high-copy Gimbozu landraces and Nipponbare. In this study, it was discovered that the vast majority (about 90%) of the analyzed insertions were into single copy regions and near transcribed genes. However, there was a significant difference in the genic proximity of the insertions between high-copy Gimbozu and Nipponbare: about 70% of the newer *mPing* insertions of high-copy Gimbozu were located within 5 kb from a cDNA coding sequence but only about 45% of the older insertions of Nipponbare were like that. Moreover, a lower-than-expected number of *mPing* insertions were found in exons and introns of genes in the high-copy Gimbozu, suggesting that negative selection had already acted to remove most insertions in rice genes.

mPing is a nonautonomous element, which requires transposases provided by an autonomous partner for its mobility. Two putative *trans*-acting autonomous elements, *Ping* and *Pong*, have been identified (Jiang et al. 2003). *Ping*, discovered by sequence similarity search to

mPing, is a 5341 bp element and its terminal sequences of 252 bp and 178 bp are identical to those of *mPing* except for a single base-pair mismatch, clearly indicating that *mPing* is a recent deletion derivative of *Ping* (Jiang et al. 2003). However, *mPing* was shown to be actively transposing in rice *indica* cell culture that lacks the *Ping* element. This revealed that an additional transposase source necessary for *mPing* mobilization in C5924 cell culture. Blast searches using *Ping* as a query led to the discovery of the *Pong* element. *Pong* is 5166 bp in length and its two ORFs share >80% identity with *Ping* ORFs (Jiang et al. 2003). In spite of their clear-cut sequence similarity to *mPing*, there was no experimental evidence directly connecting *Ping* and/or *Pong* to *mPing* mobility until the recent data that successfully demonstrated that rice endogenous *Ping* and *Pong* both can transpose *mPing* in Arabidopsis (Yang et al. in press). This data indicates that *Ping* and *Pong* both encode fully functional transposases necessary for the *mPing* mobilization.

As mentioned earlier, most TEs are tightly regulated in the genome by host defense mechanism including epigenetic silencing, but silenced TEs have been shown to regain their activity under various stressful conditions, which implies genome-restructuring role of TEs against genome shock (McClintock 1984; Wessler 1996; Madlung and Comai 2004). So far, *mPing* has been activated by long-term tissue culture (Jiang et al. 2003), anther culture (Kikuchi et al. 2003a), hydrostatic pressurization (Lin et al. 2006), and hybridization with wild rice (*Zizania latifolia*)(Shan et al. 2005). In the cultivar Gimbozu, *mPing* insertions caused mutations in the gamma-ray-irradiated lines (Nakazaki et al. 2003). However, rapid and independent accumulation of *mPing* from about 50 to over 1,000 observed in closely related Gimbozu landraces suggests that the burst of *mPing* was a very recent event. Moreover, the fact that some

of *mPing* insertions in high-copy Gimbozu landraces were heterozygous implied that *mPing* may be still transposing in these strains under normal growth condition.

In this study, transpositional activity of *mPing* was tested in the non-irradiated Gimbozu landraces. A different level of activity was detected in high and low-copy strains, providing an important clue in understanding the amplification process of *mPing* during the domestication of rice. Furthermore, analysis of newly transposed *mPing* insertion sites, combined with the previously characterized *mPing* insertions, implies that negative selection acted upon the *mPing* elements that were inserted into or near rice genes.

MATERIALS AND METHODS

Plant Materials and genomic DNA extraction

Seeds of the rice (*Oryza sativa*) subspecies *japonica* cultivar Nipponbare, and the cultivar Aikoku and Gimbozu landraces (A101, A119, A127, A157, G190 and EG4) were obtained from the GenBank project of the National Institute of Agrobiological Science, Ibaraki, Japan. Seeds were dehusked, cleaned with 95% EtOH and 40% bleach, and germinated at 27°C in petri dishes with sterilized ddH_2O. Three-week old seedlings were transferred into soil. To generate single-seed derived materials, plants were self-pollinated and seeds harvested from one inflorescence were planted and used as progeny. To generate hybrid plants between different cultivars or landraces, female parental lines were emasculated manually by removing anthers from flowers in the evening, and were pollinated by dehiscent male parental plants in the next morning. Pollinated flowers were covered by bags until seeds were harvested. Genomic DNA was

extracted from leaf tissues of single plants using DNeasy Plant Mini Kit (Qiagen Inc., Valencia, CA) following the manufacturer's instructions.

Transposon display

Transposon display was performed as described with modifications (Casa et al. 2000). *mPing* specific primers for PCR reactions were designed based on the subterminal sequences of *mPing* (Jiang et al. 2003). Total genomic DNA (50-500 ng) was digested with *Bfa*I or *Mse*I and was ligated to an *Bfa*I-adapter in a total volume of 50 μl containing 1 × NEBuffer 4 (20 mM Tris-acetate, 50 mM potassium acetate, 10 mM Magnesium acetate, 1mM Dithiothreitol at pH 7.9), 1 × T4 DNA ligase reaction buffer (50 mM Tris-HCl, 10 mM $MgCl_2$, 1 mM ATP, 10 mM Dithiothreitol, 25 μg/ml BSA at pH 7.5), 1 × BSA, 50 pmol of *Bfa*I-adapter (5'-GACGATGAGTCCTGAG-3' and 5'-TACTCAGGACTCAT-3'), 5 units of *Bfa*I or *Mse*I (New England Biolabs Inc., Beverly, MA), and 1 unit of T4 DNA ligase (New England Biolabs., Beverly, MA) overnight at 37°C. Aliquots of the restriction/ligation reactions were visualized on 1% agarose gels to check the quality of DNA restriction and were diluted 4-fold with ddH_2O.

Preselective amplification was done by using a primer complementary to the adapter (*Bfa*I + Ø, 5'-GACGATGAGTCCTGAGTAG-3' or *Mse*I + Ø, 5'-GACGATGAGTCCTGAGTA-3') and another primer specific to *mPing* (*mPing* P1, 5'-TGTGCATGACACACCAGTG-3'). Reactions were performed in a total volume of 20 μl containing 1.5 μl of the diluted restriction/ligation reactions, 10 pmol of each primer, 1 × GeneAmp PCR buffer II (Perkin-Elmer/ABI, Foster City, CA), 2 mM dNTPs, 2.5 mM $MgCl_2$, and 1 unit of Ampli*Taq* DNA polymerase (Perkin-Elmer/ABI, Foster City, CA). PCR conditions were as follows: 72°C/2 min; 94°C/3 min; 30 cycles of 94°C/45 sec, 58°C/45 sec, and 72°C/45 sec, and a final cycle of 72°C/5

min. After visualizing aliquots of each PCR on 1% agarose gels stained with ethidium bromide, the remaining volumes were diluted 10 to 100-fold with ddH_2O.

Selective amplification for radioactive detection was performed in 10 μl containing 1 μl of the diluted preselective amplification products, 10 pmol of *Bfa*I or *Mse*I adapter primer with zero to three selective bases, 10 pmol of ^{33}P-labeled element specific primer (*mPing* P2: 5'-CAGTGAAACCCCCATTGTGAC-3'), 1 × GeneAmp PCR buffer II (Perkin-Elmer/ABI, Foster City, CA), 2 mM dNTPs, 2.5 mM $MgCl_2$, and 1 unit of Ampli*Taq* DNA polymerase (Perkin-Elmer/ABI, Foster City, CA). A touchdown protocol was used where the initial annealing temperature was 66°C and decreased by 1°C with each cycle to 58°C (Vos et al. 1995). PCR conditions were as follows: 94 °C/3 min; 9 cycles of 94°C/45 sec, 66-58°C/45 sec, and 72°C/45 sec; 30 cycles of 94°C/45 sec, 58°C/45 sec, and 72°C/45 sec, and a final cycle of 72°C/5 min.

For radioactive detection, 10 μl of loading-denaturing buffer (98% deionized formamide, 0.025% bromophenol blue, and 10 mM EDTA) was added to the PCR products. Samples were denatured at 95°C for 5 min and were placed on ice immediately, and 5 μl of mixture was loaded on the 6% denaturing (7.5 M urea) acrylamide-bisacrylamide (19:1) gel in 1 × TBE buffer (890 mM Tris, 890 mM boric acid, 20 mM EDTA at pH 8.3). After electrophoresis for 3 hours at 35 mA, the gel was transferred to filter paper, dried, and exposed to an X-ray film for 36 hours before development.

Recovery of gel bands

To recover individual bands in transposon display (TD) gels, a needle-scratching technique was used as described (Stumm et al. 1997). DNA fragments were excised from radioactive gels by scratching the dried gel with tips, placing the tip in 50 μl PCR reaction mix

with a *Bfa*I + Ø primer and an element specific *mPing* P2 primer. The PCR parameters were the same as in the preselective amplification except starting with incubation at 94°C for 5 min instead of 72°C for 2 min. PCR products were resolved on 1% agarose, excised, purified (QIAquick Gel Extraction Kit, Qiagen Inc., Valencia, CA), and cloned (TOPO TA Cloning Kit, Invitrogen, Carlsbad, CA). Cloned DNA templates were sequenced by the Molecular Genetics Instrumentation Facility (University of Georgia). Retrieved sequences from TD gels contained the one end of element sequences and its flanking sequences. To localize the element insertions in the rice genome, element sequences were trimmed out from the recovered sequences and only flanking sequences were used as a query sequence for similarity search by BLAST program against the Nipponbare genomic DNA database. Database hits that share >99% similarity were considered matches.

Annotation of *mPing* flanking sequences

To identify genes near *mPing* insertion sites, 10 kb of flanking genomic sequence (5 kb on each side of an *mPing* insertion site for each of the genomic hits) was selected from the Nipponbare genomic DNA database and used as query for a BLAST search against the full length cDNA database (http://cdna01.dna.affrc.go.jp/cDNA)(Kikuchi et al. 2003b). High-scoring pairs (HSPs) with >50 bp and >98% identity were considered matches. Position (upstream, downstream, or inside) and distances of cDNA sequences relative to *mPing* insertion sites were determined based on position information of the hits in HSPs. If a cDNA sequence hit crosses an *mPing* site (either contiguous or with a gap), the *mPing* insertion was considered to be inside of the cDNA sequence. If the start of a cDNA is downstream of an *mPing* insertion site, the insertion was considered to be upstream of the cDNA, and the distance between the insertion

site and the cDNA start was calculated. The distance of a downstream insertion from a cDNA was calculated from the 3' end of the cDNA sequence.

RESULTS

Transposition of *mPing* in high-copy Gimbozu landraces heritability

To determine whether *mPing* is transposing in non-irradiated high-copy Gimbozu landraces, *mPing* insertions among the selfed plants were compared using the transposon display (TD) technique. TD is a modified amplified fragment length polymorphism (AFLP) technique that allows the visualization of multiple copies of a transposon by amplifying their flanking sequences. More important, TD aids in the detection of transpositional activity of an element as well as providing an estimation of element copy number in the genome (Vos et al. 1995; Van den Broeck et al. 1998; Casa et al. 2000; Takagi et al. 2007).

All the Gimbozu plants tested in this study have been propagated by self-fertilization, and therefore the *mPing* loci should be identical in the progeny derived from a single plant. In addition, the TD banding pattern of *mPing* should be identical among the progeny unless *mPing* was transposing. If a band was present in a single progeny but absent in the parent or siblings, then this band was considered to be a *de novo mPing* insertion, suggesting transpositional activity of *mPing*. Subsets of putative *de novo* bands were further verified as real *mPing* insertions and not artifacts caused by technical error. For this analysis, randomly selected bands were cut from gels, reamplified, and sequenced. Then, primers were designed based on retrieved flanking sequences to amplify the insertion site in the parent and all the progeny. If a locus has an *mPing* insertion, PCR amplicons should be 430 bp (the length of *mPing*) greater than

expected size. Using this strategy, bands were counted as real *de novo mPing* insertions, if only one progeny, where the band was recovered, has longer PCR product, and the rest of progeny and parent have smaller ones (Fig. 2).

For each of three high-copy Gimbozu landraces (EG4, A119, and A157), the parent and its ~10 self-fertilized progeny were tested. *de novo* bands were detected from all four landraces (Fig. 3), which indicates that *mPing* was activated in these plants prior to radiation events.

The frequency and the inheritability of *de novo mPing* insertion in high-copy Gimbozu landraces

To estimate the total number of *de novo mPing* insertions produced per generation per plant, and to determine whether these *de novo* insertions are heritable, one high-copy Gimbozu landrace, EG4, was selected and propagated for three consecutive generations. Specifically, the eight progeny (S_0) were generated from a single selfed parent and then, the 10 progeny (S_1) were from one selfed S_0 plant (PS1), and finally, the 10 progeny (S_2) were generated from one selfed S_1 plant (PS2). Because EG4 plants have over 1,000 copies of *mPing* in its genome (Naito et al. 2006), it was impractical to visualize all copies on a single gel. Therefore, to adjust for the resolution of a gel, selective adapter primers with additional three bases were used for *mPing* TD reactions. *mPing* TDs were performed for all three generations and 16 of 64 possible primers were tested (Fig. 4). To estimate the total number of *de novo* insertions in generations S_1 and S_2, the average of *de novo* bands (per primer) was estimated and then multiplied by 64. Using this method, ~ 60 and ~ 50 *de novo* insertions per plant per generations in the S_1 and S_2 generations were determined, respectively (Table 1 and 2). The number of *de novo mPing* insertions that were inherited from the S_1 generation to the S_2 generation was estimated by counting the *de novo*

bands appearing in a parent (Fig. 4, PS2) that segregated in the S_2 progeny (Fig. 4, asterisk). Based on this analysis, >80% (35 of 43) of the *de novo* bands in parent PS2 were inherited in the S_2 progeny (Tables 1 and 2).

de novo mPing insertion sites in rice

To characterize the insertion sites of newly transposed *mPing* elements, 23 randomly selected *de novo* TD bands from the 10 progeny derived from a single EG4 plant were sequenced. The chromosomal location of each insertion was determined by BLAST searches using the flanking sequences against the rice genomic sequences from Genbank. The proximity of the insertions to rice genes was determined by the BLAST search of the extended flanking sequences (total 10 kb sequences from both sides of the insertion) against the rice full length cDNA database. All 23 insertions were located in single copy regions of the rice genome. 14 of the 23 were located within 5 kb of a rice gene and 3 of the 23 were inserted into exons (Table 3 and 4).

Transpositional activity of *mPing* in low-copy Gimbozu landraces

According to the data from the Naito et al (2006), where the copy number of *mPing* in different Gimbozu landraces was determined (Naito et al. 2006), Gimbozu landraces can be separated into two groups based on the copy number of *mPing*: high-copy Gimbozu with over one thousand copies or low-copy Gimbozu with less than fifty copies. Although two major groups exist, there is variation among the low-copy Gimbozu landraces in the number of insertions that are present (Fig. 1). In order to examine whether this copy number variation reflects the different level of *mPing* activity in these strains, transpositional activity of *mPing* in

low-copy Gimbozu landraces (A101, A127, and G190) and other *japonica* cultivars (Nipponbare and TN67) was tested in single-seed descent plant materials (a single parent and 10 selfed progeny for each landrace and cultivar) using TD technique. Based on this analysis, *de novo* insertions were detected for G190 and Nipponbare. For both G190 and Nipponbare (Fig. 5), the frequency of transposition was less than 0.5 *de novo* insertion per plant per generation. To determine whether these *de novo* insertions are also germinal, one (called, locus 4097) of *de novo* insertions in Nipponbare was selected and tested (Fig. 5, asterisk). The plant (P) with this *de novo* insertion was selfed to generate progeny (S_2) and then ten S_2 plants were tested for the presence of *mPing* at this locus by TD. PCR analysis using the primers designed for the sequences flanking this insertion was performed for both the parent and progeny to determine whether locus 4097 is homozygous or heterozygous for the *mPing* insertion. The progeny were segregated for locus 4097 (Fig. 5). One of the progeny containing a homozygous *mPing* insertion in locus 4097 was selfed to the next generation (S_3). It was found that all the tested S_2 plants have *mPing* in locus 4097, indicating that *de novo* insertion in the locus 4097 was fixed in the population.

mPing **activity in hybrid plants between high-copy and low-copy plants**

In the previous section, it was shown that high copy and low copy strains have different level of *mPing* activity. If this difference was caused by presence of any negative regulator in low-copy strains by which *mPing* or its autonomous elements are repressed, then the introduction of this factor into high-copy Gimbozu strains by crossing with low-copy strains should affect *mPing* activity presumably. To test this hypothesis, several high and low-copy strains were cross-fertilized. Because rice is generally a selfer, it was necessary to verify that the F_1 plants

were real hybrids. To this end, *mPing* itself was used as a marker for hybridization. Because each rice strain has a specific distribution of *mPing* insertions in its genome, they can be easily used to test hybridization by visualizing them on a TD gel. If a derived F_1 plant is a real hybrid, it should contain all heterozygous *mPing* insertions that were inherited from both parents. In this way, two hybrid plants were generated from A157 X G190 and EG4 X Nipponbare. The F_2 generations were produced by self-fertilization of each F_1 plant, and *mPing* TD was performed to detect any transpositional activity of *mPing* in the F_2 generations. TD bands that appear only in one of the F_2 plants but not in the parents or the F_1 were considered as *de novo* insertions. *de novo* bands were detected in the F_2 plants of both crosses, suggesting *mPing* is still transposing in these F2 plants (Fig. 6). However, the level of activity was detected at much lower than frequency in high-copy parents.

DISCUSSION

Amplification of *mPing* during recent rice domestication

mPing was discovered by searching the rice genome database for repeat families with structural features of MITEs and with very low intrafamily sequence divergence (Jiang et al. 2003). However, in this initial survey, the copy number of *mPing* was found to be low (less than 50) in most of the rice cultivars investigated; making the available rice strains unsuitable for understanding how MITEs attain very high copy numbers (Naito et al. 2006). Recently, with the availability of closely related Gimbozu landraces where both high (over thousand) and low copy numbers of *mPing* are present, it is now possible to investigate the earliest events in the amplification of *mPing* in rice. Based on the analysis of the copy number variation among

different rice cultivars, two steps of *mPing* amplification are suggested: first, from 1 to 50 copies in most *japonica* strains; then, from 50 to more than 1,000 copies in a few Gimbozu strains (Naito et al. 2006).

The first stage of *mPing* amplification can be applied to most low-copy strains including Nipponbare and the majority of Gimbozu strains. The copy number variation observed in low-copy strains can be partly explained by genomic recombination between different rice strains as a result of breeding and hybridization during the history of rice cultivation. However, it could also be caused by more recent *mPing* activity in particular rice strains. Initially, it was thought that *mPing* was inactive in most rice cultivars because their *mPing* copy numbers seemed to have remained low over a long period of time. In this study, transpositional activity of *mPing* was detected in Nipponbare and one Gimbozu low-copy strains (G190). Although the frequency of *de novo* insertions was relatively low in these strains, these *de novo* insertions could be fixed in the progeny after two generations. This suggests that a gradual increase in *mPing* copy number has probably contributed to the copy number differences among the low-copy rice strains. It is notable that Nipponbare and G190 are the rice strains where *mPing* copy number reached the first plateau at about fifty. Other strains with no *mPing* activity have less than 25 copies of *mPing* in their genome.

Since no intermediate levels of amplification have been found in other rice strains, it can be implied that *mPing* copy number has increased explosively from ~ 50 to 1,000 in a few high-copy Gimbozu landraces within a short period of time. These amplification events also occurred independently in each of the high-copy strains, which are cultivated in different geographically isolated regions of Japan (Naito et al. 2006). Although it was clearly shown in this study that some low-copy strains have active *mPing* elements, this low activity seems not enough to drive

the rapid and explosive amplification of *mPing* that has occurred in high copy strains. Here, results from this study suggest that the burst of *mPing* may be an on-going process in high-copy strains. It is because *mPing* is currently active status in these strains and transposition events occur at a much higher frequency in the high-copy strains than in the low copy strains trains. Moreover, the majority of the *de novo* insertions are inherited into the progeny, implying that *mPing* is still increasing in copy number in these strains.

Two questions are what initially triggered a burst of amplification in some Gimbozu landraces, and what causes the differences in *mPing* activity among the rice plants investigated in this study. Different *mPing* activity between low and high-copy strains could be explained by the absence of a necessary component such as an autonomous element in these strains or any negative factors present in the genome. *mPing* has two autonomous element, *Ping* and *Pong*. While multiple copies of *Pong* are widely distributed in all the rice cultivars tested, only a subset of cultivars contain the *Ping* element (Jiang et al. 2003). Furthermore, all the strains with *mPing* activity have the *Ping* element; low-copy strains (Nipponbare and G190) have only one *Ping*, while high-copy strains have more than five copies of *Ping* (unpublished data).

A cross between high and low-copy strains produced hybrid plants with low *mPing* activity. This lower than expected *mPing* activity in the F_2 plants can be interpreted in several different ways. First, there may be negative regulators in low-copy strains repressing the function of the autonomous elements. This scenario is less likely because low-copy parental strains used in the crosses were Nipponbare and G190, which contain an active *mPing*. Initially, Nipponbare and G190 were selected for low-copy parental lines because it was assumed that they do not have active *mPing* elements. However, low *mPing* activity was detected in both strains, thus excluding the possibility of presence of repressors in their genomes (Fig. 5).

Second, it is possible that a decreased number of autonomous elements could affect the frequency of transposition in hybrid plants. Based on the observation that all high-copy Gimbozu strains have more than five *Ping* copies (unpublished data), it can be hypothesized that there is a threshold in the number of *Ping* copy numbers required to accelerate *mPing* transposition. However, most hybrid plants have a lower number of *Ping* elements that is probably not enough to produce high *mPing* transpositional activity. Finally, the number of *mPing* elements present within the genome could also affect *mPing* activity because extant *mPing* elements are used as a target of their transposase. Hence, if more *mPing* elements are available for the transposase, more *de novo* insertions can be generated in the genome. Furthermore, hybrid plants that have fewer *mPing* elements than their high-copy parents, due to the reduction in *mPing* copy number, could affect the frequency of *mPing* transposition.

Although the transpositional frequency of *mPing* was estimated in both high and low-copy strains in this study, it is insufficient to estimate the rate at which the new *mPing* insertions have accumulated in the earlier stage of their amplification, because data from this study only reflects the current levels of transposition. With the observation from different landraces, it can be inferred that the activity of *mPing* must have been very low in the earliest stage and then somehow accelerated in some Gimbozu landraces, causing the burst of *mPing* amplification. However, future studies should be pursued to investigate alternate methods for the increase in transpositional activity of *mPing* in the low-copy strains.

Negative selection of *mPing* insertions working in stages

Due to their preferential insertion into genic region, it is fundamental to understand how the host can tolerate thousands of elements inserted into or near genes, before we can decipher

how MITEs are able to amplify to very high copy numbers. The analysis of *de novo mPing* insertions in EG4, combined with previously characterized newer insertions of EG4 and older insertions of Nipponbare, suggests that two stages of selection have occurred. The first stage is the rapid removal of insertions from exons and introns of the rice genes. As shown in table 4, 3 of 23 *de novo* insertions were inserted into exons, which is much higher than the ratio of new insertions of EG4 (1 of 201 analyzed insertions). This indicates that *mPing* could preferentially target rice genes, but potentially deleterious insertions are quickly eliminated from the population by negative selection. The second stage of selection occurs at a much slower rate, eliminating *mPing* insertions close to rice genes (within 5 kb) as seen by comparing EG4 and Nipponbare insertion profiles: 70 % of new insertions in EG4 were inserted within 5 kb of a gene, while about 45% of presumably ancient insertions in Nipponbare were located near genes (Naito et al. 2006).

mPing **as a tagging tool**

In this study, I have shown that *mPing* is transpositionally active in a few rice strains under natural growth conditions. Since *mPing* was discovered several years ago as the first active DNA transposon in rice (and the first active MITE in any organism), it has been shown to be activated by various conditions such as tissue culture (Jiang et al. 2003), anther culture (Kikuchi et al. 2003a), hydrostatic pressurization (Lin et al. 2006), and hybridization with wild rice (Shan et al. 2005). In the cultivar Gimbozu, *mPing* was shown to be active *in planta* in irradiated strains, causing the mutations in the *Rurm1* gene and the *Hd1* gene (Yano et al. 2000; Nakazaki et al. 2003). However, in this study, it was found that *mPing* was transposing in non-

irradiated Gimbozu lines indicating that *mPing* was already activated in these lines prior to radiation events. This suggests that *mPing* could be developed as a gene tagging tool.

Transposon tagging is a technique for gene discovery and functional studies in both plants and animals (Martienssen 1998; Klinakis et al. 2000; Springer 2000). Since the recent availability of the rice genome sequences, substantial attention has recently been paid to developing a tagging system in rice (Hirochika et al. 2004; Leung and An 2004; Takagi et al. 2007). The endogenous retrotransposon *Tos17* as well as foreign elements, such as T-DNA or maize DNA transposons, have been employed to generate insertional mutations on a large scale in rice, facilitating the isolation of several important genes (Miyao et al. 2003; Hirochika et al. 2004; Kurusu et al. 2004; Leung and An 2004). However, these tagging systems require tissue culture processes for either the introduction of the foreign elements into the rice genome (An et al. 2005) or the activation of the silenced *Tos17* element in the genome (Hirochika et al. 1996), in which somaclonal variation associated with the tissue culture is inevitable (Larkin and Scowcroft 1981; Kaeppler et al. 2000). For this reason, an *in planta* transposon-tagging method that does not require the tissue culture process is highly desirable. Successful endogenous DNA transposons have extensively been utilized for gene tagging in maize, snapdragon, petunia, and morning glory (May and Martienssen 2003; Takagi et al. 2007). Although it remains a mystery how *mPing* was initially activated in these lines, the continued activity of *mPing* in specific rice lines, can provide the first TE based rice gene tagging system that does not require tissue culture. Also, the availability of high-copy and low-copy landraces in which *mPing* is active and the continued activity of *mPing* in hybrids between high and low-copy strains will allow us to further test various landraces and hybrids, and to eventually generate the optimal tagging population for the development of a powerful gene-tagging system in rice.

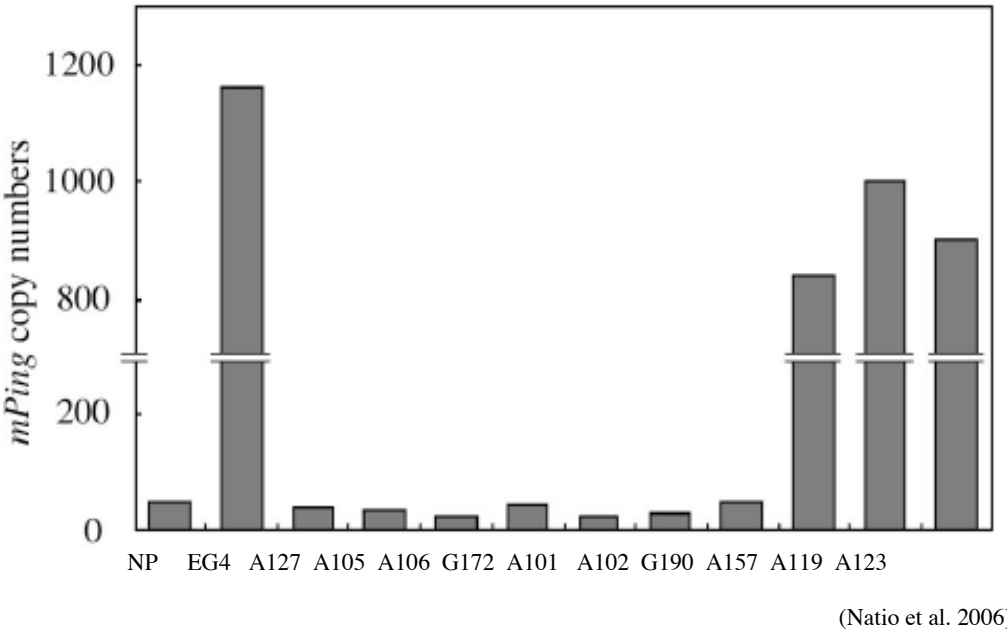


Figure 1. Copy number of *mPing* elements in Nipponbare (NP) and Gimbozu strains.

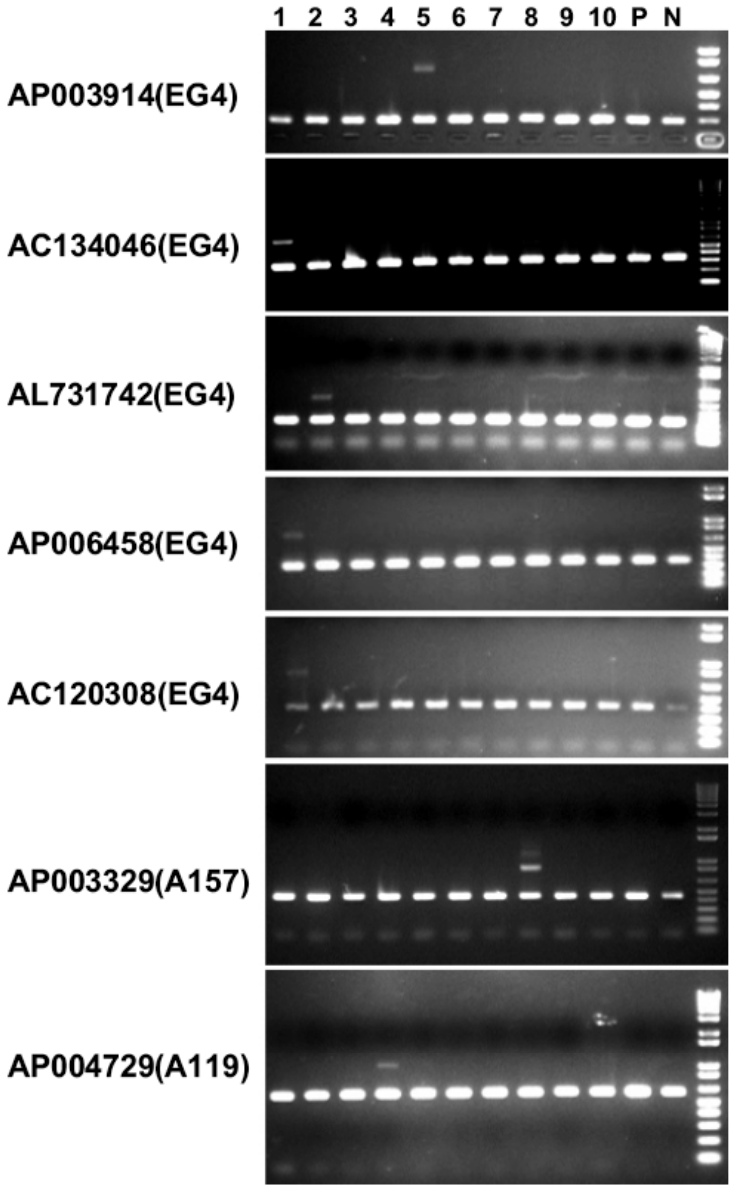


Figure 2. Verification of *de novo* insertions by locus-specific PCRs using genomic DNAs from a parent (P) and 10 selfed progeny of EG4, A119, or A157. Nipponbare (N) was used as a negative control for the absence of *mPing* in the locus. The accession numbers of BAC clones containing the *de novo* insertions are next to the strain names. The upper bands are 430 bp longer than the lower bands due to the presence of *mPing*.

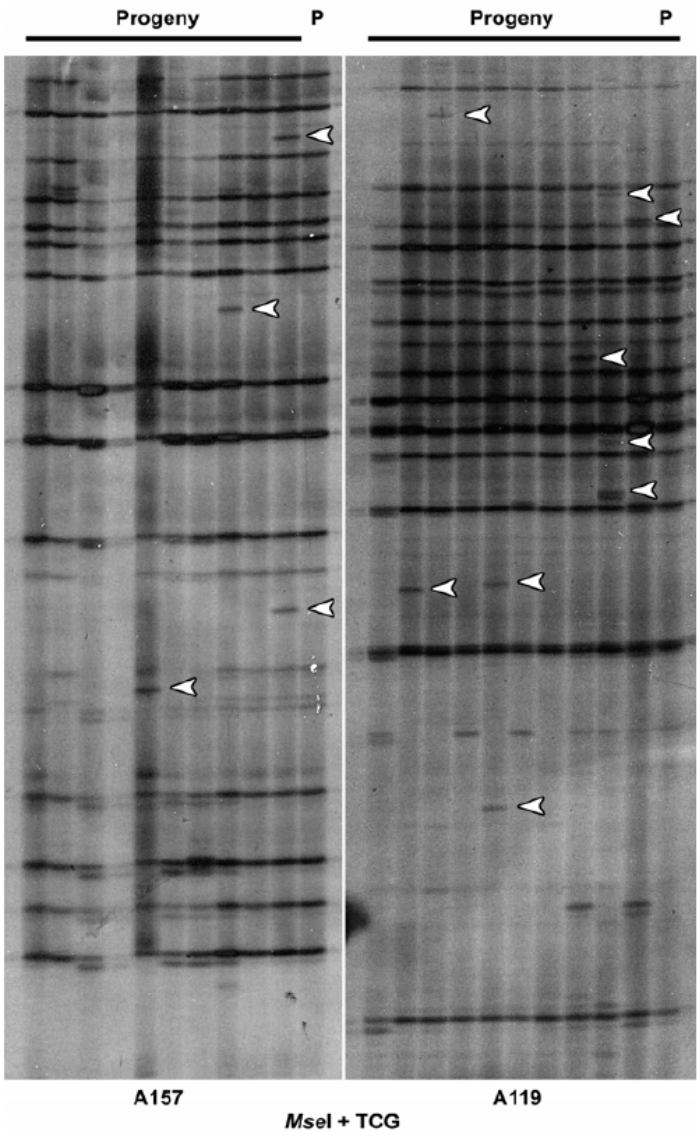


Figure 3. Transposon display of *mPing* in high-copy strains (A119 and A157) using genomic DNAs from parent (P) and 10 selfed progeny. *de novo* bands detected in progeny are noted by arrowheads. See Figure 4 for the transposon display gel of EG4.

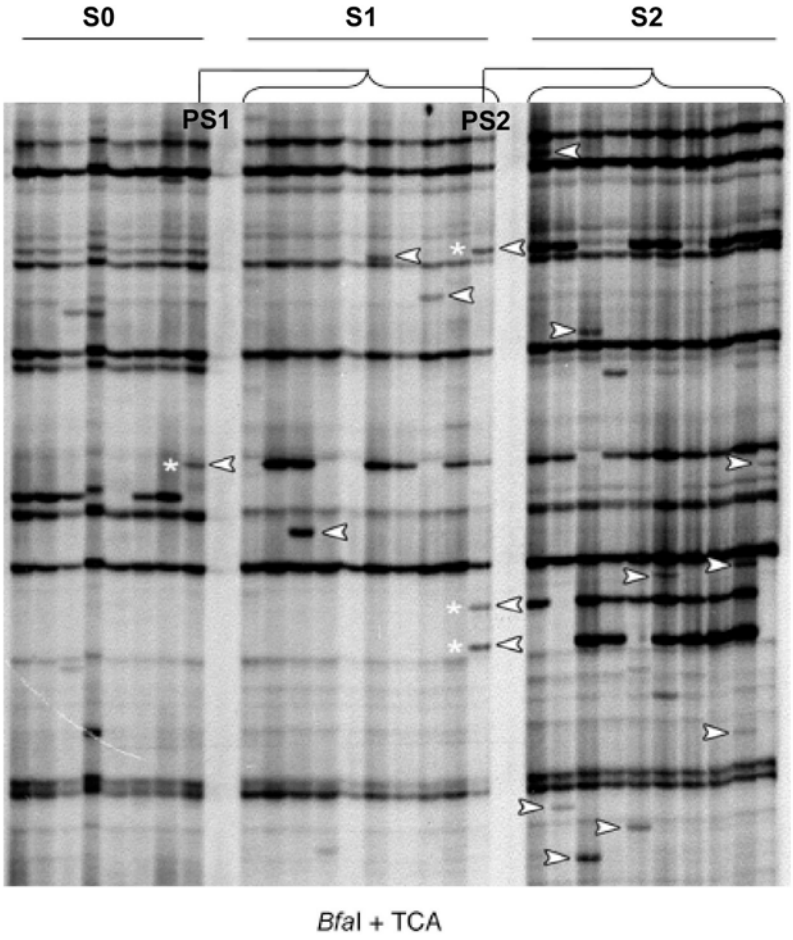


Figure 4. Transposon display analysis of *mPing* insertions for the EG4 generational materials. *de novo* bands detected in progeny are noted by arrowheads and inherited *de novo* insertions are noted by asterisks.

Table 1. *de novo* insertions of *mPing* in the S_1 generation of EG4.

Adapter Primer	S_1 plants									
	1	2	3	4	5	6	7	8	9	10 (PS1*)
*Bfa*I + AAA	0	1	1	1	0	0	0	0	1	1 (1)†
*Bfa*I +ACG	6	2	0	3	0	0	0	0	0	2 (1)
*Bfa*I +AGT	0	0	0	0	1	0	1	1	0	3 (3)
*Bfa*I +ATC	2	0	0	0	1	2	1	0	0	3 (3)
*Bfa*I +CAT	2	1	1	2	0	2	0	3	1	1 (1)
*Bfa*I +CCC	1	0	0	0	1	3	1	1	1	2 (2)
*Bfa*I +CGA	1	0	1	1	0	0	0	0	2	1 (1)
*Bfa*I +CTG	3	0	2	1	1	1	1	0	1	6 (4)
*Bfa*I +GAC	1	4	0	1	0	0	1	2	1	0 (0)
*Bfa*I +GCT	1	0	1	0	0	0	0	0	1	5 (4)
*Bfa*I +GGG	1	1	1	4	0	0	1	0	0	1 (1)
*Bfa*I +GTA	0	1	2	0	1	0	2	0	1	4 (2)
*Bfa*I +TAG	2	0	0	2	1	1	0	1	2	1 (1)
*Bfa*I +TCA	0	1	1	1	0	1	0	2	1	4 (4)
*Bfa*I +TGC	1	1	0	0	0	1	2	3	0	4 (3)
*Bfa*I +TTT	0	1	0	0	1	0	1	0	1	5 (4)
Total	21	13	10	16	7	11	11	13	13	43 (35)
Average per primer	1.31	0.81	0.63	1.00	0.44	0.69	0.69	0.81	0.81	2.69
Total per plant	84	52	40	64	28	44	44	52	52	172
Average per generation						63.2				

*S_1 plant (parent) used to generate S_2 plants by self fertilization.

†Numbers in parenthesis are the heritable insertions (bands segregating in S_2 plants).

Table 2. *de novo* insertions of *mPing* in the S_2 generation of EG4

Adapter Primer	S_2 plants									
	1	2	3	4	5	6	7	8	9	10
*Bfa*I + AAA	0	0	0	0	0	0	1	1	0	0
*Bfa*I +ACG	1	1	1	0	2	0	2	1	0	1
*Bfa*I +AGT	0	2	0	0	1	2	2	0	0	2
*Bfa*I +ATC	0	1	0	2	0	2	1	0	0	1
*Bfa*I +CAT	2	0	0	1	1	2	0	0	0	0
*Bfa*I +CCC	1	2	0	1	0	1	0	0	3	0
*Bfa*I +CGA	0	0	2	1	0	0	0	2	0	0
*Bfa*I +CTG	0	1	1	0	1	0	0	2	0	0
*Bfa*I +GAC	1	0	2	2	1	2	0	2	1	3
*Bfa*I +GCT	0	1	1	2	2	1	1	1	1	1
*Bfa*I +GGG	0	2	0	0	1	2	1	0	1	0
*Bfa*I +GTA	1	1	0	0	1	0	1	0	0	0
*Bfa*I +TAG	0	0	0	0	0	1	1	0	1	1
*Bfa*I +TCA	0	1	2	2	2	3	2	0	3	2
*Bfa*I +TGC	3	0	0	0	1	0	1	1	1	1
*Bfa*I +TTT	1	2	0	2	0	2	0	0	0	0
Total	10	14	9	13	13	18	13	10	11	12
Average per primer	0.63	0.88	0.56	0.81	0.81	1.13	0.81	0.63	0.69	0.75
Total per plant	40	56	36	52	52	72	52	40	44	48
Average per generation	49.2									

Table 3. *de novo mPing* insertion sites in EG4

BAC	Site	cDNA	Distance	Coding protein
AP003339	12036	AK120073	253	*Oryza sativa* RIP1 mRNA for ribosome-inactivating protein 1
AP003254	25123	AK058518	2256	NPK1-related protein kinase-like protein
AP003261	83788	AK073884	-1734	Amino acid permease 6
AP003232	150997	AK069400	-1121	Plectin-related protein-like
AP003266	36188	AK073211	0	PIR7a and PIR7b gene
AP004121	6154			
AP005849	106090	AK099853	4406	Putative Cgi67 serine protease
AP004048	45607	AK107997	0	Unknown protein
AC113930	28923			
AC107619	67639	AK063182	-4114	Full-length cDNA
AL662992	12708	AK071063	447	Full-length cDNA
AL662999	50722			
AC117265	38040	AK011268	2650	Putative 60S ribosomal protein L36
AP003539	16894	AK071563	0	Full-length cDNA
AP006458	106710			
AP005165	113586			
AP005755	55862	AK121134	3935	Full-length cDNA
AP005579	81939	AK108547	-131	Full-length cDNA
AC134046	136472			
AC120308	24484			
AL731742	143358			
AP005516	68380			
AP003914	47055	AK100665	3206	Full-length cDNA

Table 4. Characteristics of *mPing* insertion sites

Strain	# of analyzed insertions	Insert into:						
		Single copy(%)[a]	Exon	Intron	0-1kb[b]	1-3kb[b]	3-5kb[b]	>5kb[b]
NB	50	42(84)	1	3	9	7	2	20
EG4	221	201(91)	1	10	38	69	44	39
EG4 (*de novo*)	23	23(100)	3	0	3	4	4	9

a The number of insertions into single copy sequences. The number in parenthesis is the percentage of the number of insertions analyzed.
b The distance of the insertion sites from cDNA coding sequences.

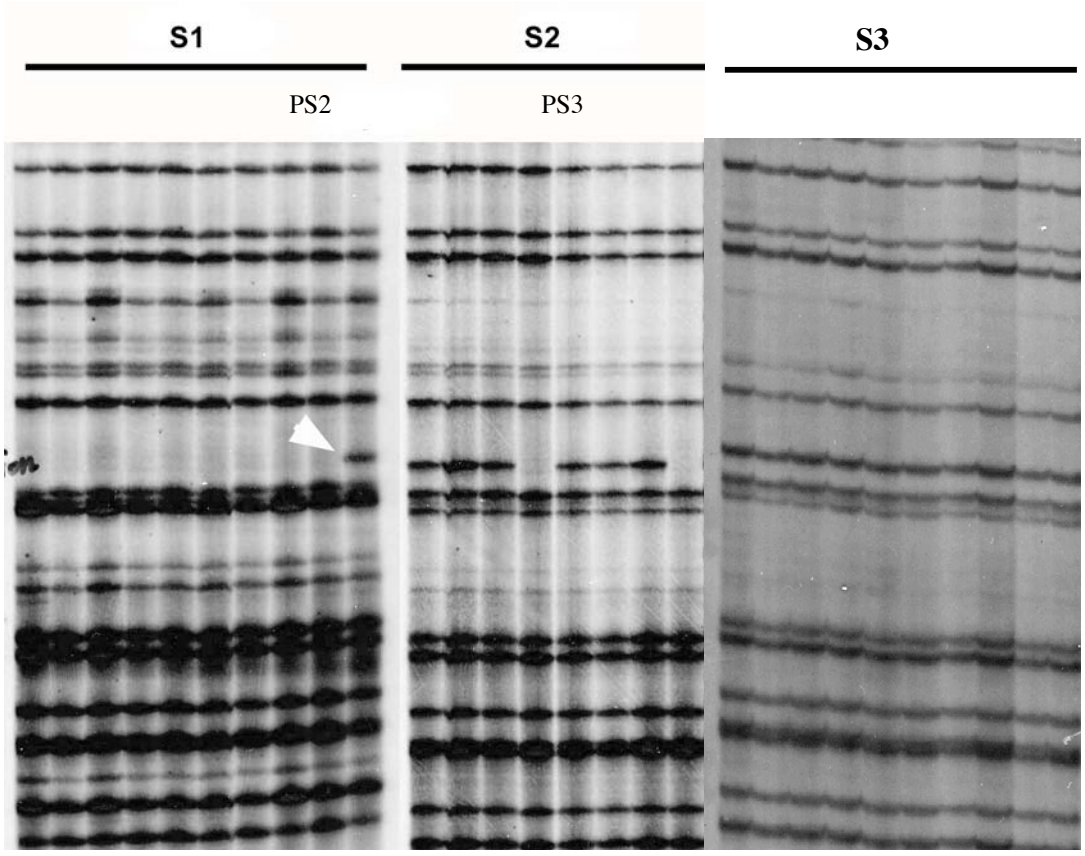
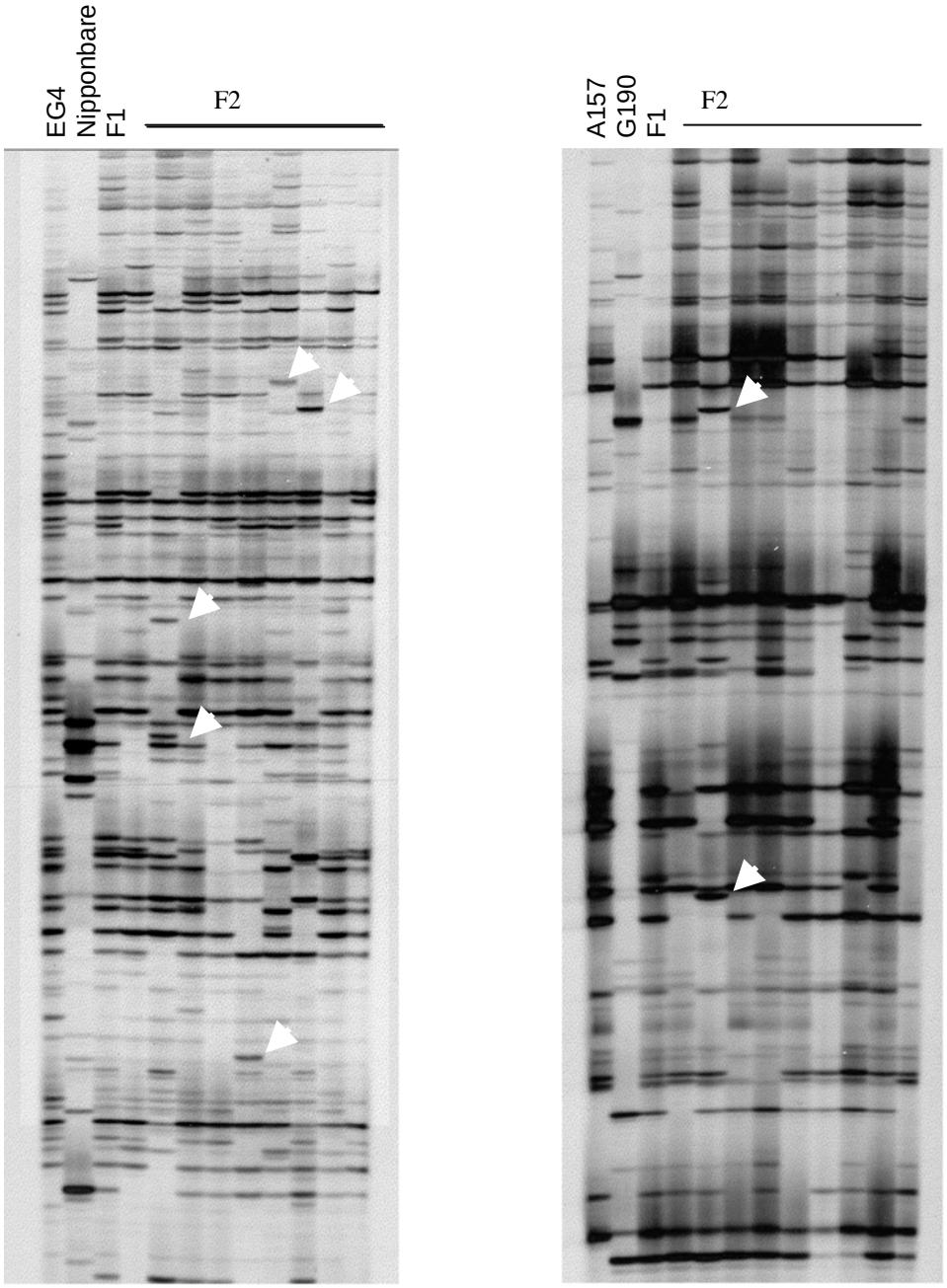


Figure 5. Transposon display analysis of *mPing* insertions of generational material from strain Nipponbare. The S2 plants were derived from one of the S1 plants (PS2) and the S3 plants were derived from one of the S2 plants (PS3) by self-fertilization. White arrowhead indicates a *de novo* insertion.



*Bfa*I + AA
EG4 (female) X Nipponbare (male)

*Bfa*I + TA
A157 (female) X G190 (male)

Figure 6. Transposon display of *mPing* using DNA from hybrid plants between high and low-copy landraces. White arrowhead indicates *de novo* insertions.

CHAPTER 3

CONCLUSION

MITEs are the most abundant group of TEs in the rice genome where ~ 90,000 MITEs are present. Because MITEs are preferentially inserted into genic regions of the genome, it has been of interest how they are able to amplify to such high copy numbers and how host genomes react to those potentially deleterious insertions. In this study, the analysis of an active rice MITE, *mPing*, in different Gimbozu strains shows that *mPing* is still actively transposing in some Gimbozu strains where *mPing* copy number already reached over a thousand. Potentially harmful *de novo mPing* elements inserted that into genes were rapidly eliminated from population and then, selection against the insertion near genes were acting slowly over time. This study also showed that *mPing* was activated in low-copy strains, but the frequency of transposition was extremely low. This suggests that *mPing* copy number had been increased gradually in most rice cultivars but somehow the transposition rate had been accelerated in some Gimbozu landraces, resulting in the explosive amplification of *mPing* elements. Lower level of *mPing* activity in hybrid plants between high and low-copy strains implies that several factors are probably involved in determining the transposition rate of *mPing*. Before this study, *mPing* was thought to be activated by radiation event in Gimbozu. However, this study showed that *mPing* was activated prior to radiation and it is still active under normal growth condition. This is significant results to provide the possibility of *mPing* as a tagging tool for rice gene discovery.

REFERENCES

An G, Jeong D-H, Jung K-H, Lee S (2005) Reverse genetic approaches for functional genomics of rice. Plant Molecular Biology 59(111-123).

Bennetzen JL (2000) Transposable element contributions to plant gene and genome evolution. Plant Mol Biol 42(1): 251-269.

Bureau TE, Wessler SR (1992) *Tourist*: a large family of inverted-repeat element frequently associated with maize genes. Plant Cell 4: 1283-1294.

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

Burr B (2002) Mapping and sequencing the rice genome. Plant Cell 14(3): 521-523.

Capy P, Bazin C, Higuet D, Langin T (1998) Dynamics and evolution of transposable elements; Bioscience L, editor. Austin, Texas: Springer-Verlag.

Capy P, Gasperi G, Biemont C, Bazin C (2000) Stress and transposable elements: co-evolution or useful parasites? [In Process Citation]. Heredity 85(Pt 2): 101-106.

Casa AM, Brouwer C, Nagel A, Wang L, Zhang Q et al. (2000) The MITE family heartbreaker (Hbr): molecular markers in maize. Proc Natl Acad Sci U S A 97(18): 10083-10089.

Coen ES, Carpenter R, Martin C (1986) Transposable elements generate novel spatial patterns of gene expression in *Antirrhinum majus*. Cell 47: 285-296.

Craig NL, Craigie R, Gellert M, Lambowitz AM (2002) Mobile DNA II. Washington, D.C.: American Society for Microbiology Press.

Doolittle RF, Feng D-F, Johnson Ms, McClure MA (1989) Origins and evolutionary relationships of retroviruses. The Quarterly Review of Biology 64: 1-30.

Feschotte C, Jiang N, Wessler SR (2002a) Plant transposable elements: where genetics meets genomics. Nat Rev Genet 3(5): 329-341.

Feschotte C, Zhang X, Wessler S (2002b) Miniature inverted-repeat transposable elements (MITEs) and their relationship with established DNA transposons. In: Craig NL, Craigie R, Gellert M, Lambowitz AM, editors. Mobile DNA II. Washington, DC: American Society for Microbiology Press. pp. 1147-1158.

Feschotte C, Swamy L, Wessler SR (2003) Genome-wide analysis of *mariner*-like transposable elements in rice reveals complex relationships with *Stowaway* MITEs. Genetics *163*: 747-758.

Flavell AJ, Pearce SR, Kumar A (1994) Plant transposable elements and the genome. Current Opinions in Genetics and Development 4: 838-844.

Hirochika H (1993) Activation of tobacco retrotransposons during tissue culture. EMBO J 12: 2521-2528.

Hirochika H, Sugimoto K, Otsuki Y, Kanda M (1996) Retrotransposons of rice involved in mutations induced by tissue culture. Proc Natl Acad Sci 93: 7783-7788.

Hirochika H, Guiderdoni E, An G, Hsing Y-I, Eun MY et al. (2004) Rice mutant resources for gene discovery. Pland Molecular Biology 54: 325-334.

Jiang N, Feschotte C, Zhang XY, Wessler SR (2004) Using rice to understand the origin and amplification of miniature inverted repeat transposable elements (MITEs). Current Opinion in Plant Biology 7(2): 115-119.

Jiang N, Bao Z, Zhang X, McCouch SR, Eddy SR et al. (2003) An active DNA transposon in rice. Nature 421: 163-167.

Jin YK, Bennetzen JL (1994) Integration and nonrandom mutation of a plasma membrane proton ATPase gene fragment within the Bs1 retroelement of maize. Plant Cell 6(8): 1177-1186.

Kaeppler SM, Kaeppler HF, Rhee Y (2000) Epigenetic aspects of somaclonal variation in plants. Plant Molecular Biology 43: 179-188.

Kalendar R, Tanskanen J, Immonen S, Nevo E, Schulman AH (2000) Genome evolution of wild barley (*Hordeum spontaneum*) by *BARE*-1 retrotransposon dynamics in response to sharp microclimatic divergence. Proc Natl Acad Sci 97: 6603-6607.

Kidwell MG, Lisch DR (2001) Perspective: transposable elements, parasitic DNA, and genome evolution. Evolution Int J Org Evolution 55(1): 1-24.

Kidwell MG, Lisch D (2002) Transposable elements as sources of genomic variation. In: Craig NL, Craigie R, Gellert M, Lambowitz AM, editors. Moblie DNA II. Washington, DC: American Society for Microbiology Press. pp. 59-90.

Kidwell MG, Kidwell JF, Sved JA (1977) Hybrid dysgenesis in Drosophila melanogaster: syndrome of aberrant traits including mutation, sterility, and male recombination. . Genetics 86: 813-833.

Kikuchi K, Terauchi K, Wada M, Hirano HY (2003a) The plant MITE mPing is mobilized in anther culture. Nature 421(6919): 167-170.

Kikuchi S, Satoh K, Nagata T, Kawagashira N, Doi K et al. (2003b) Collection, mapping, and annotation of over 28,000 cDNA clones from japonica rice. Science 301: 376-379.

Klinakis AG, Zagoraiou L, Vassilatis DK, Savakis C (2000) Genome-wide insertional mutagenesis in human cells by the Drosophila mobile element Minos. EMBO Rep 1(5): 416-421.

Komatsu M, Shimamoto K, Kyozuka J (2003) Two-Step Regulation and Continous Retrotransposition of the Rice LINE-Type Retrotransposon *Karma*. The Plant Cell 15: 1934-1944.

Kumar A, Bennetzen JL (1999) Plant retrotransposons. Annu Rev Genet 33: 479-532.

Kunze R, Saedler H, W.E. L (1997) Plant transposable elements. Advances in Botanical Research 27: 331-470.

Kurusu T, Sakurai Y, Miyao A, Hirochika H, Kuchitsu K (2004) Identification of a putative voltage-gated Ca2+ -permeable channel (OsTPC1) involved in Ca2+ influx and regulation of growth and development in rice. Plant and Cell Physiology 45(693-702).

Labrador M, Corces VG (1997) Transposable element-host interactions: regulation of insertion and excision. Annu Rev Genet 31: 381-404.

Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC et al. (2001) Initial sequencing and analysis of the human genome. Nature 409(6822): 860-921.

Larkin PJ, Scowcroft WR (1981) Somaclonal variation - a novel source of variability from cell cultures for plant improvement. Theoretical and Applied Genetics 60: 197-214.

Leung H, An G (2004) Rice functional genomics: large-scale gene discovery and applications to crop improvement. Advances in Agronomy 82: 55-111.

Lin X, Long L, Shan X, Zhang S, Shen S et al. (2006) *In planta* mobilization of *mPing* and its putative autonomous element *Pong* in rice by hydrostatic pressurization. Journal of Experimental Botany 57(10): 2313-2323.

Londo JP, Chiang Y, Hung K (2006) Phylogeography of Asian wild rice, Oryza rufipogon, reveals multiple independent domestications of cultivated rice, Oryza sativa. Proceedings of the National Academy of Sciences of the United States of America 103(25): 9578-9583.

Madlung A, Comai L (2004) The Effect of Stress on Genome Regulation and Structure. Annals of Botany 94: 481-495.

Martienssen RA (1998) Functional genomics: Probing plant gene function and expression with transposons. Proceedings of the National Academy of Sciences of the United States of America 95: 2021-2026.

May BP, Martienssen RA (2003) Transposon mutagenesis in the study of plant development. Critical Review in Plant Sciences 22: 1-35.

McClintock B (1951) Chromosome organization and genic expression. Cold Spring Harbor Symposium on Quantitative Biology 16: 13-47.

McClintock B (1984) The significances of responses of the genome to challenge. Science 226: 792-801.

Meyers BC, Tingey SV, Morgante M (2001) Abundance, distribution, and transcriptional activity of repetitive elements in the maize genome. Genome Res 11(10): 1660-1676.

Miyao A, Tanaka K, Murata K, Sawaki H, Takeda S et al. (2003) Target site specificity of the Tos17 retrotransposon shows a preference for insertion within genes and against insertion in retrotransposon-rich regions of the genome. The Plant Cell 15(1771-1780).

Naito K, Cho E, Yang G, Campbell MA, Yano K 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(47): 17620-17625.

Nakazaki T, Okumoto Y, Horibata A, Yamahira S, Teraishi M et al. (2003) Mobilization of a transposon in the rice genome. Nature 421(6919): 170-172.

SanMiguel P, Bennetzen JL (1998) Evidence that a recent increase in maize genome size was caused by the massive amplification of intergene retrotransposons. Ann Bot 81: 37-44.

Santiago N, Herraiz C, Goni JR, Messeguer X, Casacuberta JM (2002) Genome-wide Analysis of the Emigrant Family of MITEs of Arabidopsis thaliana. Mol Biol Evol 19(12): 2285-2293.

Sassaman DM, Dombroski BA, Moran JV, Kimberland ML, Naas TP et al. (1997) Many human L1 elements are capable of retrotransposition. Nat Genet 16(1): 37-43.

Shan X, Liu Z, Dong Z, Wang Y, Chen Y et al. (2005) Mobilization of the active MITE transposons mPing and Pong in rice by introgression from wild rice (Zizania latifolia Griseb.). Molecular Biology and Evolution 22: 976-990.

Smith B (1998) The Emergence of Agriculture. New York: Scientific American Library, A Division of HPHLP.

Springer PS (2000) Gene traps: tools for plant development and genomics. The Plant Cell 12: 1007-1020.

Stumm GB, Vedder H, Schlegel J (1997) A simple method for isolation of PCR fragments from silver-stained polyacrylamide gels by scratching with a fine needle. Technical Tips Online 1: 17.

Takagi K, Ishikawa N, Maekawa M, Tsugane K, Iida S (2007) Transposon display for active DNA transposons in rice. Genes & Genetic Systems 82(2): 109-122.

Van den Broeck D, Maes T, Sauer M, Zethof J, De Keukeleire P et al. (1998) *Transposon Display* identifies individual transposable elements in high copy number lines. The Plant Journal 13: 121-129.

Vicient CM, Suoniemi A, Anamthawat-Jonsson K, Tanskanen J, Beharav A et al. (1999) Retrotransposon BARE-1 and Its Role in Genome Evolution in the Genus Hordeum. Plant Cell 11(9): 1769-1784.

Vos P, Hogers R, Bleeker M, Reijans M, van de Lee T et al. (1995) AFLP, a new technique for DNA fingerprinting. NAR 23: 4407-4414.

Wessler SR (1996) Turned on by stress. Plant retrotransposons. Curr Biol 6(8): 959-961.

Wessler SR (2006) Eukaryotic Transposable elements: Teaching Old Genomes New Tricks; Caporale L, editor. New York: Oxford Univeristy Press.

Wessler SR, Bureau TE, White SE (1995) LTR-retrotransposons and MITEs: important players in the evolution of plant genomes. Curr Opin Genet Dev 5: 814-821.

Xiong Y, Eickbush TH (1990) Origin and evolution of retroelements based upon their reverse transcriptase sequences. Embo J 9(10): 3353-3362.

Yang G, Zhang F, Hancock C, S.R.Wessler (In press) Transposition of the rice miniature inverted repeat transposable element mPing in Arabidopsis thaliana. Proceedings of the National Academy of Sciences of the United States of America.

Yano M, Katayose Y, Ashikari M, Yamanouchi U, Monna L et al. (2000) Hd1, a major photoperiod sensitivity quantitative trait locus in rice, is closely related to the Arabidopsis flowering time gene CONSTANS. Plant Cell 12(12): 2473-2484.

Zhang Q, Arbuckle J, Wessler SR (2000) Recent, extensive and preferential insertion of members of the miniature inverted-repeat transposable element family *Heartbreaker (Hbr)* into genic regions of maize. Proc Natl Acad Sci USA 97: 1160-1165.

Zhang X, Jiang N, Feschotte C, Wessler SR (2004) Distribution and evolution of *PIF*- and *Pong*-like transposons and their relationships with *Tourist*-like MITEs. Genetics 166: 971-986.