DNA METHYLATION ANALYSIS OF THE RICE TRANSPOSOBALE ELEMENTS MPING AND PING

by

TIANLE CHEN

(Under the Direction of Susan Wessler)

ABSTRACT

Transposable elements are widely distributed in eukaryotes where they usually comprise the largest fraction of the genome. The *Tourist*-like MITE *mPing* was the first active DNA transposon discovered in rice. *mPing* copy number varies dramatically in different rice cultivars from less than 50 to over 1,000 copies. In this study we analyzed the methylation patterns of *mPing* and *Ping* from which it was derived by internal deletion in the high-copy strain Gimbozu EG4 and the low-copy strain Nipponbare. Our results demonstrate that *mPing* is heavily methylated overall in both strains, and that methylation of *mPing* does not prevent its movement. There are seven *Ping*s in EG4 and one in Nipponbare. The terminal ends of *Ping* are heavily methylated similar to *mPing*. The low and variable methylation pattern of *Ping* ORF1 promoter region suggested a correlation between methylation and *Ping* TPase transcription, which may be the control point of *mPing* transposition. More experiments are ongoing to unravel the mystery of *mPing* activity.

INDEX WORDS: Transposable elements, MITE, rice, transposition, DNA methylation

DNA METHYLATION ANALYSIS OF THE RICE TRANSPOSOBALE ELEMENTS MPING AND PING

by

TIANLE CHEN

B.S., Fudan University, China, 2004

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

MASTER OF SCIENCE

ATHENS, GEORGIA

2008

DNA METHYLATION ANALYSIS OF THE RICE TRANSPOSOBALE ELEMENTS MPING AND PING

by

TIANLE CHEN

Major Professor: Susan Wessler

Committee: Kelly Dawe
Michelle Momany

Electronic Version Approved:

Maureen Grasso
Dean of the Graduate School
The University of Georgia
August 2008

DEDICATION

To My Parents.

ACKNOWLEDGEMENTS

I am deeply grateful to my major professor, Dr.Susan Wessler, for her support and guidance. I also thank my committee members, Dr.Kelly Dawe and Dr.Michelle Momany for their valuable advice and help. I thank my colleagues in the Wessler lab: Eunyoung Cho, Yujun Han, Nathan Hancock, Dawn Holligan, Eleanor Kuntz, Ken Naito, Guojun Yang and Feng Zhang, for their help and friendship. Finally, I thank my parents, Yan Chen and Runzhong Zhang, for their unconditional love and support.

TABLE OF CONTENTS

Page

LIST OF TABLES

Page

LIST OF FIGURES

Page

CHAPTER 1

INTRODUCTION AND LITERATURE REVIEW

Transposable elements

Transposable elements (TEs) are genetic entities that are able to transpose from one chromosomal locus to another in their host genome. Initially discovered by Barbara McClintock in maize in the 1940s (McClintock, 1948, 1949, 1951), TEs have been found in the genome of all characterized organisms. TEs can be divided into two major classes based on their mechanism of transposition. Class I elements (also called retrotransposons) transpose via a RNA intermediate (Capy et al., 1997; Capy et al., 1998). A double stranded cDNA copy is produced by reverse transcription of the RNA intermediate of the element and then inserted into a new position in the host genome (Bingham and Zachar, 1989; Boeke and Corces, 1989). There are two types of Class I elements: LTR (long terminal repeats) retrotransposons and non-LTR retrotransposons. LTR retrotransposons can be further divided into *Ty1/copia*-like and *Ty3/gypsy*-like elements based on difference of the proteins they encode (Doolittle et al., 1989; Xiong and Eickbush, 1990; Kumar and Bennetzen, 1999). Non-LTR elements include LINEs (long interspersed nuclear elements) and SINEs (short interspersed nuclear elements) (Deininger, 1989; Hutchison et al., 1989).

Class II elements (also called DNA transposons) transpose via a DNA intermediate and the element-encoded transposase (TPase) catalyzes the excision and reinsertion of the element (Craig et al., 2002). Target site duplications (TSDs) are produced upon insertion. DNA

transposons are organized into superfamilies based on the sequence homology in their TPases. Because TPases usually bind at or near the ends of elements, superfamilies also have similar terminal inverted repeats (TIRs). Examples of DNA TE superfamilies are *Tc1/mariner* (Feschotte and Wessler, 2002; Plasterk and van Luenen, 2002), *CACTA* (Rubin et al., 2001; Kunze and Weil, 2002), *Mutator* (Lisch et al., 1999; Mao et al., 2000; Yu et al., 2000; Lisch, 2002) and *hAT* (Kunze and Weil, 2002). There are autonomous elements and nonautonomous elements in each family. Autonomous elements encode functional TPase that is necessary for their transposition, while nonautonomous elements are usually defective versions of autonomous elements and lack the coding capacity, thus transposition of nonautonomous elements requires TPase encoded by autonomous elements (Capy et al., 1998; Feschotte and Wessler, 2002).

Miniature inverted-repeat transposable elements (MITEs)

MITEs are a special type of nonautonomous DNA elements. Numerous MITEs have been found in a wide range of plant genomes, as well as animal and fungal genomes (Bureau and Wessler, 1992; Morgan, 1995; Bureau et al., 1996; Oosumi et al., 1996; Tu, 1997; Casacuberta et al., 1998; Izsvak et al., 1999; Tu, 2001; Feschotte and Wessler, 2002). General characteristics of MITEs are their small size (~100-500bp), high copy number (thousands or tens of thousands per family), and insertion site preference of genic regions (5' and 3' untranslated regions and introns of genes) (Feschotte and Wessler, 2002; Feschotte et al., 2002a). A unique feature of MITEs is that their high copy numbers were usually arisen from the amplification of a small number of elements (Feschotte et al., 2002b; Santiago et al., 2002).

MITEs can be classified into two major superfamilies based on the similarity of their TIRs and TSDs (Jiang et al., 2004). *Tourist*-like MITEs preferentially insert into the trinucleotide TAA or TTA and are associated with the *PIF/Harbinger* superfamily. Direct functional evidence of this association is that the rice *Tourist*-like MITE *mPing* can be transposed by TPases encoded by the autonomous element *Ping* or *Pong* (Yang et al., 2007). *Stowaway*-like MITEs preferentially insert into the dinucleotide TA and are associated with the *Tc1/mariner* superfamily (Feschotte et al., 2003). Transposition assays in yeast and *Arabidopsis* showed that *Stowaway*-like MITEs from rice are functionally linked to *Osmar* elements since they can be cross-mobilized by TPase from *Osmar*s (Yang et al., 2006).

Epigenetic regulation of transposable elements

Because TEs are able to replicate and increase their copy numbers in host genomes, they are directly or indirectly involved in genome restructuring, including ectopic recombination, chromosome translocation and inversion, which provides a source for genetic diversity (Kidwell and Lisch, 2001; Kidwell and Lisch, 2002; Wessler, 2006). Although host genomes may benefit from TE activity in the long-term, more deleterious effects can be caused due to mutations they cause in the genome(Bennetzen, 2000). The host appears to have evolved several mechanisms to prevent the amplification of transposable elements. Deletions and mutations of TE sequences are important mechanisms for TE inactivation (Jin and Bennetzen, 1994; Bennetzen, 2000). For instance, LTR-retrotransposons in many organisms were found to have undergone deletions by unequal recombination and illegitimation, which generate solo-LTRs (Devos et al., 2002; Vitte

and Panaud, 2003). However, these recombinations have no effect on TEs that are accumulated in regions with low recombination rates. Another example is the repeat-induced point mutation (RIP) in *Neurospora crassa*, which specifically targets the duplicated sequences in the genome during the sexual phase of the life cycle, and thus inactivates TEs by mutations (Freitag et al., 2002; Galagan et al., 2003; Selker et al., 2003; Galagan and Selker, 2004).

Although in many cases TEs are silenced by mutations and deletions, there are still some TEs that remain intact in the host genomes. However, most of these intact TEs are silenced, suggesting that their activities are suppressed by epigenetic silencing mechanisms, such as DNA methylation, a common DNA modification that defends genome against TE transposition (Martienssen, 1998; Feschotte and Wessler, 2002; Kato et al., 2003; Chan et al., 2005; Freitag and Selker, 2005; Slotkin and Martienssen, 2007; Weil and Martienssen, 2008). There are three types of DNA methylation in plants. CG methylation is established and maintained by the mammalian DNMT1 homolog METHYLTRANSFERASE1 (MET1). The DNMT3a/b homologs DOMAINS REARRANGED METHYLASE 1 and 2 (DRM1/2) are responsible for the maintenance of CHH methylation and also the establishment of *de novo* methylation in all sequence contexts. This pathway requires the active targeting of siRNA (Tran et al., 2005b; Zhang, 2008). The plant specific CHROMO-METHYLASE3 (CMT3) was found to maintain CNG methylation under the direction of the H3K9me2 pathway (Jackson et al., 2002; Jackson et al., 2004; Zhang et al., 2007; Zhang, 2008). DNA methylation in promoters is correlated with transcriptional repression whereas DNA methylation in the transcribed regions of genes usually does not affect transcription (Zhang et al., 2006; Zilberman et al., 2007).

Based on the transposition mechanism, DNA methylation affects the activity of DNA TEs in two ways, either silencing the TPase expression transcriptionally, or impairing the binding affinity of TPase to TE ends. Abundant evidence has been found for a correlation between DNA methylation and the inactive state of TEs. McClintock's *Activator* (McClintock, 1951) and *Suppressor-Mutator* (McClintock, 1957) and Robertson's *Mutator* TEs (Martienssen et al., 1990) are known to undergo cycles of active phase and inactive phase. There is a decreased level of DNA methylation of these autonomous TEs during their active phase, especially in the 5' end of the TE near the TPase transcription start site. This is usually the promoter region of TPase and demethylation is associated with TPase transcription and TE transposition (Chandler and Walbot, 1986; Banks et al., 1988; Brutnell and Dellaporta, 1994; Fedoroff, 1999; Cui and Fedoroff, 2002; Lisch et al., 2002). In addition, *in vitro* assays on *En/Spm* and *Ac/Ds* showed that unmethylated or hemi-methylated transposon DNA has a much higher binding affinity for TPase than fully-methylated DNA, suggesting that DNA methylation at TE TIRs or sub-TIRs might inhibit TPase binding and affect TE transposition (Kunze and Starlinger, 1989; Wang et al., 1996; Ros and Kunze, 2001; Cui and Fedoroff, 2002; Hashida et al., 2006).

For both DNA and RNA elements, loss of DNA methylation often reactivates TE transcriptionally (Kato et al., 2003). For example, the rice retrotransposon *Tos17* is transcribed and reactivated after the treatment with an inhibitor of DNA methylation, 5-azacytidine (Cheng et al., 2006). The silencing of *Mutator* elements in Arabidopsis is relieved in several DNA methyltransferase mutants, and the transposition rate of the elements is increased by two to three fold in *met1* mutant, five to six fold in *cmt3* mutant, 20 fold in *ddm1* mutant and *met1/cmt3*

double mutant (Singer et al., 2001; Kato et al., 2003; Lippman et al., 2003). Therefore, DNA methylation plays an important role in regulating TE transposition.

CHAPTER 2

DNA METHYLATION ANALYSIS OF THE RICE TRANSPOSOBALE ELEMENTS MPING AND PING

BACKGROUND AND RATIONALE

mPing is a *Tourist*-like MITE discovered by rice genomic database analyses. It was first found to be transposing in three kinds of plant materials in different labs independently: in long-term cell culture (Jiang et al., 2003), in newly derived anther culture (Kikuchi et al., 2003) and in a gamma-ray-irradiated mutant line derived from the japonica cultivar Gimbozu EG4 (Nakazaki et al., 2003). Later it was found to be actively transposing in Gimbozu EG4 prior to gamma-ray irradiation and in related landraces (Naito et al., 2006).

Structurally, *mPing* is 430bp of length, with 15bp terminal inverted repeats (5'-GGCCAGTCACAATGG-3') and a target site duplication of TAA or TTA (Jiang et al., 2003). Database searches of the annotated Nipponbare genome revealed that *mPing* is a deletion derivative of *Ping*, an autonomous element of 5341bp. *mPing* shares the 253bp of the 5' end and 177bp of the 3' end with *Ping* (Jiang et al., 2003). There is only a single copy of *Ping* in the Nipponbare genome. Database searches also showed that *Ping* shares over 85% sequence identity with another element called *Pong*. Both *Ping* and *Pong* have two open reading frames (ORFs) and they share >80% similarity of internal sequences corresponding to the two ORFs (Fig.1) (Jiang et al., 2003).

As a nonautonomous element, transposition of *mPing* relies on TPase provided *in trans*. *Ping* was found to be co-activated with *mPing* in the anther culture derived from Nipponbare (Kikuchi et al., 2003), and *Pong* was co-activated with *mPing* in the long-term culture derived from an *indica* strain C5924 where there was no *Ping* (Jiang et al., 2003). Recent data demonstrated that both *Ping* and *Pong* are able to mobilize *mPing* in transgenic *Arabidopsis* (Yang et al., 2007). Both ORF1 and ORF2 of *Ping* or *Pong* are required for the transposition event. ORF2 encodes the TPase, while the function of ORF1 is still unknown (Yang et al., 2007).

A survey of *mPing* abundance suggested that *mPing* copy number is highest in the temperate *japonica*s (about 70 copies) and lowest in the tropical *japonica*s (less than 10 copies) (Jiang et al., 2003). Later studies showed that the *mPing* copy number varies among different temperate *japonica*s. Strikingly, *mPing* has reached over 1,000 copies in landrace EG4 of Gimbozu cultivar, which was generated from Aikoku cultivar by a single mutation event. Recent studies on Gimbozu EG4 showed that *mPing* is still actively transposing at a rate of 50 new insertions per generation per plant. In contrast, in Nipponbare, a closely related cultivar that were bred from common ancestors of Gimbozu, there are less than 50 *mPing* copies and the transposition frequency is very low (Naito et al., 2006).

The significant difference in *mPing* copy number between Gimbozu EG4 and Nipponbare prompted people to examine the status of their autonomous partners *Ping* and *Pong*. *Ping* copy number correlates with *mPing* copy number, with only one copy of *Ping* in Nipponbare and 7 copies in EG4. In contrast, *Pong* copy number is about the same in the two strains and their insertion sites are conserved, suggesting that *Pong* has been inactive in these

strains (Naito et al., 2006). These facts suggested that *Ping*, not *Pong*, is primarily responsible for the high frequency of *mPing* transposition in Gimbozu.

The recent and ongoing transposition of *mPing* and the drastic difference in *mPing* copy numbers between the two extremely closely related cultivars provide a unique opportunity to study the mechanisms that affect MITE amplification. A random sample of *mPing* in Nipponbare and EG4 was sequenced and all of them are same. DNA sequences of *Ping* copies in the two cultivars are identical, indicating that *Ping* TPase is not mutated. Although *mPing* is not active in Nipponbare, it can be reactivated in the anther culture derived from Nipponbare (Kikuchi et al., 2003), which means that all the factors required for *mPing* transposition are present in Nipponbare. Based on these facts, we came up with two possible models for the regulation of *mPing* transposition. In the first model we suspected that *mPing* transposition is regulated by DNA methylation. Methylation of *mPing* may affect the TPase binding affinity to the element, and/or methylation of *Ping* may prevent TPase transcription. Either of these two ways may be the cause pf the distinct *mPing* transposition in EG4 and Nipponbare. The alternative model is that *mPing* transposition is not correlated with the methylation status of *mPing* or *Ping*. Each *Ping* may be transcribed regardless of their methylation status and *mPing* transposition depends on the dosage of TPase in the genome. In order to test these models, we need to investigate the methylation status of *mPing* and *Ping* to see whether they show difference in the two strains and whether they are associated with *mPing* transposition. In this study, we performed the methylation analysis of *mPing* and *Ping* in EG4 and Nipponbare, which provides an important clue in understanding the regulation of *mPing* transposition.

MATERIALS AND METHODS

Plant Materials and genomic DNA extraction

Seeds of the rice (*Oryza sativa*) subspecies *japonica* cultivar Nipponbare, and the cultivar Gimbozu landrace EG4 were obtained from the GenBank project of the National Institute of Agrobiological Science, Ibaraki, Japan. Seeds were dehusked, sterilized with 95% ethanol and 40% bleach, and germinated at 27°C on plates with sterilized ddH_2O. Three-week old seedlings were transferred into soil. Genomic DNA was extracted from leaf tissues of mature plants using CTAB method as described in (Porebski, 1997).

Genomic Bisulfite Sequencing

About 300ng genomic DNA was digested with 25 unit *HindIII* overnight, and then purified using phenol:chloroform precipitation. Bisulfite conversion was done using EpiTect Bisulfite Kit (Qiagen, Inc., Valencia, CA) following the manufacturer's instructions. After elution of the treated samples, 2μL of the elute was used for subsequent PCR amplification with AmpliTaq Gold DNA polymerase (Foster City, CA). PCRs were performed according to the manufacturer's recipe with an optimized $MgCl_2$ concentration of 4mM. The cycling parameters were as follows: 94℃ for 2 min; 4 cycles of 94℃ for 1 min, 57℃ for 2 min, and 72℃ for 4 min; 39 cycles of 94℃ for 45 s, 57℃ for 2 min, and 72℃ for 2 min; and 72℃ for 10 min.

Degenerate primers (listed in Tables.1 and 2) were used to amplify the sequences of interest. PCR products were cloned using the TOPO TA cloning kit (Invitrogen, Carlasbad, CA).

20 individual clones containing inserts of the correct size were sequenced using T7 primer.

Methylation level calculation and methylation map construction

ClustalX was used for the alignment of the 20 clones sequenced. At each cytosine site, the number of methylated cytosines out of 20 clones was counted and the methylation level was calculated using this number divided by 20. Methylation map was constructed in the way that one black box represents 10% of methylation level at each cytosine site. Total methylation level of CG, CNG or CHH was calculated by taking the average level of all the cytosines in CG, CNG or CHH context in the sequence of interest respectively.

RESULTS

Methylation patterns of *mPing* in EG4 and Nipponbare

To gain insight into the methylation status of *mPing* in high copy and low copy strains, we used bisulfite genomic sequencing to detail the 5-methylcytosine distribution of *mPing*. Bisulfite treatment of genomic DNA converts unmethylated cytosines to uracils, while 5-methylcytosines remained unchanged. We used degenerate primers within the *mPing* terminal so that the PCR products were a mixture derived from different *mPing* templates. Twenty different clones of both EG4 and Nipponbare were randomly chosen and sequenced and the result can roughly represent the average methylation levels of *mPing* in EG4 and Nipponbare.

Methylation maps of *mPing* in EG4 and Nipponbare are shown in Fig.2A and 2B. Different cytosine types are drawn in different colors and methylation levels are represented by black boxes. Comparison of the two maps indicates that the methylation patterns are consistent at most cytosine sites within *mPing*. In EG4, *mPing* shows a total level of 96.4% CG, 61.6%CNG, and 45.3% CHH methylation (see Methods for an explanation of this calculation). In Nipponbare, the total level is 93.8% CG, 53.5%CNG, and 32.7% CHH methylation (Fig.2C.). Thus, despite an almost 20-fold difference in copy number, *mPing* elements in both EG4 and Nipponbare are heavily methylated with essentially identical methylation patterns.

Methylation pattern of *de novo mPing* insertions in EG4

Our data did not rule out the possibility that EG4 harbored a small number of *mPing*

elements that are hypomethylated and that these elements are able to transpose. A prior study found that there are, on average, 50 new *mPing* insertions per plant per generation. Because these so-called *de novo* insertions contain newly transposed elements, we decided to test this hypothesis by analyzing the methylation patterns of two *de novo mPing* insertions in two progeny of an EG4 plant (GB#6). The 2 *mPing* elements (called NEW1 and NEW3) were inserted into genic regions, with NEW1 in the exon of a gene (GenBank accession AP004048.3) and NEW3 131bp upstream of another gene (GenBank accession AP005579.3). Flanking primer sets were used to amplify the two elements and 20 different clones of each insertion were sequenced.

Methylation maps of NEW1 and NEW3 are shown in Fig.3A and 3B. All the CG sites are methylated at a level of over 90% and 8 out of 9 CNG sites are methylated over 70%. In addition, all the cytosines in the 15bp TIR are densely methylated (over 60%). However, methylation levels vary greatly among CHH sites in each insertion, some of which are 100% methylated while some are not methylated at all. Several CHH sites also show greatly different level between the two new insertions. Overall, the total methylation levels of the two new insertions are similar to the older insertions in EG4, indicating that *de novo* insertions are also heavily methylated (Fig 3C).

Methylation pattern of flanking genic regions of *de novo mPing* insertions

The availability of new *mPing* insertions that are highly methylated allowed us to address the issue of whether methylation spreads from *mPing* inserts by comparing the

methylation status of flanking sequences before and after insertion. To this end we utilized primer sets of one flanking primer and one *mPing*-internal primer, and analyzed the methylation patterns of about 400bp upstream and downstream of NEW1 and NEW3 before (in the parental plant GB#6) and after *mPing* insertion. The results show that the methylation levels of all three cytosine types in these regions are very low, with a total level of 5% and that there is no big difference in methylation patterns between the parent plant and the progeny (Table.3). Thus, newly inserted *mPing* elements are highly methylated and this methylation does not spread to flanking genomic sequences.

Comparison of the shared terminal ends of *Ping* and *mPing*

Our results demonstrate that the difference in *mPing* transposition between EG4 and Nipponbare does not reflect differences in *mPing* methylation. As *Ping* is the likely source of TPase, we turned our attention to the *Ping* elements in EG4 and Nipponbare. Perhaps the autonomous element *Ping*, rather than the MITE *mPing*, is epigenetically regulated and this difference is responsible for the drastic difference in *mPing* copy number between EG4 and Nipponbare and for the continued activity of this family of elements in EG4.

There are 7 *Ping*s in EG4 (EG-1 though EG-7) and one *Ping* in Nipponbare, all of which are identical and all are located in intergenic regions on different chromosomes (Table.4). Bisulfite sequencing was performed on the terminal ends of the 8 *Ping* elements using primer sets that included one flanking primer and one *Ping*-internal primer. The 5' terminal 253bp and 3' terminal 177bp regions were compared with the identical regions in *mPing* (for all except the 3'

end of EG-1 which could not be isolated). The 253bp of the 5' ends show a great difference between *Ping* and *mPing*. In this region, CG methylation of *Ping* copies varies from about 20% to 50%, which is much lower than the over 90% CG methylation of *mPing*. The CNG methylation level is moderate and similar between *mPing* and *Ping*, while the CHH methylation level of some *Ping* copies even reaches 60% or higher, which is strikingly high for asymmetric cytosines (Fig.4A). The methylation level of the 3' ends is similar among all *mPing* and *Ping* copies tested, with over 90% CG methylation and moderate levels of CNG and CHH methylation (Fig.4B.).

When comparing the methylation maps of *mPing* and *Ping* in Nipponbare, we found a region in the 5' end with a distinct methylation pattern (Fig.5A and 5B.). Specifically, the methylation level of cytosines in the underlined region is much lower in *Ping* than that in *mPing*, while other cytosines in the 5' end show similar patterns.

Methylation analysis of *Ping* 5' terminal region

In order to compare the methylation patterns of the 5' ends of each of the 8 *Ping* elements, we employed a forward primer in the unique flanking DNA and a reverse primer inside *Ping* ORF1. A gene model of the regions analyzed is shown in Figure 6. Transcription of a full-length *Ping* mRNA (GenBank accession AK068363) starts at 296bp (Yang et al., 2007) while position 218bp is the presumed start site of the ORF1 promoter (predicted by Neural Network Promoter Prediction with a score of 0.65). For this analysis, the terminal 554bp are divided into three parts: Part I (1-217bp), Part II (218-295bp) and Part III (296-554bp). Pink

regions in Figure 6 are shared with *mPing*. Part II is the putative promoter of *Ping* ORF1, and it can be further divided into IIa (218-253bp), which is shared with *mPing*, and IIb (254-295bp), which is not shared with *mPing*. Part III contains the start of the *Ping* transcribed region, and includes part of ORF1 (454-554bp).

Methylation of each part was analyzed in detail and compared among *Ping* elements. Like *mPing*, Part I is heavily methylated for all three types of methylation. The level of EG-4 is the highest at 80%, while the levels of the other copies vary from 50% to 70% (Fig.7A). Part II shows a low but variable methylation level among the 8 *Ping* copies (Fig.7B). When we looked at IIa and IIb separately, we found that the difference in IIa is more dramatic. The methylation level of IIa in EG-4 and EG-5 is the highest, with about 20% symmetric methylation and over 30% asymmetric methylation. EG-7 is the least methylated among all the *Ping*s in region IIa at less than 5%. The methylation levels of other 4 *Ping*s also vary within a range from about 5% to 15% (Fig.7C). The methylation level of Part IIb is as low as less than 7% for all the *Ping*s (Fig.7D). The methylation map of Part II shows that the methylation level of each cytosine site is consistent with the overall level of each *Ping* copy (Fig.8). For part III, which contains the 5' end of *Ping* mRNA, all 8 *Ping* copies show a low methylation level of 4%, regardless of the methylation type (Fig.7E). Thus, the methylation level of *Ping* ORF1 promoter is more variable than that of the other regions among the 8 *Ping*s.

Methylation pattern of *Ping* flanking regions

5' and 3' flanking sequences of some *Ping* copies were successfully amplified and

analyzed (Fig.9). Same as sequences flanking *mPing* insertion sites, most cytosines in the flanking regions of *Ping* show a very low methylation level, although some cytosines located very close to Ping ends show a higher methylation level. Several CG sites in the 3' flanking region of EG-7 are methylated as high as over 90%, while methylation levels of the CNG and CHH sites in this region is very low at around 10%.

DISCUSSION

DNA methylation of *mPing* does not prevent its transposition

Our data show that average methylation level of *mPing* is very similar in EG4 and Nipponbare, and that the methylation distribution is consistent between the two strains. However, it is possible that only a few copies in EG4 are less methylated and capable of moving. Therefore, two newly moved *mPing* elements, NEW1 and NEW3 were chosen for methylation analysis. Their methylation patterns were found to be similar to the average level of the older insertions. All of the symmetric cytosine sites display a similar densely methylated pattern between the two *de novo* insertions, while the levels of a small proportion of asymmetric sites vary greatly (Fig.3).

Previous studies showed that CG methylation is maintained by MET1 (Tran et al., 2005a; Zhang, 2008) and CNG methylation is maintained by plant specific CMT3 (Jackson et al., 2002; Jackson et al., 2004; Zhang, 2008). Methylation on asymmetric cytosines (CHH) cannot be inherited and thus needs to be established *de novo*. It is suggested that TEs at different sites in the genome produce transcripts from nearby promoters in both directions, which form dsRNA and promote DNA methylation of the element (Mette et al., 2000; Bender, 2004). Since methylation on asymmetric sites has to be reestablished under the direction of dsRNA after each cell division, it is reasonable to see such a variation among asymmetric sites of individual *mPing* copies and this variation does not account for the difference in transposition. Therefore, we conclude that the

methylation patterns of *mPing* are essentially identical in high copy and low copy strains and methylation of *mPing* does not prevent its transposition.

Since the transposition of DNA transposons requires TPase binding to transposon ends (Craig et al., 2002), we believe that *mPing* also recruits TPase to its ends for transposition. Although no biochemical assay has been performed on *mPing*/TPase interaction, a recent work on the transposition of a reconstructed *PIF/Harbinger* element in Zebra fish (Sinzelle et al., 2008) provided an attractive model that may also apply to *mPing* since the two systems share great similarities. The two ORFs of *Harbinger3_DR* encode two proteins, the TPase and a Myb-like protein of unknown function. *Harbinger3N_DR* is a non-autonomous short element derived from *Harbinger3_DR* and it has about 1,000 copies in the genome. Biochemical assay showed that the Myb-like protein binds to six sites in the transposon subterminal regions and recruits the TPase to transposon ends by protein-protein interaction. Based on this model, we assume that both *mPing* TIRs and internal regions are involved in TPase complex binding.

Since the TIRs of *de novo mPing*s are densely methylated, and most symmetric sites in *mPing* internal part also show a very high methylation level (Fig.3), we suspect that DNA methylation may not prevent the binding of *Ping* TPase complex into *mPing* ends, which is contrary to the situations on some other DNA TEs. Demethylation at transposon ends where DNA/TPase interactions occur is required for the transposition of *CACTA* elements (Kato et al., 2003; Bender, 2004). *Ac* TPase has a methylation-dependent binding capacity for *Ds* ends. Holo-methylated *Ds* has a very low TPase binding affinity and is unable to be excised (Wang et al., 1996; Ros and Kunze, 2001). A similar binding affinity was found for the *Spm*-encoded

protein TnpA, which binds more efficiently to hemi- or unmethylated *Spm* termini than fully methylated ones (Cui and Fedoroff, 2002). Therefore we conclude that *mPing* transposition is regulated differently from these elements and DNA methylation may not prevent the binding of *Ping* TPase complex, which is a reasonable explanation to the situation that both old and *de novo* *mPing* copies are densely methylated.

mPing does not alter the low methylation pattern of flanking genic regions

It was shown previously that *mPing* prefers to integrate into genic regions (Naito et al., 2006; Yang et al., 2007). The two *de novo* insertions we studied were located in or very close to genes. Our data showed that the methylation levels of their flanking genic regions are as low as about 5% (Table.3). We also checked the flanking sequences of the empty allele in the parental plant, and no change of the methylation patterns was found before and after insertion. Therefore, we conclude that *mPing* integration does not alter the low methylation level of the flanking genic regions upon insertion.

Previous studies showed that some SINEs are capable of spreading DNA methylation into hypomethylated flanking sequences. Human *Alu* elements were proposed to promote *de novo* methylation in the juxtaposed CG islands and cause the silencing of tumor suppressor genes in neoplasia (Graff et al., 1997). The plant SINE S1 elements in *Brassica* have been shown to preferentially integrate into hypomethylated regions where they induce *de novo* methylation of upstream and downstream flanking sequences (Arnaud et al., 2000). Although MITEs and SINEs share some features including small size and a lack of coding capacity, they are in different TE

classes that utilize distinct transposition mechanisms. One possible explanation for this difference is that retrotransposons transpose via an RNA intermediate, so transcription of the element is required. Transcripts may form siRNAs by folding, or form dsRNAs between complementary transcripts derived from different strands. These dsRNAs could induce the RNA-dependent DNA methylation (RdDM) along the element into flanking DNA. Because MITEs utilize a “cut and paste” mechanism, which does not require transcription of the elements, fewer transcripts may be produced, which reduces the efficiency of the RdDM pathway, and may result in the differential spreading of methylation between SINEs and MITEs. An analysis of the flanking regions of several TE families in rice also showed that surrounding methylation level of retrotransposons is higher than some DNA TEs that don’t have well-matched transcripts in the database (Takata et al., 2007), which agrees with our explanation.

Ping **activity is not associated with its flanking methylation pattern**

As discussed above, the methylation pattern of *mPing* does not correlate with its ability to transpose. Therefore, the autonomous element *Ping*, which encodes the proteins required for *mPing* transposition, is more likely to be the key factor in *mPing* movement.

One conclusion we can draw from the methylation analysis of *Ping* is that DNA methylation is restricted inside the elements and does not spread into the flanking sequences. Our study shows that the 7 *Ping*s in EG4 and 1 *Ping* in Nipponbare are located in intergenic regions on different chromosomes. They are several kbs far away from the nearest genes both upstream and downstream (except EG-7 5’end) (Table.4). The methylation patterns of some *Ping* flanking

regions were analyzed and are shown in Fig.9. We can see that the methylation levels of most *Ping* flanking regions are low, in sharp contrast to the heavy methylation of *Ping* terminal ends. For example, most cytosines in the 73bp 5' flanking region and 60bp 3' flanking region of *Ping* EG-6 show a methylation level lower than 10% (Fig.9), while the level of its 217bp 5' terminal and 177bp 3' terminal regions are around 70% (Fig.7A) and 80% (Fig.4B) respectively. This is consistent with what we found about *mPing*, suggesting that DNA methylation of *mPing* and *Ping* does not spread into the flanking sequences.

Another conclusion we draw here is that the methylation pattern of *Ping* flanking sequences is not likely to affect *Ping* activity. The methylation level of *Ping* EG-7, especially in its ORF1 promoter region, is the lowest among all the *Ping*s (Fig.7 and 8), while the methylation level of its flanking regions is higher than that of other *Ping*s (Fig.9), suggesting that the methylation pattern of the flanking regions does not correlate with the methylation pattern of *Ping* itself. On the other hand, the flanking regions of NH, the only *Ping* in Nipponbare, showed the lowest methylation level both upstream and downstream. The inactivity of *mPing* and *Ping* in Nipponbare suggests that lack of methylation in *Ping* flanking regions does not promote *Ping* activity.

Ping ORF1 promoter region displays a low and variable methylation pattern

The 5' end 554bp sequences of 8 *Ping*s in EG4 and Nipponbare were analyzed in detail. We divided it into three parts (Fig.6) and each part showed differential methylation patterns. Part I is the terminal end shared with *mPing*, and it shows a high methylation level for all *Ping*s

ranging from 50% to 80% (Fig.7A). Part III is where the transcription starts and it includes part of *Ping* ORF1 sequence. This part of all *Ping*s is almost unmethylated in the majority of the cells since the methylation level of this part is as low as about 3% (Fig.7E). Part II, the putative ORF1 promoter region, is suspected to be the most interesting part. Although the function of the Myb-like protein encoded by *Ping* ORF1 is still unknown, both ORF1 and ORF2 are required for *mPing* transposition (Yang et al., 2007). The model of *Harbinger3_DR* in zebra fish illustrated that the Myb-like protein binds to the transposon and then recruits TPase to transposon ends (Sinzelle et al., 2008). Because of the great similarity between the two elements, we assume that *Ping* ORF1 product also anchors the TPase to transposon ends by protein-protein interactions.

Previous study of the maize *Spm* element showed that methylation of its promoter and GC-rich downstream control region (DCR) correlates with the inactivation of TPase transcription and TE transposition (McClintock, 1958, 1961; Banks et al., 1988; Banks and Fedoroff, 1989; Fedoroff, 1989; Schlappi et al., 1993; Schlappi et al., 1994). Reactivation of *Ac* is also associated with the decrease of methylation near the transcription start site (Brutnell and Dellaporta, 1994). Since *Ping* TPase is very likely to function in a complex interacting with the Myb-like protein, the transcription of both ORF1 and ORF2 is important for *mPing* transposition. Therefore, the methylation pattern of *Ping* ORF1 promoter may correlate with transposition of elements in the family and, as such, is worthy of future experimental analysis.

The methylation level of ORF1 promoter is low and variable (Fig.7B). The beginning part (218-253bp) shows a more dramatic variation among the 8 *Ping*s, ranging from 3% in EG-7 to about 30% in EG-4 and 5 (Fig.7C). This variation prompts us to suspect that *Ping*

transcription is also regulated epigenetically. Because *Ping* is transcribed in both EG4 and Nipponbare (Naito, PhD thesis, 2007), and we cannot distinguish the transcripts from different *Ping*s in EG4, it is impossible to tell whether a particular *Ping* is more transcriptionally active or whether the activity correlates with methylation level. Ongoing rice RIL analysis is designed to solve problem. The RILs were originally generated from the hybrid of EG4 and Nipponbare and then selfed for 10 generations. Individual F10 plants carry different numbers of *Ping* elements, which are almost certainly homozogous in the genome. Our expectation is that by profiling *Ping* transcription and *mPing* transposition in individual RILs, one should be able to correlate *mPing* movement with the presence of one or more *Ping* copy. Comparisons between active and inactive RIL lines should reveal whether *Ping* methylation levels correlate inversely with *Ping* transcription.

Unfortunately we cannot rule out other possibilities or draw additional conclusion until the RILs are analyzed. For example, it is possible that *mPing* activity cannot be correlated to certain functional *Ping*s by RIL profiling. Instead, all 8 *Ping*s may be expressed equally and *mPing* activity is dosage-dependent. When the amount of TPase reaches a threshold, *mPing* movement is induced.

CONCLUSION

mPing is an active MITE discovered in rice. The copy number of *mPing* varies dramatically among temperate *sativa*s and its activity is also distinct between high-copy and low-copy strains. Therefore, it has been of interest to determine how *mPing* activity is differentially regulated in different strains. Epigenetic mechanisms are likely to be involved in this process since *mPing* can be transpositionally activated in Nipponbare anther culture while this transposition is normally repressed. In this study, DNA methylation patterns of *mPing* and its autonomous ancestor *Ping* were analyzed and compared in high and low copy rice strains. The results showed that *mPing* is always heavily methylated overall in both high- and low-copy strains, suggesting that the distinct activities are not associated with differential *mPing* methylation. De novo insertions of *mPing* are also heavily methylated in both their TIRs and internal sequences, suggesting that DNA methylation does not prevent *mPing* transposition, possibly because TPase complex binding is not prevented by methylation of *mPing*. In addition, methylation does not spread significantly from *mPing* and *Ping* into flanking regions in the genome, unlike the situation with SINEs in plants and animals. The *mPing*-shared terminal ends of *Ping* are also highly methylated, while the *Ping* ORF1 promoter regions show a low and variable pattern among different *Ping*s. This observation suggests an attractive model that DNA methylation in the promoter sequences affects element transcription, and that the transcription level determines *mPing* activity. More experiments are being performed to test this model and, hopefully, unravel the mystery of *mPing* transposition in rice.

Table.1 Degenerate primers for amplification of *mPing* used in this study

mPing	general[a]	F	5’-GGYYAGTYAYAATGGGGGT-3’
		R	5’-RRCCARTCACAATRRCTA-3’
NEW1[b]	element[d]	F	5’-TGYAYATGTAGTGGTGGT-3’
		R	5’-ARRCCAAATCAAACAACT-3’
	5'-flanking[e]	F	5’-AAGAAAGTYGAYYTYYTYG-3’
		R	5’-TTRRRTAGCCRTRCAATRA-3’
	3'-flanking[f]	F	5’-GYATGAYAYAYYAGTGAAA-3’
		R	5’-TCRARTRRATRTCCATTTR-3’
	Parent[g]	F	5’-AAGAAAGTYGAYYTYYTYG-3’
		R	5’-TCRARTRRATRTCCATTTR-3’
NEW3[c]	element	F	5’-AGAGYGYAGAAGYAGAAT-3’
		R	5’-ARTRAACCARCCGTAATC-3’
	5'-flanking	F	5’-TGTATAGGGTATATYYGT-3’
		R	5’-TTTRCTTARTCTTRRAAAC-3’
	3'-flanking	F	5’-GYATGAYAYAYYAGTGAAA-3’
		R	5’-CTTRACACACARACCAATRAT-3’
	parent	F	5’-TGTATAGGGTATATYYGT-3’
		R	5’-CTTRACACACARACCAATRAT-3’

a. *mPing* general primers are designed inside *mPing* TIRs in order to amplify from any *mPing* copies in the genome.
b. NEW1 is a *de novo mPing* insertion in the plant GB#6-5.
c. NEW3 is a *de novo mPing* insertion in the plant GB#6-2. GB#6-5 and GB#6-2 are two siblings derived from GB#6, a Gimbozu EG4 plant.
d. The element primer set is designed from flanking regions of this element in order to amplify the specific *mPing* copy.
e. The 5’ flanking primer set contains one forward primer in the 5’ flanking region of the element and one *mPing*-internal reverse primer to amplify the 5’ flanking region of the specific element.
f. The 3’ flanking primer set contains one reverse primer in the 3’ flanking region of the element and one *mPing*-internal forward primer to amplify the 3’ flanking region of the specific element.
g. The parent primer set is designed to amplify the corresponding flanking region in GB#6 of the *mPing* de novo insertion in GB#6-5.

Table.2. Degenerate primers for amplification of *Ping* terminal sequences used in this study

NH	5'-end[a]	F	5’-AAGAYAGGGATAAAAAAAYTTTGTTG-3’
		R	5’-CCTCTTCAATCCTARTTRCTTRCCC-3’
	3'-end[b]	F	5’-GTTAYGGYTGGAAGTTGTGTAAGTTTG-3’
		R	5’-CAACCTAAATTAACCATACTATCCCT-3’
EG-1	5'-end	F	5’-AAATTTTTGTTAGGCCAGTCACAATGGA-3’
		R	5’-CCTCTTCAATCCTARTTRCTTRCCC-3’
	3'-end	F	5’-GTTAYGGYTGGAAGTTGTGTAAGTTTG-3’
		R	5’-ATTTTTCACTCRCRCRCTRCTAAR-3’
EG-2	5'-end	F	5’-ATTGTGGAGAGAGGAGYTTYAYG-3’
		R	5’-CCTCTTCAATCCTARTTRCTTRCCC-3’
	3'-end	F	5’-GTTAYGGYTGGAAGTTGTGTAAGTTTG-3’
		R	5’-AAAARCCACTCTCTCTCTACTCCCC-3’
EG-3	5'-end	F	5’-TTYGTTATTGTAYYATTATTAGGYYAGT-3’
		R	5’-CCTCTTCAATCCTARTTRCTTRCCC-3’
	3'-end	F	5’-GTTAYGGYTGGAAGTTGTGTAAGTTTG-3’
		R	5’-TARTTCTTTTTCTAARRCCARTCACAA-3’
EG-4	5'-end	F	5’-GTAGAGAAAAAGGGAYYTAATYGAGG-3’
		R	5’-CCTCTTCAATCCTARTTRCTTRCCC-3’
	3'-end	F	5’-GTTAYGGYTGGAAGTTGTGTAAGTTTG-3’
		R	5’-TTCTAAAATCTACTAACATRACAACTCARAAA-3’
EG-5	5'-end	F	5’-AAAGGAGAGAGAAAAAAAYAAATAATAATTGG-3’
		R	5’-CCTCTTCAATCCTARTTRCTTRCCC-3’
	3'-end	F	5’-GTTAYGGYTGGAAGTTGTGTAAGTTTG-3’
		R	5’-TTCTTCTTCTTCTTCTTCTTCTACTCTCRC-3’
EG-6	5'-end	F	5’-TTGATGYATGYAATTGAATGGAGTG-3’
		R	5’-CCTCTTCAATCCTARTTRCTTRCCC-3’
	3'-end	F	5’-GTTAYGGYTGGAAGTTGTGTAAGTTTG-3’
		R	5’-AAAAATCTRATCCTTTCACTCTCRATC-3’
EG-7	5'-end	F	5’-GAAAAGGTGAATTGGATAYATGTYATTA-3’
		R	5’-CCTCTTCAATCCTARTTRCTTRCCC-3’
	3'-end	F	5’-GTTAYGGYTGGAAGTTGTGTAAGTTTG-3’
		R	5’-CTAARCAAGCCCRATCTACCATCA-3’

a. The 5’-end primer set contains one forward primer from the 5’ flanking region of the element and one *Ping*-internal reverse primer to amplify the *Ping* 5’ end 454bp region.
b. The 3’-end primer set contains one reverse primer from the 3’ flanking region of the element and one *Ping*-internal forward primer to amplify the *Ping* 3’ end 195bp region.

Table.3.
Methylation levels of the flanking regions of *de novo mPing* insertions in EG4.

		CG[a]	CNG[b]	CHH[c]
NEW1	GB#6	7.6%	2.7%	4.2%
	GB#6-5	5.3%	4.2%	4.7%
NEW3	GB#6	7.1%	8.3%	5.0%
	GB#6-2	2.5%	5.0%	5.9%

a. The total level of all the cytosines in CG context in the sequence of interest.
b. The total level of all the cytosines in CNG context in the sequence of interest.
c. The total level of all the cytosines in CHH context in the sequence of interest.

Table.4. Location information of the 8 *Ping* copies in EG4 and Nipponbare.

	chromosome	location[a]	insertion strand	feature	distance[b]
EG-1	1	38085880bp	plus	intergenic	7kb
EG-2	1	4219010bp	plus	intergenic	3kb
EG-3	3	27819764bp	minus	intergenic	2kb
EG-4	7	26406727bp	minus	intergenic	3kb
EG-5	9	13681490bp	minus	intergenic	>20kb
EG-6	9	16635961bp	plus	intergenic	14.5kb
EG-7	9	10810267bp	plus	intergenic	0.46kb
NH	6	23137988bp	plus	intergenic	>10kb

a. Location of each *Ping* copy is based on NCBI RAP3 Build 3 and the numbers are the insertion sites of *Ping* in corresponding chromosomes.
b. Distance from the 5' end of *Ping* to the nearest upstream gene

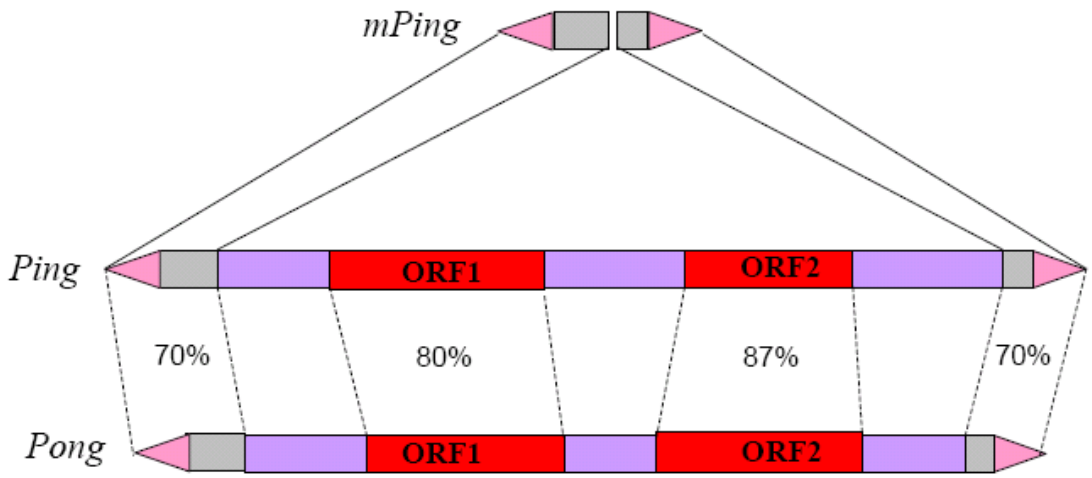
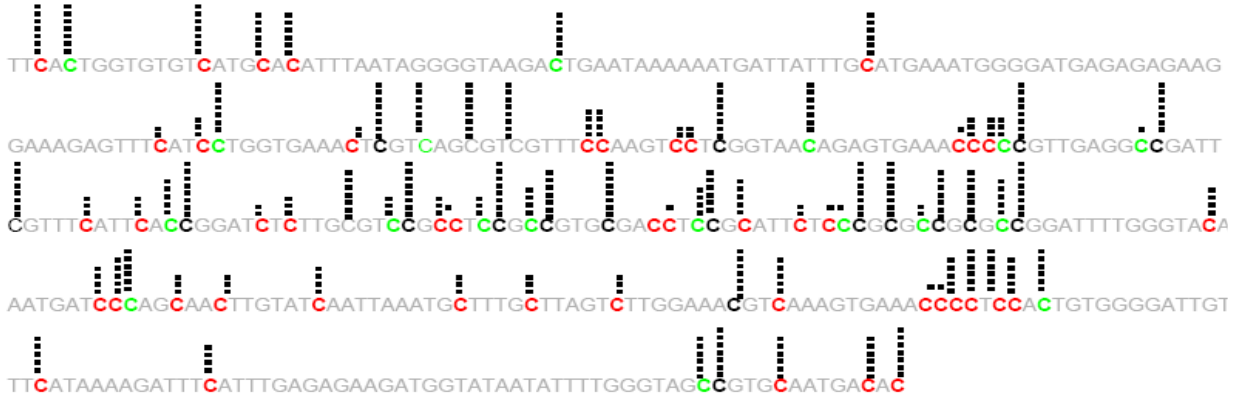
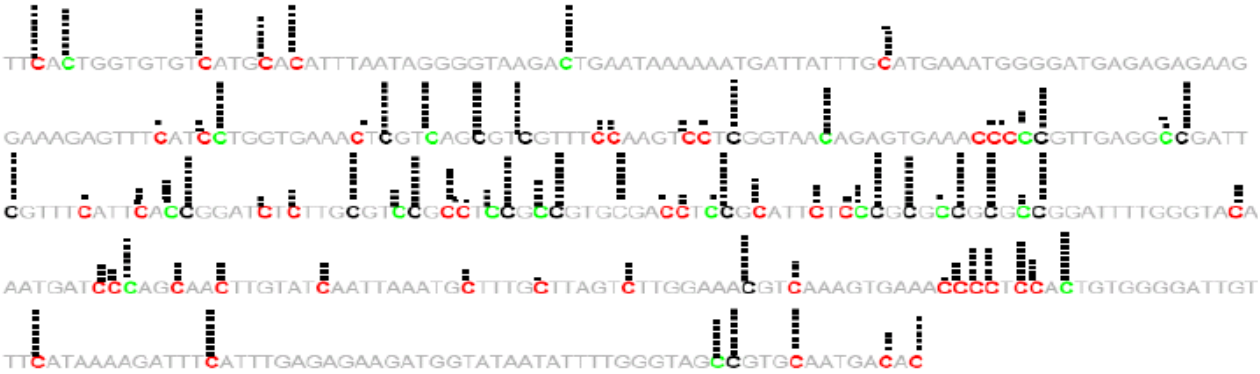


Figure.1. Structures of *mPing*, *Ping* and *Pong*. Pink triangles indicate TIRs. Red rectangles indicate ORFs. Solid lines indicate the shared terminal regions of *mPing* and *Ping*. Dotted lines and percentages indicate the similarity of the terminal regions and two ORFs between *Ping* and *Pong* respectively.

A



B



C

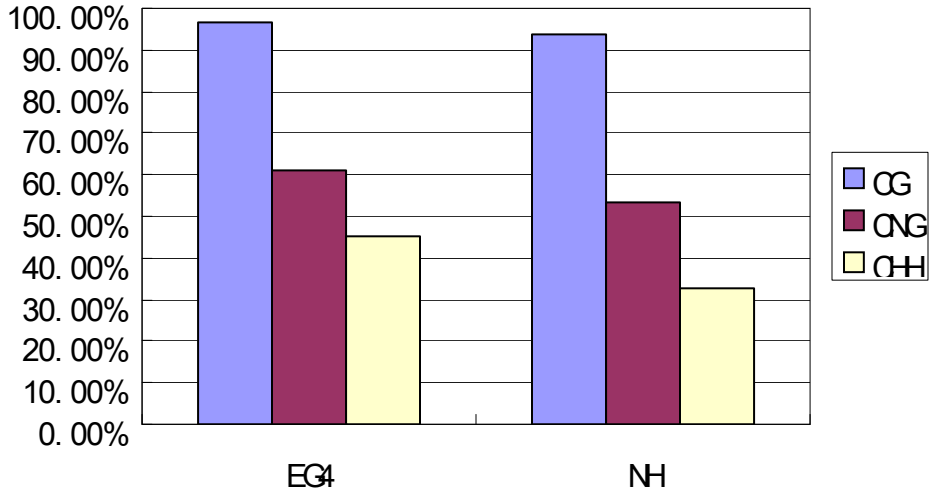
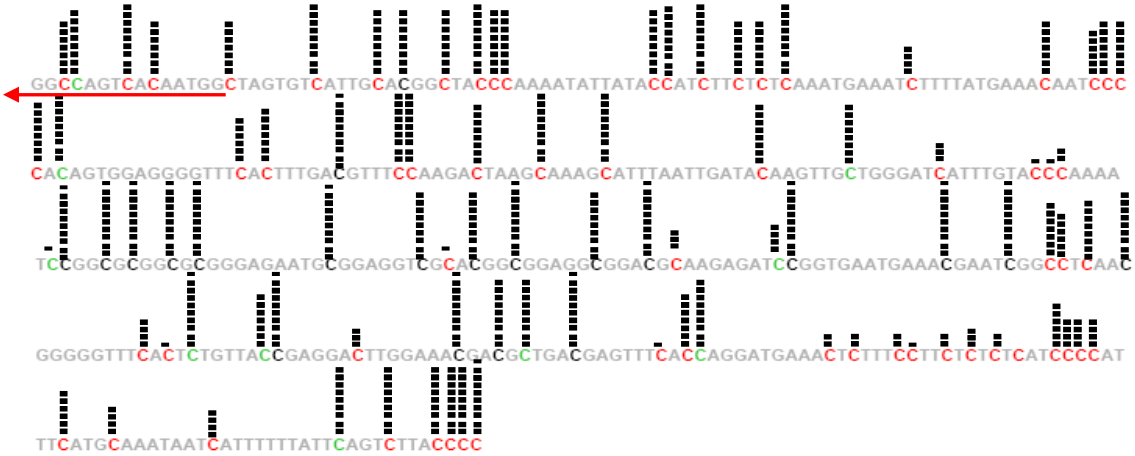
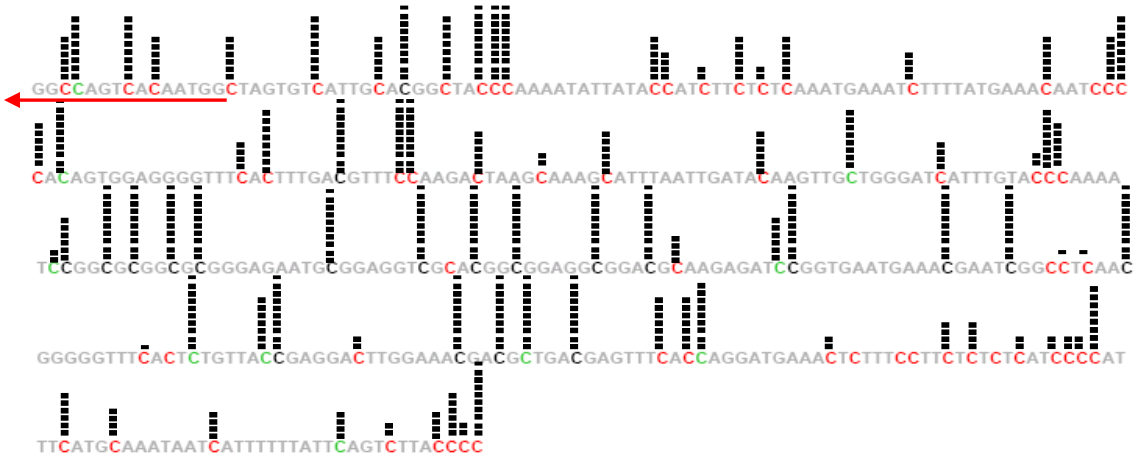

100. 00%
90. 00%
80. 00%
70. 00%
60. 00%
50. 00%
40. 00%
30. 00%
20. 00%
10. 00%
0. 00%
EG4
NH
CG
CNG
CHH

Fig.2. Methylation patterns of *mPing* insertions in EG4 and Nipponbare (NH). A. Methylation map of “average” *mPing* in EG4. C. Methylation map of “average” *mPing* in Nipponbare. Cytosines in CG, CNG and CHH contexts are drawn in black, green and red respectively. Methylation level is represented by black boxes and each box represents 10% of methylation. C. Total methylation levels of each cytosine type in EG4 and Nipponbare.

A



B



C

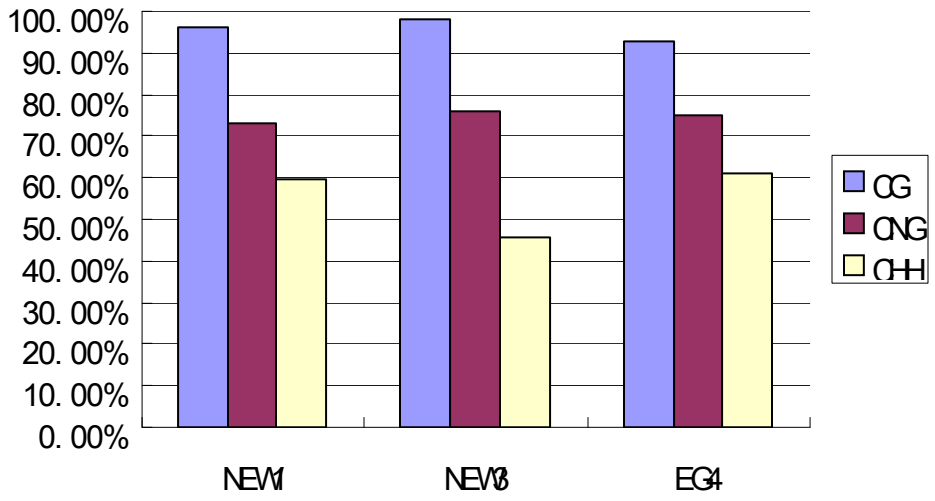
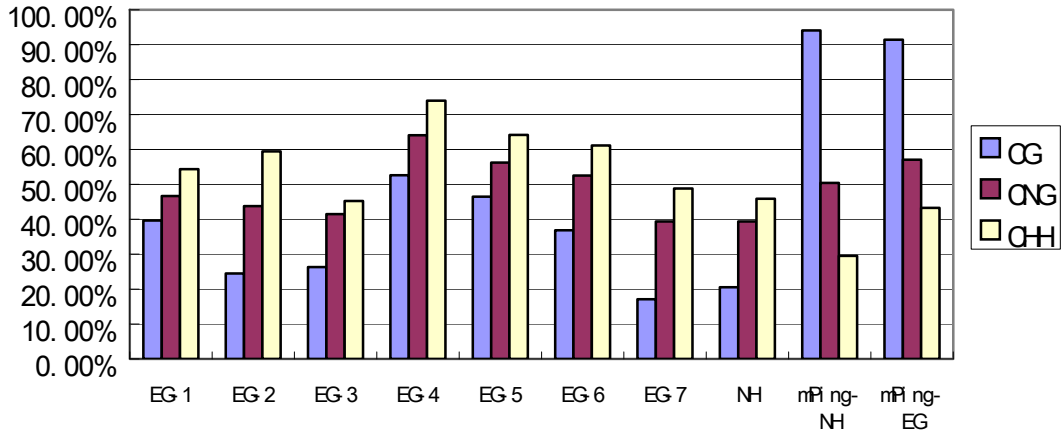


Fig.3. Methylation patterns of *de novo mPing* insertions in EG4. A. Methylation map of *mPing* NEW1 insertion (minus strand). B. Methylation map of *mPing* NEW3 insertion (minus strand). Red arrows in A and B represent 15bp 5' TIR sequence. (See Fig.2 legend for the representation of methylation level in the map). C. Total methylation levels of each cytosine context of the two *de novo* insertions and *mPing* average (minus strand) in EG4.

A



B

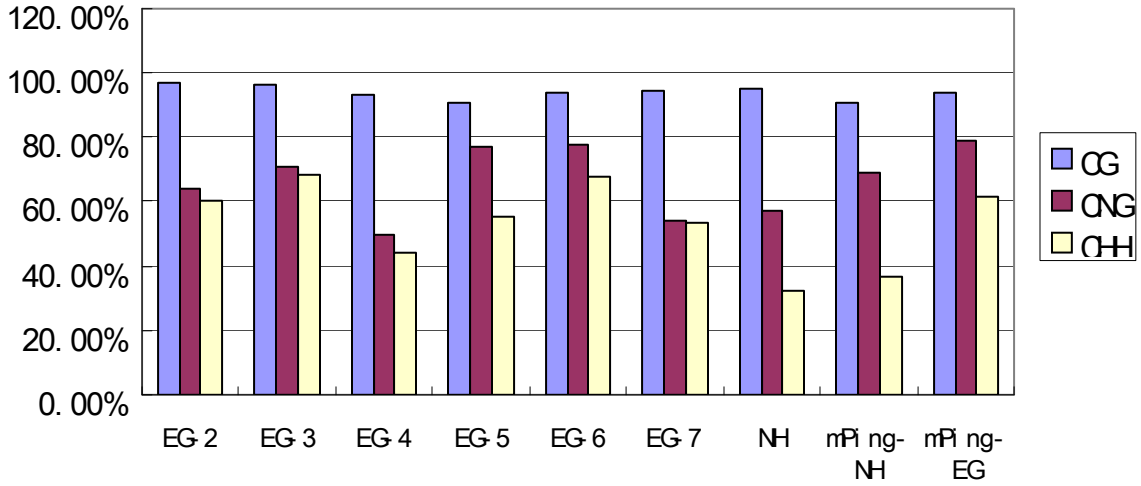
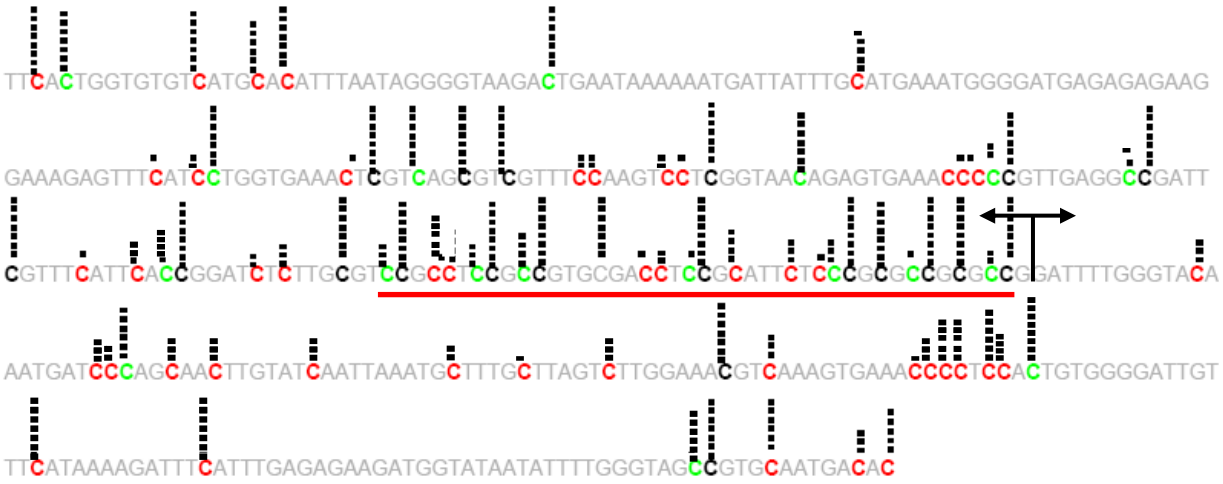


Fig.4. Comparison of the shared terminal ends of *mPing/Ping* in EG4 and Nipponbare. The 7 *Ping* copies in EG4 are represented by EG-1 to EG-7. NH stands for the *Ping* in Nipponbare. *mPing*-NH and *mPing*-EG are the average levels of *mPing* in Nipponbare and EG4, respectively. A. Methylation level of 5' end 253bp region. B. Methylation level of 3' end 177bp region (No data for EG-1).

A



B

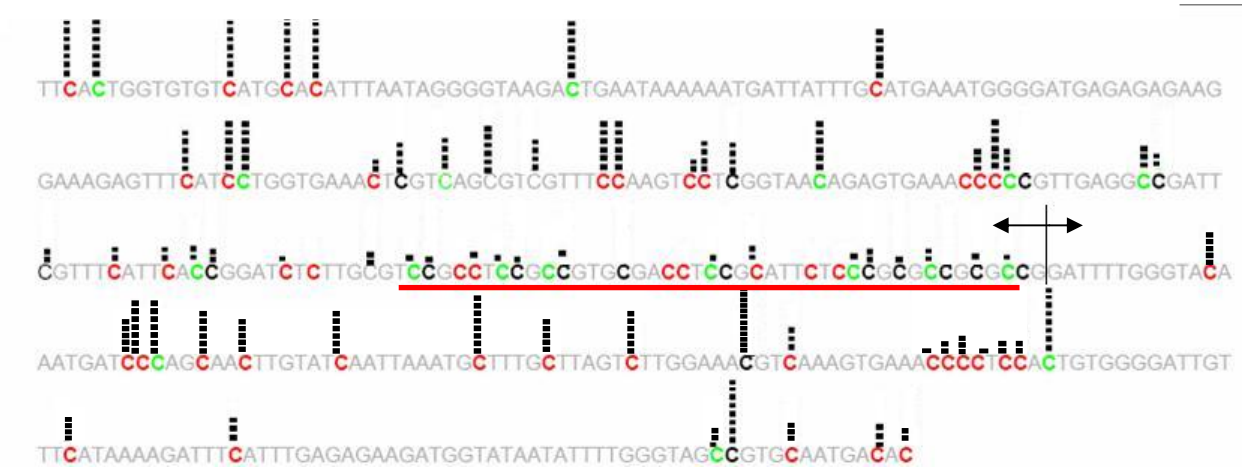


Fig.5. Methylation of *Ping* and *mPing* in Nipponbare. A. Methylation map of "average" *mPing* in Nipponbare. B. Methylation map of *Ping* in Nipponbare including the shared sequences with *mPing*. The arrows in A and B indicate the break point of two shared terminal ends and the internal part of *Ping* is omitted here. The red underline indicates the region where there is a significant difference in methylation patterns

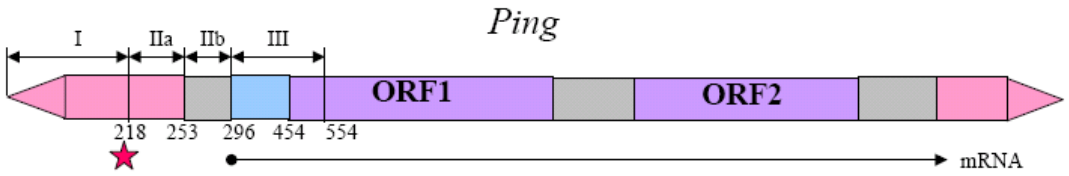
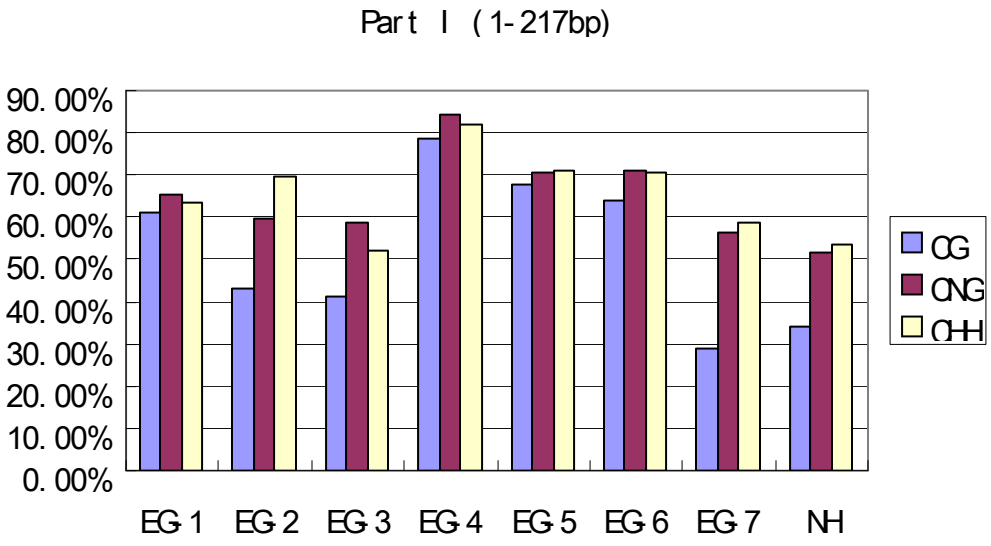
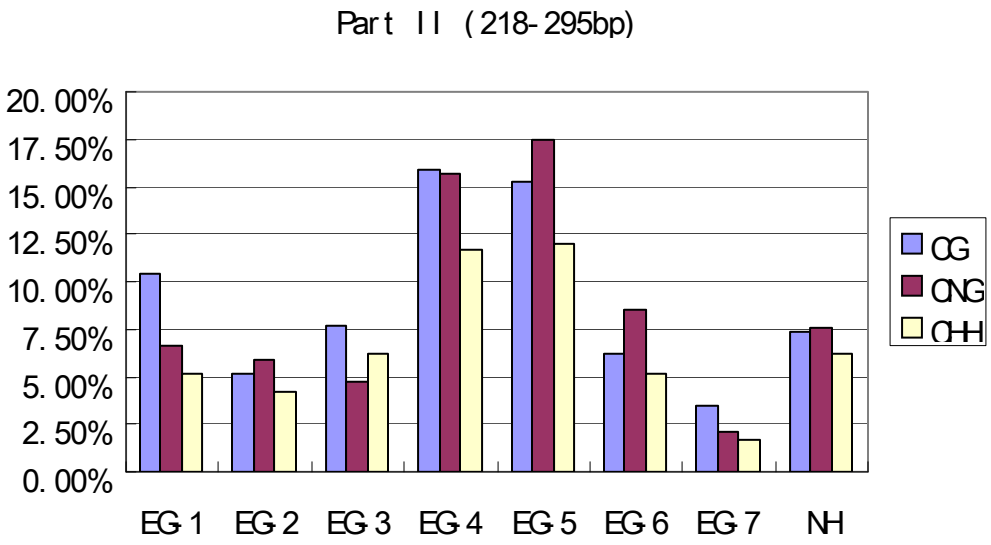


Figure. 6. Gene model of *Ping* 5’ end. The 5’ end 554bp region is divided into three parts as indicated by roman numbers: Part I 1-217bp, Part II 218-295bp (IIa 218-253bp and IIb 254-295bp), Part III 296-554bp. The pink regions are the terminal regions shared with *mPing*. The violet regions are the two ORFs. The red star represents the start of the predicted ORF1 promoter. The black dot indicates the transcription start site which is supported by GenBank accession AK068363.

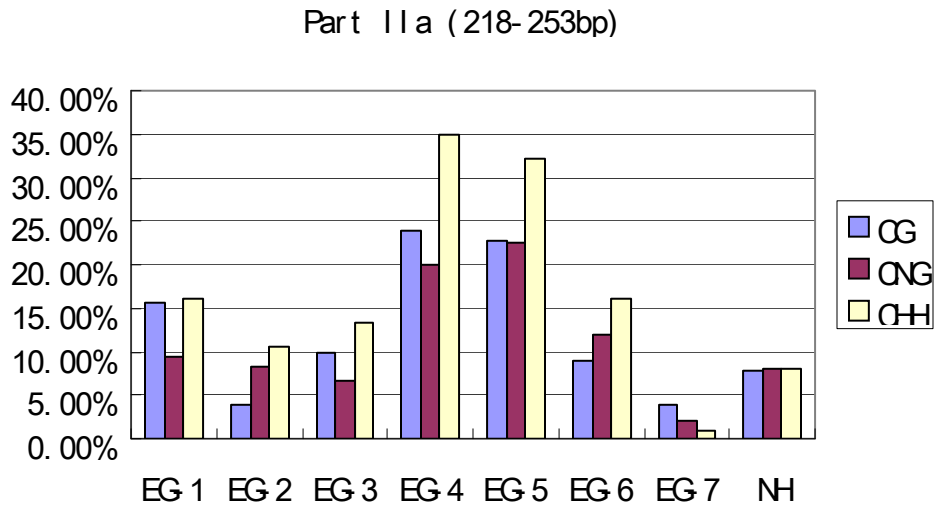
A


Part I (1-217bp)
90.00%
80.00%
70.00%
60.00%
50.00%
40.00%
30.00%
20.00%
10.00%
0.00%
EG-1
EG-2
EG-3
EG-4
EG-5
EG-6
EG-7
NH
CG
CNG
CHH

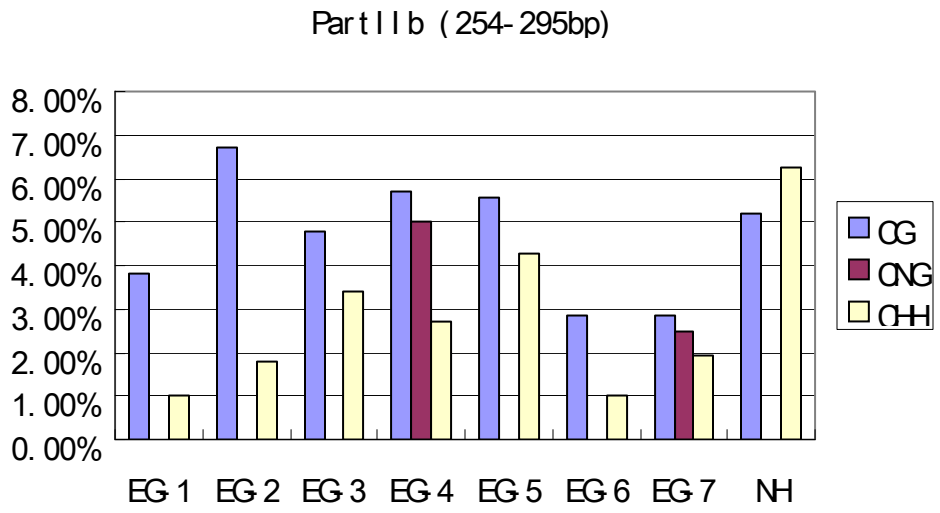
B


Part II (218-295bp)
20.00%
17.50%
15.00%
12.50%
10.00%
7.50%
5.00%
2.50%
0.00%
EG-1
EG-2
EG-3
EG-4
EG-5
EG-6
EG-7
NH
CG
CNG
CHH

C


Part IIa (218-253bp)
40.00%
35.00%
30.00%
25.00%
20.00%
15.00%
10.00%
5.00%
0.00%
EG-1
EG-2
EG-3
EG-4
EG-5
EG-6
EG-7
NH
CG
CNG
CHH

D


Part IIb (254-295bp)
8.00%
7.00%
6.00%
5.00%
4.00%
3.00%
2.00%
1.00%
0.00%
EG-1
EG-2
EG-3
EG-4
EG-5
EG-6
EG-7
NH
CG
CNG
CHH

E

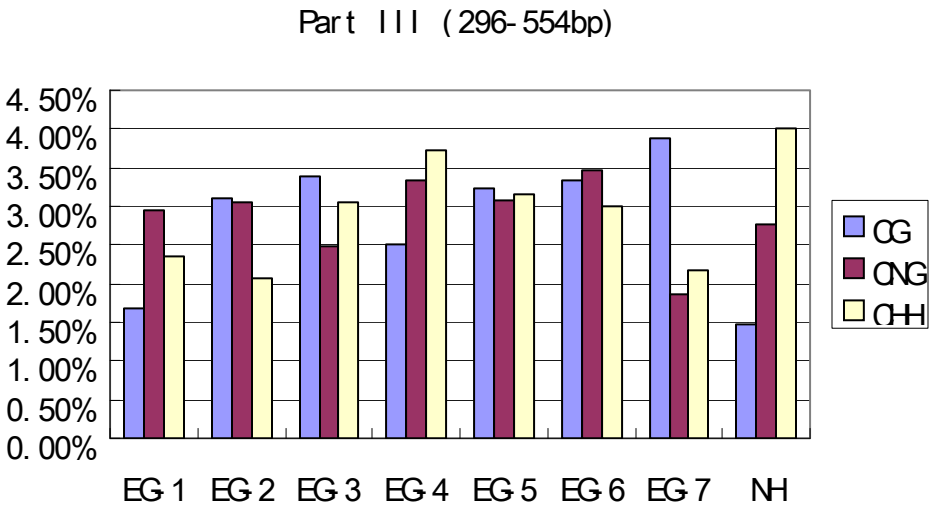


Figure. 7. Methylation analysis of *Ping* 5' terminal regions. A. Methylation level of Part I (1-217bp), the region before the predicted ORF1 promoter. B. Methylation level of Part II (218-295bp), the region starting from the predicted ORF1 promoter and ending before the transcription start site. C. Methylation level of Part IIa (218-253bp), which is shared with *mPing*. D. Methylation level of Part IIb (254-295bp), which is not shared with *mPing*. E. Methylation level of Part III (296-554bp), the start of ORF1 transcript.

Figure.8. Methylation map of *Ping* promoter regions. The upper 8 rows are *Ping* copies and the last two rows are average *mPing* elements in EG4 and Nipponbare. The vertical line indicates the break point of Part IIa (218-253bp, *mPing*-shared) and IIb (254-295bp) (see Figure 6).

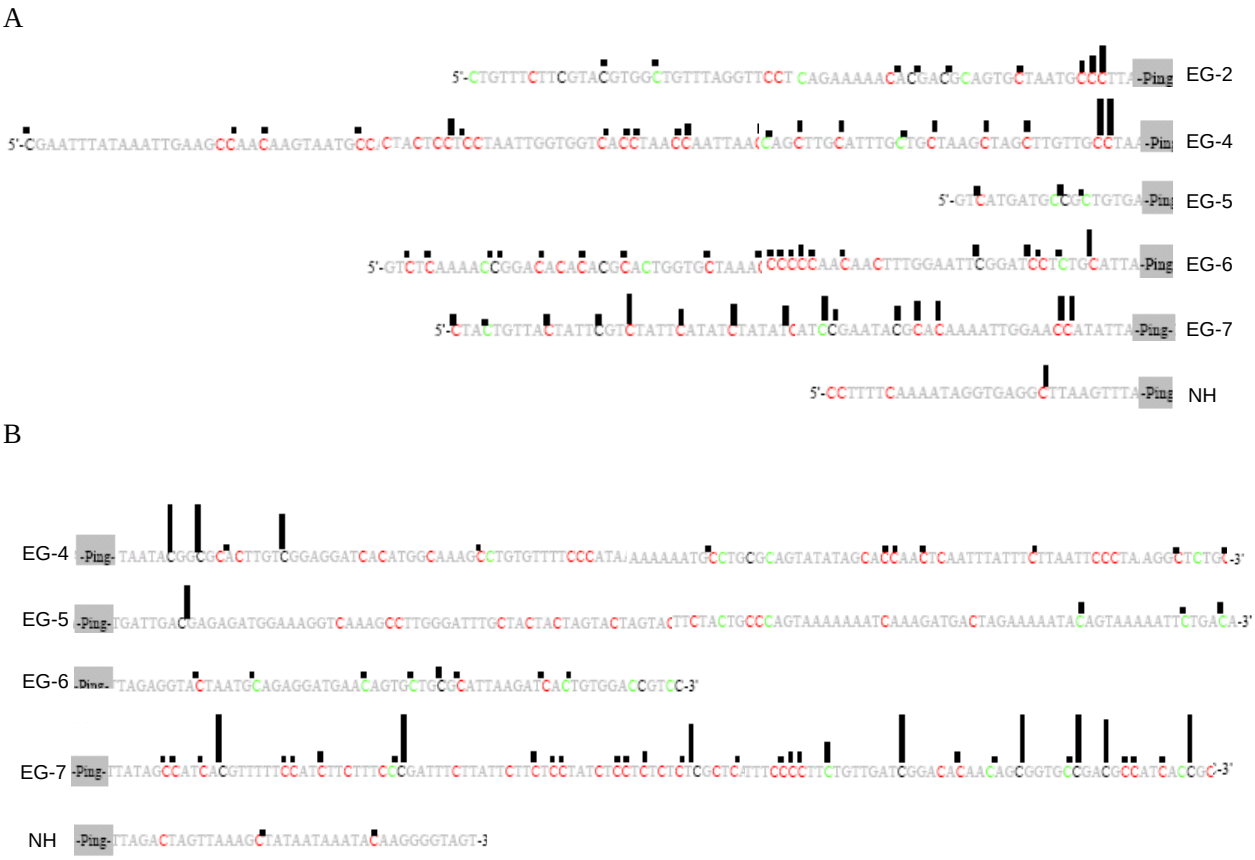


Fig.9. Methylation patterns of *Ping* flanking regions. A. Methylation map of 5' flanking regions. B. Methylation map of 3' flanking regions. Gray boxes indicate the positions of *Ping*. Black bars represent the methylation level of each cytosine and the height is proportional to methylation level.

REFERENCES

Arnaud, P., Goubely, C., Pelissier, T., and Deragon, J.M. (2000). SINE retroposons can be used in vivo as nucleation centers for de novo methylation. Mol Cell Biol **20,** 3434-3441.

Banks, J.A., and Fedoroff, N. (1989). Patterns of Developmental and Heritable Change in Methylation of the Suppressor-Mutator Transposable Element. Develop Genetics **10,** 425-437.

Banks, J.A., Masson, P., and Fedoroff, N. (1988). Molecular mechanisms in the developmental regulation of the maize Suppressor-mutator transposable element. Genes Dev **2,** 1364-1380.

Bender, J. (2004). DNA methylation and epigenetics. Annu Rev Plant Biol **55,** 41-68.

Bennetzen, J.L. (2000). Transposable element contributions to plant gene and genome evolution. Plant Mol Biol **42,** 251-269.

Bingham, P.M., and Zachar, Z. (1989). Retrotransposons and the FB transposon from *Drosophila melanogaster*. In Mobile DNA, D.E. Berg and M.M. Howe, eds (Washington, D.C.: American Society for Microgiology), pp. 485-502.

Boeke, J.D., and Corces, V.G. (1989). Transcription and reverse transcription of retrotransposons. Annu Rev Microbiol **43,** 403-434.

Brutnell, T.P., and Dellaporta, S.L. (1994). Somatic inactivation and reactivation of *Ac* associated with changes in cytosine methylation and transposase expression. Genetics **138,** 213-225.

Bureau, T.E., and Wessler, S.R. (1992). *Tourist*: a large family of inverted-repeat element frquently associated with maize genes. Plant Cell **4,** 1283-1294.

Bureau, T.E., Ronald, P.C., and Wessler, S.R. (1996). A computer-based systematic survey reveals the predominance of small inverted-repeat elements in wild-type rice genes. Proc Natl Acad Sci U S A **93,** 8524-8529.

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

Capy, P., Langin, T., Higuet, D., Maurer, P., and Bazin, C. (1997). Do the integrases of LTR-retrotransposons and class II element transposases have a common ancestor? Genetica **100,** 63-72.

Casacuberta, E., Casacuberta, J.M., Puigdomenech, P., and Monfort, A. (1998). Presence of miniature inverted-repeat transposable elements (MITEs) in the genome of *Arabidopsis thaliana*: characterisation of the *Emigrant* family of elements. Plant J. **16,** 79-85.

Chan, S.W., Henderson, I.R., and Jacobsen, S.E. (2005). Gardening the genome: DNA methylation in Arabidopsis thaliana. Nat Rev Genet **6,** 351-360.

Chandler, V.L., and Walbot, V. (1986). DNA modification of a maize transposable element correlates with loss of activity. Proc Natl Acad Sci U S A **83,** 1767-1771.

Cheng, C., Daigen, M., and Hirochika, H. (2006). Epigenetic regulation of the rice retrotransposon Tos17. Mol Genet Genomics **276,** 378-390.

Craig, N.L., Craigie, R., Gellert, M., and Lambowitz, A.M. (2002). Mobile DNA II. (Washington, D.C.: American Society for Microbiology Press).

Cui, H., and Fedoroff, N.V. (2002). Inducible DNA demethylation mediated by the maize Suppressor-mutator transposon-encoded TnpA protein. Plant Cell **14,** 2883-2899.

Deininger, P.L. (1989). SINEs: Short Interspersed Repeated DNA Elements in higher eukaryotes. In Mobile DNA, D.E. Berg and M.M. Howe, eds (Washington, DC: American Society for Microbiology), pp. 619-636.

Devos, K.M., Brown, J.K., and Bennetzen, J.L. (2002). Genome size reduction through illegitimate recombination counteracts genome expansion in Arabidopsis. Genome Res **12,** 1075-1079.

Doolittle, R.F., Feng, D.-F., Johnson, M.s., and McClure, M.A. (1989). Origins and evolutionary relationships of retroviruses. The Quarterly Review of Biology **64,** 1-30.

Fedoroff, N.V. (1989). The heritable activation of *cryptic Suppressor-mutator* elements by an active element. Genetics **121,** 591-608.

Fedoroff, N.V. (1999). The suppressor-mutator element and the evolutionary riddle of transposons. Genes Cells **4,** 11-19.

Feschotte, C., and Wessler, S.R. (2002). *Mariner*-like transposases are widespread and diverse in flowering plants. Proc. Natl. Acad. Sci. USA **99,** 280-285.

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

Feschotte, C., Jiang, N., and Wessler, S.R. (2002b). Plant transposable elements: where

genetics meets genomics. Nat Rev Genet **3,** 329-341.

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

Freitag, M., and Selker, E.U. (2005). Controlling DNA methylation: many roads to one modification. Curr Opin Genet Dev **15,** 191-199.

Freitag, M., Williams, R.L., Kothe, G.O., and Selker, E.U. (2002). A cytosine methyltransferase homologue is essential for repeat-induced point mutation in Neurospora crassa. Proc Natl Acad Sci U S A **99,** 8802-8807.

Galagan, J., and Selker, E. (2004). RIP: the evolutionary cost of genome defense. Trends Genet **20,** 417 - 423.

Galagan, J.E., Calvo, S.E., Borkovich, K.A., Selker, E.U., Read, N.D., Jaffe, D., FitzHugh, W., Ma, L.J., Smirnov, S., Purcell, S., Rehman, B., Elkins, T., Engels, R., Wang, S., Nielsen, C.B., Butler, J., Endrizzi, M., Qui, D., Ianakiev, P., Bell-Pedersen, D., Nelson, M.A., Werner-Washburne, M., Selitrennikoff, C.P., Kinsey, J.A., Braun, E.L., Zelter, A., Schulte, U., Kothe, G.O., Jedd, G., Mewes, W., Staben, C., Marcotte, E., Greenberg, D., Roy, A., Foley, K., Naylor, J., Stange-Thomann, N., Barrett, R., Gnerre, S., Kamal, M., Kamvysselis, M., Mauceli, E., Bielke, C., Rudd, S., Frishman, D., Krystofova, S., Rasmussen, C., Metzenberg, R.L., Perkins, D.D., Kroken, S., Cogoni, C., Macino, G., Catcheside, D., Li, W., Pratt, R.J., Osmani, S.A., DeSouza, C.P., Glass, L., Orbach, M.J., Berglund, J.A., Voelker, R., Yarden, O., Plamann, M.,

Seiler, S., Dunlap, J., Radford, A., Aramayo, R., Natvig, D.O., Alex, L.A., Mannhaupt, G., Ebbole, D.J., Freitag, M., Paulsen, I., Sachs, M.S., Lander, E.S., Nusbaum, C., and Birren, B. (2003). The genome sequence of the filamentous fungus Neurospora crassa. Nature **422,** 859-868.

Graff, J.R., Herman, J.G., Myohanen, S., Baylin, S.B., and Vertino, P.M. (1997). Mapping patterns of CpG island methylation in normal and neoplastic cells implicates both upstream and downstream regions in de novo methylation. J Biol Chem **272,** 22322-22329.

Hashida, S.N., Uchiyama, T., Martin, C., Kishima, Y., Sano, Y., and Mikami, T. (2006). The temperature-dependent change in methylation of the Antirrhinum transposon Tam3 is controlled by the activity of its transposase. Plant Cell **18,** 104-118.

Hutchison, C.A., Hardies, S.C., Loeb, D.D., Shehee, W.R., and Edgell, M.H. (1989). LINES and related retrotransposons: long interspersed repeated sequences in the eukaryotic genome. In Mobile DNA, D.E. Berg and M.M. Howe, eds (Washington, D.C.: American Society of Microbiology), pp. 593-617.

Izsvak, Z., Ivics, Z., Shimoda, N., Mohn, D., Okamoto, H., and Hackett, P.B. (1999). Short inverted-repeat transposable elements in teleost fish and implications for a mechanism of their amplification. J. Mol. Evol. **48,** 13-21.

Jackson, J.P., Lindroth, A.M., Cao, X., and Jacobsen, S.E. (2002). Control of CpNpG DNA methylation by the KRYPTONITE histone H3 methyltransferase. Nature **416,** 556-560.

Jackson, J.P., Johnson, L., Jasencakova, Z., Zhang, X., PerezBurgos, L., Singh, P.B., Cheng,

X., Schubert, I., Jenuwein, T., and Jacobsen, S.E. (2004). Dimethylation of histone H3 lysine 9 is a critical mark for DNA methylation and gene silencing in Arabidopsis thaliana. Chromosoma **112,** 308-315.

Jiang, N., Feschotte, C., Zhang, X.Y., and Wessler, S.R. (2004). Using rice to understand the origin and amplification of miniature inverted repeat transposable elements (MITEs). Current Opinion in Plant Biology **7,** 115-119.

Jiang, N., Bao, Z., Zhang, X., McCouch, S.R., Eddy, S.R., and Wessler, S.R. (2003). An active DNA transposon in rice. Nature **421,** 163-167.

Jin, Y.K., and Bennetzen, J.L. (1994). Integration and nonrandom mutation of a plasma membrane proton ATPase gene fragment within the Bs1 retroelement of maize. Plant Cell **6,** 1177-1186.

Kato, M., Miura, A., Bender, J., Jacobsen, S.E., and Kakutani, T. (2003). Role of CG and Non-CG Methylation in Immobilization of Transposons in Arabidopsis. Curr Biol **13,** 421-426.

Kidwell, M.G., and Lisch, D.R. (2001). Perspective: transposable elements, parasitic DNA, and genome evolution. Evolution Int J Org Evolution **55,** 1-24.

Kidwell, M.G., and Lisch, D. (2002). Transposable elements as sources of genomic variation. In Moblie DNA II, N.L. Craig, R. Craigie, M. Gellert, and A.M. Lambowitz, eds (Washington, DC: American Society for Microbiology Press), pp. 59-90.

Kikuchi, K., Terauchi, K., Wada, M., and Hirano, H.Y. (2003). The plant MITE mPing is mobilized in anther culture. Nature **421,** 167-170.

Kumar, A., and Bennetzen, J.L. (1999). Plant retrotransposons. Annu Rev Genet **33,** 479-532.

Kunze, R., and Starlinger, P. (1989). The putative transposase of transposable element Ac from *Zea mays* L. interacts with subterminal sequences of Ac. EMBO J. **8,** 3177-3185.

Kunze, R., and Weil, C.F. (2002). The hAT and CACTA superfamilies of plant transposons. In Mobile DNA II, N.L. Craig, R. Craigie, M. Gellert, and A.M. Lambowitz, eds (Washington DC: American Society for Microbiology Press), pp. 565-610.

Lippman, Z., May, B., Yordan, C., Singer, T., and Martienssen, R. (2003). Distinct Mechanisms Determine Transposon Inheritance and Methylation via Small Interfering RNA and Histone Modification. PLoS Biol **1,** E67.

Lisch, D. (2002). Mutator transposons. Trends in Plant Science **7,** 497 - 504.

Lisch, D., Girard, L., Donlin, M., and Freeling, M. (1999). Functional analysis of deletion derivatives of the maize transposon MuDR delineates roles for the MURA and MURB proteins. Genetics **151,** 331-341.

Lisch, D., Carey, C.C., Dorweiler, J.E., and Chandler, V.L. (2002). A mutation that prevents paramutation in maize also reverses Mutator transposon methylation and silencing. Proc Natl Acad Sci U S A **99,** 6130-6135.

Mao, L., Wood, T.C., Yu, Y., Budiman, M.A., Tomkins, J., Woo, S., Sasinowski, M., Presting, G., Frisch, D., Goff, S., Dean, R.A., and Wing, R.A. (2000). Rice transposable elements: a survey of 73,000 sequence-tagged-connectors. Genome Res. **10,** 982-990.

Martienssen, R. (1998). Transposons, DNA methylation and gene control. Trends Genet **14,** 263-264.

Martienssen, R., Barkan, A., Taylor, W.C., and Freeling, M. (1990). Somatically heritable switches in the DNA modification of Mu transposable elements monitored with a suppressible mutant in maize. Genes Dev **4,** 331-343.

McClintock, B. (1948). Mutable loci in maize. Carnegie Institution of Washington Yearbook **47,** 155-169.

McClintock, B. (1949). Mutable loci in maize. Carnegie Institution of Washington Yearbook **48,** 142-154.

McClintock, B. (1951). Mutable loci in maize. Carnegie Institution of Washington Yearbook **50,** 174-181.

McClintock, B. (1957). Genetic and cytological studies of maize. Carnegie Institution of Washington Yearbook **56,** 393-401.

McClintock, B. (1958). The *Suppressor-mutator* system of control of gene action in maize. Carnegie Institution of Washington Yearbook **57,** 415-429.

McClintock, B. (1961). Further studies of the *suppressor-mutator* system of control. Carnegie Institution of Washington Yearbook **60,** 469-476.

Mette, M.F., Aufsatz, W., van der Winden, J., Matzke, M.A., and Matzke, A.J. (2000). Transcriptional silencing and promoter methylation triggered by double-stranded RNA. Embo J **19,** 5194-5201.

Morgan, G.T. (1995). Identification in the human genome of mobile elements spread by DNA-mediated transposition. J. Mol. Biol. **254,** 1-5.

Naito, K., Cho, E., Yang, G., Campbell, M.A., Yano, K., Okumoto, Y., Tanisaka, T., and

Wessler, S.R. (2006). Dramatic amplification of a rice transposable element during recent domestication. Proc Natl Acad Sci U S A **103,** 17620-17625.

Nakazaki, T., Okumoto, Y., Horibata, A., Yamahira, S., Teraishi, M., Nishida, H., Inoue, H., and Tanisaka, T. (2003). Mobilization of a transposon in the rice genome. Nature **421,** 170-172.

Oosumi, T., Garlick, B., and Belknap, W.R. (1996). Identification of putative nonautonomous transposable elements associated with several transposon families in *Caenorhabditis elegans*. J. Mol. Evol. **43,** 11-18.

Plasterk, R.H.A., and van Luenen, H.G. (2002). The Tc1/*mariner* family of transposable elements. In Mobile DNA II, N.L. Craig, R. Craigie, M. Gellert, and A.M. Lambowitz, eds (Washington D.C.: American Society for Microbiology Press), pp. 519-532.

Porebski, S., Bailey, L.G., and Baum, B.R. (1997). Modification of a CTAB DNA Extraction Protocol for Plants Containing High Polysaccharide and Polyphenol Components. Plant Molecular Biology Reporter **15 (1),** 8-15.

Ros, F., and Kunze, R. (2001). Regulation of activator/dissociation transposition by replication and DNA methylation. Genetics **157,** 1723-1733.

Rubin, E., Lithwick, G., and Levy, A.A. (2001). Structure and evolution of the hAT transposon superfamily. Genetics **158,** 949-957.

Santiago, N., Herraiz, C., Goni, J.R., Messeguer, X., and Casacuberta, J.M. (2002). Genome-wide Analysis of the Emigrant Family of MITEs of Arabidopsis thaliana. Mol Biol Evol **19,** 2285-2293.

Schlappi, M., Smith, D., and Fedoroff, N. (1993). Tnpa trans activates methylated maize *suppressor mutator* transposable elements in transgenic tobacco. Genetics **133,** 1009-1021.

Schlappi, M., Raina, R., and Fedoroff, N. (1994). Epigenetic regulation of the maize *Spm* transposable element - novel activation of a methylated promoter by *tnpa*. Cell **77,** 427-437.

Selker, E.U., Tountas, N.A., Cross, S.H., Margolin, B.S., Murphy, J.G., Bird, A.P., and Freitag, M. (2003). The methylated component of the Neurospora crassa genome. Nature **422,** 893-897.

Singer, T., Yordan, C., and Martienssen, R.A. (2001). Robertson's Mutator transposons in A. thaliana are regulated by the chromatin-remodeling gene Decrease in DNA Methylation (DDM1). Genes Dev **15,** 591-602.

Sinzelle, L., Kapitonov, V.V., Grzela, D.P., Jursch, T., Jurka, J., Izsvak, Z., and Ivics, Z. (2008). Transposition of a reconstructed Harbinger element in human cells and functional homology with two transposon-derived cellular genes. Proc Natl Acad Sci U S A **105,** 4715-4720.

Slotkin, R.K., and Martienssen, R. (2007). Transposable elements and the epigenetic regulation of the genome. Nat Rev Genet **8,** 272-285.

Takata, M., Kiyohara, A., Takasu, A., Kishima, Y., Ohtsubo, H., and Sano, Y. (2007). Rice transposable elements are characterized by various methylation environments in the genome. BMC Genomics **8,** 469.

Tran, R.K., Henikoff, J.G., Zilberman, D., Ditt, R.F., Jacobsen, S.E., and Henikoff, S. (2005a). DNA methylation profiling identifies CG methylation clusters in Arabidopsis genes. Curr Biol **15,** 154-159.

Tran, R.K., Zilberman, D., de Bustos, C., Ditt, R.F., Henikoff, J.G., Lindroth, A.M., Delrow, J., Boyle, T., Kwong, S., Bryson, T.D., Jacobsen, S.E., and Henikoff, S. (2005b). Chromatin and siRNA pathways cooperate to maintain DNA methylation of small transposable elements in Arabidopsis. Genome Biol **6,** R90.

Tu, Z. (1997). Three novel families of miniature inverted-repeat transposable elements are associated with genes of the yellow fever mosquito, *Aedes aegypti*. Proc. Natl. Acad. Sci. USA **94,** 7475-7480.

Tu, Z. (2001). Eight novel families of miniature inverted repeat transposable elements in the African malaria mosquito, *Anopheles gambiae*. Proc. Natl. Acad. Sci. USA **98,** 1699-1704.

Vitte, C., and Panaud, O. (2003). Formation of solo-LTRs through unequal homologous recombination counterbalances amplifications of LTR retrotransposons in rice Oryza sativa L. Mol Biol Evol **20,** 528-540.

Wang, L., Heinlein, M., and Kunze, R. (1996). Methylation pattern of Activator transposase binding sites in maize endosperm. Plant Cell **8,** 747-758.

Weil, C., and Martienssen, R. (2008). Epigenetic interactions between transposons and genes: lessons from plants. Curr Opin Genet Dev **18,** 188-192.

Wessler, S.R. (2006). Transposable elements and the evolution of eukaryotic genomes. Proc Natl

Acad Sci U S A **103,** 17600-17601.

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

Yang, G., Weil, C.F., and Wessler, S.R. (2006). A rice Tc1/mariner-like element transposes in yeast. Plant Cell **18,** 2469-2478.

Yang, G., Zhang, F., Hancock, C.N., and Wessler, S.R. (2007). Transposition of the rice miniature inverted repeat transposable element mPing in Arabidopsis thaliana. Proc Natl Acad Sci U S A **104,** 10962-10967.

Yates, P.A., Burman, R.W., Mummaneni, P., Krussel, S., and Turker, M.S. (1999). Tandem B1 elements located in a mouse methylation center provide a target for de novo DNA methylation. J Biol Chem **274,** 36357-36361.

Yu, Z., Wright, S.I., and Bureau, T.E. (2000). Mutator-like Elements in Arabidopsis thaliana. Structure, diversity and evolution. Genetics **156,** 2019-2031.

Zhang, X. (2008). The epigenetic landscape of plants. Science **320,** 489-492.

Zhang, X., Clarenz, O., Cokus, S., Bernatavichute, Y.V., Pellegrini, M., Goodrich, J., and Jacobsen, S.E. (2007). Whole-genome analysis of histone H3 lysine 27 trimethylation in Arabidopsis. PLoS Biol **5,** e129.

Zhang, X., Yazaki, J., Sundaresan, A., Cokus, S., Chan, S.W., Chen, H., Henderson, I.R., Shinn, P., Pellegrini, M., Jacobsen, S.E., and Ecker, J.R. (2006). Genome-wide High-Resolution Mapping and Functional Analysis of DNA Methylation in Arabidopsis. Cell **126,** 1189-1201.

Zilberman, D., Gehring, M., Tran, R.K., Ballinger, T., and Henikoff, S. (2007). Genome-wide analysis of Arabidopsis thaliana DNA methylation uncovers an interdependence between methylation and transcription. Nat Genet **39,** 61-69.