Transposition of a Rice Mutator-Like Element in the Yeast Saccharomyces cerevisiae

Dongyan Zhao, Ann Ferguson, and Ning Jiang1
Department of Horticulture, Michigan State University, East Lansing, Michigan 48824
ORCID IDs: 0000-0002-2080-8416 (D.Z.); 0000-0002-2776-6669 (N.J.)

Mutator-like transposable elements (MULEs) are widespread in plants and are well known for their high transposition activity as well as their ability to duplicate and amplify host gene fragments. Despite their abundance and importance, few active MULEs have been identified. In this study, we demonstrated that a rice (Oryza sativa) MULE, Os3378, is capable of exciting and reinserting yeast (Saccharomyces cerevisiae), suggesting that yeast harbors all the host factors for the transposition of MULEs. The transposition activity induced by the wild-type transposase is low but can be altered by modification of the transposase sequence, including deletion, fusion, and substitution. Particularly, fusion of a fluorescent protein to the transposase enhanced the transposition activity, representing another approach to manipulate transposases. Moreover, we identified a critical region in the transposase where the net charge of the amino acids seems to be important for activity. Finally, transposition efficiency is also influenced by the element and its flanking sequences (i.e., small elements are more competent than their large counterparts). Perfect target site duplication is favorable, but not required, for precise excision. In addition to the potential application in functional genomics, this study provides the foundation for further studies of the transposition mechanism of MULEs.

INTRODUCTION

Transposable elements (TEs) are genomic sequences that can move from one position to another within a genome. Due to their repetitive nature, TEs frequently constitute large fractions of eukaryotic genomes (Chénais et al., 2012). TEs are divided into two major classes based on whether the intermediate of transposition is DNA or RNA. Class I TEs or retrotransposons transpose via an RNA intermediate, which is transcribed as an mRNA from the genomic copy, followed by reverse transcription of the mRNA into cDNA and integration into a new genomic location. Class II or DNA TEs transpose via a DNA intermediate, which is excised from the TE’s original position (donor site) and integrates into another position (target site). An exception is Hei1itrons, which transpose through a rolling-circle mechanism (Kapitonov and Jurka, 2001). Both TE classes consist of autonomous and nonautonomous elements; autonomous TEs encode proteins (transposases) responsible for the transposition of themselves and their corresponding nonautonomous elements.

Mutator (Mu)-like transposable elements (MULEs) are class II TEs that were discovered in maize (Zea mays) because of their high mutagenicity (Robertson, 1978). The founder elements include the autonomous element MuDR and the nonautonomous elements (Mu1 to Mu8 and Mu10 to Mu13; MuDR generates the transposition machinery for the movement of all these elements (Bennetzen et al., 1984; Feehling, 1984; Chen et al., 1987; Talbert et al., 1989; Feenor et al., 1990; Chomet et al., 1991; Hershberger et al., 1991; Qin et al., 1991; Lisch et al., 1995; Dietrich et al., 2002; Lisch and Jiang, 2009; Tan et al., 2011). All Mu elements share high sequence similarity in their terminal inverted repeats (TIRs; ~220 bp) while containing heterogeneous internal sequences. Upon integration into a new position, they usually generate a 9-bp duplication of the target site sequence, which is called the target site duplication (TSD).

In addition to Mu elements in maize, MULEs have been found in a wide range of organisms, including other plants, fungi, and animals (Yu et al., 2000; Lisch et al., 2001; Singer et al., 2001; Chalvet et al., 2003; Rossi et al., 2004; Neuvéglise et al., 2005; Lisch and Jiang, 2009; Marquez and Pritham, 2010; Gao, 2012). In organisms whose TEs have been examined systematically, MULEs can reach thousands of copies. For example, both rice (Oryza sativa) and tomato (Solanum lycopersicum) contain ~30,000 copies of MULEs (Ferguson and Jiang, 2012). Pack-MULEs, a special group of MULEs, have captured genes or gene fragments (Jiang et al., 2004). Rice has nearly 3000 Pack-MULEs (Ferguson et al., 2013), and virtually all the nonautonomous Mu elements in maize are Pack-MULEs. As such, Pack-MULEs likely play a unique role in genome evolution (Hanada et al., 2009). Although important, the elucidation of the mechanism of Pack-MULE formation is hindered partly by the lack of a MULE transposition system (or a species) that can be readily manipulated.

Despite their abundance in plants, only a few active MULEs have been discovered, and little is known about their transposition mechanism. MuDR, the first active autonomous MULE element discovered, encodes two proteins, MURA and MURB (Lisch, 2002). The MURA transposase catalyzes the excision of Mu elements, and MURB was suggested to be involved in new insertions (Lisch, 2002). Subsequent to the discovery of MuDR, a number of active MULEs were reported, such as Jittery and TED in maize (Xu et al., 2004; Li et al., 2013), Os3378 in rice (Gao, 2012), and AtMu1 in Arabidopsis thaliana (Singer et al., 2001).
Transposition of MULEs has only been observed in their native hosts, and it has been suggested that host factors are involved in the transposition of Mu elements in maize (Zhao and Sundaresan, 1991). Establishment of a transposition system in a manipulable organism would facilitate studies of MULEs, especially the mechanism by which Pack-MULEs acquire genomic sequences. Yeast (Saccharomyces cerevisiae) is an ideal species for this purpose because of its small size, short life cycle, and successful recapitulation of the transposition of other TEs (Weil and Kunze, 2000; Yang et al., 2006; Hancock et al., 2010).

In this study, we show that Os3378, an active MULE from rice, is capable of transposition in yeast. The transposition frequency induced by the wild-type transposase is rather low but can be improved through a variety of modifications, including N-terminal sequence deletion and fusion of an enhanced yellow fluorescent protein (EYFP). Our study provides new insights about the function of transposases and the regulation of transposition activity.

RESULTS

The Os3378 Transposase

The Os3378 element has four copies with over 98% nucleotide identity in the reference japonica rice cv Nipponbare. All four copies are annotated as transposon proteins (LOC_Os04g28290, LOC_Os04g28350, LOC_Os05g31510, and LOC_Os11g44180) in MSU version 7 rice pseudomolecules (http://rice.plantbiology.msu.edu/). Three of the Os3378 copies are associated with a 196-bp TIR. LOC_Os11g44180 is associated with only one terminus, yet its coding region seems to be intact (Gao, 2012). To verify the gene structure, we extracted RNA and amplified the cDNA of Os3378 transposase from a japonica somaclonal line called Z418, in which the expression of Os3378 was detected previously (Gao, 2012). Comparison of cDNA from Z418 (called Os3378-Z) and that of LOC_Os05g31510 indicated the absence of point mutations between the coding regions; however, the boundaries of exon 2 and exon 4 in Os3378-Z differ from the annotation based on LOC_Os05g31510. Exon 2 of Os3378-Z is 288 nucleotides longer than that of LOC_Os05g31510, while exon 4 is 6 nucleotides shorter than that of LOC_Os05g31510 (Figure 1A; Supplemental Figure 1A). Neither of these changes results in a frame shift. The length of the Os3378-Z coding region is 886 amino acids, which is comparable to that of four known active MULE elements, MURA of MuDR (823 amino acids), JITA of Jittery (709 amino acids), TEDA of TED (785 amino acids), and that of AlMu1 (761 amino acids) (Hershberger et al., 1995; Singer et al., 2001; Xu et al., 2004; Li et al., 2013).

A functional transposase usually consists of a DNA binding domain (for binding transposon DNA) and a catalytic domain (for element excision and reinsertion). By comparing the sequence with known DNA transposases (encoded by IS256 and Hermes) (Hennig and Ziebuhr, 2010; Nesmelova and Hackett, 2010) and locating helix-turn-helix motifs, we identified the putative DNA binding domain of Os3378-Z located between amino acids 300 and 400 and a catalytic domain with a DDE motif located between amino acids 440 and 610 (Figure 1B; 445D-510D-610E). In addition, one nuclear export signal (NES) was predicted to be within the catalytic domain. Three nuclear localization signals (NLSes) were predicted, with one located upstream of the DNA binding domain and two located at the C terminus of the protein.

Os3378-Z Is Capable of Inducing Excision and Reinsertion of a Nonautonomous Element

The transposition assay was developed using two vectors: an expression vector containing the Os3378-Z transposase coding sequence downstream of the GAL1 promoter (Figure 1D) and a reporter vector with an artificial nonautonomous Os3378 element (Os3378NA469) inserted into the coding region of the ADE2 gene (Figures 1C and 1E). Os3378NA469 is composed of the TIR of Os3378, a short piece of the subterminal region (71 bp), and a linker sequence (6 bp), with a total length of 469 bp (Figure 1C). The element is flanked by a TSD (TTGTTTAAAC). Excision of Os3378NA469 and recovery of a functional ADE2 gene allow the growth of yeast on a culture medium lacking adenine. In general, the excision frequency was low (0.47 events/10⁶ cells on 2% galactose). To test whether the growth of yeast is due to factors other than excision, a control experiment was conducted using an empty expression vector. No cell growth was observed among 20 replicates, suggesting that the reversion of ADE2 function is due to the excision induced by the transposase. Transposon display was used to identify new sites of insertion of Os3378NA469 after excision. Analysis of two reinsertion sites indicated that both inserted into 35S rRNA genes, but at different locations. One insertion was in the internal transcribed spacer region (ITS1-1) and the other was in the 18S rRNA sequence (RDNN18-1). Both insertions were associated with a 9-bp TSD (AAAATTTAA and TTGAAAAAA, respectively). Thus, Os3378-Z is capable of inducing both excision and reinsertion of Os3378NA469 in yeast, albeit at a low frequency.

The Effect of the N-Terminal Deletion of Os3378-Z Transposase on Excision Frequency

Previous studies showed that deletion of the N-terminal region of Ac (Kunze et al., 1995) and phage Mu (MuA) transposases (Kim and Morrison, 2009) resulted in increased transposition activity. Analysis of secondary structures indicated that all three transposases harbor a helix-turn-helix domain at their N termini (Ac transposase, amino acids 10 to 50; MuA, amino acids 1 to 77; Os3378-Z transposase, amino acids 40 to 100). This prompted us to test whether deletion of the N-terminal region would enhance the activity of Os3378-Z. To this end, five truncated transposases were constructed, Os3378-Z-80, Os3378-Z-105, Os3378-Z-130, Os3378-Z-161, and Os3378-Z-168, which contain deletions of the N-terminal 79, 104, 129, 160, and 167 amino acids, respectively (Figure 2A). The excision frequency varied dramatically among different forms of the Os3378-Z transposases. As shown in Figure 2B, there were few excisions associated with Os3378-Z-168. Os3378-Z-105 demonstrated similar activity to the wild-type transposase, while Os3378-Z-80 and Os3378-Z-161 were less efficient than Os3378-Z. Os3378-Z-130 was the only form of transposase with significantly enhanced excision frequency, which was ~3-fold higher than that of
Figure 1. Schematic Structures of Os3378 and Constructs Used in This Study.

(A) Structural comparison between the annotated transposase (LOC_Os05g31510) and Os3378-Z. Black triangles, TIRs; empty boxes, exons; lines linking empty boxes, introns; arrows, transcription start sites; dashed lines, splicing site variations between LOC_Os05g31510 and Os3378-Z.

(B) Structural comparison between LOC_Os05g31510 and Os3378-Z at the amino acid level. Shaded regions are in common between the two proteins. HTH, helix-turn-helix.
Os3378-Z (2.02 versus 0.47 events/10^6 cells; P < 0.0001). As such, deletion of the N-terminal 129 amino acids of the Os3378-Z transposase significantly increased the excision frequency.

**Excision Frequency Altered by Substitutions of Amino Acids within 105 to 130 Amino Acids of Os3378-Z Transposase**

The differential excision frequency caused by Os3378-Z-130 and Os3378-Z-105 prompted us to study the amino acid composition of the 105- to 129-amino acid region, the only sequence/structural difference between the two truncated transposases. Multiple sequence alignment revealed that this region is not conserved among different known MULE transposases, including MURA, TEDA, JITA, and that of AtMu1. The region contains nine acidic residues and only two basic residues, and hydropathy analysis showed that the entire region is very hydrophilic (Kyte and Doolittle, 1982). To determine whether the amino acid composition of the 105- to 129-amino acid region affected the excision activity, we mutated the amino acids in this region to produce three mutant transposases (Figure 3A). In Os3378-Z-105Basic, all the acidic residues were substituted with basic amino acids; in Os3378-Z-105Ala, most of the acidic residues were substituted with alanine; in Os3378-Z-105Neutral, all the acidic residues were substituted by their corresponding neutral amino acids. Excision assays showed that both Os3378-Z-105Ala and Os3378-Z-105Neutral resulted in much higher activity than the original Os3378-Z-105 (P < 0.01, t test). Os3378-Z-105Ala was associated with the highest activity, which was comparable to that induced by Os3378-Z-130 (P = 0.6679, t test) (Figure 3B), suggesting that chemical and physical properties of the amino acids in this region (amino acids 105 to 129) are critical to the activity of the Os3378-Z transposase.

**Other Factors That Influence Excision Frequency and the Presence of Footprints**

Despite the fact that N-terminal deletion and substitution significantly improved excision frequency, the activity remained relatively low compared with that induced by other transposases. For instance, a mutated form of Pong transposase induced 16 excision events per 10^6 cells in a similar system (Hancock et al., 2010). In an attempt to further improve the excision frequency, we tested the effect of galactose concentration, the size of the nonautonomous elements, and the presence of a TSD on excision frequency.

**Galactose Concentration**

As mentioned above, expression of the Os3378-Z transposase is under the control of the GAL1 promoter, which is induced by galactose. Therefore, culture media with various galactose levels (0.05, 0.1, 0.5, 1, and 2%) were used to modulate the expression, and immunoblot analysis using His antibody (see below) was conducted to determine the protein levels of the transposase. Not surprisingly, increasing galactose levels elevated protein levels of both Os3378-Z and Os3378-Z-130, with the wild-type Os3378-Z being more responsive to high galactose levels (Figures 4A and 4B). For wild-type transposase, the excision frequency was rather low on medium with 0.05% galactose (0.09 events/10^6 cells) and reached the highest level with 2% galactose (0.47 events/10^6 cells). No significant difference was observed for other galactose concentrations (0.35 events/10^6 cells). For Os3378-Z-130, the excision frequency was largely the same at all galactose concentrations, suggesting saturation at very low concentrations.

**The Size of Elements**

In maize, the 1.4-kb Mu1 element is highly active (Benettzen, 1996). To test whether the size of Os3378NA469 is a limiting factor, we constructed another nonautonomous element, a larger deletion derivative of Os3378, which was 1485 bp in length, about three times the size of Os3378NA469 (Figure 1C). However, the excision frequency from Os3378NA1485 was even lower than that of Os3378NA469, suggesting that the excision is more efficient with smaller elements or that the larger element contains sequence features that inhibit excision.

**The Role of TSD**

To test the effect of TSD on excision frequency, we mutated the first 3 bp of the 3' side of the TSD sequence (TCT to AGC, both encode serine; Figure 5B) in the reporter vector so that the element is no longer flanked by a perfect TSD. This modified vector would encode the same protein as the original reporter vector unless the breakpoint during excision is within the AGC (i.e., between A and G or G and C) at the 3' end and the breakpoint at the 5' end is not in-frame. As shown in Figure 5A, for Os3378NA469, the lack of TSD (or nonperfect TSD) resulted in the reduction of excision frequency to 6% (2.02 versus 0.12 events/10^6 cells; P < 0.0001), and this value is 3% for Os3378NA1485 (0.257 versus 0.007 events/10^6 cells; P < 0.0001). This suggests that a perfect TSD is very important for precise excision, especially for large elements.

**Footprint of Excision**

The excision events were further characterized by amplifying the donor site following excisions using primers in the flanking region of Os3378NA and sequencing the amplicons (Supplemental Table 2). Thirty-nine amplicons were sequenced for reporter vectors with or without TSDs. All excision sites from reporter vectors
with TSD represented precise excisions (i.e., one element plus one side of the TSD sequence was removed and the original ADE2 sequence was recovered) (type I). One exception, derived from Os3378NA1485, was detected in which only the element was excised and both sides of the TSD were retained. This would result in the generation of three additional amino acids in the coding region of ADE2. For the reporter vector without the TSD (non-perfect TSD), seven nonprecise events were detected (Figure 5B). One event resembled that from the vector with the TSD, where TSD sequence on both sides was retained and the element was excised (type VI). Five events represented excisions of the element plus a 3-bp sequence immediately flanking the element, with different breakpoints (Figure 5B, types III to V). In one case, a single point mutation (in addition to excision of the element and 3-bp flanking sequence) caused an amino acid change from valine to phenylalanine (Figure 5B, type V). In addition, one event resembled precise excision except for a single nucleotide difference (T versus C, a silent mutation) compared with the original donor site sequence, likely because the breakpoint is located between G and C of the mutated portion of the TSD (type II). Thus, it appears that a perfect TSD increases the ratio of precise excisions.

Cellular Localization and Excision Frequency of Os3378-Z Transposases with N- or C-Terminal EYFP Fusion

To determine whether the increased excision rate of Os3378-Z-130 was due to its cellular localization, we fused an EYFP tag to the N or C terminus of the transposase (Supplemental Figure 1B) and examined whether the fusion affected the excision frequency. Interestingly, the effect of EYFP fusion on excision frequency varied dramatically among different forms of transposase and the position of fusion. For the wild-type Os3378-Z, the EYFP fusion at either terminus led to a reduction in excision frequency (Figure 6A), suggesting that the fusion interfered with the function of the transposase. C-terminal fusion of EYFP to the two truncated transposases, Os3378-Z-105 and Os3378-Z-130, did not significantly alter the excision frequency. By contrast, N-terminal fusion of EYFP to Os3378-Z-105 enhanced its activity, whereas the same fusion significantly suppressed the activity of Os3378-Z-130 (Figure 6A). Since the N-terminal fusion of EYFP to Os3378-Z-105 demonstrated a positive effect on transposition, the amino acid sequence of EYFP that is directly attached to the transposase was examined. The 25-amino acid sequence upstream of the transposase (C-terminal sequence of the EYFP) is hydrophilic, containing four basic amino acids and no acidic amino acids. By contrast, the sequence upstream of amino acid 105 (amino acids 80 to 104) in Os3378-Z is slightly hydrophobic, containing six acidic amino acids and only one basic amino acid.

Images from confocal microscopy showed that transposases with N- and C-terminal EYFP fusions had different cellular localizations. When the fusion was at the N terminus, both EYFP-Os3378-Z and EYFP-Os3378-Z-105 were concentrated in the nucleus, as revealed by the colocalization of the 4',6-diamidino-2-phenylindole (DAPI)-stained nuclear DNA and the EYFP-tagged transposases (Figure 6C). By contrast, EYFP-Os3378-Z-130 protein was almost evenly distributed in the nucleus and cytoplasm. When the fusion was at the C terminus, all fusion proteins formed aggregates in the cytoplasm (Figure 6C).

The distinct cellular localizations of EYFP-Os3378-Z-105 and EYFP-Os3378-Z-130 (with EYFP at the N terminus) raised the question of whether the amino acid region 105 to 129 contains a cryptic NLS. To this end, we fused this region to an EYFP tag and observed its cellular localization. As shown in Figure 6C, the resulting EYFP fusion protein was present over the entire cells, suggesting that this region alone does not act as an NLS.

Protein Levels of Different Forms of Os3378-Z Transposase Are Not Correlated with Excision Frequency

To determine whether excision frequency induced by wild-type and N-terminal truncated transposases was correlated with
protein levels in yeast, immunoblot analysis was conducted. For such analysis, a FLAG-His\(_6\) dual tag (Zanetti et al., 2005) was fused to the C terminus of the transposase (CFH), which we refer to as Os3378-ZCFH, Os3378-Z-105CFH, Os3378-Z-105AlaCFH, and Os3378-Z-130CFH. Analysis of the excision frequency revealed no significant difference between the transposases with and without the CFH tag (e.g., Os3378-Z-130 versus Os3378-Z-130CFH; \(P = 0.7122, t\) test) (Figure 7A). Immunoblot analysis indicated that protein levels of Os3378-ZCFH and Os3378-Z-105AlaCFH were higher than those of Os3378-Z-105CFH and Os3378-Z-130CFH at the same galactose level (2%) (Figure 7B). In general, no correlation was observed between protein levels and excision frequency among different forms of Os3378-Z transposase (\(P = 0.9367\); Supplemental Figure 2).

### Figure 3. Os3378-Z-105 Transposases with Amino Acid Substitutions and Their Activity.

(A) Amino acid substitutions between amino acids 105 and 130 of Os3378-Z-105.

(B) Relative excision frequency induced by Os3378-Z-105 and its mutant forms on 2% galactose. The excision frequency by Os3378-Z-130 was set as 100%. SE values were calculated from at least 20 replicates for each form of the Os3378-Z transposase.

A Single Os3378 Terminus Is Immobile

Talbert and Chandler (1988) proposed that the presence of gene sequences inside \(Mu\) elements could be caused by the mobility of \(Mu\) termini. Specifically, a single \(Mu\) terminus may transpose and two separate but similar \(Mu\) termini may fortuitously insert into adjacent genic regions. If the two termini are close enough and in the right orientation, the entire sequence (two termini and the genes inside) could be moved together, thus forming a Pack-MULE. To determine whether a single Os3378 terminus was capable of excision, a reporter construct containing only the 5' terminus of Os3378 in the coding region of the \(ADE2\) gene was transformed into yeast along with the expression construct containing Os3378-Z-130. Out of 30 replicates, no \(ADE2\) revertant was observed on culture medium lacking adenine, suggesting that no excision was induced by the transposase. This indicates that a single Os3378 terminus is unlikely to be mobile.

### Reinsertion of Os3378NA

To determine whether the factors influencing excision frequency also affect the reinsertion frequency, DNA gel blot analysis was conducted using total DNA extracted from \(ADE2\) revertant yeast colonies. Our first attempt employed two restriction enzymes that did not cut the element, yet no signal was observed. This was likely due to the annealing of the TIR of individual elements so that the probe failed to hybridize to them; thus, in subsequent
work, EcoRI, with a single restriction site in the element, was used. In this case, signals from two fragments were expected from each insertion. In Figure 8A, the hybridization signals (indicated by horizontal arrows) shared by most samples represent the Os3378NA in the reporter vector, and other signals correspond to reinserted Os3378NA elements. The absence of common signals among a few samples indicated that the excision occurred prior to the first cell division of the relevant colonies. In some samples (indicated by vertical arrows), it was obvious that multiple reinsertions occurred, suggesting multiple excision and reinsertion events within a single colony. No significant variation in the number of reinsertions was observed between those induced by Os3378-Z and Os3378-Z-130 ($\chi^2 = 0.0286$, $P = 0.8657$). The insertion frequency from the reporter vector with TSD is slightly lower than that from the vector without TSD, but the difference is not statistically significant ($\chi^2 = 2.1942$, $P = 0.1358$). Surprisingly, there were many fewer reinsertions associated with Os3378NA1485 compared with Os3378NA469 (5 events out of 58 colonies versus 11 out 28 colonies; $\chi^2 = 11.7261$, $P = 0.0006$), suggesting low transposition frequency for the larger element or a stronger selection against its insertion in the genome.

To determine the locations of new insertions, we cloned 44 reinsertion sites derived from Os3378NA469. Of these sites, 39 have significant matches with the yeast genomic sequence (Supplemental Table 1). The majority of insertions (62%) are located in gene-rich regions, where equal numbers ($n = 7$) of insertions are located in the 5‘ and 3‘ regions of protein-coding genes (Figure 8B). Additionally, 10 insertions are located in the 5‘ region of one gene and the 3‘ region of another. Compared with the genomic fraction of intergenic sequence (~24%), reinsertions in these regions are ~2-fold higher than expected if the insertion is random in the genome ($\chi^2 = 24.9328$, $P < 0.0001$) (Figure 8C; Supplemental Figure 3). On the contrary, insertions within gene bodies are only one-fourth of that expected based on the fraction occupied by genes in the genome. This is likely because of deleterious effects of insertions inside genes. Hence, reinsertions in genes are likely higher than the observed value. In addition, five insertions (13%) are located in rRNA genes, which is 68-fold higher than their genomic fraction ($\chi^2 = 13.3194$, $P = 0.0003$). We also detected insertions in autonomously replicating sequences ($n = 2$) and telomere ($n = 1$). However, no insertion was found in long terminal repeat retrotransposons, although they occupy 1.5-fold as much genomic space as the sum of autonomously replicating sequences and telomeres. For those mapped reinsertion sites, the 9-bp TSD is highly TA-rich, with an average TA content of 87%, much higher than the genomic average (62%) of yeast. This target site preference of TA-rich sequence is similar to what was observed for Os3378 in rice (~86%) (Gao, 2012). Nevertheless, the overrepresentation of insertions in rRNA genes is likely due to the functional redundancy and not the TA content, because the average TA content of rRNA genes is similar to the genomic average (61.56% versus 61.85%).

Figure 4. Protein Levels of Os3378-Z Transposases and Their Activity at Various Galactose Concentrations.

(A) and (B) Protein levels of Os3378-ZCFH (A) and Os3378-Z-130CFH (B) at various galactose levels. Partial images of the total protein are shown at the bottom.

(C) Relative excision frequency induced by Os3378-Z and Os3378-Z-130 at various galactose concentrations. The excision frequency by Os3378-Z-130 on 2% galactose was set as 100%.
A previous study showed that Relevant Transposase Likely Due to the Low Transposition Activity of the Low Copy Number of Os3378 (et al., 1998). 

The yeast genome only harbors retrotransposons and yeast factors are conserved between rice and yeast, despite the facts Os3378 either does not require host-specific factors or these factors are conserved between rice and yeast, despite the facts that the yeast genome only harbors retrotransposons and yeast and rice share a common ancestor 1500 million years ago (Kim, 2012). The transcripts of Os3378 were readily detectable in the reproductive tissues of Z418 (Gao, 2012). Nevertheless, additional new insertions of Os3378 were never detected in Z418 populations (A. Ferguson and N. Jiang, unpublished data). The detectable expression level combined with the lack of new transposition events seems to be puzzling. In this study, we observed that the wild-type Os3378-Z transposase is not highly competent in inducing transpositions in yeast, and such low competency cannot be dramatically altered through its expression level. This may explain the uniformly low copy number of Os3378 among the wide range of species surveyed (Gao, 2012).

Wild-type transposases are often suboptimal for transposition (Kunze et al., 1995; Lampe et al., 1999; Weil and Kunze, 2000; Zayed et al., 2004; Pledger and Coates, 2005; Kim and Morrisson, 2009; Yusa et al., 2011; Lazarow et al., 2012). For example, mutation of three amino acids in the transposase of Hismar1 led to a 4- to 50-fold increase in transposition activity (Lampe et al., 1999). Given those instances, it is conceivable that high transposition activity might have been selected against, because transposon insertions are more often linked to deleterious than to favorable effects. As a result, low transposition activity and low copy numbers, as demonstrated by Os3378, might be one of the strategies for long-term success. On the other hand, transposition activity could be enhanced by a variety of mutations, such as point mutations or truncations. Nevertheless, the fact that Os3378 is not amplified in any rice cultivar and the wild relatives tested suggests either that such mutations are rare in nature or are selected against.

**DISCUSSION**

In this study, we show that a rice MULE, Os3378, transposed in yeast. The recapitulation of transposition in yeast suggests that Os3378 either does not require host-specific factors or these factors are conserved between rice and yeast, despite the facts that the yeast genome only harbors retrotransposons and yeast and rice share a common ancestor 1500 million years ago (Kim et al., 1998).

**The Low Copy Number of Os3378 in Rice and Its Relatives Is Likely Due to the Low Transposition Activity of the Relevant Transposase**

A previous study showed that Os3378 is present in low copy numbers (<10) in most rice cultivars and wild relatives (Gao, 2012). The transcripts of Os3378 were readily detectable in the reproductive tissues of Z418 (Gao, 2012). Nevertheless, additional new insertions of Os3378 were never detected in Z418 populations (A. Ferguson and N. Jiang, unpublished data). The detectable expression level combined with the lack of new transposition events seems to be puzzling. In this study, we observed that the wild-type Os3378-Z transposase is not highly competent in inducing transpositions in yeast, and such low competency cannot be dramatically altered through its expression level. This may explain the uniformly low copy number of Os3378 among the wide range of species surveyed (Gao, 2012).

Wild-type transposases are often suboptimal for transposition (Kunze et al., 1995; Lampe et al., 1999; Weil and Kunze, 2000; Zayed et al., 2004; Pledger and Coates, 2005; Kim and Morrisson, 2009; Yusa et al., 2011; Lazarow et al., 2012). For example, mutation of three amino acids in the transposase of Hismar1 led to a 4- to 50-fold increase in transposition activity (Lampe et al., 1999). Given those instances, it is conceivable that high transposition activity might have been selected against, because transposon insertions are more often linked to deleterious than to favorable effects. As a result, low transposition activity and low copy numbers, as demonstrated by Os3378, might be one of the strategies for long-term success. On the other hand, transposition activity could be enhanced by a variety of mutations, such as point mutations or truncations. Nevertheless, the fact that Os3378 is not amplified in any rice cultivar and the wild relatives tested suggests either that such mutations are rare in nature or are selected against.

**A Single Transposase of Os3378 Catalyzes Both Excision and Reinsertion Events**

In maize, the autonomous element MuDR encodes two proteins, with MURA being the transposase and sufficient to induce excision. The role of MURB remains enigmatic, and it was suggested that it may be involved in element insertions (Lisch et al., 1999; Raizada and Walbot, 2000; Woodhouse et al., 2006). Sequences homologous to MURA have been found in a wide range of species; however, no MURB sequences have been found in any species other than those in the Zea genus (Lisch, 2002). Jittery, AtMu1, TED, and Os3378 all encode a single protein; except Jittery, whose reinsertion was not observed, both AtMu1 and TED can induce excision and reinsertion, even though no MURB-related sequence was found (Singer et al., 2001; Xu et al., 2004; Gao, 2012; Li et al., 2013). Similarly, Os3378 is also capable of inducing excision and reinsertion without a MURB-like protein. From a maximum parsimony point of view, it is likely that the requirement of MURB for reinsertion was derived later for the Mu system in maize and that the ancient MULE may only encode a single protein (i.e., the transposase).

**Comparison of Target Specificity between Rice and Yeast**

For the TSDs of Os3378 in rice, the average TA content is ~89%, which is similar to that of AtMu1 in Arabidopsis (89%) (Singer et al., 2001). This preference is also shown by Os3378 in yeast, where the average TA content of the TSDs is 87%. Considering the genome-wide average TA content (rice, ~56%
Figure 6. Cellular Localization and Activity of N- or C-Terminal EYFP-Tagged Transposases.

(A) Relative excision frequency induced by different forms of Os3378-Z on 2% galactose. SE values were calculated from at least 20 replicates for each form of the transposase.

(B) Cellular localization of EYFP with and without the Os3378-Z amino acids 105 to 130.

(C) Cellular localization of various forms of Os3378-Z with EYFP tag at the N termini (left panel) and C termini (right panel). DAPI is a fluorescent stain that binds to DNA.
In maize, large collections of transposon insertion lines have been developed using MuDR/Mu lines, providing materials for reverse genetics studies (Fernandes et al., 2004; Settles et al., 2007). Unlike Os3378, Mu elements preferentially target GC-rich sequences (Liu et al., 2009). With highly active forms generated in this study and different sequence specificity, Os3378 may serve as an additional tagging tool to complement MuDR/Mu.

### The Effect of Element Size and TSD on Transposition Efficiency

Nonautonomous DNA elements outnumber their autonomous counterparts, likely due to the fact that autonomous elements are often much larger than nonautonomous elements. A previous study showed that a small defective IS256 element (166 bp) excised six times more frequently than the full-length element (1.3 kb). Likewise, we observed dramatic differences in both excision and reinsertion frequency between the two nonautonomous elements with different sizes. The sizes of the two elements used in this study correspond to two different groups of elements in plant genomes: one (469 bp) resembles the size of miniature inverted repeat TEs, and the other (1485 bp) is similar to that of Pack-MULEs. Apparently, the larger element is not competitive compared with its small counterpart (the overall amplification rate is 1:36; see last section of Results). Although one cannot rule out that the low efficiency is due to selection against large insertions or that certain sequences inside the large element have an inhibitory effect on transposition, it is likely that element size plays an important role in transposition, especially given the fact that the same phenomenon was observed for IS256. For example, some transposons form a synaptic complex prior to excision as two transposon ends on the same molecule are brought together via transposase-mediated oligomerization (Sakai et al., 1995; Savilahti et al., 1995). If this is the case, it might be more difficult to bring the two ends from a single element together during transposition if they are far apart. However, this seems to be inconsistent with the fact that Mut1, which is 1.4 kb, is highly active in some maize lines; therefore, this suggests that different element families may have distinct optimal sizes for transposition.

The role of TSD in transposition is controversial. Studies on P-elements indicated that the TSD did not play a role in forward transposition (reinsertion) (Mullins et al., 1989) and was not necessary for excision (Takasu-Ishikawa et al., 1992). Furthermore, analysis of IS256, a TE distantly related to Mutator (Eisen et al., 1994), showed that no precise excision was detected when the TSD homology was reduced from 8 to 6 bp, and it was proposed that precise excision was induced through a transposase-independent mechanism. This could be achieved...
through replication slippage, whereby the TSD is used as short homologous DNA stretches (Hennig and Ziebuhr, 2008), and 8 bp is the minimum length required for this process. In this study, we indeed observed similar reinsertion rates with or without TSD, which confirmed the notion that forward transposition was not influenced by the presence of the TSD. Nevertheless, precise excisions were still observed with 6 bp homology in the TSD, although with a significantly reduced frequency. This suggests that a perfect TSD promotes but is not required for precise excisions. Moreover, no excision was observed in the absence of Os3378-Z, which was in addition to the fact that different transposases were associated with distinct excision frequencies, indicating that transposase is indispensable for precise excisions. Moreover, the slippage theory may not apply to elements with very short TSDs, such as Mariner-like (2-bp TSD) and PIF-like (3-bp TSD) elements, which show high excision frequencies (Yang et al., 2006; Hancock et al., 2010). The mechanism underlying how TSDs impact the excision process requires further investigation.

A Critical Region for the Modification of Transposition Activity through Deletion and Substitutions

Previous studies have shown that the removal of selected regions of transposases can enhance their transposition activity. For example, deletion of the first 102 amino acids of the maize Ac transposase led to increased excision activity, while further deletion (up to 188 amino acids) aborted the transposase activity (Kunze et al., 1993). Besides Ac transposase, deletion of the first 77 amino acids of MuA (transposase of phage Mu) also

Figure 8. Reinsertions of Os3378NA.

(A) DNA gel blot analysis of reinsertions of Os3378NA from different reporter vectors.
(B) Distribution of reinsertion sites of Os3378NA with regard to genomic sequence features.
(C) Ratios of observed distribution of reinsertion sites to that of the fraction of each genomic feature in yeast (S288C).

5′ region, reinsertions upstream of the transcription start site of protein-coding genes on both sides of the reinsertion site (genes are head to head); 3′ region, Os3378NA insertions downstream of the transcription termination site of protein-coding genes on both sides of the reinsertion site (genes are tail to tail); 5′ and 3′ regions, reinsertions located upstream of the transcription start site of one gene and downstream of the transcription termination site of another (genes are head to tail); ARS, autonomously replicating sequence; intergenic, insertions within 1 kb flanking sequence of genes.
resulted in enhanced transposition activity (Kim and Morrison, 2009). The consequence of the N-terminal sequence deletion of Os3378 resembles that of Ac transposase (i.e., the effective region for deletion is located around amino acids 105 to 129). No significant alteration on transposition frequency was observed if the deletion was too short (104 amino acids) or too long (160 amino acids). As a result, it is likely that a short deletion does not significantly change the property of the transposase while a long deletion may lead to the loss of essential components of the transposition machinery.

Since deletion of the first 129 amino acids does not abort the transposition activity, the role of this region in transposition is intriguing. Within this region, a 25-amino acid peptide sequence (amino acids 105 to 129) seems to be critical for both transposition activity and cellular localization of the transposase. With this peptide, the transposase is localized in the nucleus (EYFP-Os3378-Z, EYFP-Os3378-Z-105, and EYFP-Os3378-Z-105Ala), whereas the transposase is localized in both nucleus and cytoplasm when this peptide is removed (EYFP-Os3378-Z-130). Since this peptide alone fails to direct the EYFP protein to the nucleus, it is clear that this peptide does not serve as an NLS (Figure 6B). In this case, the likelihood is that the absence of this peptide influences the conformation of the remainder of the transposase so that the NES is exposed or the NLS is buried, making it less effective in the transport process.

In addition to the impact on cellular localization, deletion or replacement of the short peptide may lead to a structural change in the DNA binding domain or the catalytic domain, resulting in a more efficient transposase. The original peptide (amino acids 105 to 129) contains many acidic residues and is hydrophilic. During protein folding, hydrophobic residues tend to form internal regions while hydrophilic residues are often located on the surface of the protein (Kyte and Doolittle, 1982). This implies that, in vivo, this peptide is negatively charged and likely located in the external regions of the protein molecule. As a result, either the negative charges carried by the peptide interfere with the function (e.g., the interaction between transposase and transposon ends) or the presence of this peptide at the surface of the protein prevents the exposure of the reaction center. Substitution of some (Os3378-Z-105Ala) or all (Os3378-Z-105Neutral) of the acidic residues with neutral ones resulted in enhanced excision frequency, whereas substitution of acidic residues with basic residues was not as effective. This suggests that the presence of any significant net charge in this region might be detrimental to transposition activity.

Protein Levels and Transposition Activities of Various Forms of the Os3378-Z Transposase

For wild-type Os3378-Z, the excision frequency is positively correlated with galactose levels in the medium, which modulate protein levels (Figure 4). This is different from the behavior of Ac transposase, for which the excision activity plateaued at very low protein levels and higher protein levels did not further increase its transposition activity (Kunze et al., 1993). This discrepancy is likely due to the fact that Ac transposase forms insoluble aggregates at high protein levels (Heinlein et al., 1994); thus, the amount of transposase in its active form may become saturated at low protein levels. By contrast, aggregation due to high protein levels was not observed for Os3378-Z except when EYFP was fused to its C terminus (see below). For Os3378-Z-130, the excision frequency is not responsive to galactose, although its protein level increased moderately as more galactose was applied. This implies that little Os3378-Z-130 protein is required to achieve a maximum level of transposition activity or that the transport of Os3378-Z-130 protein to the nucleus is the limiting step because it is located in both the nucleus and cytoplasm.

Although the protein level is related to the excision activity for Os3378-Z, it does not explain the variation of excision activity among different forms of Os3378-Z transposase, which are associated with different levels of proteins at 2% galactose concentration (Figure 7). This suggests that the N-terminal sequence likely plays a role in the expression or stability of the transposase, which is consistent with the “N-end rule” (i.e., the N-terminal amino acids influence the turnover of the protein) (Bachmair et al., 1986; Varshavsky, 1997). Nevertheless, the enhancement of transposition activity through deletion and mutation is unlikely due to the change in protein levels, since both Os3378-Z-105CFH and Os3378-Z-130CFH are associated with low levels of proteins but Os3378-Z-130CFH induced significantly more excisions than Os3378-Z-105CFH (Figure 7). Instead, the chemical/structural properties of the protein may play a major role. On the other hand, the protein level of Os3378-Z-105AlaCFH (with high activity) was much greater than that of Os3378-Z-105CFH (with low activity), so it is unclear whether the enhanced activity of Os3378-Z-105AlaCFH is due to protein level or structural/chemical properties. Collectively, the high excision activity achieved by different forms of transposase in this study could be due to varying mechanisms.

It is worth mentioning that the protein for immunoblot was collected from cells growing on medium with adenine in order to obtain sufficient cell growth, while excision had to be detected on medium without adenine. Hence, we cannot rule out the possibility that the protein dynamic for different transposases might be different between the cells on different media, and the relationship between protein level and excision frequency should be interpreted cautiously.

The Effect of EYFP Fusion on Transposition Activity and Cellular Localization

Fluorescent proteins, including EYFP, have been widely used to characterize protein trafficking, lipid metabolism, protein-protein interactions, and so forth. Adverse effects of the fluorescent protein on target proteins were reported, such as changes of cellular localization and the formation of dysfunctional protein (Wiedenmann et al., 2009). In this study, we demonstrated that the fusion of another protein (EYFP) to a transposase may improve or inhibit the transposition activity. Again, the region around amino acids 105 to 129 is critical in the alteration of transposase activity through fusion. Fusion of EYFP to the N terminus of Os3378-Z-130 largely offsets the positive effect on transposition activity through deletion (Figure 6). As discussed above, significant net charge around amino acids 105 to 129 seems to be detrimental to transposition activity. If EYFP is
fused to Os3378-Z-105, the introduction of four additional basic amino acids, together with the two basic amino acids originally present at amino acids 105 to 129, may render this region neutral at physiological pH, which may cause the enhanced transposition activity. By contrast, if EYFP is fused to Os3378-Z-130, it introduces a net charge to this region via basic amino acids, which may inhibit the transposition activity. As a result, it is possible that the fusion of EYFP accidentally alters net charges in this region, thereby altering the transposition activity of the resulting fusion protein.

Unlike N-terminal EYFP-tagged transposases, fusion of EYFP to the C termini did not change the activity of Os3378-Z-105 and Os3378-Z-130 but resulted in reduced excision frequency for Os3378-Z. This is likely due to aggregation of the transposases in the cytoplasm, resulting in reduced Os3378-Z protein in the nucleus. As discussed above, the activity of Os3378-Z is responsive to protein level, so its activity declines. By contrast, the activity of Os3378-Z-130 is not responsive to protein level, so its activity remains the same. Since EYFP protein alone did not form aggregates, as indicated by the even distribution of EYFP in yeast cells (Figure 6C), it is likely that the fusion at C termini causes structural changes in the regions/residues that trigger aggregation of the transposase.

Nevertheless, it is obvious that the fusion did not fully abolish the function of the NLS (Figure 6), in that all transposases are still associated with a considerable level of excision activity when the C terminus is tagged. Given this fact, the location of the C-terminal EYFP-tagged proteins in the cytoplasm is more likely due to aggregation, which prevents the transport of the fusion proteins into the nucleus, or blocks the exposure of the NLS, or promotes the exposure of the NES. Although the underlying mechanism remains unclear, this and other studies suggest that there are multiple ways to induce the aggregation of transposases that could readily trigger the suppression of their activity (Heinlein et al., 1994; Essers et al., 2000). This may be attributed to posttranslational control of transposon activity.

In summary, we have developed a tractable system for the study of MULE transposition. MULEs represent an important superfamily of TEs with many unique characteristics that distinguish them from other DNA TEs. Previous studies have largely focused on the genetics and evolution of MULEs, but little is known about the biochemical mechanism of their transposition. The recapitulation of MULE transposition in yeast will allow us to effectively dissect all aspects of the transposition of these elements. It will be of particular interest to explore the unique mechanisms of gene capture by MULEs and to test whether such processes are associated with transposition.

METHODS

Cloning of the Coding Sequence of Os3378

Expression of Os3378 was detected in a rice (Oryza sativa) somaclonal mutant, Z418 (Gao, 2012). These rice plants were grown in a growth chamber at 28/24°C (day/night) and a photoperiod of 14/10 h (light/dark). Two pairs of primers (Supplemental Table 2) were used to amplify the coding sequence of Os3378. Total RNA was extracted from Z418 young panicles using TRizol reagent (Invitrogen) followed by DNase treatment (Qiagen) and purified using the RNeasy Mini Kit (Qiagen). RNA (4 µg) was reverse transcribed into cDNA using the GoScript Reverse Transcription system (Promega). The cDNA was used as a template to amplify the Os3378 transcript using the Platinum Taq DNA Polymerase High Fidelity Kit (Invitrogen). PCR was performed with an initial denaturing step at 95°C for 2 min, followed by 30 cycles of 94°C for 30 s, 58°C for 30 s, and 68°C for 1 min, and a final elongation step at 68°C for 5 min. The resulting amplicons were sequenced at the Research Technology Support Facility of Michigan State University. The sequences of the two PCR products were concatenated to obtain the entire coding sequence of Os3378, which was referred to as Os3378-Z.

Construction of Reporter and Expression Constructs

All the restriction sites used in this study are shown in Supplemental Figure 1C. Sequences of all the constructs were verified by sequencing at the Research Technology Support Facility of Michigan State University.

Reporter Constructs

The initial reporter vector, pWL89a, was kindly provided by Nathan Hancock and Sue Wessler. Four reporter vectors were constructed using pWL89a; each contains an artificial nonautonomous Os3378 element (two Os3378NAAs with different sizes [469 and 1485 bp], with and without a TSD) using different pairs of primers (Supplemental Table 2). The perfect and nonperfect TSDs for Os3378NA were obtained through the addition of a 9–bp sequence immediately flanking the element, which was accomplished through incorporating the sequence in the primer. Os3378NA is composed of the TIR (196 bp) and partial subterminal sequences (71 bp for Os3378NA496 and 1158 bp for Os3378NA1485) of the Os3378 copy on chromosome 5 in cv Nipponbare (Gao, 2012), using Z418 genomic DNA as a template, and a linker sequence (6–bp EcoRI restriction site). The amplicons were cloned into TA vectors (prepared by the Jiang laboratory) and verified by sequencing. The Os3378NA was integrated into the coding region of ADE2 through an HpaI restriction site. This construct was obtained by digesting the TA vectors containing the 5’ end or 3’ end of the Os3378NA using Hpal and EcoRI followed by ligating the two released fragments to pWL89a that was digested using Hpal. To construct the reporter vector with a single terminus of Os3378, the 5’ terminal sequence was amplified with a pair of primers with Hpal restriction sites in them. The resulting amplicons and pWL89a were digested with Hpal and ligated together to form pWL89aOs3378singleTIRHpaI.

Expression Constructs for the Os3378-Z Transposase with the N-Terminal EYFP Tag

A Gateway vector, pAG423GAL-EYFP-ccdB (Addgene plasmid 14341), was used to build the expression constructs. The coding sequence of the transposase obtained above was synthesized by Genscript and transferred to an entry vector (pENTR-D-TOPO; Invitrogen) through BR recombination reaction using the Gateway BP Clonase Enzyme Mixes (Invitrogen). The final transfer of the transposase from the entry vector to the expression vector was accomplished through LR recombination reaction using the Gateway LR Clonase Enzyme Mixes (Invitrogen). The resulting construct was referred to as pAG423GAL-EYFP-Os3378-Z. In this construct, EYFP and Os3378-Z formed a fusion protein, whereby its transcription was controlled by the GAL1 promoter.

The constructs with N-terminal deleted Os3378-Z transposase were obtained by modifying pAG423GAL-EYFP-Os3378-Z. The construct was digested using SacI and SpeI. SacI recognizes a site between the EYFP and the transposase, and SpeI makes a single cut at 508 bp (170 amino acids) within the transposase (Supplemental Figure 1C). The larger
Expression Constructs for the Os3378-Z Transposase without the EYFP Tag

Expression constructs without the EYFP tag were obtained by removing the EYFP tag in the constructs prepared above. First, the constructs were digested with XmaI (5 bp upstream of the transcription start codon of EYFP) and Sacll (between EYFP and Os3378-Z) to release the EYFP tag (Supplemental Figure 1C). Second, the sticky ends of the vectors were repaired using the Quick Blunting Kit (New England Biolabs) and a self-ligation reaction was incubated at 4°C overnight. Standard bacteria (DH5α) transformation was conducted, and positive clones were identified through restriction enzyme digestion. The resulting constructs were referred to as pAG423GAL-Os3378-Z, pAG423GAL-Os3378-Z-105, pAG423GAL-Os3378-Z-130, and pAG423GAL-Os3378-Z-161, which lack the N-terminal 74, 104, 129, and 160 residues of the Os3378-Z transposase, respectively.

To prepare a construct containing EYFP with Os3378-Z-105-130 peptide (25 amino acids), pAG423GAL-EYFP-Os3378-Z was digested with Sacll (between EYFP and Os3378-Z) and XmaI (7 bp downstream of the Os3378-Z stop codon) restriction enzymes to remove Os3378-Z transposase. The vector backbone was purified. Primers containing restriction sites (Sacll in the forward primer and XmaI in the reverse primer) were used to amplify the Os3378-Z-105-130 peptide, and amplicons were digested with Sacll and XmaI followed by ligation to the purified vector fragment.

Expression Constructs for the Os3378-Z Transposase with the C-Terminal EYFP Tag

Os3378-Z transposase with the C-terminal EYFP tag was obtained by transferring the EYFP sequence from the N terminus to the C terminus of the transposase in each construct. The construction of pAG423GAL-Os3378-Z-EYFP (wild-type transposase with C-terminal EYFP) was used as an example to explain the detailed procedures. To initialize the construction, pAG423GAL-EYFP-Os3378-Z was digested using XmaI and Sacll to create the EYFP sequence and the rest of the construct. A filler DNA was used to seal the staggered ends of the construct and prevent the recovery of the XmaI site (Supplemental Table 2). The vector DNA was self-ligated, which resulted in pAG423GAL-Os3378NoXmaI. The original stop codon at the end of the transposase was removed, and new XmaI and Sacll sites were created downstream of the transposase by amplifying a segment of the transposase (2471 to 2658 bp) with a forward primer specific to 2471 to 2505 bp with an EcoRV site and a reverse primer complementary to the last 27 bp of the transposase (no stop codon, with an additional linker containing XmaI, Sacll, and Ascll sites). The resulting PCR product and pAG423GAL-Os3378NoXmaI were digested using EcoRV and Ascll following by ligating them together. The resulting sequence was digested by XmaI and Sacll, and the EYFP sequence was ligated into it (pAG423GAL-Os3378-ZCEYFP). Other C-terminal EYFP constructs (e.g., pAG423GAL-Os3378-Z-105CEYFP and pAG423GAL-Os3378-Z-130CEYFP) were prepared similarly.

Expression Constructs Containing Amino Acid Substitutions in Os3378-Z 105 to 129

Amino acid substitutions in the amino acids 105 to 129 region were made so that the resulting peptide had different chemical and physiological properties (Figure 3A). These are Os3378-Z-105Ala, with five acidic and two hydrophobic residues mutated to alanine, Os3378-Z-105Neutral, with nine acidic residues changed to neutral amino acids, and Os3378-Z-105Basic, with nine acidic residues changed to basic amino acids (Figure 3A). Codons for these amino acids were chosen according to their corresponding usage frequency in the wild-type Os3378-Z transposase. For each construct, a pair of primers was used to amplify the amino acids 105 to 170 segment. The forward primer (~110 nucleotides) contained a start codon (ATG), an alanine residue, mutated amino acids, and a Sacll restriction site, while the reverse primer contained a XmaI restriction site (Supplemental Table 2). pAG423GAL-Os3378-Z was used as the template for PCR amplification, and amplicons were digested with Sacll and XmaI. Meanwhile, the expression constructs without the EYFP tag were digested using the same restriction enzymes. The digested expression constructs and the amplicons were ligated, which resulted in three constructs, pAG423GAL-Os3378-Z-105Ala, pAG423GAL-Os3378-Z-105Neutral, and pAG423GAL-Os3378-Z-105Basic.

Expression Constructs with the C-Terminal FLAG-His6 Dual Tag

The expression constructs (without tag) generated above were digested using EcoRV and Ascll (Supplemental Figure 1C), which cut off part of the C-terminal sequence of the transposases. This sequence plus the FLAG-His6 tag was amplified using a pair of primers (Zanetti et al., 2005), with the forward primer containing an EcoRV restriction site and the reverse primer containing the FLAG-His6 sequence and an Ascll site (Supplemental Table 2). The resulting amplicons were digested with EcoRV and Ascll to generate corresponding ends, which were ligated to the digested (with EcoRV and Ascll) expression constructs.

Computational Characterization of Functional Domains of the Os3378-Z Transposase

The coding sequence of the Os3378-Z transposase was translated into protein sequence using the ExPaSy translation tool (http://web.expasy.org/translate/). NLS was predicted using the PSORT II Prediction program (http://psort.hgc.jp/form2.html); the seven-residue NLSs are presented in Figure 1B. The DNA binding domain was defined through a combination of secondary structure analysis (Jpred3; http://www.cmpbio.dundee.ac.uk/www-jpred/) and software prediction for a helix-turn-helix DNA binding motif (http://npsa-pbil.ibcp.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_hh.html). To locate the catalytic domain, multiple sequence alignment was performed on MURA-related transposases (i.e., MURA, JTA, and Os3378-Z), and the conserved amino acid triad was obtained through comparison with the defined DDE in the known transposases. NES was identified by manually checking the Os3378-Z transposase for the conserved pattern of sequence [(L-x(2,3)-[LIVFM]-x(2,3)-L-x-x]-[L]] (Bogerd et al., 1996).

Transformation of Reporter and Expression Constructs into Yeast and Selection for ADE2 Revertants

The reporter construct was first transformed into the yeast (Saccharomyces cerevisiae) haploid strain DG2523 (MATalpha ura3-167 trp1-hisG) followed by ligating them together. The resulting sequence was digested by XmaI and Sacll, and the EYFP sequence was ligated into it (pAG423GAL-Os3378-ZCEYFP). Other C-terminal EYFP constructs (e.g., pAG423GAL-Os3378-Z-105CEYFP and pAG423GAL-Os3378-Z-130CEYFP) were prepared similarly.
leu2-hisG his3-del200 ade2-hisG (provided by Nathan Hancock and Sue Wessler). Yeast colonies growing on synthetic defined medium without uracil (SD/-Ura) were verified for the presence of the reporter vector by PCR. Competent cells of the confirmed yeast transformants were prepared using the Frozen-EZ Yeast Transformation II kit (Zymo Research) followed by transforming the expression constructs. As a control, the empty expression vector was also transformed at the same time.

Transformed yeast cells were grown on SD plates lacking uracil and histidine with 3% raffinose as the carbon source. After recovery for 6 to 7 d, colonies were plated on SD medium without uracil, histidine, and adenine (SD/-Ura-His-Ade) and supplied with 3% raffinose. Yeast colonies were plated on SD/-Ura-His-Ade medium and supplied with 3% raf transposase. To measure the excision frequency, yeast colonies were suspended in 50 μL of sterile deionized water, where 49 μL was plated on SD/-Ura-His-Ade medium and 1 μL was diluted 10-fold and 49 μL was plated on YPD medium (1% yeast extract, 1% peptone, and 2% glucose) plates to obtain the total number of viable cells in each colony. Excision of Os3378NA and restoration of a functional ADE2 gene allowed the growth of yeast cells on the SD/-Ura-His-Ade medium; the surviving colonies were called ADE2 revertants. Differences of excision frequency among different forms of Os3378-Z transposase were determined using the t test in SAS/9.3.

Determination of the Sequences of the Donor Sites following Excisions

Primers flanking the insertion sites of Os3378NA in ADE2 were used to amplify the sequences of the donor sites following excisions (Supplemental Table 2). All PCR products were cloned and sequenced.

Extraction of Yeast Total Protein and Determination of Protein Levels of Transposases

Yeast colonies containing expression (those with FLAG-His$_6$ tag) and reporter constructs were streaked on SD/-Ura-His plates supplied with various galactose levels. After growing for 5 d, the cells were collected and lysed using extraction buffer (200 mM Tris, pH 8.0, 150 mM ammonium sulfate, 10% glycerol, 1 mM EDTA, 0.2 mM DTT, 10 mM phenylmethylsulfonyl fluoride, and 1× Yeast/Fungal Protease Arrest [BD Biosciences]). Glass beads were added into the mixture followed by vigorous vortexing to break the cells. Centrifugation was used to precipitate cell debris, and the supernatant contained the total crude proteins. For transposases whose transcription was induced by various galactose concentrations, 2.5 μg of total protein extract for Os3378-Z and 5 μg for Os3378-Z-130 was used for immunoblot analysis. For all forms of transposases on 2% galactose, 2.5 μg of protein extract was used. To ensure the usage of similar amounts of total protein, the protein extracts were separated on a 4 to 10% SDS-PAGE gel followed by transfer to a polyvinylidene difluoride membrane (Millipore). The membrane was first probed using the first antibody, THE His Tag Antibody (Mouse) (Genscript), and then the second antibody, anti-mouse IgG, horseradish peroxidase-conjugated (Cell Signaling Technology). Signal detection was performed using the SuperSignal West Pico Chemiluminescent Substrate (Thermo Scientific) and exposed to an x-ray film. Relative protein levels (relative density) of various forms of transposases were quantified using ImageJ software (http://imagej.nih.gov/ij/).

Determination of the Cellular Localization of Os3378-Z Transposases

The cellular localization of the transposases was determined using confocal microscopy in the Center for Advanced Microscopy at Michigan State University. Yeast cells containing both expression (those with EYFP tag) and reporter constructs were cultured in SD/-Ura-His medium containing 2% galactose at 30°C and 250 rpm shaking overnight. The culture was diluted to ~0.4 OD$_{600}$ units and incubated at 30°C at 250 rpm for 3 h before collecting the cells by centrifugation (3000 rpm). Cells were washed using 1× PBS buffer (137 mM NaCl, 2.7 mM KCl, 10 mM Na$_2$HPO$_4$, and 2 mM KH$_2$PO$_4$) and pelleted before treating cells with 1× PBS containing 1 μg/mL β-D API, a fluorescent stain for the nucleus. Confocal images were acquired using the Olympus FluoView FV1000 laser scanning confocal microscope configured on an automated IX81 inverted microscope with a 100× UPLSAPO (numerical aperture 1.4) oil objective. EYFP fluorescence was excited with the 515-nm argon laser line, while emission was collected with a 535- to 565-nm band-pass filter. Nuclei were counterstained blue with DAPI excited with a 405-nm diode laser, while emission was collected with a 430- to 470-nm band-pass filter.

Analysis of Reinsertions of Os3378NA following Excisions

DNA gel blot assay was conducted to determine reinsertions of Os3378NA after excisions. For each combination of Os3378-Z transposase and a nonautonomous element, at least 19 yeast colonies were tested. Genomic DNA of ADE2 revertant yeast colonies was extracted using Zymolase (Zymo Research) digestion of yeast cell wall, phenol:chloroform:isoamyl alcohol, and the isopropanol precipitation method. Genomic DNA (300 ng) was digested with EcoRI and XhoI, where the former recognizes a sequence within Os3378NA. The digested DNA was resolved on a 1% agarose gel for 3 h, followed by transferring DNA to a nylon membrane (GE Healthcare) using capillary flow. The digoxigenin-labeled probe specific to the 5′ TIR of Os3378 was used to bind to the fragments of Os3378NA, which was detected by anti-digoxigenin-AP, Fab fragments (Roche Applied Science).

To determine the reinsertion sites, transposon display was conducted (Van den Broeck et al., 1998). Specifically, genomic DNA of ADE2 revertant yeast colonies was digested using HinP1II followed by ligating adapters to these genomic fragments. Nested PCRs were conducted using one primer specific to the TIR sequence of Os3378 and another specific to the adapter sequence, which resulted in amplicons consisting of part of the TIR and flanking sequences of the reinsertion sites. The resulting products were resolved on a polyacrylamide gel, and polymorphic fragments were recovered by PCR and sequenced. The flanking sequences were mapped to the yeast genome (S288C, release R64-1-1; http://yeastgenome.org/), and the reinsertion sites were determined with regard to the closest genomic features (i.e., genes; rRNAs based on the yeast genome (S288C, release R64-1-1; http://yeastgenome.org/), and the reinsertion sites were determined with regard to the closest genomic features (i.e., genes; rRNAs based on the file containing genomic sequence features, coordinates, and annotations, which was downloaded in February 2011 from http://yeastgenome.org/). Insertions that are within 1 kb from genes were considered to be in flanking regions. If the insertions are between two genes that were in the opposite transcription orientations (head-to-head or tail-to-tail), the insertions were considered to be in either the 5′ or 3′ region. Otherwise, they were assigned to the group of 5′ and 3′ regions (head-to-tail or tail-to-head).

Accession Numbers

The GenBank accession number for the cDNA sequence of Os3378-Z transposase is KJ865693.

Supplemental Data

Supplemental Figure 1. Os3378-Z Transposase.

Supplemental Figure 2. Relative Excision Frequency as a Function of Protein Levels (Relative Density of Hybridization Signals) of Various Forms of Os3378-Z Transposase with C-Terminal FLAG-His$_6$ Tag (GFP).
**Supplemental Figure 3.** Fractions of Genomic Sequence Features of the Yeast (S288C).

**Supplemental Table 1.** Reinsertion Sites.

**Supplemental Table 2.** Primers Used in This Study.

**ACKNOWLEDGMENTS**

We thank Sue Wessler and Nathan Hancock (University of Georgia) for providing the vector (pWL89a), Cornelius Barry, Stefan Cerbin, Rebecca Grumet, Shihuan Shiu, and Dechun Wang (Michigan State University) for critical reading of the manuscript, and Jason Miller (Pennsylvania State University) and Dongmei Yin (Michigan State University) for providing technical assistance. This study was supported by the National Science Foundation (Grants DBI-0607123 and MCB-1121650).

**REFERENCES**


Hancock, C.N., Zhang, F., and Wessler, S.R. (2010). Transposition of the Tourist-MITE mPing in yeast: An assay that retains key features of catalysis by the class 2 PIF/Harbinger superfamily. Mob. DNA 1: 5.


Transposition of a Rice Mutator-Like Element in the Yeast *Saccharomyces cerevisiae*
Dongyan Zhao, Ann Ferguson and Ning Jiang
*Plant Cell*; originally published online January 13, 2015;
DOI 10.1105/tpc.114.128488

This information is current as of June 23, 2017

| Supplemental Data | /content/suppl/2015/01/12/tpc.114.128488.DC1.html  
/content/suppl/2015/02/04/tpc.114.128488.DC2.html |
|-------------------|----------------------------------------------------|
| eTOCs             | Sign up for eTOCs at:  
http://www.plantcell.org/cgi/alerts/ctmain |
| CiteTrack Alerts  | Sign up for CiteTrack Alerts at:  
http://www.plantcell.org/cgi/alerts/ctmain |
| Subscription Information | Subscription Information for *The Plant Cell* and *Plant Physiology* is available at:  
http://www.aspb.org/publications/subscriptions.cfm |

© American Society of Plant Biologists
ADVANCING THE SCIENCE OF PLANT BIOLOGY