Breeding Science 59: 297–307 (2009)

Analysis of distribution and proliferation of *mPing* family transposons in a wild rice (*Oryza rufipogon* Griff.)

Shanta Karki[1], Takuji Tsukiyama[1], Yutaka Okumoto[1], Govinda Rizal[1], Ken Naito[2], Masayoshi Teraishi[1], Tetsuya Nakazaki[1] and Takatoshi Tanisaka*[1]

[1] *Laboratory of Plant Breeding, Graduate School of Agriculture, Kyoto University*, Oiwake, Kitashirakawa, Sakyo, Kyoto 606-8502, Japan

[2] *Department of Plant Biology, University of Georgia*, Athens, GA 30602, USA

The distribution and proliferation among *Oryza* species of miniature *Ping* (*mPing*), the only miniature inverted-repeat transposable element (MITE) that is active in all organisms, were investigated through analysis of 78 accessions encompassing nine of the 10 genomic constitutions of the genus *Oryza*. The *mPing* family transposons, *mPing*, *Ping*, and *Pong*, were detected in only two of the AA genome species, *Oryza sativa* and its direct ancestor *O. rufipogon*, whereas the inactive rice MITE *Kiddo* was detected in the AA, BB, CC, BBCC and CCDD genome species. Furthermore, the insertion sites and copy numbers of *mPing* were considerably different among *O. rufipogon* accessions. A phylogenetic analysis showed that the *O. rufipogon* accessions used in this study could be grouped into four clusters which were not associated with the copy number of *mPing* or the presence of *Ping*. These results suggest that the *mPing* family was present in the ancestor of *O. rufipogon* and that *mPing* proliferated independently in each accession. Based on these results, one possible evolutionary history of the distribution and proliferation of *mPing* family in *O. rufipogon* is discussed.

Key Words: *mPing*, wild rice, *Oryza rufipogon*, transposable elements.

Introduction

Transposable elements (TEs) are mobile DNA sequences which can change their positions in the genome by inserting themselves into new sites. TEs are classified into two classes according to the mode of propagation: retrotransposons (class I elements) and DNA transposable elements (transposons in a narrow sense, class II elements). While the former transposes through RNA intermediate by the action of reverse transcriptase, the latter transposes in a DNA form through a cut-and-paste mechanism. For a long time, TEs were dismissed as selfish DNA. Recently, however, TEs have been revealed as major players in genomic evolution, because they cause genome rearrangements and alter the structure and regulation of individual genes (Bennetzen 2000, Biemont and Vieira 2006). Most eukaryotic genomes, in fact, have been shaped by amplification and dispersion of TEs (Le *et al.* 2000).

In rice (*Oryza sativa* L.), TEs account for at least 35% of the genome (International Rice Genome Sequencing Project 2005). The numerically predominant type of TE is the miniature inverted-repeat transposable elements (MITEs) (Bureau *et al.* 1996, Mao *et al.* 2000, Turcotte *et al.* 2001, Feng *et al.* 2002, Jiang *et al.* 2004). MITEs are reminiscences of non-autonomous DNA transposons that are distinguished from other TEs because of their small size (<600 bp) and presence of short terminal inverted repeats (TIRs). For over 300 protein-coding genes in rice, coding sequences, poly(A) sites, transcription start sites, and splicing sites overlap with MITEs (Oki *et al.* 2008). Moreover, 4069 of the rice genes contain MITE insertions within a 0.5 kb upstream region (Oki *et al.* 2008). These facts suggest that MITEs have significantly contributed to rice genome evolution.

Miniature *Ping* (*mPing*) is the first active MITE discovered in the rice genome (Jiang *et al.* 2003, Kikuchi *et al.* 2003, Nakazaki *et al.* 2003). Because *mPing* is a 430 bp sequence with 15 bp TIRs which lacks an open reading frame (ORF), its mobilization depends on the transposase encoded by the ORF2 in the autonomous elements *Ping* and *Pong* (Jiang *et al.* 2003, Kikuchi *et al.* 2003, Yang *et al.* 2007). Both *Ping* and *Pong* also have the ORF1, in addition to the ORF2, coding a protein with weak similarity to the DNA-binding domain of *myb* transcription factor (Jiang *et al.* 2003). Although the copy number of *mPing* is low compared with those of other characterized MITEs in plants (Feschotte *et al.* 2002), it varies among the rice species; temperate *japonica* cultivars possess large numbers of *mPing* (≈50 copies), whereas tropical *japonica* and *indica* cultivars possess small numbers of *mPing* (<10 copies) (Jiang *et al.* 2003, Kikuchi *et al.* 2003). Moreover, *mPing* are preferentially found in single-copy regions, and *mPing* insertions into promoter regions directly affect gene regulation and expression

---

Communicated by R. Ohsawa
Received April 30, 2009. Accepted July 31, 2009.
*Corresponding author (e-mail: tanisaka@kais.kyoto-u.ac.jp)

level (Naito *et al.* 2006). Based on these observations, *mPing* family transposons, which consists of *mPing*, *Ping* and *Pong*, have been considered to play important roles in the genome evolution of *O. sativa*. However, very little is known about the evolutionary history of the *mPing* family transposons in *O. rufipogon*, the direct ancestor of *O. sativa*, although the *mPing* family is known to be present not only in *O. sativa* but also in *O. rufipogon* (Hu *et al.* 2006).

Here we show the insertional polymorphisms and the diversity in copy number of *mPing* in *O. rufipogon* accessions. Furthermore, we describe a possible evolutionary history of the *mPing* family in *O. rufipogon*, based on a phylogenetic analysis using nucleotide sequences around *qSH1* (Konishi *et al.* 2006), which is the major quantitative trait locus (QTL) responsible for seed shattering.

Materials and Methods

Plant materials

A total of 78 accessions of 19 wild *Oryza* species together with two *O. sativa* varieties (ssp. *japonica* variety Nipponbare and ssp. *indica* variety IR36) were used in all experiments. In phylogenetic analysis, *O. sativa* ssp. *indica* variety Kasalath was used in addition to the above 78 varieties/accessions. Wild rice accessions were provided from the National Institute of Genetics. All *O. rufipogon* accessions used in this study were selected from core collection rank 1, 2, and 3, because these collection covered typical species-specific phenotypes and wide distribution in the habitat. AA genome species in the core collections are distinguished by SINE code which is the criteria of species based on the presence or absence of retrotransposon p-SINE insertions (Ohtsubo *et al.* 2004). The genomic constitution, species name, habitat and SINE code of each accession are presented in Table 1.

DNA extraction

Genomic DNA was extracted from young leaves, either by the modified cetyltrimethylammonium bromide (CTAB) method of Murray and Thompson (1980) or with a DNeasy Plant Mini Kit (Qiagen, Tokyo, Japan) according to the manufacturer's protocol.

PCR amplification

To determine whether *mPing* was present in each accession, PCR was carried out using the following two *mPing*-specific primer pairs: *mPing*-F1 (5′-GGCCAGTCACAATG GGGGTT-3′) and *mPing*-R1 (5′-GGCCAGTCACAATGG CTAGTG-3′), and *mPing*-F2 (5′-GGGATGAGAGAGAAG GAAAGAG-3′) and *mPing*-R2 (5′-AACAATCCCCACA GTGGAG-3′). The *mPing*-F1 and *mPing*-R1 were designed on the 5′ and 3′ ends of *mPing*, respectively, and thus to amplify the whole *m-Ping* sequence (430 bp), whereas the *mPing*-F2 and *mPing*-R2 were designed in the internal regions of *mPing*.

To determine whether *Ping* and *Pong* were present in each genome, we used the following specific primer pairs designed to complement *Ping*-ORF1, *Ping*-ORF2, *Pong*-ORF1, and *Pong*-ORF2, respectively: *Ping*-ORF1-F1 (5′-ACAAGCGGATACTCCGAC-3′) and *Ping*-OFR1-R1 (5′-GAGGACAATGCCTTCCATAAC-3′); *Ping*-ORF2-F1 (5′-AAGGTATTTGCTAAGTCATAT-3′) and *Ping*-ORF2-R1 (5′-ATGAGCCAAACGATCTCTCATACTA-3′); *Pong*-ORF1-F1 (5′-TCTACTCCACCACCAACACC-3′) and *Pong*-ORF1-R1 (5′-TCAGCCTTGTTTTTGTCCTTC-3′), and *Pong*-ORF2-F1 (5′-CTCATACGAAGACCTCCTCC-3′) and *Pong*-ORF2-R1 (5′-CGCCTAAGATACCTCTCACC-3′).

To detect the MITE *Kiddo*, the following primer sequences were used: *Kiddo*-F (5′-GGGGCTGTTTGGTTCC CAGCCA-3′) and *Kiddo*-R (5′-TTTGGTTGCAAGCTAC ACTTTG-3′).

The genomic DNA of each accession was used as the template for PCR with Ex-*Taq* DNA polymerase (TaKaRa, Shiga, Japan). The PCR conditions consisted of predenaturation for 3 min at 94°C followed by 30 cycles of polymerization reaction, each consisting of a denaturation step for 10 s at 98°C, an annealing step for 45 s at 60°C, and an extension step for 45 s at 72°C, with a final extension step for 7 min at 72°C. PCR products were visualized under UV light following electrophoresis on 1.5% agarose gel in 0.5x TBE buffer.

Transposon display

The copy numbers of *mPing* and *Kiddo* were determined by transposon display (TD) according to the protocols of Casa *et al.* (2000) and Takata *et al.* (2005), respectively. The genomic DNA of each *O. rufipogon* accession, Nipponbare, and IR36 was digested with *Mse*I or *Msp*I, and ligated to an adapter (*mPing*-Adapter-top: 5′-GACGATGAGTCCTGAG-3′ and *mPing*-Adapter-bottom: 5′-TACTCAGGACTCAT-3′, and *Kiddo*-Adapter-top: 5′-GTAATACGACTCACTA TAGGGCACGCGTGGGTGCTTGATGCTTGAAAA-3′ and *Kiddo*-Adapter-bottom: 5′-CGTTTTCAAGCATCAA-aminolink-3′).

Primary PCR was carried out using an adapter-specific primer (*Mse*I-1: 5′-GACGATGAGTCCTGAGTAA-3′ or *Msp*I-1: 5′-GTAATACGACTCACTATAGGGC-3′) and an MITE-specific primer (*mPing*-P1: 5′-AATGTGCATGACA CACCAG-3′ or *Kiddo*-P1: 5′-CAAGTAATGGTAAAGTG TGGCTGGG-3′). Secondary PCR was carried out using an adapter-specific primer (*Mse*I-1 or *Msp*I-2: 5′-GTGCTTGA TGCTTGAAAACGG-3′) and another MITE-specific primer (*mPing*-P2: 5′-CAGTGAAACCCCCATTGTGAC-3′ or *Kiddo*-P2: 5′-AATGGTAAAGTGTGGCTGGGAAC-3′) labeled with TAMRA fluorescence reagent.

Primary PCR conditions were as follows: predenaturation for 2 min at 72°C and subsequent predenaturation for 3 min at 94°C, followed by 30 cycles of polymerization reaction, each consisting of a denaturation step for 30 s at 95°C, an annealing step for 45 s at 58°C, and an extension step for 45 s at 72°C, with a final extension step for 7 min at 72°C. PCR products were diluted between 20-

Table 1. List of the plant materials used in the present study

S.N.	Genome	Acc. no.	Species	Country	SINE Code
1	AA	Nipponbare	*O. sativa (japonica)*	Japan	
2	AA	IR36	*O. sativa (indica)*	Philippines	
3	AA	W0106	*O. rufipogon*	India	AA110000
4	AA	W0120	*O. rufipogon*	India	AA110000
5	AA	W1294	*O. rufipogon*	Philippines	AA110000
6	AA	W1866	*O. rufipogon*	Thailand	AA110000
7	AA	W1921	*O. rufipogon*	Thailand	AA110000
8	AA	W2003	*O. rufipogon*	India	AA110000
9	AA	W1625	*O. meridionalis*	Australia	AA100010
10	AA	W1635	*O. meridionalis*	Australia	AA100010
11	AA	W0652	*O. barthii*	Sierra Leone	AA101000
12	AA	W1588	*O. barthii*	Cameroon	AA101000
13	AA	W1169=W1165	*O. glumaepatula*	Cuba	AA101100
14	AA	W2145	*O. glumaepatula*	Brazil	AA101100
15	AA	W2199	*O. glumaepatula*	Brazil	AA101100
16	AA	W1413	*O. longistaminata*	Sierra Leone	AA100001
17	AA	W1508	*O. longistaminata*	Madagascar	AA100001
18	BB	W1514	*O. punctata*	Kenya	
19	BBCC	W1024	*O. punctata*	Ghana	
20	BBCC	W1213	*O. minuta*	Philippines	
21	BBCC	W1331	*O. minuta*	Philippines	
22	CC	W1527	*O. eichingeri*	Uganda	
23	CC	W1805	*O. eichingeri*	Sri Lanka	
24	CC	W0002	*O. officinalis*	Thailand	
25	CC	W1361	*O. officinalis*	Malaysia	
26	CC	W1830	*O. officinalis*	Unknown	
27	CCDD	W0017	*O. alta*	Surinam	
28	CCDD	W1182	*O. alta or O. latifolia*	British Guinea	
29	CCDD	W0613	*O. grandiglumis*	Brazil	
30	CCDD	W1194	*O. grandiglumis*	Brazil	
31	CCDD	W2220	*O. grandiglumis*	Brazil	
32	CCDD	W1166	*O. latifolia*	Mexico	
33	CCDD	W1197	*O. latifolia*	Colombia	
34	CCDD	W2200	*O. latifolia*	Brazil	
35	EE	W0008	*O. australiensis*	Australia	
36	EE	W1628	*O. australiensis*	Australia	
37	FF	W1401	*O. brachyantha*	Sierra Leone	
38	FF	W1711	*O. brachyantha*	Cameroon	
39	GG	W0003	*O. granulata*	India	
40	GG	W0067(B)	*O. granulata*	Thailand	
41	GG	W1356	*O. meyeriana*	Malaysia	
42	HHJJ	W1220	*O. longiglumis*	Dutch New Guinea	
43	HHJJ	W0001	*O. ridleyi*	Thailand	
44	HHJJ	W0604	*O. ridleyi*	Malaya	
45	AA	W0630	*O. rufipogon*	Burma	AA110000
46	AA	W1236	*O. rufipogon*	Australian New Guinea	AA110000
47	AA	W1807	*O. rufipogon*	Sri Lanka	AA110000
48	AA	W1945	*O. rufipogon*	China	AA110000
49	AA	W2051	*O. rufipogon*	Bangladesh	AA110000
50	AA	W2078	*O. rufipogon*	Australia	AA110000
51	AA	W2263	*O. rufipogon*	Cambodia	AA110000
52	AA	W0107	*O. rufipogon*	India	AA110000
53	AA	W0108	*O. rufipogon*	India	AA110000
54	AA	W0137	*O. rufipogon*	India	AA110000
55	AA	W0180	*O. rufipogon*	Thailand	AA*10000
56	AA	W0593	*O. rufipogon*	Malaya	AA110000
57	AA	W0610	*O. rufipogon*	Burma	AA110000
58	AA	W1230	*O. rufipogon*	Dutch New Guinea	AA110000

Table 1. (continued)

S.N.	Genome	Acc. no.	Species	Country	SINE Code
59	AA	W1235	*O. rufipogon*	New Guinea	AA100010
60	AA	W1238	*O. rufipogon*	New Guinea	AA110000
61	AA	W1239	*O. rufipogon*	New Guinea	AA100010
62	AA	W1551	*O. rufipogon*	Thailand	AA110000
63	AA	W1666	*O. rufipogon*	India	AA110000
64	AA	W1669	*O. rufipogon*	India	AA110000
65	AA	W1681	*O. rufipogon*	India	AA110000
66	AA	W1685	*O. rufipogon*	India	AA110000
67	AA	W1690	*O. rufipogon*	Thailand	AA110000
68	AA	W1715	*O. rufipogon*	China	AA110000
69	AA	W1852	*O. rufipogon*	Thailand	AA110000
70	AA	W1865	*O. rufipogon*	Thailand	AA110000
71	AA	W1939	*O. rufipogon*	Thailand	AA110000
72	AA	W1981	*O. rufipogon*	Indonesia	AA110000
73	AA	W2014	*O. rufipogon*	India	AA110000
74	AA	W2080	*O. rufipogon*	Australia	AA100010
75	AA	W2109	*O. rufipogon*	Australia	AA110000
76	AA	W2265	*O. rufipogon*	Laos	AA110000
77	AA	W2266	*O. rufipogon*	Laos	AA110000
78	AA	W2267	*O. rufipogon*	Laos	AA110000

and 50-fold with 1/10 TE buffer and used as templates for secondary PCR (selective amplification). Selective amplification cycling started with pre-denaturation for 3 min at 94°C followed by 7 cycles of touchdown PCR. The first touchdown cycle consisted of a denaturation step for 30 s at 95°C, an annealing step for 45 s at 64°C, and an extension step for 45 s at 72°C; starting with the second cycle, temperature at the annealing step was lowered by 1.0°C per cycle until it reached 58°C. Touchdown PCR was followed by 30 cycles of polymerization reaction, each consisting of a denaturation step for 30 s at 95°C, an annealing step for 45 s at 58°C, and an extension step for 45 s at 72°C, with a final extension step for 7 min at 72°C.

Phylogenetic analysis

Nucleotide polymorphisms around *qSH1*, a major quantitative trait locus for seed shattering in rice (Konishi *et al.* 2006), were investigated in Nipponbare, IR36, Kasalath, and *O. rufipogon* accessions using the following four primer pairs: Con24-13U: 5′-CAATGGAAAAGCTGATG-3′ and Con24-13L: 5′-CGTTGCATGAATTGTAGCAC-3′, RBEL-E1-U: 5′-ATCATGCAGCAAGTGACCAC-3′ and RBEL-E1-L2: 5′-TCACAACCTAGAGATGAGGC-3′, 106.5k-U: 5′-CAATACATGCATGGATGCGT-3′ and 107.2k-L: 5′-TCCATATGTGTGTAGGACCC-3′, and 108.9k-U: 5′-ACAGGGTGATCCCAACAGTT-3′ and 109.8k-L: 5′-TAACCGGTGATGGTTGTGCA-3′. The PCR conditions consisted of pre-denaturation for 3 min at 94°C followed by 30 cycles of polymerization reaction, each consisting of a denaturation step for 10 s at 98°C, an annealing step for 45 s at 55–58°C, and an extension step for 1 min at 72°C, with a final extension step for 7 min at 72°C. PCR products were sequenced using ABI PRISM DNA sequencing system (Applied Biosystems, Foster City, CA). After the sequence data obtained from the four primer pairs were combined, the sequence alignment was determined using the online computer program CLUSTAL W (http://align.genome.jp/). A non-rooted phylogenetic tree was constructed using MEGA version 3.1 (Kumar *et al.* 2004). The neighbor-joining method (Saitou and Nei 1987) was conducted with Kimura's 2-parameter distances (Kimura 1980).

Results

mPing is a specific element in O. sativa and O. rufipogon

Recently, the *mPing* family was shown to be distributed only in the cultivated rice *O. sativa* and the wild rice *O. rufipogon* (Hu *et al.* 2006). Because its distribution and proliferation have not previously been substantially discussed, we set out to investigate the distribution of *mPing* among wild rice species by applying PCR to 78 accessions covering nine of the ten genomic constitutions of rice. In the PCR analysis using *mPing*-F1 and *mPing*-R1 primers which were designed on the 5′ and 3′ ends of *mPing*, respectively, only *O. sativa* and *O. rufipogon* accessions yielded PCR products with almost the same size as *mPing* (430 bp), whereas other species yielded products far longer than 430 bp (Fig. 1). Sequencing analysis showed that the PCR products with sequences of approximately 430 bp all shared very high similarity (98–100%) with the *mPing* sequence. In previous studies, the sequences nearly identical to *mPing* but with a few nucleotide polymorphisms have been regarded as subtypes of *mPing* (Jiang *et al.* 2003, Kikuchi *et al.* 2003, Hu *et al.* 2006). We also concluded that the PCR products of approximately 430 bp obtained when using *mPing*-F1 and *mPing*-R1 primers were subtypes of *mPing*,

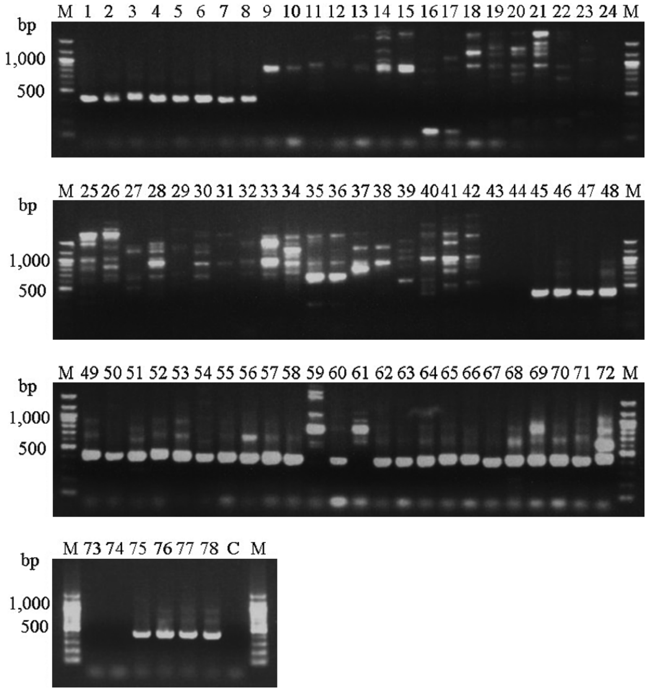

Fig. 1. Determining the presence or absence of *mPing* in wild *Oryza* species by PCR analysis using *mPing*-F1 and *mPing*-R1 primers. M, DNA size marker (100 bp ladder, Nacalai, Kyoto, Japan); Lanes 1–2, *Oryza sativa*; lanes 3–8, *O. rufipogon*; lanes 9–10, *O. meridionalis*; lanes 11–12, *O. barthii*; lanes 13–15, *O. glumaepatula*; lanes 16–17, *O. longistaminata*; lanes 18–19, *O. punctata*; lanes 20–21, *O. minuta*; lanes 22–23, *O. eichingeri*; lanes 24–26, *O. officinalis*; lanes 27–28, *O. alta*; lanes 29–31, *O. grandiglumis*; lanes 32–34, *O. latifolia*; lanes 35–36, *O. australiensis*; lanes 37–38, *O. brachyantha*; lanes 39–40, *O. granulate*; lane 41, *O. meyeriana*; lane 42, *O. longiglumis*; lanes 43–44, *O. ridleyi*; lanes 45–78, *O. rufipogon*; lane C, negative control (distilled water). Lane numbers represent the serial numbers of the accessions as listed in Table 1.

because they were not completely identical to the *mPing* sequence. Similar results were obtained in the PCR analysis with *mPing*-F2 and *mPing*-R2 primers, which were designed at the internal position of *mPing* (data not shown). It was therefore confirmed that *mPing* is deployed only in *O. sativa* and *O. rufipogon* as reported by Hu *et al*. (2006). Four (W1235, W1239, W2014, and W2080) of the 40 *O. rufipogon* accessions did not yield any PCR products corresponding to *mPing*, indicating that these accessions do not harbor *mPing* in their genomes. According to the morphological characteristics, these four accessions are classified as *O. rufipogon*. Nevertheless, three of them, W1235, W1239, and W2080, have been assigned the SINE code of *O. meridionalis* (Table 1). Taken together with our results, we consider them to have been misclassified as *O. rufipogon*. We have treated these three accessions as *O. meridionalis* in our further analyses.

Distribution of the autonomous elements Ping and Pong in genus Oryza

mPing has been thought to be a deletion derivative of *Ping*, because the terminal sequences of *mPing* are identical to those of *Ping* except for a single base-pair mismatch. It is known, however, that not only *Ping* alone but also *Pong* alone can mobilize *mPing* (Jiang *et al.* 2003, Kikuchi *et al.* 2003, Yang *et al.* 2007). This observation underscores the necessity of investigating whether the wild species with *mPing* also possess *Ping* and/or *Pong* elements.

We conducted PCR analyses with *Ping*- and *Pong*-specific primer pairs. Among 37 *O. rufipogon* accessions, 17, like Nipponbare (*O. sativa* ssp. *japonica*), yielded *Ping*-ORF1 (264 bp) and *Ping*-ORF2 (369 bp) as PCR products; two (W2078 and W1690) yielded only *Ping*-ORF2; and the remaining 18 yielded neither of them (Fig. 2). In PCR of *Ping*-ORF1, out of other species, *O. officinalis* (W0002, W1361, and W1830) produced a single band of around 500 bp, but sequencing of these bands confirmed that they were not the specific product of *Ping*-ORF1 (data not shown). We detected no amplified product of *Ping*-ORF2 in all species except for *O. sativa* and *O. rufipogon*. Along with Nipponbare and IR36 (*O. sativa* ssp. *indica*), all *O. rufipogon* accessions produced both *Pong*-ORF1 (500 bp) and *Pong*-ORF2 (586 bp) as PCR products (Fig. 3). On the other hand, other species yielded products far longer than 500 and 586 bp in *Pong*-ORF1 and *Pong*-ORF2, respectively. These results show that *Ping* and *Pong* are present only in *O. sativa* and *O. rufipogon*, but also that the distribution of *Ping* is limited, even within *O. rufipogon,* to certain accessions (Table 2).

Insertional polymorphism and copy number of mPing in O. rufipogon genomes

TD was employed to elucidate the insertional polymorphisms and copy numbers of *mPing* in *O. rufipogon* genomes. The banding patterns of *mPing* varied considerably among the accessions, and no bands common to all accessions were detected (Fig. 4). This indicates that there are no *mPing* insertion sites common between *O. rufipogon* accessions.

The copy number of *mPing* was estimated by counting all the bands amplified when using combinations of the *mPing*- and adapter-specific primers. The copy number of *mPing* in Nipponbare estimated by this method (51 copies) was consistent with the actual number of *mPing* in the database, suggesting that the copy numbers of *mPing* in *O. rufipogon* were also accurately evaluated by this method. The copy numbers of *mPing* in *O. rufipogon* accessions ranged from zero to 30 (Table 2). Nine accessions had less than 10 copies of *mPing*, 20 accessions had 10–20 copies, and 8 accessions had 21–30 copies. The *O. rufipogon* accession W2014 produced no band not only in PCR analysis but also in TD, suggesting that this accession harbored no *mPing* copy in the genome. The average copy numbers of *mPing* among all accessions, among accessions with *Ping*, and among accessions without *Ping* were 14.3, 14.9, and 13.7, respectively. According to Jiang *et al.* (2003), almost all *O. sativa* ssp. *japonica* varieties have ca. 50 copies. We concluded, therefore, that the copy number of *mPing* in *O. rufipogon* is lower than that in *O. sativa* ssp. *japonica*.

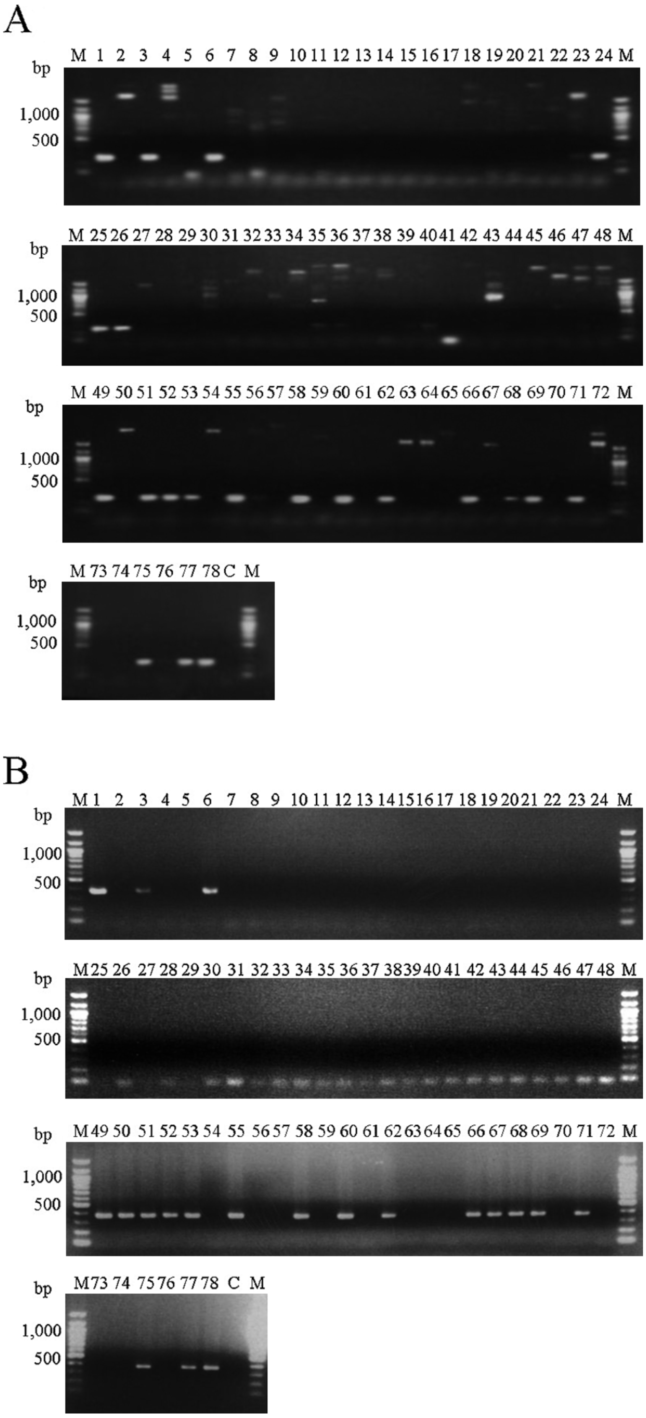

Fig. 2. PCR amplifications of *Ping*-ORF1 and *Ping*-ORF2. (A) Amplifications of *Ping*-ORF1 using *Ping*-ORF1-F1 and *Ping*-ORF1-R1 primers (expected product size, 264 bp) (B) Amplifications of *Ping*-ORF2 using *Ping*-ORF2-F1 and *Ping*-ORF2-R1 primers (expected product size, 369 bp). Lane numbers represent the serial numbers of the accessions as listed in Table 1. Lane M, DNA size marker (100 bp ladder, Nacalai); lane C, negative control (distilled water).

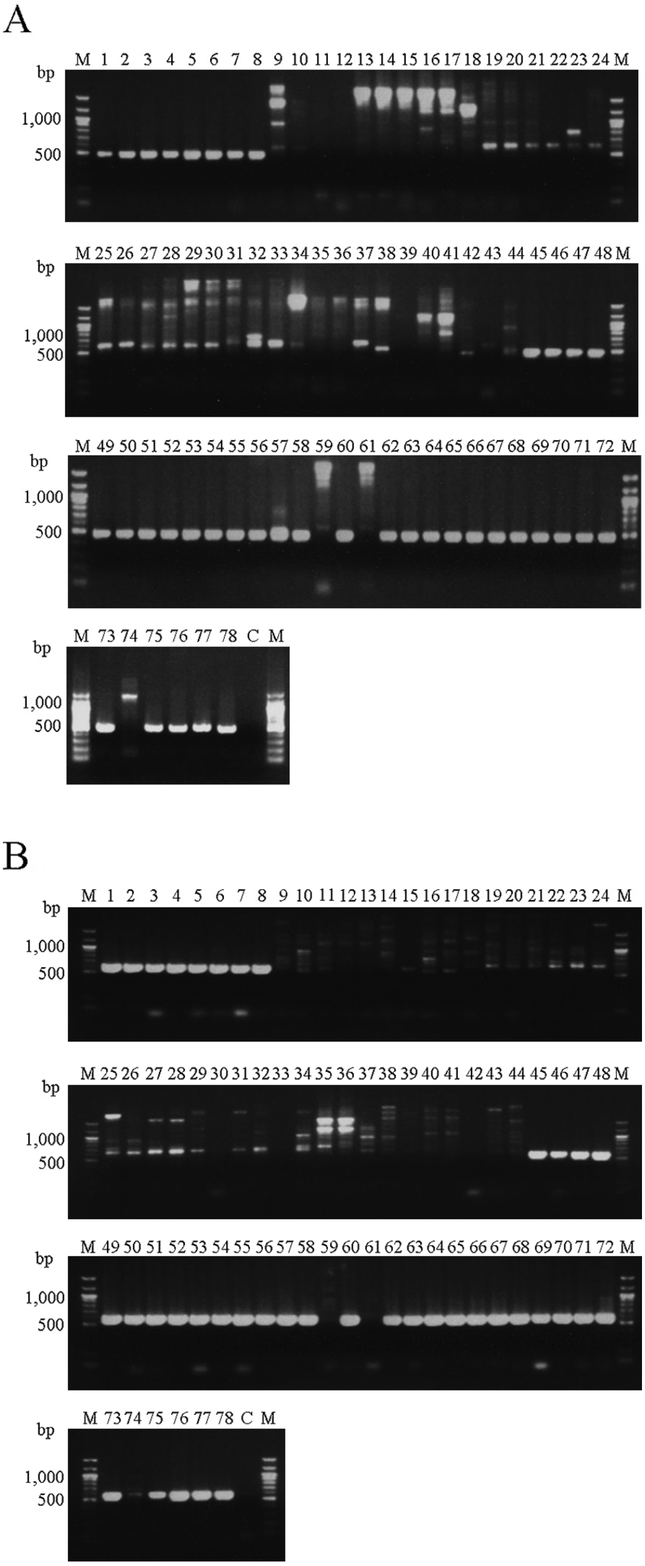

Fig. 3. PCR amplifications of *Pong*-ORF1 and *Pong*-ORF2. (A) Amplifications of *Pong*-ORF1 using *Pong*-ORF1-F1 and *Pong*-ORF1-R1 primers (expected product size, 500 bp) (B) Amplifications of *Pong*-ORF2 using *Pong*-ORF2-F1 and *Pong*-ORF2-R1 primers (expected product size, 586 bp). Lane numbers represent the serial numbers of the accessions as listed in Table 1. Lane M, DNA size marker (100 bp ladder, Nacalai); lane C, negative control (distilled water).

Distribution of Kiddo in genus Oryza

There were no *mPing* insertion site common between *O. rufipogon* accessions. To determine whether *mPing* is the

Table 2. Distribution pattern and copy number of *mPing* family in *O. rufipogon* accessions

S.N.	Acc.	*mPing*	*Ping*		*Pong*	
		Total copies	ORF1	ORF2	ORF1	ORF2
1	Nipponbare	50	+	+	+	+
2	IR36	11			+	+
3	W0106	12	+	+	+	+
4	W0120	15			+	+
5	W1294	4			+	+
6	W1866	16	+	+	+	+
7	W1921	10			+	+
8	W2003	8			+	+
9	W0630	13			+	+
10	W1236	19			+	+
11	W1807	9			+	+
12	W1945	25			+	+
13	W2051	15	+	+	+	+
14	W2078	8		+	+	+
15	W2263	21	+	+	+	+
16	W0107	13	+	+	+	+
17	W0108	28	+	+	+	+
18	W0137	14			+	+
19	W0180	6	+	+	+	+
20	W0593	16			+	+
21	W0610	15			+	+
22	W1230	16	+	+	+	+
23	W1238	11	+	+	+	+
24	W1551	21	+	+	+	+
25	W1666	13			+	+
26	W1669	18			+	+
27	W1681	5			+	+
28	W1685	6	+	+	+	+
29	W1690	6		+	+	+
30	W1715	30	+	+	+	+
31	W1852	16	+	+	+	+
32	W1865	22			+	+
33	W1939	20	+	+	+	+
34	W1981	13			+	+
35	W2014	0			+	+
36	W2109	15	+	+	+	+
37	W2265	14			+	+
38	W2266	14	+	+	+	+
39	W2267	22	+	+	+	+

+, Presence of ORF

only MITE to show this exclusive distribution pattern, we investigated the distribution and proliferation of other rice-specific MITEs. *Kiddo* is one of the rice-specific MITEs that have been identified in *O. sativa* ssp. *indica* variety (Yang *et al.* 2001). It is known that the copy number of *Kiddo* in the Nipponbare genome (42 copies), like that of *mPing* (50 copies), is lower than those of the other rice MITEs (Takata *et al.* 2007). Moreover, many of *Kiddo* elements, like *mPing* elements, were inserted in the gene regions (Yang *et al.* 2001) and showed insertional polymorphism between *japonica* and *indica* rices (Yang *et al.* 2005). For these reasons, we investigated the distribution, copy number, and insertional polymorphisms of *Kiddo* in *O. rufipogon*, although no autonomous element for *Kiddo* has been identified so far.

When the *Kiddo*-specific primer pair was used, PCR products of prospective size (269 bp) were amplified in all the AA genome species (Fig. 5). Moreover, *O. punctata* of BB and BBCC genome, *O. minuta* of BBCC genome, *O. eichingeri* and *O. officinalis* of CC genome, and *O. alta*, *O. grandiglumis* and *O. latifolia* of CCDD genome also yielded PCR products of 269 bp (Fig. 5). Sequencing analysis showed that these products shared high homology (>88%) with *Kiddo* sequence in the AA genome species. Thus it was found that *Kiddo* is present not only in the AA genome but also in the BB, CC, BBCC and CCDD genome species.

To investigate the insertional polymorphisms and copy numbers of *Kiddo* in *O. rufipogon* genomes, TD was carried out using a combination of *Kiddo*- and adapter-specific primers. The copy number was estimated by the method described above. Many bands were detected that were shared by several accessions (Fig. 6). This indicates that these accessions have *Kiddo* elements at the same chromosomal positions. The copy number of *Kiddo* ranged from 43 to 55.

Phylogenetic relationships between O. rufipogon accessions

Jiang *et al.* (2003) found that the copy number of *mPing* apparently varied between two Oryza subspecies, temperate *japonica* (≈50 copies) and *indica* (>10 copies), and suggested that *mPing* might have proliferated under extreme environmental conditions during domestication of temperate *japonica* varieties. This indicates that the distribution pattern of the *mPing* family might differ among *O. rufipogon* accession clusters, if any. We first investigated the phylogenetic relationships among *O. rufipogon* accessions using DNA sequences of four genomic regions around *qSH1* (Konishi *et al.* 2006). *qSH1* is a major QTL responsible for seed shattering that has been identified in the F_2 population of the cross between *indica* and *japonica* varieties, and nucleotide polymorphisms around *qSH1* have proven suitable for studying the evolutionary history of Asian rice, including *O. rufipogon* (Onishi *et al.* 2007). Then we carried out PCR using specific primer pairs for four regions around *qSH1* including the 5′ regulatory region of *qSH1*; consequently, the primer pairs of con24-13U and con24-13L, RBEL-E1-U and RBEL-E1-L, 106.5k-U and 107.2k-L, and 108.9k-U and 109.8k-L yielded 876, 520, 718, and 862 bp products in all accessions, respectively. A total of approximately 2600 bp of the four products was sequenced for each accession. Among the 37 accessions, there was no SNP single-nucleotide polymorphism (SNP) for seed shattering habit (Konishi *et al.* 2006) in the 5′ regulatory region of *qSH1* gene; all accessions had a shattering-type *qSH1*. In addition, we detected 12 polymorphisms including two deletions of 16 and 6 bp in addition to 12 out of 13 haplotypes reported by Konishi *et al.* (2006) (data not shown). The phylogenetic tree based on the sequence analysis showed that *O. rufipogon* accessions used in this study were divided into four major clusters (Fig. 7). Subspecies *japonica* (Nipponbare) and *indica* (IR36 and Kasalath) were classified into cluster I and II, respectively.

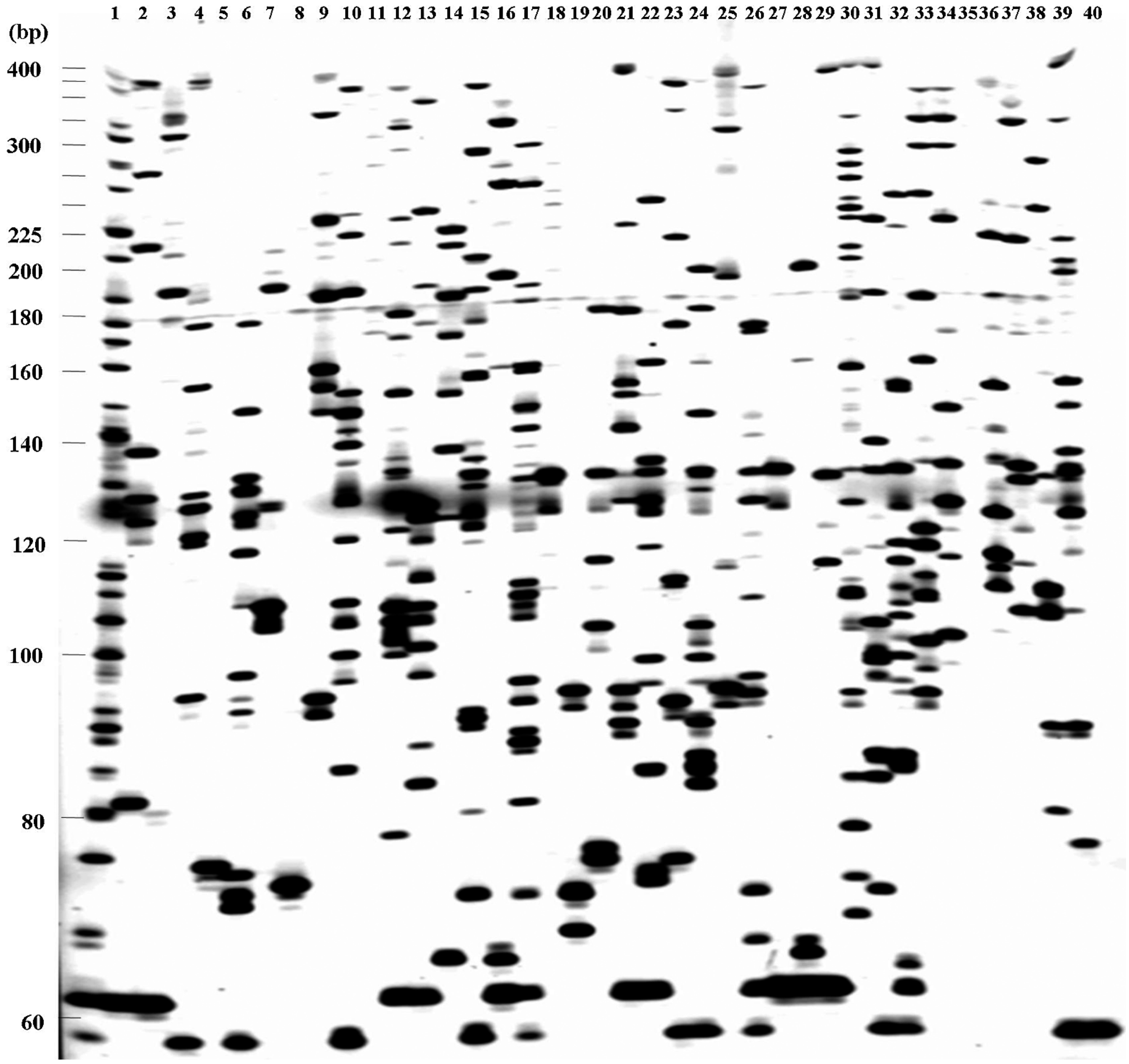

Fig. 4. Transposon display for *mPing* in *O. rufipogon* accessions. The migration of the DNA marker is indicated on the left in units of base pairs (bp). *Mse*I-digested DNA samples were ligated to an adapter and used as templates for PCR with an *mPing* specific primer and an adapter primer. Lanes 1–2, *Oryza sativa*; lanes 3–39, *O. rufipogon*; lane 40, negative control (distilled water). Lane numbers represent the serial numbers of the accessions as listed in Table 2.

The accessions without *Ping* were randomly distributed in all clusters. Although the copy number of *mPing* varied from zero to 30 in *O. rufipogon* accessions, the average copy numbers were 12.8, 17.8, 11.0, and 12.0 in clusters I, II, III, and IV, respectively.

Discussion

Recent report demonstrated that *mPing* family transposons were present only in *O. sativa* and *O. rufipogon*, and *mPing* population could be generalized into two families, *mPing*-1 and *mPing*-2, according to their sequence structure (Hu *et al.* 2006). Moreover, the proportion of *mPing*-1 and *mPing*-2 was found to be apparently different among *O. sativa* ssp. *japonica*, ssp. *indica*, and *O. rufipogon* (Hu *et al.* 2006). On the basis of this, Hu *et al.* (2006) suggested that the two extant subspecies of *O. sativa* evolved independently from corresponding ecotypes of *O. rufipogon*. However, the molecular evolution of *mPing* family has not been substantially analyzed. In the present study, we propose a possible evolutionary history of the distribution and proliferation of *mPing* family in *O. rufipogon*.

The present study shows that the *mPing* elements are deployed only in *O. rufipogon*, and are not deployed in 18 other wild rice species. *Kiddo*, in contrast, is deployed not only in AA genome species but also in BB, CC, BBCC and CCDD genome species. It has been reported that the *Tourist C* element, another MITE found in rice, is present in all four *Oryza* species complexes (Iwamoto *et al.* 1999). Moreover, *Pangrangja* elements were found to be present in AA, BB, CC, BBCC, CCDD, and EE genome species, and to be useful for studying relationships among species in the genus *Oryza* (Park *et al.* 2003a, 2003b). These observations suggest that *mPing* arose later than the other MITE elements

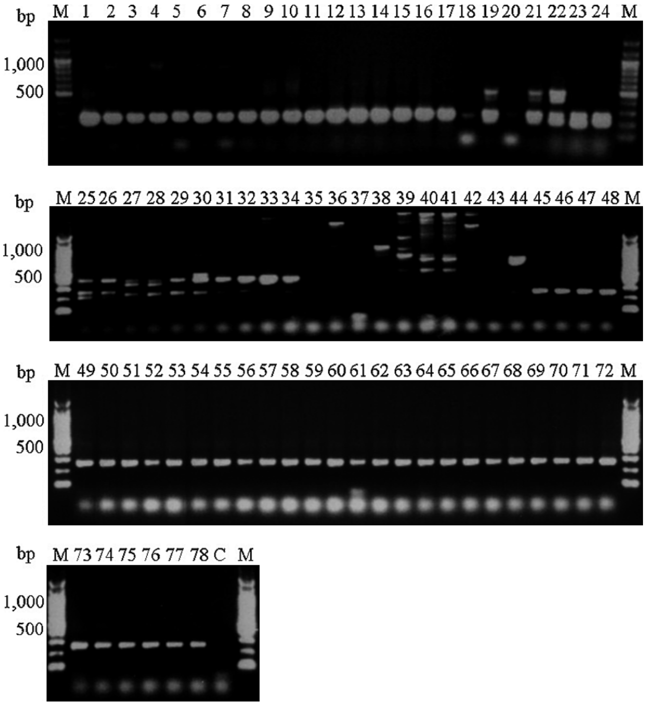

Fig. 5. Determining the presence or absence of *Kiddo* in wild *Oryza* species. PCR was carried out using *Kiddo*-F and *Kiddo*-R primers (expected product size, 269 bp). Lane numbers represent the serial numbers of the accessions as listed in Table 1. Lane M, DNA size marker (100 bp ladder, Nacalai); lane C, negative control (distilled water).

during the evolutionary history of the genus *Oryza*: in other words, *mPing* is a comparatively young MITE.

Pong was present in all of the 37 *O. rufipogon* accessions used in this study, but *Ping* was detected in only 17 of these. The phylogenetic tree based on nucleotide sequences around *qSH1* showed that the accessions with *Ping* were not grouped together in the same cluster; in fact, many of them were closely related to accessions without *Ping*. Because several *O. rufipogon* accessions used in this study were phylogenetically analyzed by other research groups, we tried to understand proliferation processes of the *mPing* family in *O. rufipogon* using the phylogenetic trees constructed from the insertional polymorphism of SINE (Cheng *et al.* 2003, Xu *et al.* 2007). Nevertheless, the accessions with *Ping* were randomly distributed in different clusters (data not shown). We therefore propose that the *mPing* family had already evolved and was present in the common ancestor of the *O. rufipogon* accessions, and that, later on, *Ping* was eliminated from the genomes of some of the accessions by random drift and/or natural selection along with the advancement of diversification in *O. rufipogon*. In *O. sativa*, *Ping* is present only in temperate *japonica* varieties, and the copy number of *mPing* in temperate *japonica* varieties is higher than those of tropical *japonica* and *indica* varieties (Jiang *et al.* 2003, Kikuchi *et al.* 2003); therefore, *Ping* has been believed to play a dominant role in the proliferation of *mPing* in *O. sativa*. In this study, however, the average copy number of *mPing* in *O. rufipogon* was not associated with the presence or absence of *Ping*. This indicates that *Pong*, as well as *Ping*, is responsible for the mobilization and proliferation of *mPing* in *O. rufipogon*.

O. rufipogon has been considered to be the direct progenitor of the cultivated rice species *O. sativa* ssp. *japonica* and *indica* (Xu *et al.* 2007). The copy number of *mPing* in *O. sativa* ssp. *japonica* was estimated by database searches to be 60–80 copies (Kikuchi *et al.* 2003) (it is actually ca. 50 copies), whereas that of *O. sativa* ssp. *indica* was estimated to be less than 10 copies (Jiang *et al.* 2003). Accordingly, it was suggested that the first amplification of *mPing* might have occurred during the domestication of *O. sativa*, and that, subsequently, *mPing* was activated again in *japonica* varieties, where it proliferated from 10 to 50 copies (Jiang *et al.* 2003). In the present study, we found that the copy numbers of *mPing* were markedly different even among *O. rufipogon* accessions, and that the accessions with different copy numbers were randomly distributed in different phylogenetic clusters. Like the case of *Ping* described above, the accessions with high copy numbers of *mPing* were not closely related to each other, even when using phylogenetic trees proposed by Cheng *et al.* (2003) and Xu *et al.* (2007) (data not shown). Moreover, we could not find any relationship between the copy number of *mPing* and the geographical distribution of *O. rufipogon* accessions. Hence, we concluded that the amplification of *mPing* occurred independently and randomly in each *O. rufipogon* accession.

In *mPing*-TD analysis, *O. rufipogon* accessions yielded only polymorphic bands. Moreover, the copy number of *mPing* varied considerably among the accessions. Estimates of the outcrossing rates of Asian wild rice species have ranged from 5 to 60%, though cultivated rice species are predominantly selfing (Morishima and Barbier 1990). This suggests that *O. rufipogon* strains easily outcross with each other; this is one way in which insertional polymorphisms of *mPing* may occur. Yet outcrossing cannot fully explain the differences in insertional polymorphisms and copy numbers of *mPing* among *O. rufipogon* accessions, because these accessions have several *Kiddo* elements at the same chromosomal positions. *mPing* is still actively transposing in some *japonica* cultivars under natural conditions (Naito *et al.* 2006). It is therefore considered that *mPing* may be currently active even in *O. rufipogon*.

Acknowledgements

We thank the National Institute of Genetics and the National Bioresource Center, Japan, for providing the plant materials of the wild *Oryza* species. We are very grateful to Dr. S. Konishi and Dr. T. Izawa for providing genetic information about *qSH1*. We are also thankful to Dr. Y. Kishima for his valuable advice. This study was supported by grants from the Ministry of Education, Culture, Sports and Technology of Japan in the form of Grants-in-Aid for Scientific Research, 1780003 and 18380005.

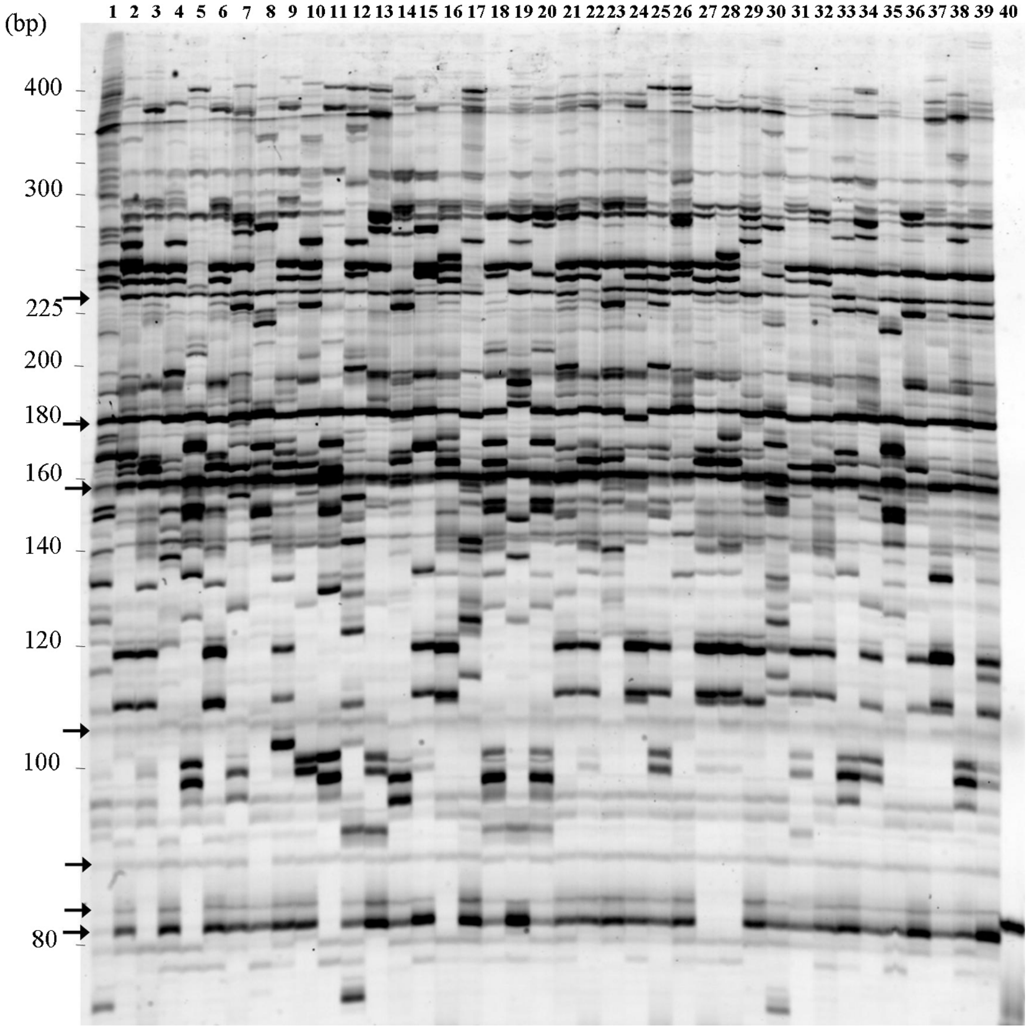

Fig. 6. Transposon display for *Kiddo* in *O. rufipogon* accessions. The migration of DNA markers is indicated on the left in units of base pairs (bp). *Msp*I-digested DNA samples were ligated to an adapter and used as templates for PCR with a *Kiddo* specific primer and an adapter primer. Lanes 1–2, *Oryza sativa*; lanes 3–39, *O. rufipogon*; lane 40, negative control (distilled water). Lane numbers represent the serial numbers of the accessions as listed in Table 2. Arrowheads indicate the bands common to many accessions.

Literature Cited

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

Biemont, C. and C. Vieira (2006) Genetics: Junk DNA as an evolutionary force. Nature 443: 521–524.

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

Casa, A.M., C. Brouwer, A. Nagel, L. Wang, Q. Zhang, S. Kresovich and S.R. Wessler (2000) The MITE family *Heartbreaker* (*Hbr*): Molecular markers in maize. Proc. Natl. Acad. Sci. USA 97: 10083–10089.

Cheng, C., R. Motohashi, S. Tsuchimoto, Y. Fukuta, H. Ohtsubo and E. Ohtsubo (2003) Polyphyletic origin of cultivated rice: Based on the interspersion pattern of SINEs. Mol. Biol. Evol. 20: 67–75.

Feng, Q., Y. Zhang, P. Hao, S. Wang, G. Fu, Y. Huang, Y. Li, J. Zhu, Y. Liu, X. Hu *et al.* (2002) Sequence and analysis of rice chromosome 4. Nature 420: 316–320.

Feschotte, C., N. Jiang and S.R. Wessler (2002) Plant transposable elements: where genetics meets genomics. Nat. Rev. Genet. 3: 329–341.

Hu, H., J. Mu, H.J. Zhang, Y.Z. Tao and B. Han (2006) Differentiation of a miniature inverted transposable element (MITE) system in Asian rice cultivars and its inference for a diphyletic origin of two subspecies of Asian cultivated rice. J. Integr. Plant Biol. 48: 260–267.

International Rice Genome Sequencing Project (2005) The map-based sequence of the rice genome. Nature 436: 793–799.

Iwamoto, M., H. Nagashima, T. Nagamine, H. Higo and K. Higo (1999) A *Tourist* element in the 5′-flanking region of the catalase gene *Cat*A reveals evolutionary relationships among *Oryza* species with various genome types. Mol. Gen. Genet. 262: 493–500.

Jiang, N., Z. Bao, X. Zhang, H. Hirochika, S.R. Eddy, S.R. McCouch and S.R. Wessler (2003) An active DNA transposon family in rice. Nature 421: 163–167.

Jiang, N., C. Feschotte, X. Zhang and S.R. Wessler (2004) Using rice to

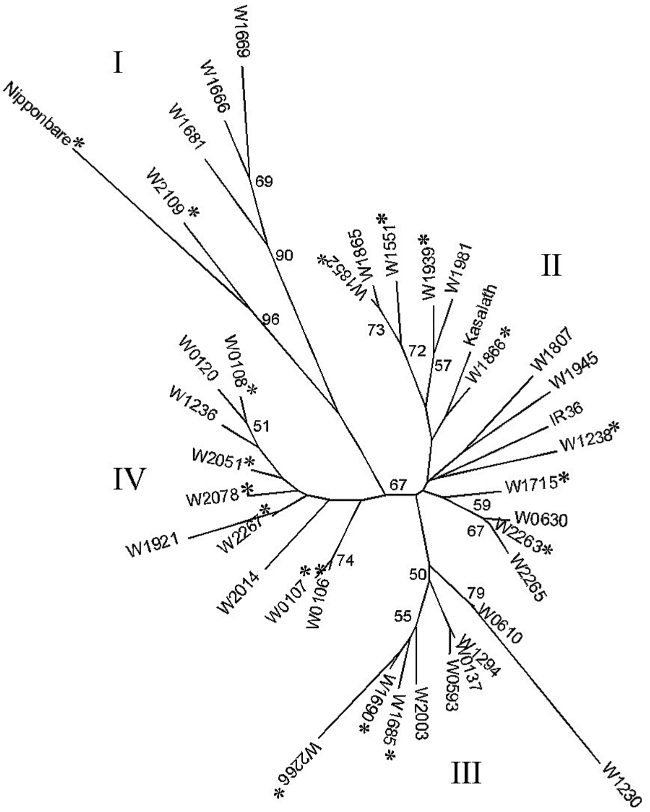

Fig. 7. Neighbor-joining tree of *O. rufipogon* accessions generated from sequence data around the *qSH1* region. Numbers on branches represent the bootstrap values. Roman numerals I, II, III, and IV indicate the four major clusters. Asterisks indicate the accessions harboring *Ping*.

understand the origin and amplification of MITEs. Curr. Opin. Plant Bio. l7: 115–119.

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

Kimura, M. (1980) A simple method for estimating evolutionary rate of base substitution through comparative studies of nucleotide sequences. J. Mol. Evol. 16: 111–120.

Konishi, S., T. Izawa, S. Lin, K. Ebana, Y. Fukuta, T. Sasaki and M. Yano (2006) An SNP caused loss of seed shattering during rice domestication. Science 312: 1392–1396.

Kumar, S., K. Tamura and M. Nei (2004) MEGA3: Integrated software for molecular evolutionary genetics analysis and sequence alignment. Briefings in Bioinformatics 5: 150–163.

Le, Q.H., S. Wright, Z. Yu and T. Bureau (2000) Transposon diversity in *Arabidopsis thaliana*. Proc. Natl. Acad. Sci. USA 97: 7376–7381.

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

Morishima, H. and P. Barbier (1990) Mating system and genetic structure of natural populations in wild rice, *Oryza rufipogon*. Plant Spec. Biol. 5: 31–39.

Murray, M. and W.F. Thompson (1980) Rapid isolation of high-molecular weight plant DNA. Nucleic Acids Res. 8: 4321–4325.

Naito, K., E.Y. Cho, G. Yang, M. Campbell, K. Yano, Y. Okumoto, T. Tanisaka and S.R. Wessler (2006) Dramatic amplification of a rice transposable element during recent domestication. Proc. Natl. Acad. Sci. USA 103: 17620–17625.

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

Ohtsubo, H., C. Cheng, I. Ohsawa, S. Tsuchimoto and E. Ohtsubo (2004) Rice retroposon *p-SINE1* and origin of cultivated rice. Breed. Sci. 54: 1–11.

Oki, N., K. Yano, Y. Okumoto, T. Tsukiyama, M. Teraishi and T. Tanisaka (2008) A genome-wide view of miniature inverted-repeat transposable elements (MITEs) in rice, *Oryza sativa* ssp. *japonica*. Genes Genet. Syst. 83: 321–329.

Onishi, K., K. Takagi, M. Kontani, T. Tanaka and Y. Sano (2007) Different patterns of genealogical relationships found in the two major QTLs causing reduction of seed shattering during rice domestication. Genome 50: 757–766.

Park, K.C., N.H. Kim, Y.S. Cho, K.H. Kang, J.K. Lee and N.S. Kim (2003a) Genetic variations of AA genome *Oryza* species measured by MITE-AFLP. Theor. Appl. Genet. 107: 203–209.

Park, K.C., J.K. Lee, N.H. Kim, Y.B. Shin, J.H. Lee and N.S. Kim (2003b) Genetic variation in *Oryza* species detected by MITE-AFLP. Genes Genet. Syst. 78: 235–243.

Saitou, N. and M. Nei (1987) The neighbor-joining method: A new method for reconstruction phylogenetic trees. Mol. Biol. Evol. 4: 406–425.

Takata, M., Y. Kishima and Y. Sano (2005) DNA methylation polymorphisms in rice and wild rice strains: Detection of epigenetic markers. Breed. Sci. 55: 57–63.

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

Turcotte, K., S. Srinivasan and T. Bureau (2001) Survey of transposable elements from rice genomic sequences. Plant J. 25: 169–180.

Vaughan, A.D., K. Kadowaki, A. Kaga and N. Tamooka (2005) On the phylogeny and biogeography of the genus *Oryza*. Breed. Sci. 55: 113–122.

Xu, J.H., C. Cheng, S. Tsuchimoto, H. Ohtsubo and E. Ohtsubo (2007) Phylogenetic analysis of *Oryza rufipogon* strains and their relations to *Oryza sativa* strains by insertion polymorphism of rice SINEs. Genes Genet. Syst. 82: 217–229.

Yang, G., J. Dong, M.B. Chandrasekharan and T.C. Hall (2001) *Kiddo*, a new transposable element family closely associated with rice genes. Mol. Genet. Genomics 266: 417–424.

Yang, G., Y.H. Lee, X. Shi, S. Kertbundit and T.C. Hall (2005) A two-edged role for the transposable element *Kiddo* in the rice ubiquitin 2 promoter. Plant Cell 17: 1559–1568.

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