reference to their impact on the salt marsh. *Contrib. Mar. Sci.* **23,** 25–55 (1980).
8. Zeil, J. & Layne, J. *Crustacean Experimental Systems in Neurobiology* (ed. Wiese, K.) 227–247 (Springer, Heidelberg, 2002).
9. Zeil, J., Nalbach, G. & Nalbach, H.-O. Eyes, eye stalks, and the visual world of semi-terrestrial crabs. *J. Comp. Physiol. A* **159,** 801–811 (1986).
10. Krapp, H. G., Hengstenberg, B. & Hengstenberg, R. Dendritic structure and receptive-field organization of optic flow processing interneurons in the fly. *J. Neurophysiol.* **79,** 1902–1917 (1998).
11. Wehner, R. 'Matched filters'—neural models of the external world. *J. Comp. Physiol. A* **161,** 511–531 (1987).
12. Schall, R. Estimation in generalized linear models with random effects. *Biometrika* **78,** 719–727 (1991).
13. Zeil, J. & Al-Mutairi, M. M. The variation of resolution and of ommatidial dimensions in the eyes of the fiddler crab *Uca lactea annulipes* (Ocypodidae, Brachyura, Decapoda). *J. Exp. Biol.* **199,** 1569–1577 (1996).
14. Land, M. & Layne, J. The visual control of behaviour in fiddler crabs I. Resolution, thresholds and the role of the horizon. *J. Comp. Physiol. A* **177,** 81–90 (1995).

Acknowledgements We thank P. Dixon, L. Trott and L. Howlett for help; J. Wood for advice; and M. R. Ibbotson and M. V. Srinivasan for comments on the manuscript. The work was supported by a postdoctoral fellowship from the Swiss National Foundation to J.M.H. and in part by an Human Frontier Science Program (HFSP) grant.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to J.M.H. (e-mail: jan.hemmi@anu.edu.au).

An active DNA transposon family in rice

Ning Jiang*†, Zhirong Bao†‡, Xiaoyu Zhang*, Hirohiko Hirochika§, Sean R. Eddy‡, Susan R. McCouch‖ & Susan R. Wessler*

* *Department of Plant Biology, University of Georgia, Athens, Georgia 30602, USA*
‡ *Howard Hughes Medical Institute and Department of Genetics, Washington University School of Medicine, St Louis, Missouri 63110, USA*
§ *National Institute of Agrobiological Resources, Tsukuba, Ibaraki 305, Japan*
‖ *Department of Plant Breeding, Cornell University, Ithaca, New York 14853, USA*
† *These authors contributed equally to this work*

The publication of draft sequences for the two subspecies of *Oryza sativa* (rice), *japonica* (cv. Nipponbare) and *indica* (cv. 93-11)[1,2], provides a unique opportunity to study the dynamics of transposable elements in this important crop plant. Here we report the use of these sequences in a computational approach to identify the first active DNA transposons from rice and the first active miniature inverted-repeat transposable element (MITE) from any organism. A sequence classified as a *Tourist*-like MITE of 430 base pairs, called *miniature Ping* (*mPing*), was present in about 70 copies in Nipponbare and in about 14 copies in 93-11. These *mPing* elements, which are all nearly identical, transpose actively in an *indica* cell-culture line. Database searches identified a family of related transposase-encoding elements (called *Pong*), which also transpose actively in the same cells. Virtually all new insertions of *mPing* and *Pong* elements were into low-copy regions of the rice genome. Since the domestication of rice *mPing* MITEs have been amplified preferentially in cultivars adapted to environmental extremes—a situation that is reminiscent of the genomic shock theory for transposon activation[3].

Rice is the most important crop for human nutrition in the world. At 430 megabase pairs (Mb), it also has the smallest genome among the agriculturally important cereals (including maize, sorghum, barley and wheat)[4]. For these reasons, rice is the focus of several genome-sequencing projects[1,2,5]. Computer-assisted analyses of rice genomic sequence indicate that, despite its small size, over 40% is repetitive DNA and most of this is related to transposable elements[1,2]. The class 1 long-terminal repeat (LTR) retrotransposons comprise the largest component of transposable elements in the rice genome (14% of the genomic DNA) but, numerically, MITEs form the largest group with over 100,000 elements divided into hundreds of families comprising about 6% of the genome[6,7]. MITEs are the predominant transposable element associated with the non-coding regions of the genes of flowering plants, especially grasses, and have been found in several animal genomes including *Caenorhabditis elegans*, mosquitoes, fish and human (reviewed in ref. 8).

Structurally, MITEs are reminiscent of non-autonomous DNA (class 2) elements with their small size (<600 base pairs) and short (10–30 bp) terminal inverted repeats (TIRs). But their high copy number (up to 10,000 copies per family) and target-site preference (for TA or TAA) distinguish them from most previously described non-autonomous DNA elements[8]. Non-autonomous elements, which make up a significant fraction of eukaryotic genomes, have been classified into families according to the transposase responsible for their mobility. But classifying MITEs in this way is problematic because no actively transposing MITE had been reported in any organism. Instead, the tens of thousands of plant MITEs have been classified into two superfamilies on the basis of the similarity of their TIRs and their target site duplication (TSD): *Tourist*-like MITEs and *Stowaway*-like MITEs[7,9,10]. Much evidence links *Tourist* and *Stowaway* MITEs with two superfamilies of transposases, *PIF*/*Harbinger* and *Tc1*/*mariner*, respectively[9–11].

Because newly transposed elements should be identical, we considered that an active MITE family would be characterized by extremely low intrafamily sequence divergence. The availability of almost half of the Nipponbare genome (187 Mb by 24 December 2001) in public databases (http://rgp.dna.affrc.go.jp) provided the possibility of searching for repeat families with the structural features of MITEs and with very low intrafamily sequence divergence. We first identified 1,257 repeat families with RECON[12], a program for the *de novo* identification of repeat families. Manual inspection of these sequences subsequently indicated that a repeat of 430 bp named *mPing* was a candidate for an active MITE. The TSDs (the trinucleotide TAA or TTA) and TIRs of *mPing* indicated that it is a *Tourist*-like MITE (Fig. 1). Of the 36 independent copies mined from 270 Mb of Nipponbare sequence, 26 were identical. By contrast, only eight complete copies of *mPing* were found in the 361 Mb of contig sequence of the *indica* cultivar 93-11 (ref. 2). From the frequency of elements recovered per megabase, we estimated the whole genome of Nipponbare and 93-11 to contain 70 and 14 copies of *mPing*, respectively (Methods).

The only rice transposable elements shown previously to be active are three families of LTR retrotransposons (*Tos10*, *Tos17*, *Tos19*) that transpose in both *japonica* (Nipponbare) and *indica* (C5924) cell culture[13]. To assess whether *mPing* elements are activated in the same cell lines, we used a modification of the AFLP technique called transposon display[14,15] to detect *mPing* insertions that might have occurred in culture. Because all of the *mPing* elements were virtually identical at their ends (Fig. 1 and Supplementary Information), we designed element-specific primers located in subterminal sequence (Methods) to amplify all family members and the flanking sequence.

The number of polymerase chain reaction (PCR) products (referred to as amplicons) amplified from DNAs isolated from Nipponbare (*japonica*) and C5924 (*indica*) plants before culture was consistent with the copy-number estimates for these cultivars (Fig. 2a); however, a differential response to cell culture was observed. Whereas the Nipponbare amplicon pattern remained the same, C5924 cells have undergone a marked increase in amplicons. Nonspecific genomic rearrangements during cell culture were ruled out as a cause of the pattern differences by repeating transposon display with primers derived from two other rice transposable elements (a *gypsy*-type LTR retrotransposon (*SZ-2*; N.J. and S.R.W., unpublished data) and a MITE (*ID-1*; ref. 7). In

contrast to the *mPing* transposon display, these amplicons were essentially identical before and after cell culture.

In several studies, MITEs have been found predominantly in non-coding genic regions[16–18]. Without an example of an actively transposing MITE, it has not been possible to determine whether this distribution reflects preferential targeting or selection against insertion into other regions of the genome, or whether it reflects the elimination of alleles containing such insertions. To address this issue, we recovered 42 amplicons from cell line C5924 from the transposon display gel and sequenced them. Insertion sites of newly transposed elements were deduced by using sequences flanking the TIR (37–268 bp) to query the 93-11 and Nipponbare sequences. Thirty-four of 42 flanking sequences matched entries from 93-11 contigs, whereas one of the sequences was found only in *japonica* (cv. Nipponbare). Of the 35 matches, 32 were single-copy sequences, and one was a two-copy sequence (Supplementary Table 1). The remaining two insertion sites were in or next to other MITEs that were themselves in single-copy sequences. Thus, 34 of 35 new insertions were in single-copy regions of the genome.

Like other MITE families, *mPing* elements have no coding capacity and, as such, are incapable of catalysing their own transposition. Thus, movement of *mPing* must be catalysed by a transposase encoded *in trans*. To identify putative autonomous elements, we used the *mPing* sequence to query all available rice genomic sequence for related, but longer elements. A single element called *Ping* was found in the Nipponbare but not in the 93-11 draft sequence. *Ping* is 5,341 bp and shares 253 bp and 177 bp, respectively, of its terminal sequences with *mPing*, suggesting that *mPing* arose very recently from *Ping* by internal deletion. Further BLAST searches using *Ping* as the query led to the identification of *Pong* (5,166 bp), with identical 15-bp TIRs and similar subterminal regions to those of *mPing* and *Ping* (Fig. 1). At least five copies of *Pong* were found in Nipponbare and at least three copies in 93-11. Six of these eight copies (five in Nipponbare and one in 93-11) are complete and nearly identical (>99% identity), whereas two are truncated.

In addition to their termini, *Ping* and *Pong* also share similarity in two blocks of internal sequence corresponding to the two main open reading frames (ORFs) of each element (Fig. 1). When used as queries in TBLASTN searches against GenBank databases, both ORFs yielded numerous hits ($E < 10^{-10}$) from several plants, as well as animals and fungi (Supplementary Table 2). ORF2 homologues are abundant in plants and found most frequently in organisms with large amounts of genomic sequence in databases: 82 hits ($E < 10^{-46}$) from rice, 56 hits ($E < 10^{-23}$) from *Arabidopsis* and over 100 hits ($E < 10^{-36}$) from *Brassica oleracea*. Most ORF1 homologues are located within 2 kb of ORF2 homologues, and several ORF pairs are flanked by TIRs and TSDs that are similar to those of *Ping* and *Pong* (data not shown). It is therefore likely that each 'pair' of ORF1 and ORF2 belongs to the same element.

The function of ORF1 is unclear. It has only very weak sequence similarity to the DNA-binding domains of Myb (*P*fam 7.3, $E = 0.002$). By contrast, ORF2 shares significant similarity with the transposase of the maize *PIF* element[11], which is also associated with a *Tourist*-like MITE family (called *mPIF*)[11]. *PIF* belongs to a superfamily of transposons with members identified in all three eukaryotic kingdoms. ORF2 encodes a putative Asp-Asp-Glu (DDE) motif (Fig. 1) containing three acidic amino acids found at the catalytic core of the transposases of bacterial insertion sequences and some eukaryotic transposons. The amino-acid residues surrounding and including the DDE motif (referred to as N2, N3 and C1)[19] are highly conserved among *Ping*/*Pong*, *PIF* and several IS*5*-like elements (Fig. 1).

Although *mPing* elements are clearly derived from *Ping*, the evidence suggests that *Ping* is not the autonomous element that mobilizes *mPing* in C5924 cell culture. *Ping* was detected only as a single copy in Nipponbare and it was absent in the draft sequence of 93-11 (~84% of the genome) and in 20 of 24 rice cultivars (8 cultivars for each of the following groups: temperate *japonica*, tropical *japonica* and *indica*) tested by PCR, including C5924 itself. Only four temperate *japonicas* were found to contain *Ping*: Nipponbare, Gihobyeo, JX 17 and Koshikari (data not shown). The apparent absence of *Ping* from all *indica* cultivars tested provides strong evidence that it could not be responsible for the movement of *mPing* elements in the *indica* cell line. *Pong*, by contrast, is present in several near-identical copies in both *indica* and *japonica*. In addition, ORF1 of *Ping* seems to be truncated at the amino terminus as compared with its homologues, and lacks at least 60 conserved amino acids (Fig. 1).

If *Pong* is the autonomous element responsible for the transposition of *mPing* elements, it should also be capable of transposition. By exploiting sequence differences between *Pong* and *mPing*, we designed PCR primers to amplify *Pong* but not *mPing* elements in a transposon display assay. The results with primers for *Pong*

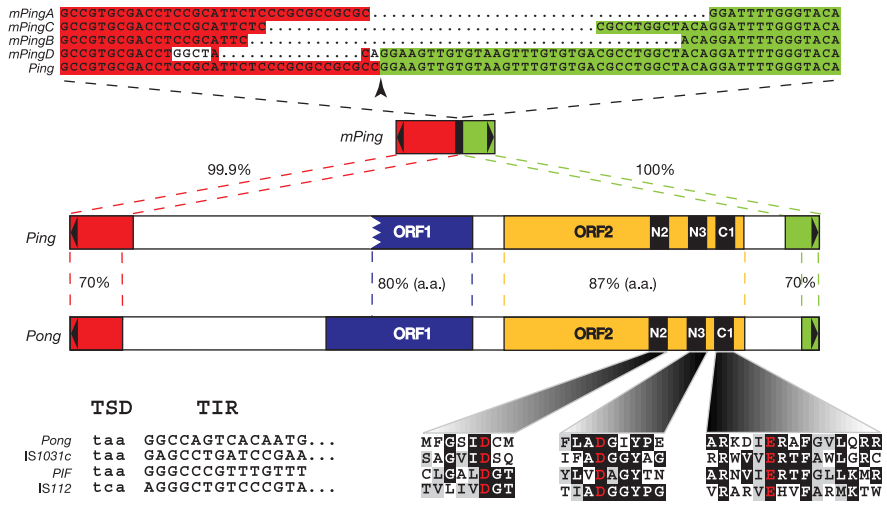

Figure 1 Comparison of *mPing*, *Ping* and *Pong* elements. Black triangles represent TIRs and black boxes represent putative N2, N2 and C1 catalytic domains (nucleotide sequences of the TIRs and TSDs, and amino-acid sequences of the catalytic domains of rice *Pong*, maize *PIF*[11] and bacterial IS*1031c* and IS*112* elements are shown). The thick vertical black line in *mPing* represents internal sequences that differ among the four subtypes derived from *Ping*. An alignment of this region is shown at the top. Arrowhead indicates the breakpoint in *Ping* where 4,923 bp of its internal sequence is not shown in alignment.

mirrored those for *mPing*; that is, the number of amplicons increased markedly in the *indica* cell line but remained virtually the same in Nipponbare (Fig. 2). The nature of the insertion sites and the inserted elements were determined as they were for *mPing*. Nine of ten insertion sites were located in single-copy sequences (Supplementary Table 1). Eight newly inserted elements were successfully amplified by PCR and all were indistinguishable in size from *Pong*.

The difference in the estimated copy number of *mPing* elements in a *japonica* (Nipponbare) and an *indica* (93-11) genome (70 versus 14) suggested the recent amplification of this MITE family, perhaps since domestication. To assess the timing of amplification, we carried out transposon display with a series of *O. sativa* DNAs to determine the approximate copy number of *mPing* and *Pong* elements. The temperate *japonicas* contain the largest number of different *mPing*-anchored amplicons, whereas the tropical *japonicas* contain the least (Fig. 3). This marked difference in *mPing* copy number between the two subgroups of *japonica* is significant because the temperate and tropical cultivars are thought to have diverged since domestication (5,000–7,000 years ago) and are more closely related to each other than either is to *indica*[20–24]. The different amplicon patterns of *Pong* elements observed in these cultivars also suggest that this element has been active since domestication. But the consistency of amplicon number across cultivars suggests that *Pong* elements have not substantially increased their copy number.

We have reported the use of a systematic computational approach to isolate active transposon elements from a host genome. So far, only a single active DNA transposon has been isolated from a vertebrate (*Tol2*)[25] and none has been isolated from mammals. Our protocol could potentially be applied to the available draft sequences of many organisms, including human, mouse and zebrafish.

With the identification of an active MITE family, long-standing mechanistic issues concerning the birth, spread and death of MITEs can now be addressed experimentally. Some issues have been resolved already; for example, we have shown that the prevalence

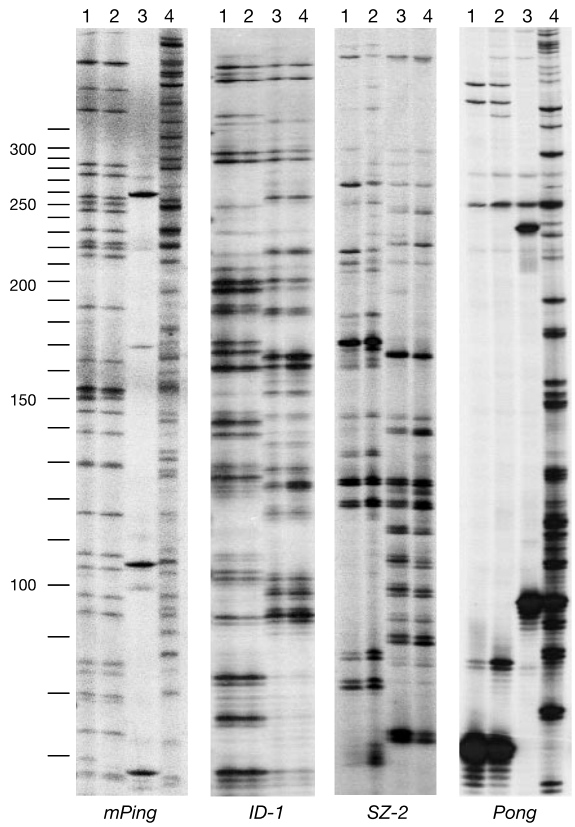

Figure 2 Autoradiograph of transposon display gels of *mPing* and *Pong* amplicons with rice genomic DNAs isolated before and after cell culture. The same genomic DNAs (digested and ligated with adapter primers) were used for each set of primers: 1, Nipponbare; 2, calli of Nipponbare; 3, C5924; 4, Oc cell lines derived from C5924 (ref. 30). The migration of DNA markers is indicated on the left in base pairs.

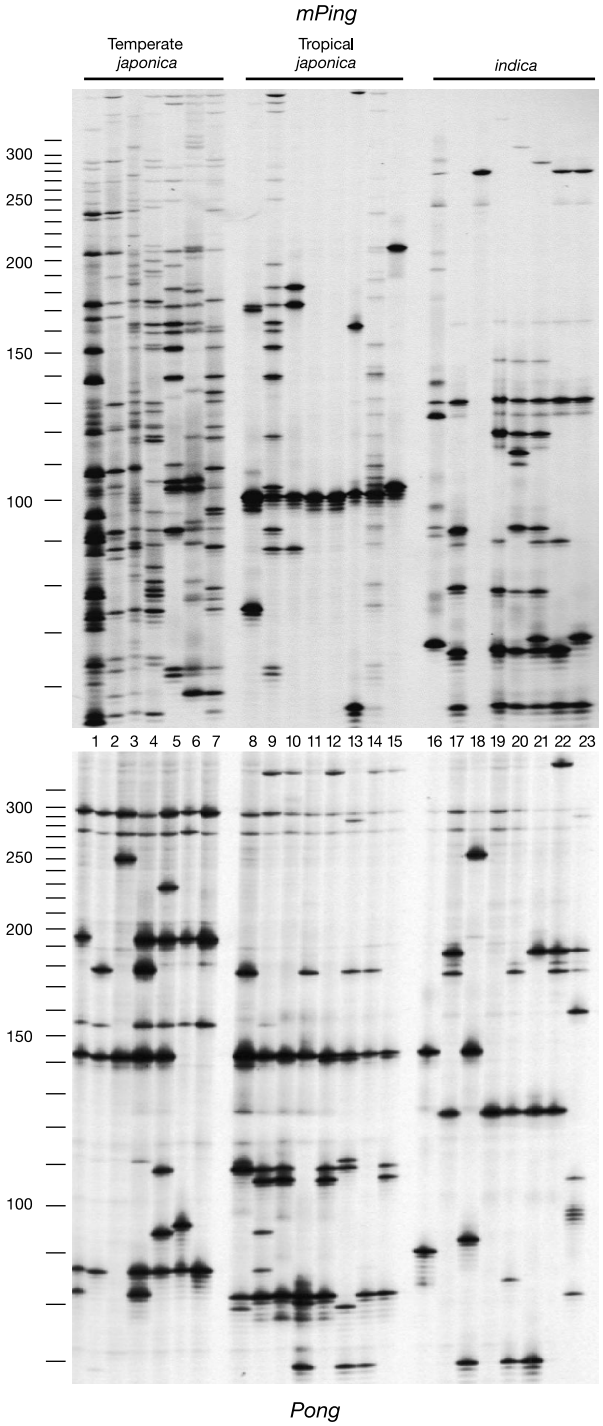

Figure 3 Autoradiograph of transposon display gels of *mPing* and *Pong*. Genomic DNAs: 1, Nipponbare; 2, Gihobyeo; 3, JX 17; 4, Koshikari; 5, Calrose; 6, Early Wataribune; 7, Shinriki; 8, Azucena; 9, Lemont; 10, Jefferson; 11, Moroberekan; 12, Rexoro; 13, Wab56-104; 14, Carolina Gold; 15, Kaybonnet; 16, C5924; 17, IR64; 18, Kasalath; 19, GuangLuAi4; 20, 93-11; 21, Tequing; 22, IR36; 23, Bs125. The migration of DNA markers is indicated on the left in base pairs.

of MITEs in single-copy regions primarily reflects targeting rather than selection. Our results also lead to some unexpected preliminary conclusions. First, *Ping*, which has clearly given rise to *mPing*, seems to be dispensable for the transposition of *mPing* in the C5924 cell line. The fact that in temperate *japonica* cultivars the presence of *Ping* correlates with *mPing* amplification suggests, however, that *Ping* may serve as a coactivator (with *Pong*, perhaps) to enhance transposition of *mPing*. Second, the requirements for transposition of *mPing* may differ in plants and in cell culture. Our data indicate that in cell culture, *Pong*, which is related but distinct from *mPing*, is the most likely source of transposase for this family. This relationship will have to be confirmed experimentally, perhaps in transgenic rice plants. At present, however, our data suggest that one reason for the success of MITEs is an ability to be cross-mobilized by related transposases.

The observation that *mPing* elements seem to have undergone amplification more extensively in temperate than in tropical *japonica* cultivars can be explained by an alternative hypothesis. As mentioned above, temperate and tropical *japonicas* are thought to have diverged from a common ancestor since domestication[20,21,23]. These two varietal groups are adapted to radically different temperature and water regimes[24,26,27]. Tropical *japonica* cultivars (previously known as *javanica*) are broadly adapted to tropical and subtropical environments, whereas temperate *japonicas* represent an evolutionary extreme, having been selected for productivity in cool, temperate zones with very short growing seasons. Thus, in a situation reminiscent of the genome shock theory[3], stress activation of *mPing* elements during the domestication of temperate *japonicas*, followed by their preferential insertion into genic regions, might have diversified these cultivars and hastened their domestication by creating new allelic combinations that might be favoured by human selection. Notably, it was speculated that an apparent γ-ray-induced insertion in an intron of the rice homologue of the *CONSTANS* gene (*Hd1*) was responsible for a quantitative change in flowering time[28]. On close inspection, this insertion is identifiable as an *mPing* element. □

Methods

Plant material

DNA from cv. Nipponbare (*japonica*), C5924 (*indica*) and derived cell cultures was obtained from the Hirochika laboratory (National Institute of Agrobiological Resources, Japan). DNA from other rice cultivars was obtained from the McCouch laboratory (Cornell University) or from G. Kochert (University of Georgia). We extracted plant DNA as described[29].

Computer-assisted identification of repeat families

Nipponbare sequence (total 187 Mb) was downloaded from http://rgp.dna.affrc.go.jp on 24 December 2001, and 93-11 sequence (361 Mb) was downloaded from http://210.83.138.53/rice/index.php on 25 February 2002. We used Nipponbare sequence for a systematic identification of repeat families using an all-versus-all comparison with WU-BLASTN2.0 (http://blast.wustl.edu; options M=5, N=-11, Q=22, R=11, -kap, E=0.0001, -hspmax 5000, -wordmask=dust, -wordmask=seg, -maskextra=50). Alignments were then clustered into repeat families using RECON (http://www.genetics.wustl.edu/eddy/recon) with default options and further examined individually using programs from the Genetics Computer Group (GCG; see below). We searched the output (1,257 repeat families) for similarity with features of known MITEs, including short TIRs, TSDs and size (N. Jiang, Z. Bao, S. R. Eddy and S. R. Wessler, manuscript in preparation). Repeat number 1,031 (called *mPing*) was identified as a *Tourist*-like MITE because of its size (430 bp), its TIR similarity to other known *Tourist* MITEs (refs 7, 9, 16 and Fig. 1) and its TSD of TTA or TAA.

Sequences analysis and copy-number estimates

DNA sequence analysis (pairwise comparisons, multiple sequence alignments, sequence assembling and formatting) was done with programs from the University of Wisconsin GCG program suite (version 10.1) accessed through Research Computing Resources at the University of Georgia. Elements related to *mPing* in Nipponbare (updated on 25 February 2002) and 93-11 were identified by BLAST search (WU-BLASTN 2.0) using a derived consensus sequence of *mPing*. Alignment of the central region of *mPing* identified distinct subtypes (*mPingA*, *-B*, *-C* and *-D*; Fig. 1 and Supplementary Table 3). Nipponbare contains only *mPingA*, whereas 93-11 contains *mPingA*, *-B* and *-C*; *mPingD* was present only in C5924.

We estimated the copy number of *mPing* as follows: 270 Mb of Nipponbare sequence (including overlaps; downloaded from the Rice Genome Project on 25 April 2002) contained 44 *mPing* elements ($E < 10^{-70}$) including 36 independent insertions and eight duplicates in regions of overlap. The copy number was estimated to be 70 ($44/270 \times 430$). In 361 Mb of 93-11 contig sequence there were eight complete *mPings* ($E < 10^{-70}$) and four incomplete *mPings* (where the element was located at the end of a contig; $E < 10^{-16}$). Comparison of the flanking sequences of incomplete copies in 93-11 and Nipponbare showed that all incomplete copies were independent insertions. Thus, the copy number in 93-11 was estimated to be 14 ($12/361 \times 430$). Because the 93-11 sequence is in a draft form and was masked before assembly, the fully masked sequence was examined for *mPing*. No *mPing* sequence was found, indicating that the masking process did not significantly affect the assembly of *mPing*. Some *mPing* fragments were found in unassembled sequence but these were not taken into consideration for copy number estimation.

Transposon display and recovery of new insertion sites

Transposon display was done as described[15] with the following modifications. Element-specific primers were designed on the basis of the subterminal sequences of *mPing*, *ID-1*, *SZ-2* and *Pong*, and used in PCR with different rice genomic DNAs. For the display in Fig. 2, the adapter primers were *Mse*I + 0 for *mPing* display, *Mse*I + AT for *ID-1* display, *Bfa*I + C for *SZ-2* display, and *Bfa*I + 0 for *Pong* display. For both displays in Fig. 3, the adapter primer was *Mse*I + 0. The other primer was specific either for *mPing* (Fig. 3, top) or *Pong* (Fig. 3, bottom). Primer sequences and PCR conditions are given in the Supplementary Methods.

We identified sequences flanking new insertions by cloning and sequencing PCR fragments from transposon display gels as described[7]. The context of the genomic sequence adjacent to the new insertion was determined using a BLAST search (WU-BLASTN 2.0) of the Nipponbare and 93-11 genomic sequence database. Single-copy sequence was defined as a query that results in no more than one hit per genome (except duplicates) with WU-BLASTN 2.0 default parameters.

Search for *Ping/Pong*-related elements

We identified homologues of ORF1 and ORF2 (in *Ping* and *Pong*) by database searches (TBLASTN) using the BLAST server from the National Center for Biotechnology Information (http://www.ncbi.nlm.nih.gov) against the non-redundant, expressed-sequence tag and Genome Survey Sequence databases.

Received 11 June; accepted 2 October 2002; doi:10.1038/nature01214.

1. Goff, S. A. *et al.* A draft sequence of the rice genome (*Oryza sativa* L. ssp. *japonica*). *Science* **296,** 92–100 (2002).
2. Yu, J. *et al.* A draft sequence of the rice genome (*Oryza sativa* L. ssp. *indica*). *Science* **296,** 79–92 (2002).
3. McClintock, B. The significances of responses of the genome to challenge. *Science* **226,** 792–801 (1984).
4. Arumuganathan, K. & Earle, E. D. Nuclear DNA content of some important plant species. *Plant Mol. Biol. Rep.* **9,** 208–218 (1991).
5. Burr, B. Mapping and sequencing the rice genome. *Plant Cell* **14,** 521–523 (2002).
6. Tarchini, R., Biddle, P., Wineland, R., Tingey, S. & Rafalski, A. The complete sequence of 340 kb of DNA around the rice *Adh1–Adh2* region reveals interrupted colinearity with maize chromosome 4. *Plant Cell* **12,** 381–391 (2000).
7. Jiang, N. & Wessler, S. R. Insertion preference of maize and rice miniature inverted repeat transposable elements as revealed by the analysis of nested elements. *Plant Cell* **13,** 2553–2564 (2001).
8. Feschotte, C., Jiang, N. & Wessler, S. R. Plant transposable elements: where genetics meets genomics. *Nature Rev. Genet.* **3,** 329–341 (2002).
9. Turcotte, K., Srinivasan, S. & Bureau, T. Survey of transposable elements from rice genomic sequences. *Plant J.* **25,** 169–179 (2001).
10. Feschotte, C. & Wessler, S. R. *Mariner*-like transposases are widespread and diverse in flowering plants. *Proc. Natl Acad. Sci. USA* **99,** 280–285 (2002).
11. Zhang, X. *et al. P Instability Factor*: an active maize transposon system associated with the amplification of *Tourist*-like MITEs and a new superfamily of transposases. *Proc. Natl Acad. Sci. USA* **98,** 12572–12577 (2001).
12. Bao, Z. & Eddy, S. R. Automated *de novo* identification of repeat sequence families in sequenced genomes. *Genome Res.* **12,** 1269–1276 (2002).
13. Hirochika, H., Sugimoto, K., Otsuki, Y. & Kanda, M. Retrotransposons of rice involved in mutations induced by tissue culture. *Proc. Natl Acad. Sci. USA* **93,** 7783–7788 (1996).
14. Van den Broeck, D. *et al.* Transposon display identifies individual transposable elements in high copy number lines. *Plant J.* **13,** 121–129 (1998).
15. Casa, A. M. *et al.* The MITE family heartbreaker (*Hbr*): molecular markers in maize. *Proc. Natl Acad. Sci. USA* **97,** 10083–10089 (2000).
16. Bureau, T. E., Ronald, P. C. & Wessler, S. R. 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 (1996).
17. Zhang, Q., Arbuckle, J. & Wessler, S. R. Recent, extensive and preferential insertion of members of the miniature inverted-repeat transposable element family *Heartbreaker* (*Hbr*) into genic regions of maize. *Proc. Natl Acad. Sci. USA* **97,** 1160–1165 (2000).
18. Mao, L. *et al.* Rice transposable elements: a survey of 73,000 sequence-tagged-connectors. *Genome Res.* **10,** 982–990 (2000).
19. Rezsohazy, R., Hallet, B., Delcour, J. & Mahillon, J. The IS4 family of insertion sequences: evidence for a conserved transposase motif. *Mol. Microbiol.* **9,** 1283–1295 (1993).
20. Ting, Y. The origin and evolution of cultivated rice in China. *Acta Agron. Sinica* **8,** 243–260 (1957).
21. Glaszmann, J. C. Isozymes and classification of Asian rice varieties. *Theor. Appl. Genet.* **74,** 21–30 (1987).
22. Wang, Z. Y., Second, G. & Tanksley, S. D. Polymorphism and phylogenetic relationships among species in the genus *Oryza* as determined by analysis of nuclear RFLPs. *Theor. Appl. Genet.* **83,** 565–581 (1992).
23. Matsuo, T., Futsuhara, Y., Kikuchi, F. & Yamaguchi, H. *Science of the Rice Plant* (Ministry of Agriculture, Forestry and Fisheries, Tokyo, 1997).
24. Morishima, H. & Oka, H. I. Phylogenetic differentiation of cultivated rice. XXVII. Numerical evaluation of the *indica–japonica* differentiation. *Jpn. J. Breed.* **31,** 402–413 (1981).
25. Kawakami, K., Shima, A. & Kawakami, N. Identification of a functional transposase of the *tol2*

element, an *Ac*-like element from the japanese medaka fish, and its transposition in the zebrafish germ lineage. *Proc. Natl Acad. Sci. USA* **97,** 11403–11408 (2000).
26. Glaszmann, J. C. & Arraudeau, M. Rice plant type variation: '*Japonica*'–'*Javanica*' relationships. *Rice Genet. Newslett.* **3,** 41–43 (1986).
27. Ueno, K. Differentiation of ecotypes in *Oryza sativa* L. 2. Characteristics of ecotypes: Japanese lowland and upland rice. *Bull. Inst. Agri. Res. Tohoku Univ.* **39,** 43–49 (1988).
28. Yano, M. *et al. Hd1*, a major photoperiod sensitivity quantitative trait locus in rice, is closely related to the *Arabidopsis* flowering time gene *CONSTANS*. *Plant Cell* **12,** 2473–2484 (2000).
29. McCouch, S. R. *et al.* Molecular mapping of rice chromosomes. *Theor. Appl. Genet.* **76,** 815–829 (1988).
30. Baba, A., Hasezawa, S. & Syono, K. Cultivation of rice protoplasts and their transformation mediated by *Agrobacterium Spheroplasts*. *Plant Cell Physiol.* **27,** 463–471 (1986).

Supplementary Information accompanies the paper on *Nature*'s website (http://www.nature.com/nature).

Acknowledgements We thank J. Edwards for help with database searches; C. Feschotte and E. Pritham for critically reading the manuscript; and D. Holligan and Y. Hu for technical assistance. This study was supported by a grant from National Science Foundation to S.R.E., S.R.M. and S.R.W.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for material should be addressed to S.R.W. (e-mail: sue@dogwood.botany.uga.edu). The *mPing*, *Ping* and *Pong* sequences have been deposited in GenBank under accession codes BK000586, BK000587 and BK000588, respectively.

The plant MITE *mPing* is mobilized in anther culture

Kazuhiro Kikuchi*, Kazuki Terauchi†‡, Masamitsu Wada*† & Hiro-Yuki Hirano§

* *National Institute for Basic Biology, Okazaki 444-8585, Japan*
† *Graduate School of Science, Tokyo Metropolitan University, Tokyo 192-0397, Japan*
§ *Graduate School of Agricultural and Life Sciences, University of Tokyo, Tokyo 113-8657, Japan*

Transposable elements constitute a large portion of eukaryotic genomes and contribute to their evolution and diversification. Miniature inverted-repeat transposable elements (MITEs) constitute one of the main groups of transposable elements and are distributed ubiquitously in the genomes of plants and animals[1] such as maize[2–5], rice[3], *Arabidopsis*[6,7], human[8], insect[9,10] and nematode[11]. Because active MITEs have not been identified, the transposition mechanism of MITEs and their accumulation in eukaryotic genomes remain poorly understood. Here we describe a new class of MITE, called *miniature Ping* (*mPing*), in the genome of *Oryza sativa* (rice). *mPing* elements are activated in cells derived from anther culture, where they are excised efficiently from original sites and reinserted into new loci. An *mPing*-associated *Ping* element, which has a putative *PIF* family[5] transposase, is implicated in the recent proliferation of this MITE family in a subspecies of rice.

Miniature inverted-repeat transposable elements are small transposable elements (usually <500 base pairs), containing short terminal inverted repeats (TIRs), that occur in high copy numbers (~1,000–15,000) in eukaryotic genomes. In plants, MITEs have a preference for target sites that yield 2-bp or 3-bp (A + T)-rich duplications on transposition. Evolutionary studies suggest that some MITEs in plant and animal genomes have spread recently[1,4,10]; however, molecular evidence for their transposition has not been documented. Doubled haploid breeding, which exploits mutations that occur with high frequency during anther culture, has been used as an efficient method to improve agronomically important crops such as rice and barley by producing useful cultivars[12]. Here we show that a rice MITE is mobilized in anther culture, suggesting that this transposition may account, in part, for the high frequency of mutations in rice.

A transposable-element-like element with TIRs, which we named *mPing*, was identified by database analyses of rice genomic sequence. *mPing* is a short 430-bp element with 15-bp TIRs and lacks an open reading frame (ORF; Fig. 1a, b). Members of this rice-specific MITE family of transposons seem to have a high-preference insertion site because each one is flanked by 3-bp (A + T)-rich duplicated sequences. Of 42 such elements examined, 22 had terminal TAA duplications and 20 had TTA duplications. The copy number of *mPing* in *O. sativa* L. spp. *japonica* was estimated to be 60–80 by database searches (Monsanto database, 60% coverage) and by Southern blot analysis (Fig. 1c). This copy number is unusually low, as the copy number of other MITEs is typically in the thousands[2–11]. *mPing* sequences are highly conserved: the nucleotide sequences of 34 members (type A1) of this family are completely identical, and the other seven members have only one base change from type A1 members (see Supplementary Information). Southern blot analysis showed that the banding patterns of *mPing* are highly polymorphic among two domesticated rice (*japonica* and *indica*) and their ancestral species, *Oryza rufipogon* (ref. 13 and Fig. 1c).

In contrast to *japonica* rice, which has a pattern with many bands, *indica* rice and *O. rufipogon* have a pattern consistent with fewer *mPing* elements. In agreement with the Southern blot analysis, a database analysis indicated that the *indica* genome (about 92% coverage)[14] contains only eight members of the 430-bp *mPing* MITE family (http://btn.genomics.org.cn/rice/). These results indicate

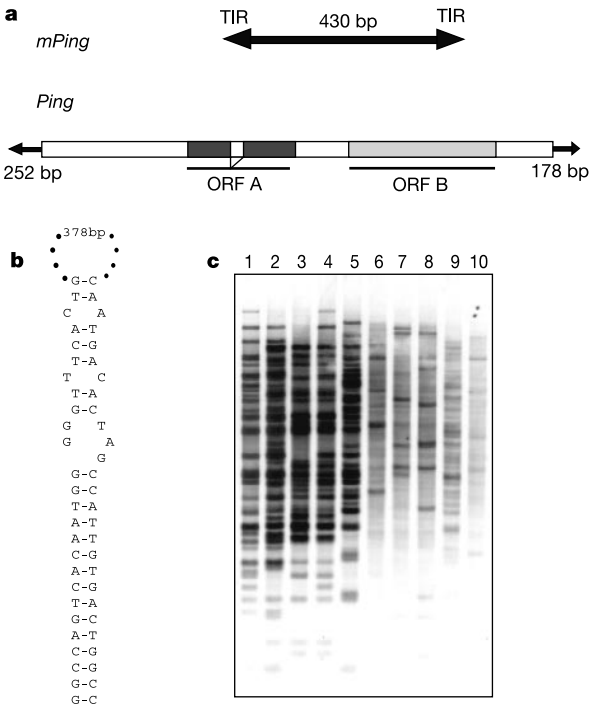

Figure 1 Structures and polymorphisms of *mPing* and *mPing*-related *Ping*. **a**, *mPing* and *Ping* structures. Sequences identical to *mPing* elements are indicated as bold arrows on the *Ping* element. **b**, Sequence of the TIR of *mPing* (complete, 15/15; incomplete 22/27). **c**, Southern blot showing band polymorphism of *mPing* in various rice cultivars and their ancestors. Lanes 1–5, *japonica* cultivars Nipponbare, Sachikaze, Koshihikari, Hitomebore and Taichung 65, respectively; lanes 6–8, *indica* cultivars Kasalath, IR36 and T0710; lanes 9 and 10, *O. rufipogon* W1987 and W0125.

‡ Present address: Graduate School of Science, Osaka University, Toyonaka, Osaka 560-0043, Japan.