

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

Hao Hu^{1, 2}, Jie Mu¹, Hong-Ji Zhang³, Yue-Zhi Tao³ and Bin Han^{1*}

(1. National Center for Gene Research, Shanghai Institutes of Biological Sciences, the Chinese Academy of Sciences, Shanghai 200233, China;

2. Graduate School of the Chinese Academy of Sciences, Beijing 100039, China;

3. Zhejiang Academy of Agricultural Sciences, Hangzhou 310021, China)

Abstract

In the present study, we report a survey on a Miniature Inverted Transposable Element (MITE) system known as mPing in 102 varieties of Asian cultivated rice (*Oryza sativa* L.). We found that mPing populations could be generalized into two families, mPing-1 and mPing-2, according to their sequence structures. Further analysis showed that these two families of mPing had significant bias in their distribution pattern in two subspecies of rice, namely *O. sativa* ssp. *japonica* and *indica*. *O. sativa japonica* has a higher proportion of mPing-1 as a general trait, whereas *O. sativa indica* has a higher proportion of mPing-2. We also examined the mPing system in a doubled haploid (DH) cross-breeding population of jingxi 17 (*japonica*) and zhaiyeqing 8 (*indica*) varieties and observed that the mPing system was not tightly linked to major subspecies-determining genes. Furthermore, we checked the mPing system in 28 accessions of Asian common wild rice *O. rufipogon* and found the mPing system in *O. rufipogon*. The distribution pattern of the mPing system in *O. rufipogon* indicated a diphyletic origin of the Asian cultivated rice *O. sativa* species. We did not find the mPing system in another 20 *Oryza* species. These results substantiated a previous hypothesis that *O. rufipogon* and *O. nivara* species were the closest relatives of *O. sativa* and that the two extant subspecies of *O. sativa* were evolved independently from corresponding ecotypes of *O. rufipogon*.

Key words: diphyletic; *indica*; *japonica*; Miniature Inverted Transposable Element (MITE); mPing; rice.

Hu H, Mu J, Zhang HJ, Tao YZ, Han B (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 Integrat Plant Biol* 48(3), 260–267.

Rice (*Oryza sativa* L.) has long been domesticated as a main staple crop from its wild progenitors. Rice feeds more than half the world's population and is among the main topics of plant research (Khush 1997; Takane et al. 1997). The genus *Oryza* (rice)

contains 21 wild species and two cultivated species. According to F₁ sterility of cross-breeding, all 23 species of the genus *Oryza* have been divided into 10 genomes (Takane et al. 1997). *O. sativa*, the Asian cultivated rice, has the AA genome with 2n = 24 chromosomes. The history of rice domestication is over 9 000 years old (Khush 1997) and people in different cultures have named the two subspecies of Asian domesticated rice in their own way (Takane et al. 1997), which are now have formally called *japonica* and *indica* varieties. The *indica* and *japonica* types are usually characterized by an association of certain diagnostic characteristics, such as KClO₃ resistance, cold tolerance, apiculus hair length, and phenol reaction (Oka 1958). The most commonly held view is that the Asian common wild rice *O. rufipogon*

Received 1 Feb. 2005 Accepted 20 Apr. 2005

Supported by Grants from the Ministry of Science and Technology of China (2002AA2Z1003 and 2003AA222091), Shanghai Municipal Commission of Science and Technology (038019315), and the National Natural Science Foundation of China (30325014).

*Author for correspondence. Tel: +86 (0)21 6484 5260; Fax: +86 (0)21 6482 5775; E-mail: <bhan@ncgr.ac.cn>.

is the ancestor of *O. sativa* and that the annual wild rice *O. nivara* is the closest relative of Asian cultivated rice (Khush 1997; Takane et al. 1997). The origin of Asian cultivated rice remains an issue of debate. Monophyletic and diphyletic origin hypotheses both have supporting evidence (Khush 1997). Kato et al. (1928) expressed the opinion that *indica* and *japonica* rice originated independently from a wild ancestor. Conversely, Ting (1957) has proposed that *japonica* is derived from *indica*. An in-depth investigation on rice genetics, including its origin and evolution, would provide a theoretical foundation for advanced breeding, leading to high-yielding crops with improved traits. Until now, the origin, timing, and subsequent process of rice domestication remains unresolved (Takane 1997).

A series of morphological and physiological standards has been adopted to identify the two subspecies of Asian rice cultivars *indica* and *japonica* (Kato et al. 1928; Matsuo 1952; Oka 1958). Progress in biochemistry and molecular biology has led to more sophisticated analyses of differentiation. The chloroplast DNA (Ishii 1986), isozyme (Chu 1967; Glaszmann 1987), and restriction fragment length polymorphism (RFLP; Wang and Tanksley 1989) have been used as alternative yardsticks to classify the subspecies.

With the sequencing of the genome of *O. sativa* (Feng et al. 2002; Stephen et al. 2002; Yu et al. 2002; The Rice Chromosome 10 Sequencing Consortium 2003), it has become possible to conduct a fine-level survey on the rice genome to reveal more about the details of molecular evolution. Recently, an active Miniature Inverted Transposable Element (MITE) with a low copy number called mPing (GenBank accession No. BK000588) was found in the rice genome (Jiang et al. 2003; Kazuhiro et al. 2003; Tetsuya et al. 2003). This 430-bp mPing was truncated from an autonomous transposon called Ping in the past transposition process and could still be catalysed to transpose and amplify by the transposase produced by Ping and Pong (another type of transposon similar to the sequence of Ping) in a very low frequency.

The MITE represents the third class of transposable elements other than class I (retrotransposons) and class II (transposons). Among all the mobile elements in genomes, retrotransposons transpose via an RNA intermediate, whereas transposons jump with a DNA intermediate. The MITE is grouped into a separate class because of its high copy number (1 000–10 000 copies every family), short length (commonly <600 bp), short terminal inverted repeats (TIR; usually 10–30 bp), secondary structure, and non-autonomous transposition mechanism. It is thought that MITE is the remainder of a transposon, inheriting its flanking sequences, so that MITE can be catalysed to transpose by transposase encoded by its forefather transposon, although MITE itself lacks the ability to encode the transposase (Craig et al. 2002; Feschotte et al. 2002).

In the present study, we performed a series of experiments on this mPing system with a collection of 102 Asian rice cultivars (*O. sativa*), a doubled haploid (DH) cross-breeding population of Jingxi 17 (*japonica*) × Zhaiyeqing 8 (*indica*), 28 accessions of Asian common wild rice (*O. rufipogon*, including *O. nivara*), 20 other

species of the genus *Oryza*, and three other cereal plants (maize, barley, and wheat). The data were used to analyze the molecular evolution of Asian cultivated rice.

Results

Variants of mPing in rice genomes could be generalized into mPing-1 and mPing-2

In previous studies, many different subtypes of mPing have been identified (Jiang et al. 2003; Kazuhiro et al. 2003; Tetsuya et al. 2003). In the present study, we assessed 102 rice samples through PCR cloning and sequencing of mPing fragments from 192 clones for each sample genome. We found a lot of variants of mPing. Most could be attributed to point mutations. We therefore characterized two major types of mPing (mPing-1 and mPing-2), representing all the variants through sequencing analysis. The mPing found in *O. sativa* ssp. *japonica* Nipponbare was taken as a representative of mPing-1, whereas mPing-2 was from *O. sativa* ssp. *indica* Guangluai 4. These two typical patterns of mPing are shown in Figure 1. The mPing-1 variant was 430 bp long, whereas mPing-2 had a deletion in its middle domain and had a typical length of 419 bp. Obviously, these two types of mPing are the remainders of transposed Ping, the autonomous transposon. Neither of them had any coding capacity. It should be noted that mPing-2 could not be produced from the transposition of the longer mPing-1, because mPing-2 was not just a curtailed mPing-1 but inherited 3 bp in the splicing sites that was absent from mPing-1. This also indicates that these two types of mPing are the products of different splicing of Ping.

The mPing-1 and mPing-2 variants showed significant distribution bias between the two subspecies of *O. sativa*

We surveyed the mPing population in 102 rice variety genomes. The PCR and sequencing results are given in Table 1 and show the percentage of the two types of mPing in each genome. The *japonica* varieties were characterized by a high proportion of mPing-1 (ranging from 72% to 100%, mostly around 90%–100%), whereas the *indica* group had a high percentage of mPing-2 (60%–96%). So, mPing-1 was predominant in the 52 *japonica* varieties, whereas mPing-2 was predominant in the 50 *indica* varieties. Therefore, we used the relative percentage of variants of mPing as a yardstick at the molecular level to classify the two subspecies of *O. sativa*.

Ping was not detected in all *indica* varieties, but was present in a few *japonica* varieties

Using PCR to amplify the specific fragments of Ping and Pong, we determined their existence in rice genomes (Table 1). Interestingly,



Figure 1. Two typical patterns of mPing, named mPing-1 and mPing-2, are aligned.

The different nucleotides are shaded. The sequence gaps in mPing-2 are filled by dashed lines. Boxed sequences were used as PCR primers to amplify the segment where mPing-1 and mPing-2 show their main difference.

none of the *indica* varieties had Ping, whereas a small proportion of *japonica* varieties had Ping. All rice cultivars had Pong in their genomes. So, Ping was a weak sign for the *japonica* subspecies and this may reflect different types of mechanism for the transposition and dissemination of mPing in the two subspecies. In the present study, even in the *japonica* group, Ping was not always commonly found. Thus, in the past years, this autonomous transposon had been driven out of most of the rice varieties and another transposon called Pong had taken its place to catalyse the transposition of mPing. Only a small proportion of any given variety's genome carries Ping that may still be active as progenitors of mPing.

A DH cross-breeding population test revealed no relationship between the distribution pattern of mPing and the diagnostic traits of the subspecies

To determine whether mPing had some impact on the specific characteristics of subspecies classification, we performed a test on the DH cross-breeding progenies of two typical rice cultivars called jingxi 17 (*japonica*) and zhaiyeqing 8 (*indica*). The results of this test are given in Table 2. The score for each progeny is the extent of bias towards either of the two typical subspecies. A higher score means bias for the *japonica* subspecies, whereas a

lower score indicates a higher resemblance to the *indica* group. With the cross-breeding, the distribution bias for mPing-1 and mPing-2 in each progeny showed no relationship with the bias of the progeny for any of the two subspecies, and neither did the existence of Ping. Therefore, we can conclude that the association between mPing distribution and subspecies diagnostic traits is not due to linkage or pleiotropy. So, as is the case for many other MITEs, mPing is another "guest" living in the rice genome that has no apparent function and has followed the host's evolutionary progress without any selective force upon itself. These results also contradict the patterns observed under natural conditions, so we assume that there have been few such inter-subspecies cross-breeding in the domestication history of rice.

Sequence and distribution pattern of mPing in *O. rufipogon*

It was believed that *O. rufipogon*, which shares the same AA-type genome with *O. sativa*, is the ancestor of modern rice cultivars. To assess this hypothesis, we identified mPings from 28 varieties of *O. rufipogon* and the results are given in Table 3. We found that the sequence and distribution pattern of mPing were similar between *O. sativa* and *O. rufipogon*. The proportion distribution of mPing patterns in *O. rufipogon* seemed more random than that in

Table 1. Percentage of the two patterns of mPing in 102 rice varieties

Variety	<i>japonica</i>		Variety	<i>indica</i>	
	mPing-1 (%)	mPing-2 (%)		mPing-1 (%)	mPing-2 (%)
Nipponbare	100	0	Guangchang'ai	4	96
Huangzhong	100	0	Changzixian	6	94
Jijing 8 ^a	100	0	Xiaobaixian	6	94
Jinghuang	100	0	Zhaiyeqing 8	9	91
Luokema 246	100	0	Jiazao 9708	10	90
Taihuqing	100	0	Zhongjian 100	10	90
Xinlewuming ^a	100	0	Nanjing 1	11	89
Xiushui 04	100	0	Xiaohongjiang	11	89
Yedihuangjin ^a	100	0	Hongmidongzhan	12	88
Yueguang	100	0	Erjiuqing	15	85
Zhenuo 2	100	0	Nanjing 11	15	85
Zhonghua 8 ^a	100	0	Wanjinxian	15	85
Guihuahuang ^a	98	2	Xiganwanqing	15	85
Hongmaodazhong	98	2	Zhenzhu'ai	16	84
Jijing 6 ^a	98	2	Zhou 903	17	83
Shidao	98	2	Teqing	17	83
Yuanjing 4	98	2	Hongjiaoman	18	82
Ballila ^a	97	3	Zhuxi 26	19	81
Huangdao	97	3	Jiayu 293	19	81
Yangjing 9538 ^a	97	3	Tianjidu	20	80
Zaodanba	97	3	Shuangzhuzhan	20	80
Liaojing 9377 ^a	96	4	Luqingzao	20	80
Yuanjing 6 ^a	96	4	Dahuangzhan	20	80
Zhonghua 11 ^a	96	4	Tuanjie 1	22	78
Luweidao	95	5	Zaoxian 99-358	22	78
Qingke	95	5	Zihong	22	78
Yongjing	95	5	Guanglu'ai 4	23	77
Xiushui 117	94	6	Aijiaonante	25	75
Jingxi 17	93	7	Erjiu'ai 4	25	75
Aihongdao	93	7	Hunanruanmi	25	75
Xiangqing	93	7	Junxie	25	75
Changhuangzhong	92	8	Baike'ai	26	74
Heijing	92	8	Heixian	26	74
Baixiangnuo	91	9	Qingganhuang	26	74
Laohudai	91	9	Shuangjia 1	27	73
Dachejing	90	10	Yedao	27	73
Wuyujing 3	90	10	9308 Hui	28	72
Xa 21	90	10	Zhulian'ai	29	71
Guozhu	89	11	Jiazao 935	30	70
Nonghong 73 ^a	89	11	Zaojian	30	70
Nonghu 6	89	11	Aizaizhan	32	68
Yuanjing 7	89	11	Zhenxian 97B	32	68
Youmangzaojing	86	14	IR8	35	65
Huangsanbai	86	14	Minghui 63	35	65
Youfeng ^a	85	15	Xiazhibai	35	65
Fengjin	83	17	Zaoxian 156	35	65
Hongxujing	82	18	Huakewan	37	63
Nongken 58 ^a	82	18	Jiayu 97104	40	60
Dalixiang	78	22	Tuanhuangzhan	40	60
Tiejingqing	78	22	Xinzaao 629	40	60
Laolaihong	75	25			
Lemont	72	28			

^aThere is a Ping in the corresponding variety.

The *japonica* varieties were characterized by a high proportion of mPing-1 (ranging from 72% to 100%, mostly around 90%–100%), whereas the *indica* group exhibited a high percentage of mPing-2 (range 60%–96%), with most around 75%–96%.

Table 2. Tests on samples of doubled haploid (DH) cross-breeding populations of jingxi 17 (*japonica*) × zhaiyeqing 8 (*indica*)

Variety	Score	mPing-1 (%)	mPing-2 (%)
Zhaiyeqing 8	4	9	91
p127 ^a	5	84	16
p63 ^a	6	83	17
p97 ^a	7	83	17
p67 ^a	8	75	25
p74 ^a	9	79	21
p65 ^a	10	79	21
p95 ^a	11	75	25
p69 ^a	12	64	36
p54	13	96	4
p52 ^a	14	71	29
p60 ^a	15	60	40
p62	16	71	29
p86	17	79	21
p71 ^a	18	73	27
p88 ^a	19	90	10
p110 ^a	20	69	31
p111 ^a	21	74	26
Jingxi 17 ^a	21	93	7

^aThere is a Ping in the corresponding variety.

Scores of 0–8 were typical for *indica*, scores of 9–13 were intermediates biased for *indica*, scores 14–17 were intermediates biased for *japonica*, and scores of 18–24 were typical of *japonica*. All samples carried at least one copy of Pong.

O. sativa.

There is no mPing system in other species of the *Oryza* genus and cereals

Using PCR with primers for mPing, we were not able to identify the mPing system from another 20 species of the *Oryza* genus (Table 4) and three cereal plants (maize, barley, and wheat). This indicated that mPing lived only in the cultivated rice *O. sativa* and the wild rice *O. rufipogon*. This transposable element invaded these two genomes after they diverged from other species of the *Oryza* genus.

Discussion

Independent origin and evolution of *japonica* and *indica*

In the present study, mPing-1 and mPing-2 were identified to be the two major types in most rice varieties through screening of 102 varieties of the Asian cultivated rice *O. sativa*. The results indicate that *japonica* and *indica* have their own, independent evolutionary route. We could assume that the common ancestor

Table 3. Tests on 28 accessions of *Oryza rufipogon*

<i>O. rufipogon</i>	mPing-1 (%)	mPing-2 (%)	Has Ping or not accession
Yunnan	83	17	No
Dongxiang	80	20	Yes
1	57	43	Yes
2	73	27	No
3	78	22	No
4	71	29	No
5	72	28	No
6	69	31	No
7	73	27	No
8	69	31	No
9	0	100	No
10	60	40	No
11	42	58	No
12	94	6	Yes
13	69	31	Yes
14	0	100	No
15	43	57	No
16	32	68	Yes
17	63	37	No
18	81	19	No
19	34	66	No
20	0	100	Yes
21	96	4	No
22	29	71	No
23	17	83	No
24	69	31	No
25	68	32	No
26	100	0	Yes

All samples carried at least one copy of Pong. They showed a similar distribution pattern of mPings with that of *O. sativa*.

of *japonica* and *indica* spread from their original locale and diverged in different ecosystems. The ancestor should have carried Ping and Pong, which are the autonomous elements of mPing, and may have had a very low copy number of mPing, or even none (meaning that Ping was a new invader into the primitive rice genome when it began to diverge). The *japonica* group initially comprised those plants inclined to live in northern areas of Asia and gradually spawned mPing-1 from Ping. Conversely, the archetype of *indica* should have dispersed around the southern regions, producing and disseminating mPing-2. Why is mPing-1 found in *japonica* and mPing-2 in *indica* and not vice versa? We propose that, at the very beginning when the first several mPings were truncated from Ping, whether mPing-1 or mPing-2, it was a stochastic (neutral) process and it just happened that mPing-1 was truncated in *japonica* and mPing-2 appeared in *indica*. Another assumption is that the transformation from Ping to mPing would happen at a very, very low frequency compared with the

Table 4. The 20 species other than *Oryza sativa* examined

Species	Genome (2n)	Source country
<i>O. rufipogon/O. nivara</i>	AA (24)	China
<i>O. glumaepatula</i>	AA (24)	Brazil
<i>O. barthii</i>	AA (24)	Tanzania
<i>O. longistaminata</i>	AA (24)	Kenya
<i>O. glaberrima</i>	AA (24)	Africa
<i>O. meridionalis</i>	AA (24)	Australia
<i>O. punctata</i>	BB (24), BBCC (48)	Cameroon
<i>O. minuta</i>	BBCC (48)	Phillipines
<i>O. rhizomatis</i>	CC (24)	Sri Lanka
<i>O. eichingeri</i>	CC (24)	Uganda
<i>O. officinalis</i>	CC (24)	Myanmar
<i>O. grandiglumis</i>	CCDD (48)	Brazil
<i>O. latifolia</i>	CCDD (48)	Central and South America
<i>O. alta</i>	CCDD (48)	Guyana
<i>O. australiensis</i>	EE (24)	Australia
<i>O. brachyantha</i>	FF (24)	Sierra Leone
<i>O. granulata</i>	GG (24)	Thailand
<i>O. meyeriana</i>	GG (24)	Southeast and South Asia
<i>O. ridleyi</i>	HHJJ (48)	Papua New Guinea
<i>O. longiglumis</i>	HHJJ (48)	Papua New Guinea

amplification of mPing with the aid of a transposase coded by Ping and Pong. This helps to explain the fixation of the predominance of a specific pattern of mPing in a specific subspecies. Then, the Ping was pumped out of the genome via some unknown mechanism so that the amplification of mPing itself became the unique way of dissemination.

Mainstream domestication involved no inter-subspecies cross-breeding

From the sampling tests on the DH cross-breeding population of jingxi 17 (*japonica*) and zhaiyeqing 8 (*indica*), we were not able to find any correlation of the percentage of mPing with the subspecies phenotype. So, the percentage variance in the specific mPing could be explained only by the introgression of genome parts from one subspecies to another. Because the domestication of rice has a history of at least 10 000 years, the present varieties should be, to some degree, different from their prototypes. From the results given in Tables 1 and 2, we can determine that inter-subspecies hybridization was not a common way of breeding before modern science started to be used in the last century. The reverse situation would result in a random concoction of mPing-1 and mPing-2 in modern rice varieties. Therefore, inter-species hybridization was not the mainstream.

O. rufipogon* (Asian common wild rice) had diverged itself before it evolved into two subspecies of *O. sativa* and the evolutionary processes for *japonica* and *indica

are independent

Some scientists believe that the Asian common wild rice (*O. rufipogon*) is the ancestor of the Asian rice cultivars (*O. sativa*) and the evidence for this has come from wide ranging research, such as morphology, physiology, ecosystems, and enzymology, indicating that *O. rufipogon* is the direct ancestor of *O. sativa* ssp. *indica*. When we checked mPing and Ping in 28 varieties of *O. rufipogon* in the present study, we found it reasonable that *O. rufipogon* was a relative of *O. sativa* because there was a common sequence of mPing between these two and they also had the same pattern of variance in mPing. We found a very high percentage for each pattern of mPing: 100% for mPing-1 and even 100% for mPing-2. We may think of the possibility of subspecies in wild rice that were the counterparts of those in cultivated rice and a set of appropriate rules may help delineate the subspecies of *O. rufipogon*.

So, our hypothesis was that, before *O. rufipogon* was domesticated into *O. sativa*, it had diverged into two primitive subspecies, namely *japonica* and *indica*, and that these two subspecies are the direct ancestors of the modern *O. sativa* ssp. *japonica* and *indica*, respectively. Whether the Asian common wild rice had diverged into *japonica* and *indica* remains a moot point.

Origin of the mPing system

In an investigation of 20 species of *Oryza* genus other than *O. sativa* and *O. rufipogon* and three cereal plants (maize, barley, and wheat), we were not able to find the existence of mPing.

Therefore, within the scope of the present study, mPing lives only in cultivated rice and Asian common wild rice. However, there is an extensive range of sequences from plant genomes, even from animal and prokaryote genomes, that shares a fairly high homology with Ping and Pong. So, Ping and Pong may be the offspring of ancient mobile elements that may have resided in other genomes. To use more sophisticated methods and to survey a wider range of plants should provide more clues for the molecular evolution of the mPing system itself.

Determining a clear molecular view of the evolution of rice provides considerable help to establishing advantageous genes for cultivated rice. A reshuffling and appropriate rearrangement of gene sets lead to high-yielding varieties and trait improvements. This is a challenge for scientists, because at the dawn of the 21st century up to half the world's population will have rice as their main source of nutrition.

Materials and Methods

Plant materials

One hundred and two cultivated rice varieties were provided by Zhejiang Academy of Agricultural Sciences and National Center for Gene Research. These samples, with typical classification characteristics, could be distinctly categorized into two subspecies of cultivated rice (Cheng 1985). The DH cross-breeding population of jingxi 17 (*japonica*) and zhaiyeqing 8 (*indica*) were supplied by China National Rice Research Institute (Hangzhou, China). Twenty-eight accessions of *O. rufipogon*, 20 other species of the genus *Oryza*, and three cereal plants (maize, barley, and wheat) were from Yangzhou University, China National Rice Research Institute and Zhejiang Academy of Agricultural Sciences, respectively.

DNA preparation and PCR

Total DNAs used as templates for PCR were isolated from leaves of rice seedlings with the DNeasy™ Plant Mini Kit (Qiagen, Germantown, MD, USA). The protocol was described in the handbook. The PCR primers were designed with the Wisconsin Package™ GCG (v. 2.0.2) program Prime (Accelrys, San Diego, CA, USA). The PCR was performed using an LA Taq™ PCR Kit (TaKaRa, Dalian, China) with the GeneAmp PCR system 9700 thermal cycler (Applied Biosystems, Foster City, CA, USA). Approximately 100 ng genomic DNA was added to a 100-μL volume reaction. Two primers were constructed to anneal to the flanking sequence of mPing (GenBank accession No. BK000588): 5'-GGGATGAGAGAGAAGGAAAGAG-3' and 5'-ACAATCCCCACAGTGGAG-3' (Figure 1). For each DNA sample, we performed the following procedure: the PCR mixture was first heated to 95 °C for 30 s and the following PCR cycle consisted of

denaturation for 30 s at 95 °C, annealing for 30 s at 56 °C, and extension for 30 s at 72 °C. This cycle was repeated 20 times. The PCR products were analyzed by electrophoresis on a 1.5% agarose in Tris-Borate-EDTA (TBE) buffer. The expected sizes of the PCR products using these primers were approximately 264 bp. Fragments of those sizes were cut from the electrophoresis gel and extracted using the MinElute™ Gel Extraction Kit (Qiagen). It should be noted that the restricted PCR cycle number (20) ensured that the reaction was performed in the exponential phase so that the final products for each sample could be compared at a quantitative level (Sambrook et al. 1989; Dieffenbach and Dveksler 1995).

To check for the presence or absence of Ping and Pong in each sample genome, we designed two pairs of primers for Ping (GenBank accession No. BK000587), namely 5'-ACAAGCGGATACTCCGAC-3'/5'-GAGGACAATGCCTTCCATAAC-3' (in the first open reading frame (ORF)) and 5'-CTTAAAACGACCAGCCCG-3'/5'-CACCACCACCAAACCTTATTCC-3' (in second ORF), and another two pairs of primers for Pong (GenBank accession No. BK000588), namely 5'-TCTACTCCACCACCAACACC-3'/5'-TCAGCCTTGTTTTGTCTTC-3' (in the first ORF) and 5'-CTCATACGAAGACCTCCTCC-3'/5'-CGCCTAAGATACCTCTCACC-3' (in the second ORF). The expected sizes of the PCR products using these primers were approximately 264, 299, 500, and 586 bp, respectively.

Cloning and DNA sequencing

In the present study, the PCR products for mPing were cloned into the pGEM-T Easy Vector plasmid (Promega, Madison, WI, USA), according to the manufacturer's instructions. In the pool of colonies on Luria-Bertani (LB) culture plates, we randomly selected 192 white (positive) colonies. The insertion of each clone was sequenced using T7 primer (5'-TAATACGACTCACTATAGGG-3'). The DNA sequencing was performed using a BigDye Terminator v3.1 cycle sequencing kit (PE Biosystem, Boston, MA, USA) and a 3730xl DNA Analyzer sequencing system (Applied Biosystems).

Computer analysis

Primary nucleotide sequences were analyzed with Chromas 2.01 (Technelysium Pty Ltd, Tewantin Qld 4565, Australia). Sequence alignment and comparisons were performed using BLAST (<http://www.ncbi.nih.gov/BLAST>). Nucleotide sequence searches of the DBJ/GenBank/EMBL databases were performed using the BLAST program (Altschul et al. 1990).

Identification and classification of variants of mPing and calculation of their relative proportions

Clone sequences were compared with mPing (GenBank accession No. BK000588) and, according to the sequence variances,

each sequence was assigned to one of the two representative groups: mPing-1 or mPing-2. For every genomic DNA, the specific number of sequences in both groups was counted and used to calculate the relative proportion of each group.

Test on a DH cross-breeding population of *japonica* and *indica*

The DH cross-breeding population of jingxi 17 (*japonica*) × zhaiyeqing 8 (*indica*) was considered to be an ideal model to check the possible links that the mPing system may have with the subspecies-determining quantitative trait loci (QTL). There was a total of 169 samples in the population, each of which was scored according to the extent of bias to one of the two subspecies. A 24-point scoring system was set up with regard to morphological and diagnostic characteristics, such as grain shape, awns, phenol reaction, etc. (Cheng 1985). Scores from 0 to 8 were typical for *indica*, scores from 9 to 13 were intermediates biased for *indica*, scores from 14 to 17 were intermediates biased for *japonica*, and scores from 18 to 24 were typical for *japonica*. We selected 19 samples from the population encompassing a scoring range from 4 to 21. For each sample, we performed the tests used as for cultivated rice varieties.

Acknowledgements

The authors thank Dr Qian Qian (China National Rice Research Institute, Hangzhou, China) and Dr Zhukuan Cheng (Institute of Genetics and Developmental Biology, the Chinese Academy of Sciences, China) for providing some of the plant materials.

References

- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990). Basic local alignment search tool. *J Mol Biol* **215**, 403–410.
- Cheng KS (1985). A statistical evaluation of the classification of rice cultivars into *hsien* and *keng* subspecies. *Rice Genet Newsl* **2**, 46–48.
- Chu YE (1967). Variations in peroxidase isozymes of *Oryza perennis* and *O. sativa*. *JPN J Genet* **42**, 233–244.
- Craig NL, Craigie R, Gellert M, Lambowitz A (2002). *Mobile DNA II*. ASM Press, NW Washington DC, USA.
- Dieffenbach CW, Dveksler GS (1995). *PCR Primer: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, NY.
- Feng Q, Zhang YJ, Hao P, Wang SY, Fu G, Huang YC et al. (2002). Sequence and analysis of rice chromosome 4. *Nature* **420**, 316–320.
- Feschotte C, Jiang N, Wessler SR (2002). Plant transposable elements: Where genetics meets genomics. *Nat Rev Genet* **3**, 329–341.
- Glaszmann JC (1987). Isozymes and classification of Asian rice. *Theor Appl Genet* **74**, 21–30.
- Goff S, Ricke D, Lan TH, Presting G, Wang RL, Dunn M et al. (2002). A draft sequence of the rice genome (*Oryza sativa* L. ssp. *japonica*). *Science* **296**, 92–100.
- Ishii T (1986). Restriction endonuclease analysis of chloroplast DNA from cultivated rice species, *O. sativa* and *O. glaberrima*. *JPN J Genet* **61**, 537–541.
- Jiang N, Bao ZR, Zhang XY, Hirochika H, Eddy SR, McCouch SR et al. (2003). An active DNA transposon family in rice. *Nature* **421**, 163–167.
- Kato S, Kosaka H, Hara S (1928). On the affinity of rice varieties as shown by fertility of hybrid plants. *Bull Sci Fac Agr Kyushu Univ* **3**, 132–147.
- Kazuhiro K, Kazuki T, Masamitsu W, Hiroyuki H (2003). The plant MITE mPing is mobilized in anther culture. *Nature* **421**, 167–170.
- Khush GS (1997). Origin, dispersal, cultivation and variation of rice. *Plant Mol Biol* **35**, 25–34.
- Matsuo D (1952). Genealogical studies on the cultivated rice. *Bull Nat Inst Agric Sci Ser D* **3**, 111.
- Nakazaki T, Okumoto Y, Horibata A, Yamahira S, Teraishi M, Nishida H (2003). Mobilization of a transposon in the rice genome. *Nature* **421**, 170–172.
- Oka H (1958). Intervarietal variation and classification of cultivated rice. *Ind J Genet Plant Breed* **18**, 79–89.
- Sambrook J, Fritsch EF, Maniatis T (1989). *Molecular Cloning: A Laboratory Manual*, 2nd edn. Cold Spring Harbor Laboratory Press, NY.
- Sasaki T, Matsumoto T, Yamamoto K, Sakata K, Baba T, Katayose Y et al. (2002). The genome sequence and structure of rice chromosome 1. *Nature* **420**, 312–316.
- Takane M, Yuzo F, Fumio K, Hikoyuki Y (1997). *Science of the Rice Plant*, Vol. 3. The Ministry of Agriculture, Forestry and Fisheries, Tokyo, Japan.
- The Rice Chromosome 10 Sequencing Consortium (2003). In-depth view of structure, activity, and evolution of rice chromosome 10. *Science* **300**, 1566–1569.
- Ting Y (1957). The origin and evolution of cultivated rice in China. *Acta Agron Sin* **8**, 243–260 (in Chinese).
- Wang ZY, Tanksley SD (1989). Restriction fragments length polymorphism in *Oryza sativa* L. *Genome* **32**, 1113–1118.
- Yu J, Hu S, Wang J, Wong KS, Li SG, Liu B et al. (2002). A draft sequence of the rice genome (*Oryza sativa* L. ssp. *indica*). *Science* **296**, 79–92.

(Managing editor: Li-Hui Zhao)