Genomics 93 (2009) 274–281

Contents lists available at ScienceDirect

Genomics

journal homepage: www.elsevier.com/locate/ygeno

Identification of a high frequency transposon induced by tissue culture, *nDaiZ*, a member of the *hAT* family in rice

Jian Huang [a,c], Kewei Zhang [a,c], Yi Shen [a,c], Zejun Huang [a,c], Ming Li [a], Ding Tang [a], Minghong Gu [b], Zhukuan Cheng [a,*]

[a] State Key Laboratory of Plant Genomics and Center for Plant Gene Research, Institute of Genetics and Developmental Biology, Chinese Academy of Sciences, Beijing 100101, China
[b] Key Laboratory of Crop Genetics and Physiology of Jiangsu Province/Key Laboratory of Plant Functional Genomics of Ministry of Education, Yangzhou University, Yangzhou 225009, China
[c] Graduate University of Chinese Academy of Sciences, Beijing 100049, China

ARTICLE INFO

Article history:
Received 14 July 2008
Accepted 14 November 2008
Available online 23 December 2008

Keywords:
Rice
nDaiZ
DaiZ
hAT
Transposable element

ABSTRACT

Recent completion of rice genome sequencing has revealed that more than 40% of its genome consists of repetitive sequences, and most of them are related to inactive transposable elements. In the present study, a transposable element, *nDaiZ0*, which is induced by tissue culture with high frequency, was identified by sequence analysis of an allelic line of the *golden hull and internode 2* (*gh2*) mutant, which was integrated into the forth exon of *GH2*. The 528-bp *nDaiZ0* has 14-bp terminal inverted repeats (TIRs), and generates an 8-bp duplication of its target sites (TSD) during its mobilization. *nDaiZ*s are non-autonomous transposons and have no coding capacity. Bioinformatics analysis and southern blot hybridization showed that at least 16 copies of *nDaiZ* elements exist in the *japonica* cultivar Nipponbare genome and 11 copies in the *indica* cultivar 93-11 genome. During tissue culture, only one copy, *nDaiZ9*, located on chromosome 5 in the genome of Nipponbare can be activated with its transposable frequency reaching 30%. However, *nDaiZ9* was not present in the 93-11 genome. The larger elements, *DaiZ*s, were further identified by database searching using *nDaiZ0* as a query because they share similar TIRs and subterminal sequences. *DaiZ* can also generate an 8-bp TSD. *DaiZ* elements contain a conserved region with a high similarity to the *hAT* dimerization motif, suggesting that the *nDaiZ–DaiZ* transposon system probably belongs to the *hAT* superfamily of class II transposons. Phylogenetic analysis indicated that it is a new type of plant *hAT*-like transposon. Although *nDaiZ* is activated by tissue culture, the high transposable frequency indicates that it could become a useful gene tagging system for rice functional genomic studies. In addition, the mechanism of the high transposable ability of *nDaiZ9* is discussed.

Introduction

Transposable elements (TEs) or transposons are DNA fragments that can insert into new chromosomal sites, and generate duplicate copies during transposition. The first transposon was discovered in maize by Barbara McClintock more than a half century ago [1], and is characterized as the famous *Ac*/*Ds* transposable system belonging to the *hAT* superfamily. As more and more genome-sequencing projects have been completed in the past years [2–5], computer-assisted analyses of these genomic sequences has indicated that TEs are major components of all eukaryotic genomes as well as most prokaryotic genomes. The only known exceptions are *Plasmodium falciparum* and several of its related species [6]. For example, over 40% of the rice genome is repetitive DNA and most of them are transposable elements. In grass genomes, transposon related sequences can account for 50–80% of their genomes [7,8]. Due to their ability to move and replicate, TEs play a significant role in genome evolution and accordingly can explain the *C*-value paradox [9].

TEs can be divided into two classes according to their transposable modes proposed by Finnegan [6]. One is the class I elements (or retrotransposon) that are mediated during reverse transcription by RNA intermediates using a “copy and paste” transposable mechanism. The other is the class II elements (or DNA transposons) that are directly transposed from one site to another in the genome using the “cut and paste” mechanism without any RNA intermediates. Challenging the view of the two class categorizing method is the discovery of non-autonomous TEs called miniature inverted repeat transposable elements (MITEs), which also employ the “copy and paste” mechanism but without RNA intermediates. Furthermore, unlike class II elements which are transposed by the double strands of DNA being cut, there also exists another type of transposon called helitron, which is transposed with a single strand being cut *via* a rolling-circle mechanism [10].

Rice is the most important food crop for human beings and also has the smallest genome (430 Mb) among the cereals including maize,

* Corresponding author: Fax: +86 10 64873428.
E-mail address: zkcheng@genetics.ac.cn (Z. Cheng).

0888-7543/$ – see front matter © 2008 Elsevier Inc. All rights reserved.
doi:10.1016/j.ygeno.2008.11.007

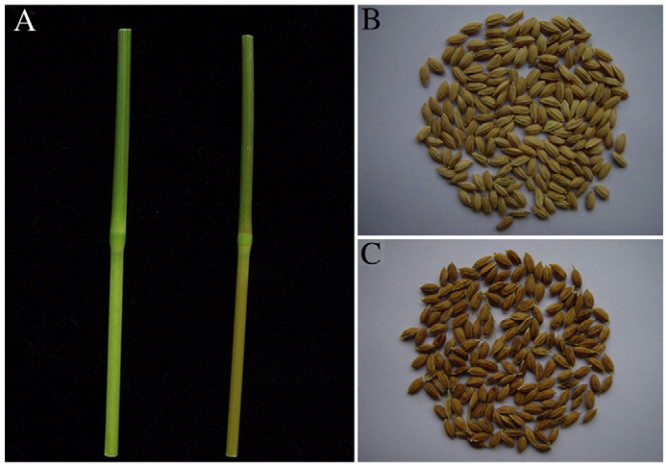

Fig. 1. Phenotype of wild type and *gh2-1* mutant. (A) The internodes of the wild type and *gh2-1* mutant at the heading stage. The wild-type internode is green (left), while the *gh2-1* internode is reddish-brown (right). (B) Seeds of the wild type in light yellow color. (C) Seeds of *gh2-1* in golden yellow color.

sorghum, barley and wheat. Therefore, rice has become a model system for cereal genomic studies, just like *Arabidopsis* is for dicotyledonous plants. Computer-assisted analyses of rice high quality sequence data reveals that nearly all of the DNA transposons (class II) identified previously in other plants also exist in the rice genome. Furthermore, MITEs are the major type of TEs comprising about 6% of the rice genome by having over 100,000 copies [11,12].

Most TEs in higher eukaryotes are inactive. One important reason for this is that most TEs are hypermethylated and cannot be detected by transposases, while only a tiny cantle is active under normal growing conditions. Active transposons, especially high transposable frequency mobile elements, are useful genetic tools for isolating genes and studying their functions. The accumulation of knowledge of transposons can lead researchers to elucidate their contributions to genome evolution studies. However, only a few active transposons have been identified so far in rice. One of them is *copia*-like retrotransposon, *Tos17*, which is silent under normal cultivated conditions but can be activated by tissue culture [13]. Another transposon in rice identified is *karma*, which can be induced by *Tos17*'s mobilization. Further studies have indicated that both of these two retrotransposons are responsible for DNA hypomethylation [14,15]. The tourist-like MITE, *mPing*, can also be induced by tissue culture, especially in anther culture processes [16]. These results indicate that tissue culture may generate a suitable environment that is beneficial for transposon activation in rice.

In this study, we report the identification of a novel DNA transposon, *nDaiZ*, from an allelic line of *gh2* mutant named as *gh2-1*, which can be easily activated by tissue culture. We cloned this transposable element and found it is a non-autonomous transposon. Database searching led us to characterize the corresponding autonomous elements with a *hAT* dimerization motif, and accordingly belonging to the *hAT* superfamily of class II transposons. Like the retrotransposon *Tos17*, the mobilization of only one active element, *nDaiZ9*, is also related to the hypomethylation status of its corresponding region. Southern blotting analysis showed that *nDaiZ* elements were restricted to a few genomes in the genus of *Oryza*, and its transposition may be constrained by DNA hypomethylation. The character of its high transposable frequency and only single copy can be activated during tissue culture make it useful as the mutagenesis tool for gene tagging.

Results

Identification of a gh2 allelic mutant gh2-1

A golden hull and internode mutant was induced by tissue culture from the *japonica* rice variety Nipponbare. The mutant exhibits a

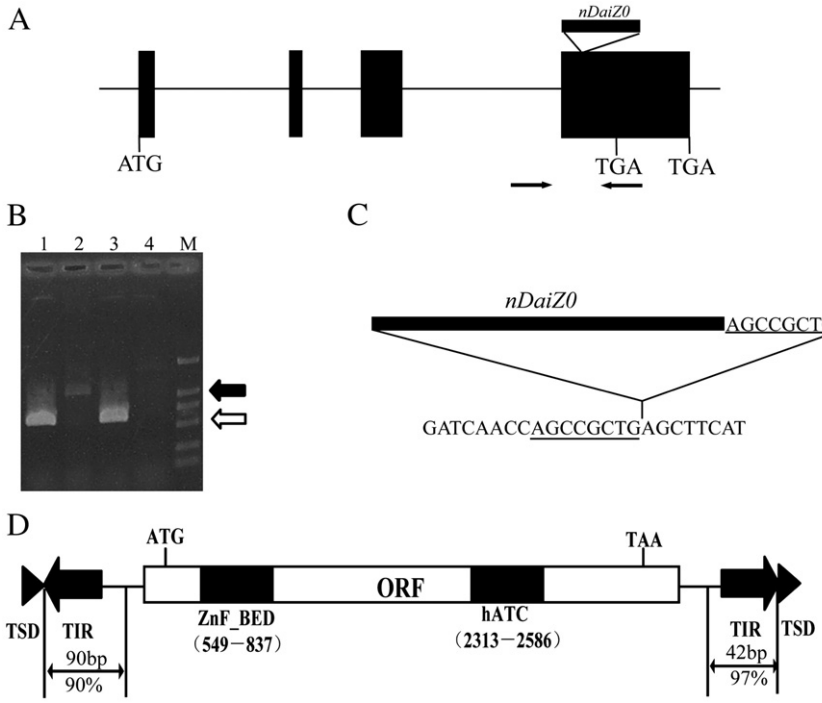

Fig. 2. Structures of the *GH2* gene in the *gh2-1* mutant and the *DaiZ* element. (A) The 528 bp insertion is detected at the forth extron of the *GH2* gene, the arrows indicate the position and orientation of the primer pair GH2. (B) The PCR products amplified with the primer pair GH2, lanes 1–4 *gh2* mutant, *gh2-1* mutant, Nipponbare and water, respectively. The filled arrow indicates the band with transposon insertion, and the open arrow indicates the wild-type bands. (C) Feature of the insertion region in *GH2*, the TSD is underlined. (D) The structure of *DaiZ* element. The region with similarity to the *hAT* dimerization motif and ZnF_BED motif are marked with black box, the number under the box shows the positions of the two motifs. The number under the TIRs shows the homologous level between the *nDaiZ9* and *DaiZ*.

reddish-brown pigment in the hull, internode and basal leaf sheath at the heading stage (Fig. 1A). The reddish-brown pigment gently becomes more intense from the basal internode to the apical internode, and the seed coloration also becomes darker, finally turning into a golden yellow color during plant maturation (Fig. 1C). We found this mutant has a similar phenotype when compared with the *gh2* mutant we studied before [17]. To test whether these two mutants are under genetic control of the same gene, the *gh2* genes in the two mutants were completely sequenced. Doing PCR with the primer pair *GH2*, a 577-bp product was amplified from the wild type Nipponbare and the original *gh2*, while the *gh2-1* mutant yielded a 1105-bp product, indicating there is an insertion in the *GH2* loci in *gh2-1* (Fig. 2B). We also verified this result using a genetic approach by crossing *gh2-1* with the original *gh2* mutant. All the F1 plants show the golden hull and internode phenotype, demonstrating that the *gh2-1* mutant is an allelic mutant of the original *gh2* mutant.

Characterization of the nDaiZ transposon

A 528-bp insertion was found in the *GH2* gene region of the *gh2-1* mutant (Fig. 2A). The sequence was inserted into the forth exon and generated an AGCCGCTG target sequence duplication. This insertion resulted in a frameshift in the coding region and generated a premature stop codon. The wild type *GH2* gene encodes a 299 amino acid cinnamyl-alcohol dehydrogenase (CAD), while the premature stop codon in *gh2-1* results in a 265 amino acid protein losing its C terminus. The insertion itself displays the structural features of a transposon with 14-bp terminal inverted repeats (TIRs) and 8-bp target sequence duplication (TSD) (Fig. 2C). The element itself also contains 11 copies of repetitive ACCC (or GGGT) tetranucleotides (Fig. 3A), which are known as the binding sequence of the transposase TAG1, a member of *hAT* superfamily [18]. We named it *nDaiZ0*, a non-autonomous DNA transposon.

BLAT analysis using the *nDaiZ0* sequence found in *gh2-1* as the query identified 16 sequences with high homology to *nDaiZ0* in the Nipponbare genome. These copies are located on all the other chromosomes except chromosomes 4, 8, 9 and 12, and named *nDaiZ1–nDaiZ16* (Fig. 3B). Among them, only two copies, *nDaiZ10* and *nDaiZ12*, are located near the telomeric or centromeric region, respectively. Furthermore, no conservation was detected in the TSDs of the 16 *nDaiZ* elements, suggesting that *nDaiZ* can randomly insert into the rice genome.

Among the 16 copies, only *nDaiZ9*, located on chromosome 5, is exactly identical to the insertion element presented in the *gh2-1* gene. Other copies either have deletions in the transposon bodies or have at least one mismatch in the TIRs (Supplementary Table 1). To test whether the transposon found in *gh2-1* loci originated from *nDaiZ9* of chromosome 5, PCR analysis for the two sites was performed in both *gh2-1* and the wild type Nipponbare, by amplifying with the primer pair nDaiZ9-F and R (Supplementary Table 2). This can amplify a product containing the whole *nDaiZ9*

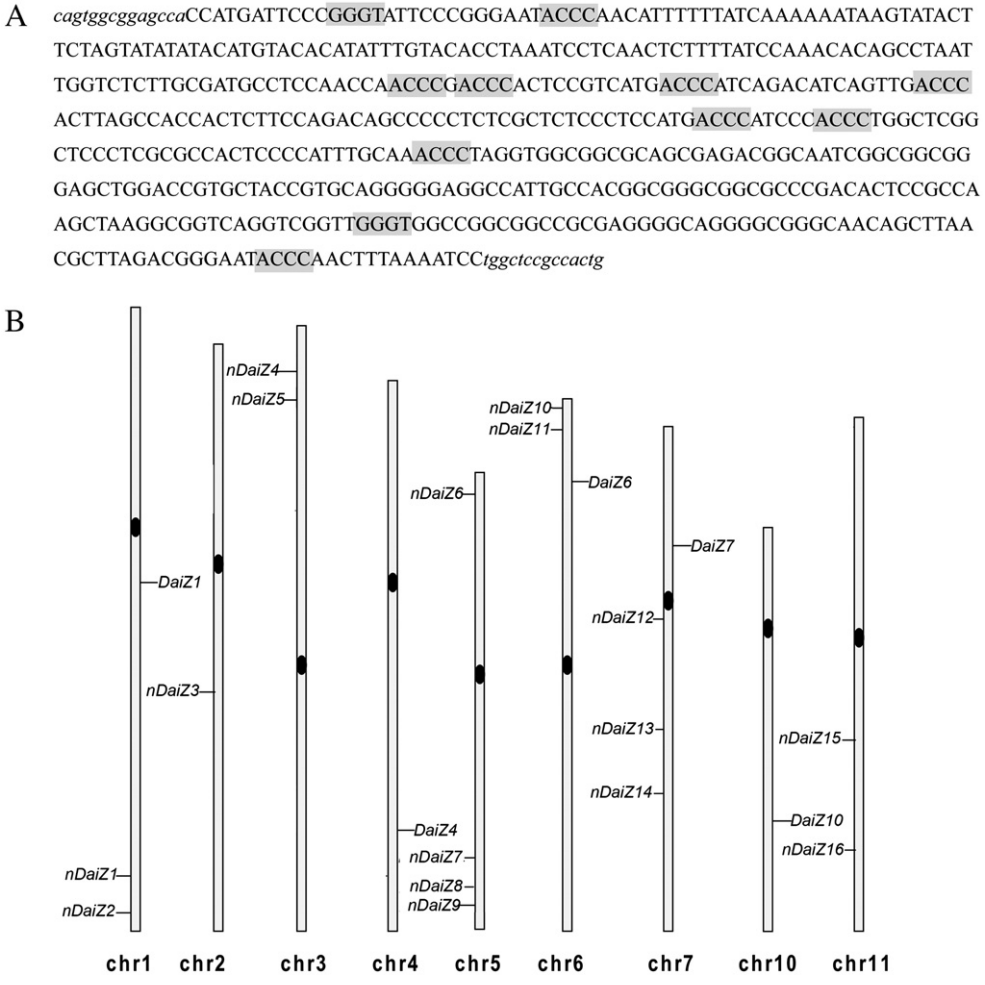

Fig. 3. Sequence and chromosomal location of *nDaiZ* elements in Nipponbare. (A) The sequence of *nDaiZ0* inserted in *GH2*. Terminal inverted repeats are represented by lower case letters and italics. Putative TAG1 (a member of the *hAT* superfamily) binding sites ACCC and its complementary sequence GGGT are marked by light gray boxes. (B) The locations of 16 *nDaiZ* elements and 5 *DaiZ* elements in Nipponbare genome. Longitudinal bars indicate chromosomes. Centromeres are marked by black circles.

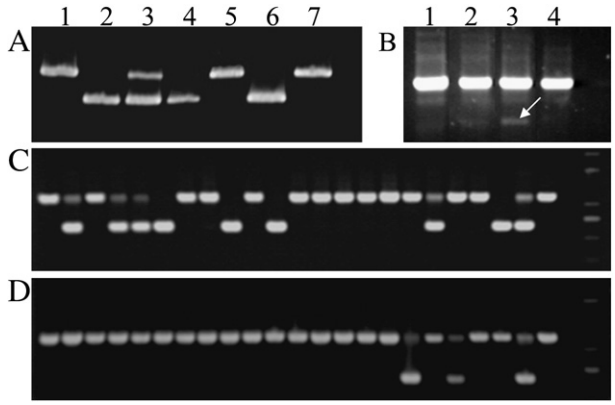

Fig. 4. Stability analysis of *nDaiZ9*. (A) Status of *nDaiZ9* in six individuals (lanes 1–6) of *gh2-1* and Nipponbare (lane 7). The upper bands indicate the original bands containing the *nDaiZ9*, the nether bands indicate the transposed *nDaiZ9*. (B) Status of *nDaiZ9* in three different generations of the calli. Lanes 1–3 are J1–J3. Arrow shows the band without *nDaiZ9*. Lane 4 is Nipponbare. (C, D) Status of *nDaiZ9* in regenerated plants induced from scutellem calli and anther calli, respectively.

sequence of chromosome 5 (Fig. 4A). As a result, a 628 bp band was amplified in the *gh2-1* mutant, while an 1156 bp band was detected in the wild type Nipponbare. Sequence analysis indicated that the 528-bp transposon sequence of *nDaiZ9* was lost in the 628 bp product from the *gh2-1* mutant compared with the 1156 bp product from Nipponbare.

nDaiZ and its component transposon system DaiZ are members of the hAT superfamily

Like other non-autonomous TEs of the *hAT* superfamily, *nDaiZs* have no coding capacity and are incapable of catalyzing their own transposition. Thus, the movement of *nDaiZs* must be catalyzed by a transposase encoded *in trans*. Further database searches using the UCSC's BLAT program identified five larger elements related to *nDaiZ* in the Nipponbare genome. They are located on chromosomes 1, 4, 6, 7 and 10, and thus are named as *DaiZ1*, *DaiZ4*, *DaiZ6*, *DaiZ7* and *DaiZ10*, respectively (Fig. 3B). The different *DaiZ* sequences are all 3488 bp in length and share highly identical ORFs. They also contain 14-bp TIRs with high similarity to those of *nDaiZ* elements (only a 2-bp mismatch at the TIRs compared with the *nDaiZ0*) (Fig. 2D and Supplementary Table 1). The terminal sequences of *DaiZs* are highly homologous to *nDaiZs*. For example, the 90 bp at the left terminus and the 42 bp at the right terminus, including the TIRs, are almost identical to the termini of *nDaiZ0* (Fig. 2D), while compared with *nDaiZ5*, the 46 bp and 510 bp at the two terminal sides of *DaiZ* share a high identity to those of *nDaiZ5*. These structural features indicate that the *nDaiZ* elements are generated from the *DaiZ* element by deletion of the internal region. Moreover, *DaiZ* insertion can also generate an 8-bp TSD.

The deduced amino acid sequence of the ORF of the *DaiZ* element was submitted to the SWISS-MODEL protein server (http://swissmodel.expasy.org/SWISS-MODEL.html) [19]. The exported result showed that the protein contains a putative *hAT* dimerization motif at the C terminus and a ZnF_BED domain near the N terminus [20,21], which

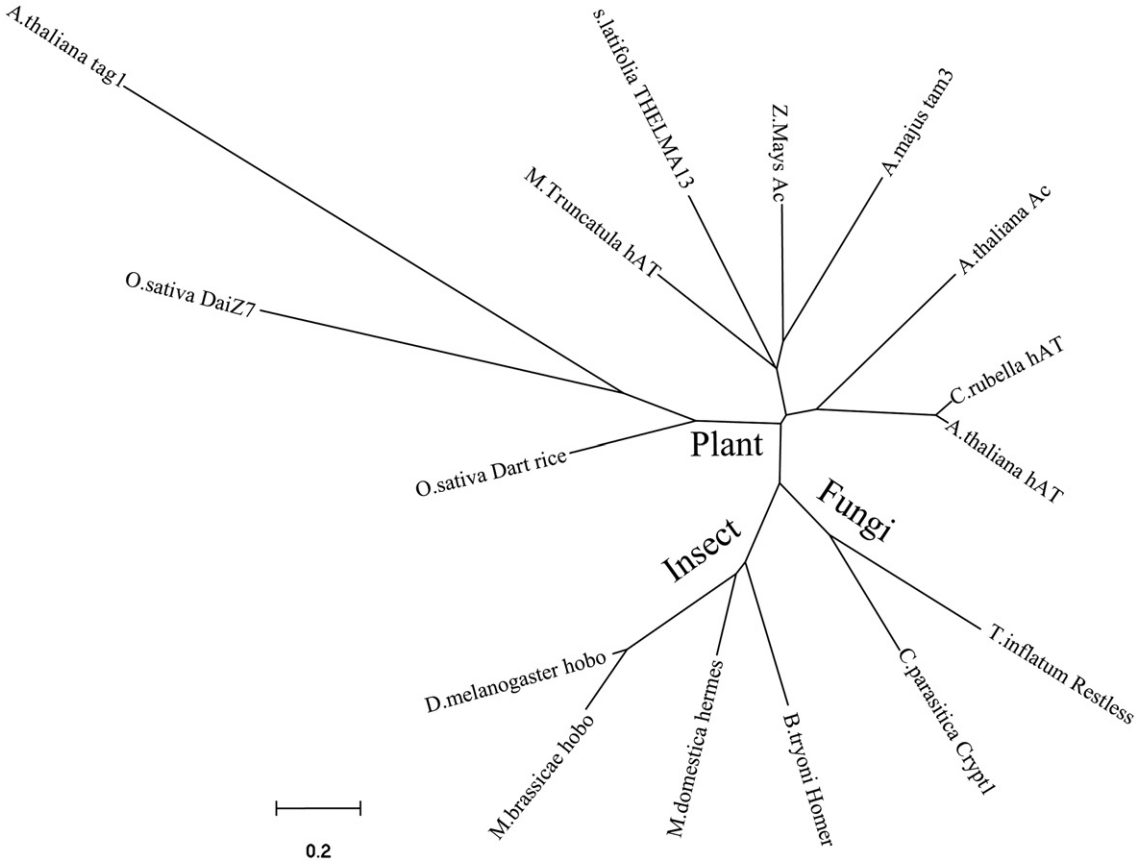

Fig. 5. Phylogenetic analyses of the *DaiZ* element and its putative transposase product in the *hAT* superfamily. The unrooted tree was constructed by the neighbor-joining method from a CLUSTALX alignment. 1000 replicates bootstrap values support the grouping of the plant, fungi and animal proteins. Accession numbers of the sequences used to build the tree are as follows: *hAT* elements CAB68118 (*Arabidopsis thaliana*), ABA18016 (*Capsella rubella*), ABD32595 (*Medicago truncatula*) *Ac* elements BAB02646 (*Arabidopsis thaliana*), P08770 (*Zea mays*), *THELMA13* AAP59878 (*Silene latifolia*), *Tam3* Q38743 (*Antirrhimum majus*), *Crypt1* AF283502 (*Cryptonectria parasitica*), Restless Z69893 (*Tolypocladium inflatum*), *hermes* U36211 (*Musca domestica*), *Homer* AF110403(*Bactrocera tryoni*), *hobo* 04705 (*Drosophila melanogaster*), *hobo* AF487501 (*Mamestra brassicae*), *Dart* BAD93710 (*Oryza sativa*), and *Tag1* AF051562 (*Arabidopsis thaliana*). The scale bar is an indicator of genetic distance based on branch length.

are highly conserved among the transposases of *hAT* that has been identified in all the eukaryotic kingdoms [22,23] (Fig. 2D). The features of the 8-bp TSD, 14-bp TIRs and the *hAT* dimerization motif reveal that *DaiZ* belongs to the *hAT* superfamily of class II transposons. Further BLASTP searches using the DaiZ protein as the query led to several putative transposases including proteins from plants, fungi and insects. All of them share the identical *hAT* dimerization motif. Phylogenetic results showed that although the *DaiZ* element belongs to the plant branch of the *hAT* superfamily, it has been differentiated pretty much among the plant *hAT* transposon superfamily (Figs. 5 and Fig. S1).

The polymorphisms of nDaiZ

Comparisons of the 16 copies found in Nipponbare revealed that they are not exactly identical. Only one of the 16 copies, *nDaiZ9*, was completely identical to *nDaiZ0*. However, although there are 11 copies of *nDaiZ* in the 93-11 genome, none of them is completely identical to *nDaiZ0*. All the other copies in Nipponbare, except *nDaiZ9*, and *nDaiZ10* have 1–4 bp changes in TIRs (Supplementary Table 1).

Southern blotting analysis using the 528-bp *nDaiZ0* as a probe showed almost identical bands among the different *japonica* rice varieties. Although many bands can also be detected in the *indica* varieties, the number of bands was fewer than in the *japonica* varieties, and the band patterns in *indica* varieties are also different from those in *japonica* varieties (Fig. 6A). The southern hybridization result was further verified by BLAT analysis between the two genomic sequences of *japonica* rice Nipponbare and *indica* rice 93-11. For example, *nDaiZ1*, *nDaiZ10* and *nDaiZ12* are always present at the same loci between the two cultivars, while *nDaiZ15* occurred at a different locus in the two cultivars. In Nipponbare, *nDaiZ15* locates to the BAC clone AC135568 on chromosome 11, while in 93-11, it situates at another site of chromosome 11 with an 8-bp (GCGGCGCC) TSD. On the contrary, at the corresponding locus of Nipponbare, the transposon is gone and left two 8-bp GCGGCGCC footprints, indicating that the transposon at this locus has been transposed during rice evolution. As a same result, the remaining *nDaiZ9* present in Nipponbare cannot be detected in the 93-11 genome.

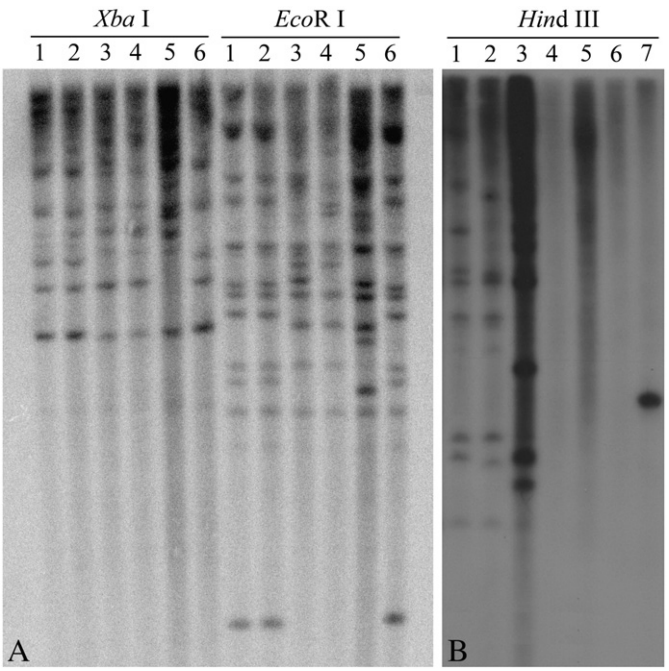

Fig. 6. Polymorphism of *nDaiZ* in different rice species. (A) Southern blot hybridized with *nDaiZ* in different *indica* and *japonica* varieties. Lanes 1–6, Nongken 57 Balilla, Nantehao, Nanjing 11, IR36 and Nipponbare respectively. Nongken57, Balilla and Nipponbare are typical *japonica* rice, while the others are typical *indica* rice. The DNA was digested with XbaI and EcoRI, and the full length of 528 bp *nDaiZ9* was used as the probe. (B) Southern blot hybridized with *nDaiZ* in different rice species. DNA were extracted from leaf tissue and digested with Hind III. Lane 1, *O. sativa*, AA genome; lane 2, *O. rufipogon* (ACC105404), AA genome; lane 3, *O. longistaminata* (ACC101776), AA genome; lane 4, *O. punctata* (ACC103896), BB genome; lane 5, *O. officinalis* (ACC101114), CC genome; lane 6, *O. rbizomatis* (ACC105659), CC genome; lane 7, *O. australiansis* (ACC100882), EE genome. The full length of *nDaiZ9* was used as the probe.

nDaiZ9 is an active transposable element induced by tissue culture

The retrotransposon *Tos17* of rice is mobilized under tissue culture [13]. Since the transposon *nDaiZ0* was first found from the regeneration line of tissue culture, we imagined that transposition of this mobile element may also be induced by tissue culture. We have already confirmed that the *nDaiZ0* found in *gh2-1* came from *nDaiZ9* on chromosome 5, so we first examined the presence or absence of *nDaiZ9* by PCR using 58 independent regenerated Nipponbare plants induced from scutellum calli and 93 induced from anther calli. As shown in Figs. 4C and D, the presence of both an upper and lower band in the same lane indicates that these plants are heterozygous at the *nDaiZ9* locus. However, we have not detected any transposition at the other 15 loci (data not shown), indicating the *nDaiZs* at those sites are quite stable. These results showed that the *nDaiZ9* element was excised from the original site of chromosome 5 with a high frequency in both scutellum and anther derived calli. Moreover, the mobilization frequency of scutellum derived calli is higher than that of anther derived calli. The excision also indicates that *nDaiZ9*, a member of *hAT* superfamily, belongs to the class II transposable elements and mobilizes by a "cut-and-paste" mechanism. The transposition of *nDaiZ9* was not found in intact plants originating from rice seeds, suggesting that it is quite stable under normal cultivation conditions.

The above results demonstrate that *nDaiZ9* is the only active *nDaiZ* element under tissue culture conditions. The transposable frequency is 29.31% (17/58) in scutellum culture derived plants and 8.60% (8/93) in anther culture derived plants.

nDaiZ elements are not existent in all rice species

To examine the presence and copy number of the *nDaiZ* element in several *Oryza* genomes, we performed southern blotting using the 528-bp *nDaiZ0* as a probe. We selected one accession from each species, including *O. rufipogon* and *O. longistaminata* with the AA genome, *O. punctata* with the BB genome, *O. officinalis* and *O. rbizomatis* with the CC genome, and *O. australiansis* with the EE genome. We cannot detect any hybridization signals in the BB or CC genome, and only one band was detected in EE genome. Many bands were detected in wild rice with the AA genome, but the copy number and band pattern are quite different among the different species (Fig. 6B). These results demonstrate that *nDaiZ* elements are flexible transposons during genome evolution and are not widely existent in all rice genomes.

The activation of nDaiZ9 is induced by hypomethylation during tissue culture

It has been demonstrated in nearly all eukaryotes that DNA methylation alterations in promoter and/or coding regions of a given gene may cause a change of its expression [24]. The retrotransposon *Tos17* and the MITE *mPing* are all able to be induced by tissue culture [16,25]. It has previously been shown that *Tos17* DNA methylation controls its transposition activity [25]. In the present study, we found *nDaiZ9* can also be activated by tissue

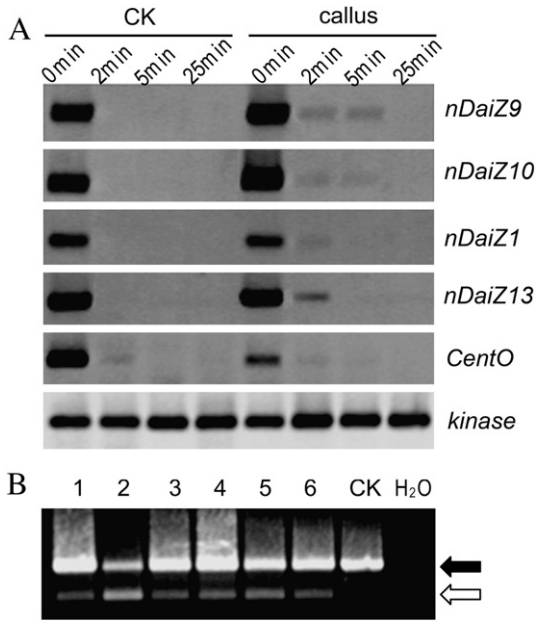

Fig. 7. Decreased DNA methylation at *nDaiZ* loci under culture condition and 5-azaC treatment. (A) *Mcr*BC-PCR analysis of DNA methylation at *nDaiZ1*, *nDaiZ9*, *nDaiZ10* and *nDaiZ13* loci. Equal amounts of the genomic DNA extracted from Nipponbare leaf tissue (CK) and scutellem induced calli were digested with *Mcr*BC for 0 min, 2 min, 5 min and 25 min, followed by PCR amplification using the primer pairs nDaiZ9, 10, 1, 13, H1035 and H1001. (B) Demethylation induced by 5-azaC treatment would promoter the *nDaiZ9* element's transposition. CK represents seeds grow under the same condition without 5-azaC treatment. Lanes 1–6 are six individual samples treated with 5-azaC. The filled arrow indicates the bands with *nDaiZ9*, and the open arrow indicates the bands without *nDaiZ9*.

culture. To examine whether the transposition activity was under the control of its DNA methylation status, we performed *Mcr*BC-PCR to check the DNA methylation level of the *nDaiZ9* locus and its franking regions. *Mcr*BC is an endonuclease which cleaves DNA containing methylcytosine on one or both strands, but will not act upon unmethylated DNA. The primer pair nDaiZ9-F and R was used to amplify the fragment containing the *nDaiZ9*. As a result, intact plants originating from seeds are highly methylated at the *nDaiZ9* locus (Fig. 7A). However, in calli where the transposition of *nDaiZ9* can be detected (Fig. 4B), the DNA methylation level is always reduced. It suggests that *nDaiZ9* transposition in intact plants originating from seeds is subject to DNA methylation-related suppression. While in callus, *nDaiZ9* transposition will be activated once its methylation level is reduced. To confirm this hypothesis, we treated Nipponbare seeds with 5-azaC, a DNA methylation inhibitor. PCR data indicated the transposition of *nDaiZ9* was activated in all 90 samples tested. Since the nether bands, which are not carrying the *nDaiZ9* sequence, are weaker, we selected six of them as templates for three more rounds of PCR amplification. As a result, clear bands were then seen (Fig. 7B). Hence, we concluded that the element *nDaiZ9* in intact plant was epigenetically silenced by DNA methylation, and demethylation during the process of tissue culture or 5-azaC treatment can promote the mobile element's transposition. Although the methylation level was also reduced in *nDaiZ1*, *10* and *13* loci after tissue culture (Fig. 7A), these *nDaiZ* elements were still stable. Except for these four loci, other *nDaiZ* elements are still highly methylated after tissue culture such as *nDaiZ8*, *nDaiZ15* and *nDaiZ16* (Fig. S2). We also examined whether the other elements except *nDaiZ9* could be activated after 5-azaC treatment. We found that all the other *nDaiZ* elements were quite stable even under the same treatment. We proposed that they either possess another mechanism to induce their transposition or have lost their transposable ability during their evolution.

Discussion

nDaiZ0 is a member of the hAT-like transposable elements

TEs are the main components of both eukaryotic and prokaryotic genomes, and play significant roles in genome evolution. The *hAT* transposon superfamily, whose name is from three well-described TE families: *hobo* from *Drosophila*, *Ac–Ds* from maize and *Tam3* from *Antirrhinum*, is ubiquitously distributed among eukaryotes. It is also considered as an ancient family and probably predates the plant–fungi–animal split 1000 million years ago [26]. Some additional elements such as *Hermes* from *Drosophila* and *TAG1* from *Arabidopsis* [18] also belong to the *hAT* superfamily. Based on the isolation of a rice *gh2-1* mutant, we have identified a new type of non-autonomous transposable element, *nDaiZ*, where its transposition capacity is highly induced by tissue culture. The element shares several characteristics with the *hAT* superfamily, such as generating an 8 bp target site duplication, and containing identical short TIRs (14 bp) and terminal sequences. The TIRs also have the (T/C)A(A/G)NG consensus that is considered a characteristic of *hAT* elements [22]. The *nDaiZ* elements comprise a number of tetranucleotide sequences, ACCC and its complementary sequences, GGGT, which are the binding motifs of the transposase TAG1, a member of the *hAT* superfamily [18]. Using the 528-bp *nDaiZ0* as a query, we identified other elements dubbed *DaiZ*, which have nearly identical TIRs and terminal sequences with an ORF between the TIRs. The predicted product of the ORF in *DaiZ* contains a 92 (771–862) amino acid *hAT* dimerization motif located at the C terminus. All these features indicate that the *nDaiZ/DaiZ* transposable system belongs to the *hAT* superfamily.

Recently, two active *hAT*-like elements *nDart* and *dTok*, have been isolated in rice. Both of them are active in intact plants, the difference is that *nDart* is active during vegetative stages [27], while *dTok* is active in the reproductive phase [28]. Nevertheless, the two transposons are not ubiquitously active in all *Oryza* species. Compared with the two transposons, *nDaiZ* does not share any similarity with them at the DNA sequence level. Phylogenetic analyses using several *hAT* transposases from different species, including plant, fungi and animal, reveals that *DaiZ* fell into the plant clade, but the differentiation was more severe compared with the other members of the plant *hAT* superfamily.

nDaiZ9 is an active TE induced by hypomethylation during tissue culture

PCR based analysis indicated that only *nDaiZ9* can be highly activated under tissue culture. Unlike the *mPing* element, which was highly induced by anther culture, the transposable frequency of *nDaiZ9* was lower in anther induced plants than that in scutellum induced plants, with a transposable frequency of 8.60% and 29.31%, respectively. The Nipponbare genome contains two copies of *Tos17*, but only the $Tos17^{chr7}$ is strongly activated by tissue culture. It was proposed that DNA methylation was a key factor controlling Tos17 activity, and another mechanism called transcriptional interference was also involved in the control of $Tos17^{chr10}$ activity [25]. We checked the methylation level of *nDaiZ1*, *8*, *9*, *10*, *13*, *15* and *16* by McrBC PCR, and found four of them (1, 9, 10, 13) were highly demethylated after tissue culture. However, only the transposition of *nDaiZ9* was detected, the others were quite stable under the same condition. The treatment of 5-azaC, a DNA methylation inhibitor, can also specifically induce the transposition of *nDaiZ9*. TIRs also play important roles in transposition. Among the 16 copies, 11 copies (*nDaiZ1*, *2*, *6*, *8*, *9*, *10*, *12*, *13*, *14*, *15*, *16*) have the (T/C)A(A/G)NG consensus sequence, perhaps the five copies that do not have the consensus sequence are unable to transpose and are probably the fossil transposons. Therefore, we proposed, besides the TIRs, the methylation status of the mobile element as well as its adjacent regions is also important for its transposition.

The nDaiZ are rice specific transposons and not widely existent in all Oryza species

BLAST analysis revealed that both *nDaiZ* and *DaiZ* have no homology in other species including *Arabidopsis*, maize, fungi and mammals. Southern blotting using the 528-bp *nDaiZ0* as the probe showed that only the AA and EE genomes of *Oryza* could be detected with the hybridization signals. Nevertheless, the band patterns observed in wild rice species with the AA genome differed from those in cultivar varieties, and only one locus was detected in the EE genome. The positions of *nDaiZ* between *japonica* and *indica* rice were also showed with many polymorphisms. For example, the *nDaiZ9* which is the only one active copy in Nipponbare doesn't exist in 93-11 genome. Another copy, which is located on chromosome 11 in the 93-11 genome with a GCGGCGCC TSD, is transposed at the same loci of the Nipponbare genome, and leaves two GCGGCGCC TSDs footprint, suggesting that *nDaiZs* were more stable in *indica* rice than in *japonica* rice during their evolution. Based on these results, we proposed that the *nDaiZ* elements are a new kind of DNA TE that had already existed before the *indica–japonica* separation.

nDaiZ9 is a potential insertion mutagen

TEs are useful tools for generating insertion mutants in a wide variety of organisms, including plants and animals [29–33]. *Ac*/*Ds* and *Tos17* have been successfully used in construction of mutant libraries in *Arabidipsis* and rice [34]. Both of these two elements can be induced by tissue culture, although the mutants derived from tissue culture are not always associated with the transposon insertion. The two transposons are widely considered an important tool for gene tagging. In the present study, we found only one copy, *nDaiZ9*, could be transposed at a high frequency under certain conditions, which led it to easily be detected after its transposition. Moreover, unlike the MITE elements which have TSD preference, the *nDaiZ* elements display random insertion into the genome. These two factors suggest that *nDaiZ* elements have potential abilities for tagging genes. Actually, we have already isolated an *nDaiZ9* insertion mutant, which shows a complete sterile phenotype in the 17 regeneration plants where *nDaiZ9* were transposed and the insertion was stable in the next several generations (unpublished data). The efforts for isolating more *nDaiZ9* insertion mutants will validate its advantages for gene tagging in rice functional genomic studies.

Materials and methods

Plant materials

A *japonica* rice variety, Nipponbare, was used in this study. We induced scutellum as well as anther calli and further obtained their regenerated plants to both the T0 and T1 generations. Several kinds of mutants were further identified from these plants.

Cloning the transposon

Total genomic DNA was extracted from leaf tissues of *gh2-1* using a modified CTAB method and purified by phenol extractions. PCR was performed using rTaq (Takara, Japan) as the enzyme and GC buffer I (Takara, Japan) as the reaction buffer. Products both from *nDaiZ0* insertion region at the *GH2* locus and from the *nDaiZ9* region were amplified using the primer pairs GH2-F and R and nDaiZ9-F and R (Supplementary Table 2), respectively. The purified products were then cloned using the pMD18-T cloning kit (Takara, Japan) and sequenced. PCR primers used to detect the stability of *nDaiZ* elements in regenerated plants and calli are listed in Supplementary Table 2.

Tissue culture and transposable frequency calculation

Calli were induced from scutellums and anthers individually. Scutellum callus induction and subculturing were carried out as described [35]. Anthers were collected from panicles at the single nuclear pollen stage. Anther callus induction and its subsequent propagation were performed as described [36]. The first generation of the calli induced directly from scutellums was named as J1, and the second generation of the calli after their subculture was named as J2, followed by the same analogy. We selected several calli from each generation and extracted their genomic DNA and used it as the template for PCR amplification. The left over calli were transferred into differentiation medium to regenerate plants. We calculated the transposable frequency as the proportion between the number of plants where transpositions occurred and the total regenerated plants obtained.

DNA gel blot analysis

10 μg of genomic DNA was digested with 50 U of the appropriate restriction enzymes (New England Biolabs) for 8 h. The digested DNAs were separated on 0.8% agarose gel and transferred onto Hybond N+ (Amersham). The *nDaiZ* DNA probe was obtained by two round PCR amplifications with the following primer pair: 5′-GTTTGGAGGGCTTTT-GACA-3′ and 5′-CCACGAACCCCACTATACG-3′, 5′-CAGTGGCGGAGC-CACCATGA-3′ and 5′-AAATCCTGGCTCCGCCACTG-3′. Then purified with the QIAquick Gel Extraction Kit (QIAGEN, USA), and further labeled with ^{32}P-dCTP using the prime-a-gene labeling system Kit (Promega, USA). The membrane was incubated in the prehybridization solution (molecular clone, version III) at 65 °C for 2 h and then hybridized with ^{32}P-labeled *nDaiZ* DNA for 24 h. Following the hybridization, the membrane was sequentially washed with 2×SSC, 0.1%SDS and 0.5×SSC, and 0.1%SDS at 65 °C for 15 min each. The radioactive membrane was then exposed to X-ray film.

5-azaC treated PCR analysis

The deglumed rice seeds were sterilized with 70% ethanol for 2 min and 4% sodium hypochlorite solution for 35 min, washed five times with sterilized water, and soaked in 1 mM sterilized 5-azaC solution at room temperature in the dark for 16 h. The treated seeds were grown on a 1/2 MS medium without any hormone or sugar, solidified with 0.25% phytogel (SIGMA) at room temperature for two weeks, and their leaf tissues were harvested. *nDaiZs*' transpositions among these plants were evaluated using the primer pairs listed in Supplementary Table 2.

McrBC-PCR

Genomic DNA was isolated from leaf tissue and calli. For *Mcr*BC-PCR analysis of *nDaiZ* elements, 500 ng of genomic DNA was digested with 20 U of *Mcr*BC restriction enzyme (New England Biolabs) for 0 min, 2 min, 5 min and 25 min, separately. Equal amount of *Mcr*BC-digested DNA was used for PCR amplification by primer pairs nDaiZ1, 8, 9, 10, 13, 15 and 16 listed in Supplementary Table 2. We used the primer pairs H1035 (U, 5′-TGT CAA TTC TAT CAA CCA TCA G-3′, L, 5′-CAA ATC ATC TAT CCT CAA GTC C-3′) and H1001 (U, 5′-CGA CTA AAC CAC TCC AAT CAT C-3′, L, 5′-CCA ATC AAA ACT TCT CCT GTA A-3′) as the positive and negative control primers. H1035 is a specific primer pair for the centromeric repeat CentO which is located in a highly methylated region [37]. H1001 is a specific primer pair for the C-Kinase substrate gene, which is situated in a region with less DNA methylation.

Database analysis

We searched for copies of *nDaiZ* from TIGR *japonica* rice genome (version 5) and BGI *indica* rice genome using UCSC's BLAT program

[38] with the following parameters: repmatch 10000, minScore 14, and minIdentity 90. Protein sequences were obtained from GenBank. The collected protein sequences were aligned by clustalw 1.8 [39]. The alignment was revised critically and used as the input for MEGA3 to construct phylogenetic tree. A neighbor-joining tree was built by MEGA3 using the Poisson correction model with gaps complete deletion. Topological robustness was assessed by bootstrap analysis with 1000 replicates [40].

The sequence of the transposon inserted in the *GH2* gene, *nDaiZ0* (or *nDaiZ9*), was deposited in GenBank under the accession no. FJ046194. We also deposited one representative copy for *DaiZ* on chromosome 7, *DaiZ7*, in GenBank as accession no. FJ167392.

Acknowledgments

We are grateful to Cory Hirsch and Jiming Jiang for critical reading of the manuscript. This work was supported by grants from the Ministry of Sciences and Technology of China (2005CB120805 and 2006AA02Z124), and the National Natural Science Foundation of China (30530070 and 30621001).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.ygeno.2008.11.007.

References

[1] C.B. Mc, The origin and behavior of mutable loci in maize, Proc. Natl. Acad. Sci. U. S. A. 36 (1950) 344–355.
[2] M.D. Adams, et al., The genome sequence of *Drosophila melanogaster*, Science 287 (2000) 2185–2195.
[3] E.S. Lander, et al., Initial sequencing and analysis of the human genome, Nature 409 (2001) 860–921.
[4] S.A. Goff, et al., A draft sequence of the rice genome (*Oryza sativa* L. ssp. *japonica*), Science 296 (2002) 92–100.
[5] J. Yu, et al., A draft sequence of the rice genome (*Oryza sativa* L. ssp. *indica*), Science 296 (2002) 79–92.
[6] T. Wicker, et al., A unified classification system for eukaryotic transposable elements, Nat. Rev. Genet. 8 (2007) 973–982.
[7] C.M. Vicient, et al., Retrotransposon BARE-1 and its role in genome evolution in the genus *Hordeum*, Plant Cell 11 (1999) 1769–1784.
[8] B.C. Meyers, S.V. Tingey, M. Morgante, Abundance, distribution, and transcriptional activity of repetitive elements in the maize genome, Genome Res. 11 (2001) 1660–1676.
[9] J. Yu, et al., Minimal introns are not "junk", Genome Res. 12 (2002) 1185–1189.
[10] V.V. Kapitonov, J. Jurka, Rolling-circle transposons in eukaryotes, Proc. Natl. Acad. Sci. U. S. A. 98 (2001) 8714–8719.
[11] R. Tarchini, et al., The complete sequence of 340 kb of DNA around the rice Adh1-adh2 region reveals interrupted colinearity with maize chromosome 4, Plant Cell 12 (2000) 381–391.
[12] N. Jiang, S.R. Wessler, Insertion preference of maize and rice miniature inverted repeat transposable elements as revealed by the analysis of nested elements, Plant Cell 13 (2001) 2553–2564.
[13] H. Hirochika, et al., Retrotransposons of rice involved in mutations induced by tissue culture, Proc. Natl. Acad. Sci. U. S. A. 93 (1996) 7783–7788.
[14] Y. Ding, et al., SDG714, a histone H3K9 methyltransferase, is involved in Tos17 DNA methylation and transposition in rice, Plant Cell 19 (2007) 9–22.
[15] M. Komatsu, K. Shimamoto, J. Kyozuka, Two-step regulation and continuous retrotransposition of the rice LINE-type retrotransposon Karma, Plant Cell 15 (2003) 1934–1944.
[16] K. Kikuchi, et al., The plant MITE mPing is mobilized in anther culture, Nature 421 (2003) 167–170.
[17] K. Zhang, et al., GOLD HULL AND INTERNODE2 encodes a primarily multifunctional cinnamyl-alcohol dehydrogenase in rice, Plant Physiol. 140 (2006) 972–983.
[18] A.M. Mack, N.M. Crawford, The *Arabidopsis* TAG1 transposase has an N-terminal zinc finger DNA binding domain that recognizes distinct subterminal motifs, Plant Cell 13 (2001) 2319–2331.
[19] N. Guex, M.C. Peitsch, SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling, Electrophoresis 18 (1997) 2714–2723.
[20] L. Essers, R.H. Adolphs, R. Kunze, A highly conserved domain of the maize activator transposase is involved in dimerization, Plant Cell 12 (2000) 211–224.
[21] M. Tsutsumi, A. Koga, H. Hori, Long and short mRnas transcribed from the medaka fish transposon Tol2 respectively exert positive and negative effects on excision, Genet. Res. 82 (2003) 33–40.
[22] E. Rubin, G. Lithwick, A.A. Levy, Structure and evolution of the hAT transposon superfamily, Genetics 158 (2001) 949–957.
[23] B.R. Calvi, et al., Evidence for a common evolutionary origin of inverted repeat transposons in *Drosophila* and plants: hobo, Activator, and Tam3, Cell 66 (1991) 465–471.
[24] A.P. Wolffe, M.A. Matzke, Epigenetics: regulation through repression, Science 286 (1999) 481–486.
[25] C. Cheng, M. Daigen, H. Hirochika, Epigenetic regulation of the rice retrotransposon Tos17, Mol. Genet. Genomics 276 (2006) 378–390.
[26] Z. Xu, H.K. Dooner, Mx-rMx, a family of interacting transposons in the growing hAT superfamily of maize, Plant Cell 17 (2005) 375–388.
[27] K. Fujino, H. Sekiguchi, T. Kiguchi, Identification of an active transposon in intact rice plants, Mol. Genet. Genomics 273 (2005) 150–157.
[28] S. Moon, et al., Identification of active transposon dTok, a member of the hAT family, in rice, Plant Cell Physiol. 47 (2006) 1473–1483.
[29] C. Martin, et al., Activity of the transposon Tam3 in *Antirrhinum* and tobacco: possible role of DNA methylation, EMBO J. 8 (1989) 997–1004.
[30] C.M. Ki, et al., Reprogramming of the activity of the activator/dissociation transposon family during plant regeneration in rice, Mol Cells. 14 (2002) 231–237.
[31] K. Kawakami, et al., A transposon-mediated gene trap approach identifies developmentally regulated genes in zebrafish, Dev. Cell. 7 (2004) 133–144.
[32] S. Ding, et al., Efficient transposition of the piggyBac (PB) transposon in mammalian cells and mice, Cell 122 (2005) 473–483.
[33] K. Tsugane, et al., An active DNA transposon nDart causing leaf variegation and mutable dwarfism and its related elements in rice, Plant J. 45 (2006) 46–57.
[34] T. Kuromori, et al., A collection of 11800 single-copy Ds transposon insertion lines in *Arabidopsis*, Plant J. 37 (2004) 897–905.
[35] Y. Hiei, et al., Efficient transformation of rice (*Oryza sativa* L.) mediated by *Agrobacterium* and sequence analysis of the boundaries of the T-DNA, Plant J. 6 (1994) 271–282.
[36] G.W. Schaeffer, F.T. Sharpe, Increased lysine and seed storage protein in rice plants recovered from calli selected with inhibitory levels of lysine plus threonine and S-(2-Aminoethyl)cysteine, Plant Physiol. 84 (1987) 509–515.
[37] K. Nagaki, et al., Sequencing of a rice centromere uncovers active genes, Nat. Genet. 36 (2004) 138–145.
[38] W.J. Kent, BLAT—the BLAST-like alignment tool, Genome Res. 12 (2002) 656–664.
[39] J.D. Thompson, D.G. Higgins, T.J. Gibson, CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice, Nucleic Acids Res. 22 (1994) 4673–4680.
[40] S. Kumar, K. Tamura, M. Nei, MEGA3: integrated software for Molecular Evolutionary Genetics Analysis and sequence alignment, Brief Bioinform. 5 (2004) 150–163.