A 5-methylcytosine DNA glycosylase/lyase demethylates the retrotransposon *Tos17* and promotes its transposition in rice

Honggui La, Bo Ding, Gyan P. Mishra, Bo Zhou, Hongmei Yang, Maria del Rosario Bellizzi, Songbiao Chen, Blake C. Meyers, Zhaohua Peng, Jian-Kang Zhu, and Guo-Liang Wang

DNA 5-methylcytosine (5-meC) is an important epigenetic mark for transcriptional gene silencing in many eukaryotes. In *Arabidopsis*, 5-meC DNA glycosylase/lyases actively remove 5-meC to counteract transcriptional gene silencing in a locus-specific manner, and have been suggested to maintain the expression of transposons. However, it is unclear whether plant DNA demethylases can promote the transposition of transposons. Here we report the functional characterization of the DNA glycosylase/lyase DNG701 in rice. DNG701 encodes a large (1,812 amino acid residues) DNA glycosylase domain protein. Recombinant DNG701 protein showed 5-meC DNA glycosylase and lyase activities in vitro. Knockout or knockdown of DNG701 in rice plants led to DNA hypermethylation and reduced expression of the retrotransposon *Tos17*. *Tos17* showed less transposition in calli derived from *dng701* knockout mutant seeds compared with that in wild-type calli. Overexpression of DNG701 in both rice calli and transgenic plants substantially reduced DNA methylation levels of *Tos17* and enhanced its expression. The overexpression also led to more frequent transposition of *Tos17* in calli. Our results demonstrate that rice DNG701 is a 5-meC DNA glycosylase/lyase responsible for the demethylation of *Tos17* and this DNA demethylase plays a critical role in promoting *Tos17* transposition in rice calli.

In eukaryotes, DNA cytosine methylation at carbon 5 of the pyrimidine ring [5-methylcytosine (5-meC)] is an important epigenetic mark that contributes to gene silencing and plays critical roles in development and genome defense against viruses, transposons, and transgenes (1–3). Heavy cytosine methylation usually occurs at heterochromatin and at regions rich in transposons and repetitive DNA (2, 4). Cytosine methylation of DNA, however, is reversible through demethylation (5, 6). DNA demethylation can be passive or active, and active DNA demethylation is catalyzed by one or more enzymes to remove methylated cytosines and can occur independently of DNA replication (7, 8). So far, several models have been proposed to explain mechanisms of DNA demethylation in animals (7, 8). One of the models suggests that active DNA demethylation is mediated at least in part by a base excision repair (BER) pathway where the AID/ApoBec family of deaminases convert 5-meC to T followed by G/T mismatch repair through the DNA glycosylase MBD4 or TDG with the involvement of Gadd45α (7–9). Recently, a new study suggested that 5-meC hydroxylase TET1 promotes DNA demethylation in mammalian cells through a process that requires the BER pathway (10). Nevertheless, many aspects of these models have not been confirmed, and how DNA demethylation is carried out in animals remains controversial (7, 8).

In contrast to the uncertainty about the mechanism of DNA demethylation in animals, mechanisms of DNA demethylation in plants are much clearer and widely accepted. In the model plant *Arabidopsis*, research showed that the 5-meC DNA glycosylase/lyase-mediated BER pathway plays critical roles in DNA demethylation process (3, 5, 6, 11, 12). There are four 5-meC DNA glycosylase/lyases [i.e., REPRESSOR OF SILENCING 1 (ROS1), DEMETER (DME), DEMETER-LIKE2 (DML2) and DEMETER-LIKE3 (DML3)] found in the *Arabidopsis* genome, and genetic and biochemical analysis revealed that all four proteins function as DNA demethylases, and are involved in genomic DNA demethylation (1, 3, 5, 6, 12–14). ROS1 is required for the prevention of hypermethylation and transcriptional silencing of a repetitive reporter gene (1). DME was reported to be necessary for endosperm gene imprinting and seed viability (3, 13). Both DML2 and DML3 are required for removing 5-meC from improperly methylated DNA in target sites (12). Biochemical evidence demonstrated that the DNA glycosylase activity of these proteins removes 5-meC from DNA backbone and then the lyase activity cleaves the DNA backbone at the abasic site by successive β- and δ-elimination reactions (3, 5, 6, 12, 14). However, compared with these well-documented 5-meC DNA glycosylase/lyases in *Arabidopsis*, no 5-meC DNA glycosylase/lyase has been well characterized in other plant species.

As one of the most important food crops in the world, rice (*Oryza sativa*, L.) has become a model monocot species for functional genomics research. Nipponbare, a completely sequenced japonica rice cultivar, is now widely used for functional analysis of genes. *Tos17*, a well-characterized rice retrotransposon, is extensively used to generate *Tos17* insertional mutants in Nipponbare (15, 16). There are two almost identical *Tos17* copies, named *Tos17chr7* and *Tos17chr10* (located on chromosomes 7 and 10, respectively) in Nipponbare (17). *Tos17chr7* transcripts (∼4 kb) solely accumulate in calli and *Tos17chr10* is inserted into a putative ABC-type transporter gene to form a fusion gene that generates a ∼7-kb readthrough transcript (17). Both *Tos17chr7* and *Tos17chr10* are highly methylated and immobilized under normal plant growth conditions (16, 17). However, in calli the transposition of *Tos17* is activated and the copy number increases gradually during prolonged callus culture (16). Several recent studies showed that the activation of transcription as well as transposition of *Tos17* in calli is associated with DNA hypomethylation (17–19). However, the


The authors declare no conflict of interest.

1To whom correspondence may be addressed. E-mail: jkzhu@purdue.edu or wang.620@osu.edu.
protein that is directly responsible for the active demethylation of Tos17 in the rice genome is as yet unknown.

In this report, we identified and characterized a rice DNA glycosylase/lyase, DNG701 (for DNA Glycosylase 701) (http://www.chromdb.org/). Enzymatic activity assays demonstrated that DNG701 is an active 5-meC DNA glycosylase/lyase in vitro. Knockout or knockdown and overexpression of DNG701 in rice calli and in planta showed that DNG701 is critical for Tos17 demethylation, expression, and transposition in rice. Our demonstration of a host plant factor promoting the transposition of Tos17 suggests that the retrotransposon Tos17 may have a beneficial function to the rice genome during evolution.

Results

**DNG701 Encodes a Rice 5-meC DNA Glycosylase/Lyase.** Among the five predicted rice DNA glycosylases (http://www.chromdb.org/), DNG701 and its close paralog DNG704 are most closely related to *Arabidopsis* DML2 and ROS1 (Fig. S1A). As with the other four rice DNA glycosylases, DNG701 contains a DNA glyco-
sylase domain immediately followed by an EndIII_4Fe-4S do-
main, and two putative nuclear localization signals are located at the N terminus of the protein (Fig. S1B) (20). Similar to the four *Arabidopsis* and other four rice DNA glycosylases, the DNG701 glycosylase domain contains a helix–hairpin–helix motif and a glycine/proline-rich loop with an invariant aspartic acid (GPD) (Fig. S1C). An extremely conserved lysine (K) is located inside one of the highly conserved helices (Fig. S1C).

A comparison between the cloned DNG701 cDNA (GenBank accession no. FJ536320) and its genomic DNA sequence revealed that the ORF of DNG701 consists of 19 exons and encodes a protein of 1,812 amino acids (Figs. S1B and S2A). RT-PCR analysis indicated that DNG701 was ubiquitously expressed in the three tissues (leaves, stems, and roots) of 3-wk-old Nip-
ponbare seedlings (Fig. S2B, Upper) and four tissues (leaves, stems, roots, and panicles) of 3-mo-old flowering Nipponbare plants (Fig. S2B, Lower). The fused GFP-DNG701 protein was predominantly localized in the nucleus, whereas the GFP con-
tral protein was distributed in both the nucleus and cytoplasm of rice cells (Fig. S2C).

The purified recombinant maltose-binding protein-DNG701 (MBP-DNG701) fusion protein had nicking activity against full-
ly methylated and hemimethylated pUBIGUS plasmids, which was generated separately by CG methyltransferase (M.SssI) and by MsPl methyltransferase, but did not have any nicking activity against unmethylated pUBIGUS (Fig. S3 A and B, Upper). The purified control protein MBP possessed no nicking activity against either methylated or unmethylated pUBIGUS (Fig. S3B, Lower). Furthermore, we conducted incision assays against 40-bp double-stranded oligonucleotides containing fully methylated and hemimethylated CmCGG or mCCGG (Fig. S3C) (5). The double-stranded oligonucleotides were produced by annealing 32P-labeled single-stranded oligonucleotides to unlabeled com-
plementary strands (Fig. S3C). Incubation of MBP-DNG701 with either CmCGG- or mCCGG-containing oligonucleotides generated one 16-nt and one 29-nt cleavage products corresponding to fully methylated and hemimethylated sites, respectively, sug-
gesting that the protein has active incision activity in vitro against fully methylated and hemimethylated oligonucleotides (Fig. 1 A and B). In contrast, MBP showed no incision activity on the two types of oligonucleotides (Fig. 1A). With an increase in the amount of MBP-DNG701, more 16-nt and 29-nt cleavage products were produced from the methylated oligonucleotides (Fig. 1B). Furthermore, four other types of oligonucleotides, containing either fully methylated or hemimethylated CmCGG or mCCGG, were incubated with MBP-DNG701 (Fig. S3C). The results again showed that MBP-DNG701 had incision activity against the fully methylated or hemimethylated CmCGG-
or mCCGG-containing oligonucleotides, but had no incision ac-
tivity against unmethylated oligonucleotides (Fig. 1C). Taken together, these data showed that DNG701 has active DNA glyco-
sylase/lyase activity against 5-meC-containing DNA in vitro.

**Knockout or Knockdown of DNG701 Causes DNA Hypermethylation of Tos17 Retrotransposons in Rice.** To further elucidate the function of DNG701 in rice, we obtained two Tos17 insertion mutants named *dng701-1* and *dng701-2* (Fig. S2A). RT-PCR analysis suggested that DNG701 is knocked out in both the mutants (Fig. S4, Left). Meanwhile, homozygous *DNG701* RNA interference (RNAi) transgenic lines PCs-5 and PCs-10 were selected as well for detailed analysis because the expression of *DNG701* was substantially down-regulated although that of *DNG704* was not affected (Fig. S4, Right). DNA gel blot analysis revealed that there were no apparent DNA methylation changes in the cen-
tromeric repeat DNA, 45S rDNA, or Ty1 retrotransposons in both the mutants and RNAi lines, indicating that knockout or knockdown of *DNG701* does not cause global DNA hyper-
methylation in these regions (Fig. S5A). However, substantial DNA hypermethylation was observed in the retrotransposon *Tos17* (Fig. S4 and Fig. S6A). These results indicated that loss of...
function of DNG701 leads to DNA hypermethylation in a locus-specific manner. Bisulfite sequencing results further verified that methylation levels at both CG and CHG (H is A, T, or C) sites of Tos17 were increased in the mutants and RNAi lines (Fig. 2B and Fig. S6B).

We then examined expression levels of Tos17 in the mutants and RNAi lines and found that the expression of Tos17chr10 transcripts (~7 kb) was substantially reduced (Fig. 2C). The Tos17chr10 transcripts (~4 kb) were not detected in either the wild-type Nipponbare or the two mutants and RNAi lines (Fig. 2C). To understand if the suppression of Tos17chr10 expression is caused by hypermethylation of its promoter, we examined the methylation level of the promoter region of ~1-kb upstream of the translation start site by bisulfite sequencing (Fig. S6A). A 202-bp fragment from ~445 to ~244 was heavily methylated in all of the genotypes and only the CHH methylation levels were slightly increased in the mutants and RNAi lines, suggesting that methylation in this region may not account for the reduction of Tos17chr10 expression (Fig. S5B). Because the LTR region is known as an enhancer-promoter of LTR retrotransposons and is usually a target of transcriptionally repressive methylation (17, 21), we determined the methylation levels of 5′ LTR regions of Tos17chr10 and Tos17chr7 and observed much higher DNA methylation levels in the mutants and RNAi lines than in the controls (Fig. 2D and E). Taken together, these data indicate that knockout or knockdown of DNG701 leads to DNA hypermethylation at 5′ LTR as well as coding regions of Tos17 and transcriptional repression of Tos17chr10, suggesting that DNG701 is an important factor for protecting Tos17 from hypermethylation and silencing. In addition to Tos17, we also found that another retrotransposon (LOC Os07g31690) was hypermethylated in its 3′ terminal region in both the dng701 mutants, further supporting that DNG701 is an active demethylase/lyase that removes 5-meC at some loci in the rice genome (Fig. S5C).

Interestingly, seeds harvested from mature panicles of both mutants fell into two categories: normal seeds and wrinkled seeds (13.66% wrinkled seeds for dng701-1, 10.51% for dng701-2, and 0.79% for wild-type Nipponbare). The observation suggests that DNG701 is involved in seed development of rice (Fig. S7).

Overexpression of DNG701 Leads to DNA Hypomethylation in Rice Calli and Transgenic Plants. To further understand the function of DNG701 in rice, we overexpressed DNG701 in Nipponbare under the control of the maize ubiquitin promoter. Two independently transformed callus lines, OX-1 and OX-2, were obtained after a 2-mo selection on hygromycin-containing media. Simultaneously, two other independent callus lines, OX-CK-1 and OX-CK-2, which were transformed with the empty vector alone, were obtained and used as controls. Next, each callus line was subdivided into a few smaller callus clones and continuously cultured for another 3 mo. RNA gel-blot analysis demonstrated that the expression levels of Tos17chr10 in the OX-1 and OX-2
callus lines were increased when DNG701 was overexpressed, and low levels of Tos17chr5 transcripts were detected in all genotypes (Fig. 3A). DNA gel-blot analysis revealed that both the OX-1 and OX-2 callus lines showed hypomethylation in Tos17 relative to the control callus line, which presumably accounts for the enhanced expression of Tos17chr10 (Fig. 3B). This hypomethylation of Tos17 was further confirmed by bisulfite sequencing analysis (Fig. S8). Interestingly, CG methylation level of the 45S rDNA also declined in the OX-1 and OX-2 callus lines (Fig. 3C). Furthermore, methylation levels of the centromeric repeat DNA and Ty1 retrotransposons were also slightly reduced in the overexpression lines (Fig. 3C).

DNG701-overexpressing transgenic plants and transgenic controls regenerated from the 2-mo-old transformed calli were used to study the impact of DNG701 overexpression on Tos17 in planta. As shown in Fig. 4A, the expression levels of Tos17chr10 were markedly increased with the overexpression of DNG701 in the OX-1 and OX-2 transgenic lines. DNA gel-blot analysis and bisulfite sequencing results revealed reduced methylation levels at its promoter region and gene body in the overexpression calli (Fig. 4B–D and Fig. S6A). The transcripts of Tos17chr7 were not detected in either overexpression line despite the hypomethylation of Tos17 (Fig. 4A). Bisulfite sequencing results of the 5′ LTR regions of both Tos17chr10 and Tos17chr7 demonstrated that CG methylation levels in Tos17chr10 and Tos17chr7, and CHG and CHH methylation levels in Tos17chr7 were reduced in the overexpression plants (Fig. 4E and F). Like the overexpression calli, the overexpression plants also showed a remarkably decreased CG methylation in 45S rDNA (Fig. S9). These data suggest that overexpression of DNG701 has extensive effects on reducing DNA methylation of the rice genome.

DNG701 Is a Critical Factor That Promotes Tos17 Transposition in Calli. Because Tos17 is hypomethylated in the OX-1 and OX-2 callus lines, we wondered if transposition of Tos17 in the two lines was affected. DNA gel-blot analysis was performed to examine Tos17 copy number in 6-mo-cultured (since callus induction) callus clones derived from the OX-1 and OX-2 callus lines and the control lines. As shown in Fig. 3D (Left), an apparent increase in the copy number of Tos17 was observed in the callus clones derived from the OX-1 and OX-2 callus lines in which an average of 6.8 and 6.3 Tos17 copies were present, respectively. In contrast, there were on average 3, 2, and 3 Tos17 copies found in the callus clones from Nipponbare, OX-CK-1 and OX-CK-2, respectively (Fig. 3D, Left).

Furthermore, independent calli from Nipponbare and the dng701-2 mutant seeds were induced and cultured for 10 mo. Eight independent calli from each material were subjected to DNA gel-blot analysis to investigate the change in Tos17 copy number. In parallel, leaf samples from both genotypes were used as the controls to assess native copy number of Tos17. The results revealed that calli derived from dng701-2 mutant seeds gained only one additional Tos17 copy relative to the leaf control (Fig. 3D, Right). In contrast, calli derived from Nipponbare gained 26 additional Tos17 copies compared with the leaf control (Fig. 3D, Right). The increased copy number in these samples was not attributable to incomplete digestion because only a single band was observed when the blot was reprobed with Tubulin (Fig. 3D, Right). Taken together, these results indicate that DNG701 is necessary and also a limiting factor for high transposition activity of Tos17 in rice calli.

Discussion

Our work demonstrated that DNG701 is an active rice 5-meC DNA glycosylase/lyase, and it is able to process fully methylated
or hemimethylated external or internal 5-meC efficiently (Fig. 1 A–C). Like the Arabidopsis ros1 mutant (4), the rice dng701 knockout mutants show hypermethylation at some transposons, indicating that loss of function of DNG701 leads to DNA hypermethylation in a locus-specific manner (Fig. 2A and Fig. S3C). We did not observe apparent methylating changes in the centromeric repeat DNA, 45S rDNA, or Ty1 retrotransposons between dng701 mutants and wild-type controls (Fig. S3A). It is likely that these regions have already been hypermethylated in the wild-type control genomes, and any methylation-enhancing effect of dng701 on these regions is difficult to detect using methylation-sensitive restriction enzymes. There is no report about the impact of overexpression of ROS1 or DML2 on global DNA methylation levels at certain repetitive sequences, although ROS1 was reported to be localized in nucleoplasm and nucleolus that contain many copies of 45S rDNA units (5, 12, 22). Overexpression of DNG701 in rice calli and plants led to marked hypomethylation of 45S rDNA and slight hypomethylation on centromeric repeat DNA and Ty1 retrotransposons, suggesting that DNG701 is targeted to these regions, and overexpression of DNG701 presumably has significant dose effect to excessively demethylate them (Fig. 3C and Fig. S9). Neither the Arabidopsis ros1 nor the dml2 mutant produces wrinkled seeds (1, 12, 14), whereas the rice dng701 mutants produce a substantial proportion of wrinkled seeds (~10–13% for mutants and ~0.8% for Nipponbare) (Fig. S7). Recently, two studies revealed that both Arabidopsis and rice endosperm are hypomethylated in all sequence contexts, and CG methylation levels are partially restored in both gene bodies and repeats in dme endosperm, suggesting the participation of DME in endosperm demethylation (23, 24). It is likely that DNG701 is also involved in endosperm demethylation in rice and, unlike its Arabidopsis counterparts (DML2 and ROS1), targets developmentally important genes. Compared with Arabidopsis, the rice genome contains a substantially higher amount of transposons and other repetitive elements (25, 26). Therefore, it is possible that more genes in rice are subject to regulatory control by repetitive elements, and thus more genes are likely targeted for demethylation by DNG701. It will be of interest to determine the developmentally important target genes of DNG701 in the future.

The relationships among DNA methylation, transcription, and transposition of transposable elements in plant kingdom have been explored extensively (19, 21, 27–29). In Arabidopsis, it has been reported that DNA methylation is involved in controlling mobilization of transposons (27–29). In the ddm1 mutant and the self-pollinated progeny, silent CACTA elements and several transposons like ATGP1 and ATGP2 are transcriptionally and transpositionally activated because of DNA hypomethylation (28, 29). In the cmt3 met1 double mutant, transcription and high-frequency transposition of the CACTA elements were detected, suggesting that CG and non-CG methylation functions to immobilize transposons (27). Likewise, the reduced DNA methylation levels in rice correlate with activation of transcription and transposition of Tos17 (17, 18). Other research found that reduced histone H3K9 dimethylation decreases DNA methylation of Tos17, leading to its active transcription and transposition in SDG714 RNAi transformants (19); however, the enzymes directly responsible for demethylation of Tos17 have remained unclear. Our work suggests that DNG701 is responsible for active DNA demethylation of Tos17. In dng701 mutants, the normally expressed Tos17chr10 is silenced, which is accompanied by elevated DNA methylation levels at the 5’ LTR region and gene body (Fig. 2A–E). In dng701 mutant calli, Tos17 transposition activity is much reduced (Fig. 3D, Right). In contrast, overexpression of DNG701 in rice calli and plants not only decreases the DNA methylation levels of Tos17 and up-regulates Tos17chr10 expression, but also promotes transposition of Tos17 in rice calli (Figs. 3 A, B, and D, Left, and 4). Interestingly, the autonomous expression of the Tos17chr10 is not affected by its DNA methylation levels, suggesting that DNA hypomethylation
is not a sole determinant for the expression of *Tos17*chr7, and there must be other factors to regulate its expression in rice plants (Fig. S4). We speculate that overexpression of *DNG701* might generate certain immediate triggers to contribute to the transposition of *Tos17*chr7. One possible trigger is the favorable chromatin state/environment. Our results showed that in the overexpression calli the methylation levels of 45S rDNA were obviously decreased, and the methylation levels of centromeric repeats and Ty1 transposons were also slightly reduced, indicating that overexpression of *DNG701* has extensive effects on the methylation levels of callus genome (Fig. 3C). Therefore, it is likely that the chromatin state/environment of the overexpression calli might have been altered because of the overexpression of *DNG701*, which may make the chromatin prone to *Tos17* integration. Another possible trigger is posttranscriptional regulation. Overexpression of *DNG701* may enhance translation/reverse transcription of *Tos17* transcripts through an unknown posttranscriptional regulation. Our results suggest that *DNG701* is required for *Tos17* demethylation and expression of *Tos17*chr10. Importantly, *DNG701* is required for the high transposition activity of *Tos17* in rice calli. Our results also suggest that *DNG701* is a limiting factor for *Tos17* transposition. The fact that plants have evolved active DNA demethylases that promote the transposition of transposons implies that the affected transposons also have beneficial effects to the plant species during evolution.

**Materials and Methods**

**Plant Materials.** All rice plants used in this study were *Oryza sativa* L. ssp. *japonica* cv. Nipponbare. Details are available in **SI Materials and Methods**.

**Cloning of *DNG701* cDNA, Subcellular Localization of *DNG701*, RNA Gel-Blot and RT-PCR Analysis, and DNA Gel-Blot Analysis.** See details in **SI Materials and Methods** and Table S1.

**Protein Purification and Enzymatic Activity Assays.** The entire *DNG701* cDNA was inserted in-frame into the pMAL-c2x vector (New England Biolabs) to obtain a *malE*-*DNG701* fusion construct. Expression and purification of the MBP-*DNG701* fusion protein and MBP protein were performed as previously described (5). For the nicking activity assays, 200 ng of each of these plasmids was added to nicking reaction buffer (10 mM Tris-HCl, pH 7.9, 50 mM NaCl, 10 mM MgCl2, 1 mM DTT, 0.2 mg/mL BSA) plus purified MBP-*DNG701* or MBP. Reactions were carried out as previously described (5). For incision activity assays, upper strands of the seven oligonucleotides (Fig. S3C) were labeled with 32P at the 5′ ends, and then annealed to the corresponding lower strands, as previously described (5). Labeled double-stranded oligonucleotides (1 pmol) were incubated with MBP-*DNG701* or MBP or MspI endonuclease in the nicking reaction buffer in a total volume of 10 μL at 37°C for 2 h. The products were separated on 17% denaturing polyacrylamide gels containing 8 M urea, and then the gels were exposed to a Phosphorimager screen. See **SI Materials and Methods** for details.

**Genomic Bisulfite Sequencing.** Two micrograms of genomic DNA was digested with 20 units of appropriate restriction enzymes (New England Biolabs) for 3 h. Next, the digested DNA was purified by ethanol, treated with a sodium bisulfite solution and desalted, as previously described (19). Subsequently, 4 μL of each recovered DNA sample was subjected to PCR amplification. The PCR products were cloned into the pGEM-T Easy cloning vector (Promega), and 15 to 20 clones were sequenced to calculate the percentage of cytosine methylation in each sample. Details are in **SI Materials and Methods**.

**Sequence Alignment and Phylogenetic Analysis.** See **SI Materials and Methods** for details.

**ACKNOWLEDGMENTS.** We thank Drs. Ko Shimamoto and Michael Goodin for providing the pANDA vector and pGDG vector, respectively; Dr. Steven E. Jacobsen for providing the genomic bisulfite sequencing protocol; Dr. Jiming Jiang for providing the centromeric repeat probe; and the Rice Genome Resource Center in Japan for providing the cDNA clones and *Tos17* insertion lines. This research is supported by National Institutes of Health Grant R01GM070795 (to J.-K.Z.), and US Department of Agriculture-Coporate State Research, Education, and Extension Service Grant 2004-05425 and National Science Foundation-Plant Genome Research Program Grants 0321437 and 0701745 (to G.-L.W.).

---

Cloning of DNG701 cDNA. To clone the DNG701 cDNA, we amplified a 2,404-bp cDNA fragment corresponding to the 3′ terminus of DNG701 (primers P1-F/P1-R) from a cDNA pool derived from the young leaves of Nipponbare. A 3,618-bp cDNA fragment corresponding to the 5′ terminus of DNG701 was amplified (primers P2-F/P2-R) from cDNA clone AK070640 obtained from KOME’s (Knowledge-based Oryza Molecular Biological Encyclopedia) cDNA library. The DNG701 cDNA was obtained by ligating the two fragments via the EcoRV restriction site. The primer sequences are listed in Table S1.

Subcellular Localization of DNG701. The entire 5,439-bp DNG701 cDNA was inserted in-frame into the 3′ terminus of the GFP gene in the pGPDG vector (2), where the fused gene was driven by the 35S promoter. This construct was then transfected into rice protoplasts prepared from Nipponbare 10-d-old seedlings by the PEG-mediated transfection method (3). The pGpdG vector was transfected simultaneously as a control. GFP fluorescence was observed with a Nikon Eclipse E600 fluorescence microscope (Nikon). Excitation and emission filters Ex450-490/DM510/BA520-560 were used for GFP fluorescence (3). Images were captured with a SPOT 2 Slide charge-coupled camera.

RNA Gel-Blot and RT-PCR Analysis. Total RNA was isolated from wild-type, mutant, and RNAi transgenic plants using TRizol reagent (Invitrogen) and then treated with DNase I (Ambion). The procedures for probe labeling, blotting, hybridization, and membrane washing were carried out as previously described (4). After hybridization was performed overnight at 65°C, the membrane was washed and then exposed to a PhosphorImager screen (Amersham). Probes for Tosi17 and DNG701 were amplified by primers P3-F/P3-R (Probe 2) (Fig. S6B) and P4-F/P4-R, respectively. For RT-PCR analysis, 1.5 μg of treated total RNA was used to synthesize the first-strand cDNA with a reverse-transcription kit (Promega), as previously described (4). Two sets of primer pairs (P4-F/P5-R for DNG701; P6-F/P6-R for DNG704) were separately applied to examine the expression levels and patterns of DNG701 and DNG704. The Ubiquitin gene was used as an internal control for the RT-PCR (primers P7-F/P7-R). All primers are listed in Table S1.

DNA Gel-Blot Analysis. Genomic DNA (20 μg) isolated from different tissues was digested with 50 units of appropriate restriction enzymes (New England Biolabs) for 16 h. The digested DNA was separated on a 1% agarose gel and then blotted on Hybond N* nylon membrane. The procedures for preparation of 32P-labeled probe, hybridization, and membrane washing were as previously described (4, 5). After hybridization was performed overnight at 65°C, the membrane was exposed to the PhosphorImager screen. The probe used to check the centromeric repeat was amplified from a cDNA clone (Genbank accession no. EB086903). The probes of Tosi17 (Probe 1) (Fig. S6B), 45S rDNA, Ty1, Tosi17 R′ probe (Promoter 3) (Fig. S6A), actin, and Tubulin were amplified by primers P8-F/P8-R, P9-F/P9-R, P10-F/P10-R, P11-F/P11-R, P12-F/P12-R, and P13-F/P13-R, respectively. All primers are listed in Table S1.

Protein Purification and Enzymatic Activity Assays. The entire DNG701 cDNA was inserted in-frame into the pMAL-c2X vector (New England Biolabs) to obtain a malE-DNG701 fusion construct. The construct was introduced into Escherichia coli strain BL21 (DE3). Expression and purification of the maltose-binding protein (MBP)-DNG701 fusion protein were performed as previously described (6). The MBP was purified in parallel by the same procedures and used as the control. The plasmid pUBIGUS was purified from the strain BL21 (DE3, dcm−) using the Maxi-plasmid purification kit (Qiagen). Fifty micrograms of plasmids were methylated in vitro by 20 units of CG methyltransferase (M.SssI) or MspI methyltransferase according to the provided protocols (New England Biolabs). Unmethylated control plasmids were treated simultaneously by the same procedure without addition of the methyltransferases. For the nicking activity assays, 200 ng of each of these plasmids was added to nicking reaction buffer (10 mM Tris-HCl, pH 7.9, 50 mM NaCl, 10 mM MgCl2, 1 mM DTT, 0.2 mg/ml BSA) plus purified MBP-DNG701 or MBP. Reactions were carried out as described (6). The resulting mixtures were checked in a 1% agarose gel.

For incision activity assays, upper strands of the seven oligonucleotides (Fig. S3C) were labeled with 32P at the 5′ ends by using T4 polynucleotide kinase (Roche), and they were then annealed to the unlabeled complementary strands as previously described (6). The labeled double-stranded oligonucleotides (1 pmol) were incubated with MBP-DNG701 or MBP or MspI endonuclease in the nicking reaction buffer in a total volume of 10 μl at 37°C for 2 h. The products were separated on 17% denaturing polyacrylamide gels containing 8 M urea, and the gels were exposed to a PhosphorImager screen.

Genomic Bisulfite Sequencing. For genomic bisulfite sequencing, 2 μg of genomic DNA was digested with 20 units of appropriate restriction enzymes (New England Biolabs) for 3 h. The digested DNA was purified by ethanol, treated by sodium bisulfite solution, and desalted with the Wizard DNA Clean-Up System (Promega), as previously described (7). Subsequently, 4 μl of each recovered DNA sample was subjected to PCR amplification as follows: 95°C for 5 min; followed by 15 cycles at 95°C for 20 s, 60°C for 3 min, and 72°C for 3 min; and finally 30 cycles at 95°C for 20 s, 50°C for 1.5 min, and 72°C for 2 min. Five sets of primer pairs (P14-F/P14-R for Tosi17; P15-F/P15-R for Tosi17 chr10 5′ LTR; P16-F/P15-R for Tosi17 chr10 3′ LTR; P17-F/P17-R for Tosi17 chr10′ promoter; and P18-F/P18-R for LOC_Os07g31690) were used for the PCR reactions. The PCR products were cloned into the pGEM-T Easy cloning vector (Promega), and 15 to 20 clones were sequenced to calculate the percentage of cytosine methylation in each sample. All primers are listed in Table S1.

Sequence Alignments and Phylogenetic Analysis. Sequences of DNA glycosylase domains were aligned using ClustalX 2.0.10 (8), and the secondary structures of these domains were predicted by PSIPRED v3.0 (http://bioinf.cs.ucl.ac.uk/psipred/). The conserved residues were highlighted using GeneDoc 2.6.02 (http://
To construct a phylogenetic tree, the aligned sequences were subjected to MEGA v.4 (9) with the neighbor-joining algorithm. Neighbor-joining analysis was done with the pairwise deletion option and with the Poisson correction set for distance model. Bootstrapping was performed with 1,000 replicates.

Fig. S2. Gene structure of DNG701 and its expression pattern. (A) Structure of DNG701. Exons and introns are denoted as black blocks and lines, respectively. The open triangle marks the insertion site of Tos17, and the arrow indicates insertion direction of Tos17. GenBank accession numbers of flanking sequences of Tos17 are indicated above the arrows. The gray blocks show positions of DNA glycosylase domains. Both dng701-1 and dng701-2 are in Nipponbare background. (B) Expression patterns of DNG701 in three different tissues of 3-wk-old Nipponbare seedlings (Upper) and in four different tissues of 3-mo-old flowering plants (Lower) analyzed by RT-PCR. Ubiquitin was used as an internal control. (C) Subcellular localization of DNG701 in rice protoplasts. (a–c) 35S:GFP controls; (d–f) 35S:GFP:DNG701; (a and d) bright fields; (b and e) green fluorescence; (c and f) merged images of bright field and green fluorescence. (Scale bars, 20 μm.)
Fig. S3. Purification of recombinant MBP-DNG701 fusion protein and nicking activity assays. (A) Purification of recombinant MBP-DNG701 fusion protein and MBP protein. The open triangle indicates purified MBP protein and the filled triangle denotes purified recombinant MBP-DNG701 protein. Proteins were run in a SDS/PAGE gel and then stained with Coomassie blue. (B) Nicking activity assays of recombinant MBP-DNG701 (Upper) and MBP (Lower) on methylated and unmethylated plasmids. From the second to fourth lanes in each panel, 0.2, 0.8, and 2.4 μg of proteins were added, respectively. CK, no protein added. “CCGG, ’CG, and Unmethylated denotes “CCGG-containing, ’CG-containing, and unmethylated pUBIGUS plasmids, respectively. (C) Sequences of oligonucleotides used for enzymatic activity assays of DNG701 in vitro. The filled stars indicate 32P-labeled 5’ ends on upper strands. Methylated cytosines are underlined. CCGG sites are highlighted with blue; 29 nt and 16 nt denote the number of nucleotides from methylated cytosines to the labeled 5’ ends.

Fig. S4. RT-PCR analysis of dng701-1 and dng701-2 knockout mutants (Left) and DNG701 RNAi lines (Right). WT CTR-1 and WT CTR-2 are sibling wild-type plants segregated from heterozygous DNG701/dng701-1 and DNG701/dng701-2 plants, respectively. WT CTR-3 is a segregated wild-type plant from a hemizygous DNG701 RNAi transgenic plant PC5-9.
Fig. S5. Methylation status of centromeric repeat DNA, 45S rDNA, Ty1 retrotransposons, Tos17chr10 promoter, and another retrotransposon. (A) Methylation status of centromeric repeat DNA, 45S rDNA and Ty1 retrotransposons in the wild-type controls, mutants, and RNAi lines. (Left) Genomic DNAs were digested by HpaII or MspI and then separately blotted to different membranes for hybridization with centromeric repeat DNA probe. (B and C) Methylation status of Tos17chr10 promoter (B) and another retrotransposon (C). The region a (Fig. S6A) at Tos17chr10 promoter and 3′ terminal region of another retrotransposon (LOC_Os07g31690) were determined by bisulfite sequencing.
Fig. S6. Structural features of Tos17. (A) Tos17chr10 and Tos17chr7. For Tos17chr10, Tos17 is inserted in a putative ABC-type transporter gene (blue line) to form a fusion gene that is driven by the ABC-type transporter gene’s promoter (blue block). A 202-bp region from −445 to −244 (region a, shown with lime color) was found being methylated as revealed by bisulfite sequencing. A 3.8-kb Probe 3 was used for DNA gel-blot analysis to check methylation state of the promoter. Black blocks in Tos17chr10 and Tos17chr7 denote LTR regions. Regions b and c indicate the 5′ LTR regions examined by bisulfite sequencing. (B) Restriction map of Tos17chr10. Restriction sites of BstXI, XbaI, and HpaII/MspI (H/M) are shown above the gene. The numbers in parentheses represent positions of the restriction sites on the gene. The green block indicates the 90-bp deletion at Tos17chr7. The red and pink lines beneath the gene represent a 3.3-kb Probe 1 and a 1.1-kb Probe 2 used for hybridization analysis, respectively. The blue line shows the 405-bp region (region d) examined by bisulfite sequencing.

Fig. S7. Seed phenotypes of knockout of DNG701 in rice. Both the dng701-1 and dng701-2 mutants produced two types of seeds: normal and wrinkled seeds. Mutants were grown in a growth chamber for 4 mo to allow panicles to reach full maturity. NPB WT, Nipponbare wild-type.

Fig. S8. Methylation levels of Tos17 in the DNG701-overexpressing calli determined by bisulfite sequencing. Tos17 region d (Fig. S6B) was subjected to the bisulfite sequencing analysis.
Fig. S9. Methylation status of centromeric repeat DNA, 45S rDNA and Ty1 retrotransposons in the wild-type controls and DNG701-overexpressing plants. For 45S rDNA, not only markedly decreased CG methylation but also slightly reduced CHG methylation were observed in the DNG701-overexpressing plants. WT CTR-4 and WT CTR-5, segregated wild-type plants from hemizygous transgenic plant OX-1-27 and OX-2-35, respectively. Genomic DNAs were digested by HpaII or MspI and then separately blotted to different membranes for hybridization with the probes.

Table S1. Primer list

<table>
<thead>
<tr>
<th>Primer name</th>
<th>5'-sequence-3'</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1-F</td>
<td>AAATCTACATGCGACACAGTTCCA</td>
<td></td>
</tr>
<tr>
<td>P1-R</td>
<td>actagtTCAGTGTTTGGCTAGTGGGC</td>
<td>Lower-case letters denote SpeI site</td>
</tr>
<tr>
<td>P2-F</td>
<td>actagtATGGATCCTCCAGGGCTCAAACCTG</td>
<td>Lower-case letters denote SpeI site</td>
</tr>
<tr>
<td>P2-R</td>
<td>GCTCCCACTGCTGAAACTCTGG</td>
<td></td>
</tr>
<tr>
<td>P3-F</td>
<td>GCCCGAGTATCTCATCTCAGTTC</td>
<td></td>
</tr>
<tr>
<td>P3-R</td>
<td>AAACCTCGCGCAACAGAAG</td>
<td></td>
</tr>
<tr>
<td>P4-F</td>
<td>CTGATTCCTCAAAATGATAATAGAT</td>
<td></td>
</tr>
<tr>
<td>P4-R</td>
<td>ATCTCCTTTATCAACTCAGTT</td>
<td></td>
</tr>
<tr>
<td>P5-F</td>
<td>GCTCTCTTGGACTAAATGAAACTGTG</td>
<td></td>
</tr>
<tr>
<td>P5-R</td>
<td>GATCGAAGACGAAAGGATGG</td>
<td></td>
</tr>
<tr>
<td>P6-F</td>
<td>ATATGGGACTAGAGCATTCGAAA</td>
<td></td>
</tr>
<tr>
<td>P6-R</td>
<td>GGAGGACTACAGATTCGAAA</td>
<td></td>
</tr>
<tr>
<td>P7-F</td>
<td>AAAGAAGCTGGAGCAGCATCAGC</td>
<td></td>
</tr>
<tr>
<td>P7-R</td>
<td>CCAGGCAAGAGTATGCGCC</td>
<td></td>
</tr>
<tr>
<td>P8-F</td>
<td>TGGGACAGGATTTCTATCACA</td>
<td></td>
</tr>
<tr>
<td>P8-R</td>
<td>TGGGACAGGATTTCTATCACA</td>
<td></td>
</tr>
<tr>
<td>P9-F</td>
<td>ATGATGAAAGACGTAACGGAATG</td>
<td></td>
</tr>
<tr>
<td>P9-R</td>
<td>GTCGGTGGCTATATGATGC</td>
<td></td>
</tr>
<tr>
<td>P10-F</td>
<td>AGAATCTGGGACCACTAGAAACTC</td>
<td></td>
</tr>
<tr>
<td>P10-R</td>
<td>GCTAGGGTCTATAAGGGTATGGGAGAG</td>
<td></td>
</tr>
<tr>
<td>P11-F</td>
<td>GACCGGCGTGACCAGATGTCGAACAT</td>
<td></td>
</tr>
<tr>
<td>P11-R</td>
<td>GTCTGGTCTGTATTAAATAGAT</td>
<td></td>
</tr>
<tr>
<td>P12-F</td>
<td>GTCTGGGTAATTATGGAACGTG</td>
<td></td>
</tr>
<tr>
<td>P12-R</td>
<td>CTGCTGGATATGCTAGAGGA</td>
<td></td>
</tr>
<tr>
<td>P13-F</td>
<td>GAGATGACAGCCCTACTTGGATT</td>
<td></td>
</tr>
<tr>
<td>P13-R</td>
<td>ACAAAATTTGACAGCCAGTTTATT</td>
<td></td>
</tr>
<tr>
<td>P14-F</td>
<td>GTTGAATTAAGGGTGTATATTGATTT</td>
<td>Lower-case letters represent C-to-T substitution</td>
</tr>
<tr>
<td>P14-R</td>
<td>GaTAAaACACAAAAaAaATACATATGTATTAAATATTT</td>
<td>Lower-case letters represent G-to-A substitution</td>
</tr>
<tr>
<td>P15-F</td>
<td>tGTGTATTATTATTATTAT</td>
<td>Lower-case letters represent C-to-T substitution</td>
</tr>
<tr>
<td>P15-R</td>
<td>CAaaTaaACaaTaaAATAAATAATAAT</td>
<td>Lower-case letters represent G-to-A substitution</td>
</tr>
<tr>
<td>P16-F</td>
<td>TGGAATTATGAAAGATTATTAT</td>
<td>Lower-case letters represent C-to-T substitution</td>
</tr>
<tr>
<td>P17-F</td>
<td>tTGGAATTTATGAAAGATTATTAT</td>
<td>Lower-case letters represent C-to-T substitution</td>
</tr>
<tr>
<td>P17-R</td>
<td>CaACaAAaCAaCAaAATTAaAAT</td>
<td>Lower-case letters represent G-to-A substitution</td>
</tr>
<tr>
<td>P18-F</td>
<td>tAAATATTATTATTAT</td>
<td>Lower-case letters represent C-to-T substitution</td>
</tr>
<tr>
<td>P18-R</td>
<td>TAAAATCAAAACACCACACAT</td>
<td>Lower-case letters represent G-to-A substitution</td>
</tr>
</tbody>
</table>