DNA demethylases remodel DNA methylation in rice gametes and zygote and are required for reproduction

Shaoli Zhou, Xue Li, Qian Liu, Yu Zhao, Wei Jiang, Anqi Wu, Dao-Xiu Zhou

DNA demethylases remodel DNA methylation in rice gametes and zygote and are required for reproduction

Shaoli Zhou[1*], Xue Li[1*], Qian Liu[1], Yu Zhao[1], Wei Jiang[1], Anqi Wu[1], & Dao-Xiu Zhou[1,2§]

[1]National Key Laboratory of Crop Genetic Improvement, Huazhong Agricultural University, 430070 Wuhan China

[2]Institute of Plant Science Paris-Saclay (IPS2), CNRS, INRAE, University Paris-Saclay, 91405 Orsay, France

*co-first authors

§corresponding author (dao-xiu.zhou@universite-paris-saclay.fr)

Short Summary: DNA methylation is a major epigenetic modification in plants. It is unclear whether and how DNA methylation is reprogramed in gametes and after fertilization in zygote. In this work we show that DNA demethylases DNG701, DNG702 and DNG704 function in rice egg, sperm and zygote cells to remodel DNA methylation at distinct genomic loci which is essential for egg and zygote gene expression and post-zygotic development.

Abstract

Fertilization constitutes a critical step in the plant life cycle during which the gamete genomes undergo chromatin dynamics in preparation for embryogenesis. In mammals, parental chromatin is extensively reprogrammed through the global erasure of DNA methylation. However, in flowering plants it remains unclear whether and how DNA methylation is remodeled in gametes and after fertilization in zygote. Here, we characterized DNA methylation patterns and investigated the function of DNA glycosylases in rice eggs, sperms, unicellular zygotes and during embryogenesis. We found that DNA methylation is locally reconfigured after fertilization and is intensified during embryogenesis. Genetic, epigenomic and transcriptomic analysis revealed that rice DNA glycosylases DNG702, DNG701 and DNG704 demethylate DNA at distinct genomic regions in the gametes and the zygote and are required for zygotic gene expression and development. Collectively, the results indicate that active DNA demethylation takes place in the gametes and the zygote to locally remodel DNA methylation which is critical for egg and zygote gene expression and reproduction in rice.

Key words: DNA methylation, DNA glycosylase, fertilization

Introduction

In plants, DNA cytosine methylation occurs in the context of CG, CHG and CHH (where H is A, C, or T) and heritably silences transposable elements (TE) and repetitive sequences to maintain genome integrity and transcriptional homeostasis (He et al., 2011; Law and Jacobsen, 2010). Similarly, DNA methylation of gene regulatory sequences, especially those near transcriptional start sites, is associated with gene silencing. However, accumulating data indicate that methylations in the promoter regions may also play a role in activating gene transcription by affecting the binding affinity of different regulators (Harris et al., 2018; Zhang et al., 2018). In plants, DNA methylation patterns in general are stably transmitted across generations (Gehring, 2019; Kawashima and Berger, 2014). The persistence of DNA methylation states of silenced TEs

and genes through generations (Cubas et al., 1999; Hauser et al., 2011) and evidence of methylation heritability in plants during many generations (Becker et al., 2011; Schmitz et al., 2011) indicate that most DNA methylation is inherited through sexual reproduction in plants. However, DNA methylation reprogramming occurs during plant sexual reproduction and development (Kawashima and Berger, 2014). For instance, recent results showed that *de novo* DNA methylation creates a cell lineage-specific epigenetic signature that is required for gene expression and splicing for meiosis (Walker et al., 2018). Studies from Arabidopsis and soybean showed that CHH methylation increases during embryogenesis (Bouyer et al., 2017; Kawakatsu et al., 2017; Lin et al., 2017; Narsai et al., 2017). In addition, cell-specific methylation patterns have been reported in Arabidopsis and rice (Calarco et al., 2012; Gehring, 2019; Ibarra et al., 2012; Kawakatsu et al., 2016; Kim et al., 2019; Li et al., 2020a).

Plant haploid gametes, sperm and egg, are generated by mitosis in multicellular haploid male and female gametophytes, respectively. Vegetative cell in male gametophyte and central cell in female gametophyte are respective companion cell of the sperm and the egg and are necessary for fertilization and seed development. The egg is fertilized by one sperm to form the zygote to initiate embryogenesis, and the binuclear central cell is fertilized by the other sperm to generate the triploid endosperm, a nutrient tissue that supports growth of the embryo or the seedling. Genome-wide DNA methylation profiling of sperm and vegetative cells of the mature pollen has revealed dynamic DNA methylation changes during male gametogenesis in Arabidopsis, rice and tomato (Calarco et al., 2012; Hsieh et al., 2016; Ibarra et al., 2012; Kim et al., 2019; Li et al., 2020a; Lu et al., 2021; Walker et al., 2018). DNA methylation analysis has been also performed on egg and central cells isolated from Arabidopsis and rice (Li et al., 2020a; Park et al., 2016). DNA demethylation is initiated in the Arabidopsis central cell (Park et al., 2016). However, it is unclear whether active demethylation takes place in gametes (egg and sperm) and zygotes.

Plant CG methylation is maintained by Methyltransferase1 (MET1) activity (Kim and Zilberman, 2014). In Arabidopsis, non-CG (i.e. CHG and CHH) methylation in short TE and long TE borders that are of euchromatin feature is maintained by the RdDM pathway involving the Domains Rearranged Methyltransferase2 (DRM2), while CHG and CHH methylation in heterochromatin is maintained respectively by Chromomethylases (CMT) CMT3 and CMT2 (Kim and Zilberman,

2014; Zemach et al., 2013). Rice non-CG methylation levels are higher than that in Arabidopsis (Niederhuth et al., 2016; Tan et al., 2016). Rice CHH methylation is low in heterochromatic long TEs but highly enriched in short TEs that are mainly located in the vicinity of genes (Niederhuth et al., 2016; Tan et al., 2016; Zemach et al., 2010b; Ma et al 2021). Rice CHH methylation is essentially dependent on *DRM2*, the mutation of which causes about 85% loss of CHH methylation in the genome (Tan et al., 2016). In Arabidopsis, DNA methylation is actively removed by a family of DNA glycosylases, including DEMETER (DME), REPRESSOR OF SILENCING1 (ROS1), and DME-LIKE (DML)2 and 3 (Zhu, 2009), which excise 5-methylcytosine that is replaced by cytosine via the base excision repair pathway. DME is expressed in the vegetative and central cells and demethylates DNA in gamete companion cells (central cell and vegetative cell), primarily at euchromatin TEs and the edges of large TEs (Choi et al., 2002; Ibarra et al., 2012; Park et al., 2017; Park et al., 2016; Schoft et al., 2011). By contrast, *ROS1* and *DML* genes are expressed primarily in vegetative tissues (e.g., roots and shoots) and at a lower level compared with *DME* in reproductive cells (Calarco et al., 2012; Penterman et al., 2007). In rice 4 *ROS1*-related DNA glycosylase genes are identified namely *DNG701*, *DNG702* (also known as *OsROS1a*), *DNG703* and *DNG704* (La et al., 2011; Ono et al., 2012; Zemach et al., 2010a). *DNG701*/*DNG704* and *DNG702*/*DNG703* are closely related to Arabidopsis *DML2* and *ROS1,* respectively (Zemach et al., 2010a), (La et al., 2011). A *DME* ortholog is not detected in rice (Zemach et al., 2010a). Previous results showed that *DNG702* (or *OsROS1a*) is indispensable in both male and female gametophytes (Ono et al., 2012). A recent study by analyzing a *dng702* heterozygous mutant (*ros1a*$^{+/-}$) showed that *DNG702* is involved in local DNA demethylation in pollen vegetative cell (Kim et al., 2019). However, it is unclear whether the plant DNA glycosylases function in the gametes and the zygote to reprogram DNA methylation and gene expression.

To investigate this question, we produced loss-of-function mutants of rice DNA glycosylase genes *DNG702*, *DNG701/DNG704* (*DNG701*/*4*) and analyzed their function in the gametes and the zygote regarding to DNA methylation, gene expression, and development. We show that DNA methylation is remodeled in zygote 6.5 h after fertilization in rice. The remodeling of DNA methylation is intensified at genomic regions enriched for 24-nt siRNA during embryogenesis.

The mutations of *dng702* and *dng701/4* altered DNA methylation locally in gamete and zygote genomes and affected zygote and/or egg transcriptomes and zygote development. We show that *DNG702* and *DNG701/4* have shared but mostly distinct genomic targets in egg and zygotes and likely play different roles in sperm DNA methylation. Collectively, the results indicate that DNA methylation is remodeled mainly at genic regions before and after fertilization in which the DNA demethylase genes *DNG702*, *DNG701* and *DNG704* play an essential role.

Results

DNA methylation dynamics during fertilization

To study DNA methylation dynamics during fertilization, we manually isolated rice (DongJin or DJ variety) sperms, eggs, and zygotes (at 6.5 hours after pollination, HAP) by dissection as previously described (Zhou et al., 2019), and obtained whole-genome DNA methylation data from 50 cells per cell type (about 10 to 40 genome coverages) with two biological replicates (Supplemental Table 1), using a bisulfite sequencing (BS-seq) protocol developed for small numbers of cells (Clark et al., 2017). Density plots (displaying density of fractional difference between comparisons, see Methods) revealed little variation between the BS-seq data and the methylomes reported in the gametes of the Nippon bare (NB)(Kim et al., 2019; Park et al., 2016), but much higher difference when compared with the seedling methylome of the same variety (Supplemental Figure 1A). Boxplots revealed overall lower TE DNA methylation levels in DJ than in NB and Kitaake (KT) accessions for which egg and sperm methylome data were available (Supplemental Figure 1B), which might be due to genomic difference of the rice varieties and/or differences in cell sampling. Multidimensional scaling analysis of the methylomes obtained in this study revealed a high reproducibility of the replicates and a cell-type specific distribution pattern, with the sperm methylome more distally apart from those of egg and zygote which were relatively close in position (Supplemental Figure 1C). Heatmap of methylation levels indicated that the differential methylations detected in zygotes versus eggs or sperms were not derived from averages of egg and sperm methylation levels (Supplemental Figure 1D). Boxplot analysis of the DJ methylomes indicated that the overall methylation levels at all sequence

contexts were comparable between eggs and zygotes (Figure 1A). In sperm cells, relatively lower mCG but higher mCHG and mCHH sites were detected in genes compared with zygote or egg (Figure 1A and Supplemental Figure 2A). The decrease of sperm mCG mainly affected moderately methylated (between 0.2 and 0.6) genes and the increases of mCHG and mCHH concerned essentially lowly methylated (0.02-0.04) genes (Supplemental Figure 2A). In addition, lower mCG and mCHG was detected in sperm TEs (Figure 1A), which mainly affected the highly methylated (>0.8) TE fraction (Supplemental Figure 2A). The data suggest that sperm displays substantially divergent overall DNA methylation (mostly in genes) compared with egg or zygote. Density plots revealed fractional difference (both gain and loss) of mCHH and mCHG and, to a lesser extent, mCG between zygote and egg (Z-E) or between zygote and sperm (Z-S) (Figure 1B), indicating that differential methylation occurs right after fertilization. Scanning of differentially methylated regions (DMR) within 50 bp windows with at least 20 informative cytosines (See Methods, Supplemental Table 1), detected more hyper than hypo CG (4,847 hyper and 1,202 hypo) and CHG (2,471 hyper and 1,372 hypo) DMRs in zygote versus egg (Z-E) (Figure 1C and 1D). Similarly, we detected more hyper than hypo CG (3,954 hyper and 1,096 hypo) and CHG (2,643 hyper and 1,102 hypo) DMRs in zygote versus sperm (Z-S). At CHH context, there were much higher but comparable numbers of hyper (14,018 in zygote versus sperm (Z-S) and 17,572 in zygote versus egg (Z-E) and hypo (11,742 in zygote versus sperm (Z-S) and 14,920 in zygote versus egg (Z-E) CHH DMRs (Figure 1C and 1D), indicating a more important remodeling of CHH methylation 6.5 h after fertilization than CG and CHG. The zygote versus sperm and egg CHH DMRs corresponded to about 1.2 to 1.5 Mbp (DMR number X 50 pb), which represented about 2 to 2.7 % of -the CHH methylated regions (in total 1,099,183 methylated bins with methylation ratio higher than 0.1, which corresponded to 55 Mbp), suggesting a locally remodeling of DNA methylation in the zygote.

To get insight into DNA methylation dynamics during early embryogenesis, we isolated DJ embryos at the globular stage (3 days after pollination, DAP) and performed BS-seq analysis (Supplemental Table 1). The globular embryo showed an increase of mCG but clear decreases of mCHG (gene) and mCHH (gene and TE) compared with the zygote (Figure 1A). Density plots revealed fractional difference with both loss and gain of mCG and mCHG, but a dramatic loss of

mCHH in the globular embryo relative to zygote (GE-Z) (Figure 1B). The gain of mCG mainly concerned TE, as the percentage of highly methylated (>0.8) TE augmented from about 40% in zygote to about 85% in globular embryo (Supplemental Figure 2A). Conversely, the percentages of highly CHG (>0.8) and CHH (>0.2) -methylated TE were reduced (resulting in an increase of CHG at 0.6-0.8 quantile) in globular embryo versus zygote (Supplemental Figure 2A). In mature embryo versus globular embryo (ME-GE), mCG (in TE) and mCHG (in both gene and TE) were decreased but mCHH (in TE) was elevated (Figure 1A). The decrease of mCG and mCHG mainly concerned highly methylated TE fractions (>0.8 for mCG and >0.6 for mCHG) (Supplemental Figure 2A). Density plots showed an overall loss of mCHG but gain of mCHH in ME-GE (Figure 1B).

DMR scanning detected much higher numbers of CG (22,880 hyper and 3,754 hypo) and CHG (11,693 hyper and 3,871 hypo) DMRs in globular embryo versus zygote (GE-Z) compared with zygote versus egg (Z-E) or sperm (Z-S) (Figure 1C). At CHH context, a much higher number of hypo (110,151) than hyper (1,363) DMRs was detected in globular embryo versus zygote (Figure 1C and Supplemental Figure 2B), indicating an important loss of mCHH in globular embryo. In mature versus globular embryo (ME-GE), the DMR numbers at all sequence contexts were largely increased compared with GE-Z (Figure 1C), indicating a more important remodeling of the DNA methylation during embryo maturation. The low CHH methylation in globular embryo was consistent with the results showing that RdDM activity is low in proliferating tissues with high numbers of dividing cells (Borges et al., 2021). In ME-GE, there were 48,670 hyper and 48,656 hypo DMRs at CG, much higher numbers of hypo (84,493) than hyper (34,890) DMRs at CHG context. By contrast, at CHH context much more hyper (159,376) than hypo (18,053) DMRs were detected in ME-GE (Figure 1C), indicating a recovery of mCHH in the mature embryo. Both long and short TEs showed reduced CHH methylation in globular embryo relative to that in zygote (Supplemental Figure 3A). However, CHH methylation levels in long TEs such as Gypsy and centromere repeats were much lower (about 4-8 times) than that in short TEs (MITEs and SINEs) (Supplemental Figure 3A). In fact, the globular embryo versus zygote (GE-Z) and mature versus globular embryo (ME-GE) CHH DMRs mainly concerned short TEs (<500 bp) (Supplemental Figure 3B) which are located near genes in the rice genome (Tan et al., 2018; Tan et al., 2016). The CHH DMRs were enriched for 24-nt small RNA (Supplemental Figure 3C and

3D), suggesting that RdDM is involved in methylome dynamics during embryogenesis. The GE-Z hypo CHH DMRs showed no significant overlap with ME-GE hyper CHH DMR (Supplemental Figure 3E), suggesting that CHH methylation loss at globular stage was not re-established during embryo development. Collectively, the data indicated that DNA methylation reconfiguration occurs 6.5 h after fertilization and is intensified during post-zygotic embryo development.

Function of rice DNA demethylation genes in reproduction

Based on the present (see below) and published transcriptomic data (Anderson et al., 2013), transcripts of rice demethylase genes *DNG701*, *DNG702*, and *DNG704*, but not *DNG703*, were detected in egg, zygote or sperm cells (Supplemental Figure 4A). Relatively lower *DNG702* expression level was detected in the vegetative cell. The gene may be expressed at higher levels at earlier pollen developmental stages as its mutation affected DNA methylation in the vegetative cell of the pollen (Kim et al 2019). To study whether the genes have a function in gametes and zygote, we produced *dng701*, *dng702* and *dng704* knockout (KO) plants in the DJ background by using the CRISPR technique (Supplemental Figure 4B). The *dng702* homozygous KO plants failed initiating embryogenesis and seed development after pollination (Figure 2A-C and Supplemental Figure 4C) confirming previous results (Liu et al., 2018; Ono et al., 2012). However, crosses of *dng702* with wild type as either male or female produced seeds with defective endosperm (Supplemental Figure 4D), as previously reported (Ono et al., 2012). The seeds could be germinated in cultural media. This would suggest that defects in *dng702* egg or sperm affected zygote viability or development, but could be compensated by wild type sperm or egg after fertilization. The differences in the gametophytic lethality of *dng702* mutations in DJ and NB (Ono et al 2012) backgrounds may be related to the difference in DNA methylation levels between the two varieties (Supplemental Figure 1B). No visible phenotype was found in *dng701* and *dng704* single mutant, which may be due to functional redundancy of the genes that are closely related (Zemach et al., 2010a). Therefore, we made *dng701* and *dng704* double (thereafter named as *dng701/4*) KO plants (Supplemental Figure 4B). The *dng701/4* double KO plants displayed a similar but less severe phenotype than *dng702* and were partially fertile (with a seed setting rate at about 50% of wild type) (Figure 2A, 2B, and 2D). Close examinations of *dng701/4* ovules before and after pollination revealed that about 50% of the fertilized eggs

were aborted and that the other 50% showed normal or retarded embryo development (Figure 2A, 2B, 2D and Supplemental Figure 4E). The seed phenotype was reproduced at T3 generation with about 48% aborted, 13% retarded, and 39% normal seeds (Supplemental Figure 4F). The results indicate that the rice DNA demethylase genes are required for gamete function and zygote development in which *DNG702* likely plays a more important role than *DNG701* and *DNG704*.

Function of DNG702 and DNG701/4 genes in DNA methylation in gametes and zygotes

To study effect of the *dng* mutations in DNA methylation of gametes and zygote, we manually isolated egg and sperm cells from *dng702* and *dng701/4* and zygotes (at 6.5 HAP) from *dng701/4* plants and performed BS-seq analysis. Because the three seed phenotypes of *dng701/4* could not be distinguished at zygote stage, the harvested zygotes represented a mixture of the three types. As tissue culture during rice transgenic plant production may affect DNA methylation, to deduce tissue culture effects we isolated eggs and zygotes from a CRISPR-negative line (callus regenerated wild type, crWT) as controls for BS-seq. Two replicates per genotype were analyzed (Supplemental Table 1 and Figure 3A). Box plots of the BS-seq data revealed that in *dng702* and *dng701/4* mutant eggs there was a general increase of methylation relative to WT or crWT at all sequence contexts (Figure 3A). The increases of mCHG and mCHH were more obvious in genes than in TEs (Figure 3A). Metagene analysis confirmed the clear increases of mCHH and mCHG in gene body and flanking regions in the mutant eggs compared with wild type (Supplemental Figure 5).

Density plots revealed differential methylations particularly with a clear hyper CHH methylation peak in *dng702* egg and *dng701/4* egg and zygote compared with the wild type cells (Supplemental Figure 6A). Similar differential methylation patterns were observed in the mutants compared with crWT egg and zygote (Supplemental Figure 6B). To deduce tissue culture effect from that of the mutations on DNA methylation, we called DMRs between crWT and wild type and subtracted tissue culture-related DMRs from those detected in the mutants relative to wild type (Supplemental Figure 6C). In the mutant eggs we detected comparable numbers of hyper and hypo DMRs at CG (3,887-6,196) and CHG (2,037-3,092) contexts.

However, at CHH context, much higher numbers of hyper (32,092 in *dng702* and 28,219 in *dng701/4*) than hypo (7,635 in *dng702* and 10,927 in *dng701/4*) DMRs were detected (Figure 3B). We also called DMRs between mutants and crWT. We found that most (51 to 79%) of the DMRs detected in mutant versus wild type were found in those detected in mutant versus crWT (Supplemental Figure 6D). The hyper CHH DMRs of the two mutants showed no significant overlap (Figure 3C and Supplemental Figure 7A). The *dng702* and *dng701/4* hyper CHH DMR-associated genes were enriched for different functional categories (Supplemental Figure 7B). However, CG and CHG hyper DMRs showed partial overlaps between the two mutants (Figure 3C). The overlapped 2,220 CG and 768 CHG hyper DRMs displayed lower methylation in both wild type and crWT egg but higher methylation in both zygote (Supplemental Figure 7C), confirming that these DMRs were produced by the mutations. The data indicated that DNG702 and DNG701/4 may target a subset of common CG and CHG sites for demethylation.

The *dng701/4* mutation in zygote produced higher numbers of CG and CHG DMRs but comparable numbers of CHH DMRs than in egg (33,228 hyper and 10,118 hypo CHH DMRs in zygote compared with 28,219 hyper and 10,927 hypo CHH DMRs in egg) (Figure 3B and Supplemental Figure 7D). However, the hyper CHH DMRs detected in *dng701/4* zygote showed no significant overlap with those identified in the mutant egg and sperm cells (Figure 3D, Supplemental Figure 7E). The analysis suggested that DNG701/4 demethylases target distinct genomic regions in zygote relative to egg. Because the *dng701/4* mutation affected zygote development, an indirect effect of the mutation on DNA methylation in zygote was not excluded.

To study whether the mutant effects were sporophytic or gametophytic, we analyzed the methylome of eggs isolated from two independent heterozygous *dng702* (+/-) lines (Supplemental Figure 4B). Boxplot analysis indicated that mCHH and mCHG levels were also augmented in the *dng702* (+/-) eggs compared with wild type or crWT (Figure 3A). The increases were less obvious compared with those in the homozygous *dng702* mutants (Figure 3A). Density plots revealed clear methylation (mostly at CHG and CHH contexts) variations between the heterozygotes and wild type and between the heterozygous and the homozygous mutants (Supplemental Figure 8A and 8B). However, the hyper CHG and CHH methylations detected in *dng702* egg were reduced in the heterozygous mutants (Supplemental Figure 8C).

In sperm, the *dng702* mutation also led to a clear increase of mCHG and mCHH in genes (Figure 3A and Supplemental Figure 5A). Increase of overall mCHH was also observed in the Nippon bare *dng702* heterozygous mutant (*ROS1a*$^{+/-}$) (Supplemental Figure 9A) (Kim et al., 2019), except in the loci that show increase of mCG in the companion vegetative cell of the heterozygous mutant pollen (Kim et al., 2019). There was little overlap between the vegetative cell-sperm hypo CG DMRs and dng702-wild type sperm hyper CHH DMRs (Supplemental Figure 9A, bottom part). By contrast, the *dng701/4* mutations led to clear decreases of mCHH and mCHG in sperm (Figure 3A and Supplemental Figure 5B). In fact, the *dng702* mutation resulted in more hyper (10,464) than hypo (7,438) CHH DMRs, while *dng701/4* produced more hypo (15,330) than hyper (7,252) CHH DMRs in sperms (Figure 3B). This suggests that *DNG702* and *DNG701/4* have a distinct function in sperm DNA methylation. The mutations resulted in much fewer but distinct DMRs at all sequence contexts in sperm compared with egg (Figure 3B and Supplemental Figure 9B). Still, some tissue-culture effect is not excluded from the DNA methylation changes observed in the mutant sperms.

Function of DNG702 and DNG701/4 in DNA methylation dynamics after fertilization

Density plots indicated important variations between wild type and *dng702* or *dng701/4* mutant eggs (Figures 4A and Supplemental Figure 10A). To get insight into function of the DNA glycosylases in methylation dynamics during fertilization, we plotted *dng702* egg hyper methylation (CG >0.5, CHG>0.3, CHH >0.1, >0.3 and >0.6) fractions against differential methylation in wild type zygote versus egg (Figure 4A and 4B). We found that the hyper methylations in the mutant egg actually corresponded to hyper methylations in wild type zygote versus egg, with the *dng702* CHG and CHH (>0.6) hyper methylation peaking at about 0.5 and that of CG at about 0.75 of wild type zygotes versus eggs (Figure 4B and 4D), which corresponded to 160 to 2,600 DMRs (representing about 0.04 % of the genome) (Figure 4C). Similar observations were made with the differential methylations in *dng701/4* (Supplemental Figure 10B and Supplemental Figure 10C). The analysis indicates that methylation in DNG702 and DNG701/4-demethylated regions in the egg become reestablished in the zygote.

Similarly, genomic loci demethylated by DNG702 in sperm which were distinct from those in egg (Supplemental Figure 9B) were also hypermethylated in zygote (Supplemental Figure 11). Thus, loci demethylated by the DNA glycosylases in egg and sperm are remethylated in zygote, suggesting that *de novo* methylation pathways are already active in zygote. Conversely, we found that hyper methylations in *dng701/4* zygote corresponded to hypo methylations in wild type zygote versus egg or sperm (Figure 5, Supplemental Figure 12 and 13). The analysis indicated that DNG*701/4* locally remove methylation from egg and sperm genomes shortly after fertilization. Collectively, the results indicated that DNA methylation is remodeled before and after fertilization in which the DNA glycosylase genes DNG702, DNG701 and DNG704 play an essential role.

Function of rice DNA glycosylases in gene expression during egg to zygote transition

In parallel to the methylomes, we analyzed transcriptomes of wild type (DJ) egg and zygote by RNA-seq (Supplemental Table 3 and Supplemental Figure 14A). The egg and zygote transcriptomes were clearly different as revealed by a multidimensional scaling analysis (Supplemental Figure 14B), and showed high correlation respectively with published rice egg or zygote transcriptomes ($r>0.75$; Supplemental Figure 14C). A large number of differentially expressed genes (DEG) (2509 down- and 2073 upregulated, >2 folds, Q-value <0.01) were detected in zygote versus egg (Supplemental Figure 14D). In addition, most of the reported zygotic (at 9 HAP) up- or downregulated genes were found to be similarly regulated in the present transcriptome (Anderson et al., 2017) (Supplemental Figure 14E). RT-qPCR validated 6 zygote-expressed genes (Supplemental Figure 14F). A majority of the DEGs showed differential methylation in zygotes versus eggs (Supplemental Figure 14G), including those involved in cell cycle, DNA metabolism and transcription (Supplemental Figure 14H and Supplemental Table 4), which were previously identified as zygotic (versus egg) DEGs (Anderson et al., 2017).

To evaluate whether differential methylations in the *dng* mutants was involved in gene expression, we analyzed the transcriptomes of *dng702* and *dng701/4* eggs and *dng701/4* zygote by RNA-seq (Supplemental Table 3). Although different between wild type and mutants, the transcriptomes from the same cell types were close in position in the multidimensional scaling

analysis (Supplemental Figure 14B). In eggs, the *dng702* and *dng701/4* mutations resulted in respectively 3898 and 1824 downregulated and 2716 and 1854 upregulated genes (>2 fold, p<0.01) (Figure 6A).Numbers of up- and down-regulated DEGs were found to associate with hyper DMRs in the mutants (Figure 6B), although no general link between methylation and expression changes gene in the mutant was observed. Relatively few of the DEGs (231 down and 186 up) were commonly affected by the *dng702 and dng701/4* mutations (Figure 6B). Clustering of hyper DMR-associated DEGs in the mutants by their expression patterns in wild type egg and zygote revealed that 265 and 160 genes that were highly expressed in wild type egg but repressed in wild type zygote, were downregulated and hypermethylated respectively in *dng702* and *dng701/4* mutant eggs (Figure 6C and Supplemental Table 5) (Anderson et al., 2017). These genes included many transcription factors such as WRKY7 (Os05g46020), a bHLH (Os08g39630), and corepressor SEUSS (Os11g10060, Os11g10070) that are highly expressed in egg but repressed in zygote in the wild type background (Supplemental Table 5 and Figure 6D). These genes may be potentially important for egg cell function. In addition, 104 genes that were highly expressed in wild type zygote (relative to egg) were upregulated and hypomethylation in *dng701/4* eggs (Fig 6C). The mutations seemed not to clearly affect their expression in the zygote (Figure 6C). The data indicate that the DNA glycosylases play important roles in gene expression in egg and during egg to zygote transition.

In zygotes, the *dng701/4* mutation resulted in upregulation of 1061 and downregulation of 1152 genes (>2 fold, p<0.01) (Figure 6E). The *dng701/4* zygote and egg DEGs showed some overlap but more than 90% of the DEGs were different (Figure 6F), supporting the above proposed hypothesis that the demethylases regulate different sets of targets in egg and zygote. Density plot analysis of transcript levels of wild type zygote-egg DEGs (Supplemental Figure 14D) in the *dng701/4* background revealed that their differential expression levels were substantially reduced by the mutations (Figure 6G). About 50% of the wild type zygote-egg DEGs were no long differentially expressed in the mutant zygote, about 30% of which were found to associate with hyper DMRs (Figure 6H). Among those downregulated and hypermethylated genes, many are involved in cell cycle, DNA metabolism, translation, and transcription (Figure 6I and 6J). The analysis indicated that the *DNG701/4* genes play an important role in zygotic gene expression.

Discussion

Methylation reprograming in zygotes and during early embryogenesis

While DNA methylation patterns in general are transmitted stably across generations, there is also evidence for methylation variations at specific sequence contexts and in specific cell types during reproduction (Gehring, 2019). We showed that DNA methylation is locally reconfigured in rice gametes and zygote and is remodeled during post-zygotic development and embryogenesis (Figure 7). We found that both loss and gain of methylation, mainly at CHH contexts, occur in rice gametes, zygote and during embryo development. The DNA demethylases that we showed to target distinct genomic regions in gametes and zygote contribute to the differential methylation between the reproductive cells and are required for reproduction. In rice, mCHH is mainly found at small TEs that are located in genic regions and has been shown to be a strong indicator of RdDM (Stroud et al., 2014; Tan et al., 2018; Tan et al., 2016; Zemach et al., 2013). Gain of mCHH at specific genomic loci in zygotes implies that RdDM is functional after fertilization to undertake *de novo* methylation which appears to be important for zygote genome activation (Figure 7). This is in line with previous data showing defects in timing of zygotic genome activation in seeds produced from crosses with one parent defective in RdDM (Autran et al., 2011; Baroux et al., 2013; Zhao et al., 2019). RdDM was shown to repress paternal genome expression in Arabidopsis 2-4 cell embryo (Autran et al., 2011). Likely, both the DNA glycosylase-mediated active demethylation and RdDM shape the methylation pattern after fertilization which may be instructive for zygotic genome activation. Methylation at CHH context is not efficiently maintained during rapid cell division of the early embryo, consistent with recent results that RdDM activity is low in rapidly dividing cells (Borges et al., 2021). This is in line with previous observation of dramatic reprogramming of non-CG methylation during Arabidopsis embryo development (Bouyer et al., 2017; Lin et al., 2017).

DNA glycosylases function in reproductive cells

Active DNA demethylation by the Arabidopsis DNA glycosylase DME has been shown to regulate DNA methylation patterns in companion (vegetative and central) cells of egg and sperm (Gehring et al., 2006; Hsieh et al., 2009; Ibarra et al., 2012; Park et al., 2016; Schoft et al., 2011).

Arabidopsis DME is expressed and demethylates DNA in central cell (or endosperm) and pollen vegetative cell (Choi et al., 2002; Ibarra et al., 2012; Park et al., 2016; Schoft et al., 2011). DME-mediated DNA demethylation at central cell TEs regulates adjacent gene expression, which can result in gene imprinting in the endosperm. DNA glycosylase-regulated demethylation in central cell and vegetative cell seems to be conserved in rice (Kim et al., 2019; Park et al., 2016; Rodrigues et al., 2013). The present data provide evidence that the rice DNG702 and DNG701/4 genes function also in the gametes and the zygote to regulate DNA methylation, gene expression and zygote development. This is consistent with their expression in egg, sperm and zygote. *DNG702* had a higher expression in the zygote than in the egg cell, which may be due to the high expression of the paternal allele (Supplemental Figure 4A) (Anderson et al., 2013). The high zygotic expression of the gene and the zygotic abortion phenotype induced by its mutation suggest that this demethylase is a major player for demethylation in zygote and during early embryogenesis where important loss of CHH methylation was detected at globular stage (Figure 7). Because the production of the mutants included the tissue culture, tissue culture effect could not be completely excluded from the *dng* mutation-induced DNA methylation changes. Arabidopsis DME RNA is detected in pollen, but not in isolated sperm cells (Schoft et al., 2011). It remains unclear whether Arabidopsis DME is expressed or functional in egg cell. DME has no orthologue in monocots (Zemach et al., 2010b). The functional or expression differences in sperm or egg between rice and Arabidopsis demethylase genes may be caused by divergent evolution of the plant DNA glycosylase family. That the *dng* mutations affected methylation and expression patterns at distinct genomic regions in the different reproductive cells implies an essential role of the enzymes before and after fertilization, but also raises the question of how the glycosylases target distinct chromatin regions in the different cells (Frost et al., 2018). One possibility would be that their recruitment is conditioned by distinct overall chromatin structures in gametes and zygotes. It was shown that Arabidopsis DME can access to euchromatin but not to heterochromatin without FACT histone chaperon (Frost et al., 2018). In addition, it was shown that DME catalytic domain can be recruited to heterochromatins but could not process demethylation effectively (Zhang et al., 2019).

Multiple sex-specific changes in chromatin structure occur during development of reproductive cells (Li et al., 2020a; Walker et al., 2018; Wang and Kohler, 2017). For instance, histone H1 and the centromere-specific histone variant CENH3 are depleted from the egg cell (Ingouff et al., 2010; Ingouff et al., 2017; She et al., 2013), while H3K27me3 is depleted from the sperm cell chromatin that is highly compact, likely related to deposition of a set of male specific histone variants (Borg and Berger, 2015; Borg et al., 2020). In addition, analysis of single cell 3D chromatin structure has revealed a compact silencing center (CSC) structure in rice eggs which is dynamically remodeled after fertilization (Zhou et al., 2019). It was shown that DNG702 (or OsROS1a) demethylates locally at CG and CHG contexts in endosperm and vegetative cells (Kim et al., 2019; Liu et al., 2018). Thus, the rice DNA glycosylase genes may function in both gametes and companion cells to reprogram DNA methylation and gene expression. During human oocyte to zygote transition, pre-zygotic genome activation (ZGA) embryos acquire permissive chromatin and widespread H3K4me3 in CpG-rich regulatory regions. By contrast, the repressive mark H3K27me3 undergoes global depletion (Xia et al., 2019). Recent results indicate that H3K27me3 is depleted from sperm chromatin in Arabidopsis (Borg et al., 2020). Whether and how histone methylations are reprogrammed after fertilization in plants remain to be analyzed.

METHODS

Generation of *dng702*, *dng701/4* homozygous lines

The rice variety DJ (DongJin) (*Oryza sativa* spp. japonica) was used in this study. Single-guide RNAs (sgRNA) of CRISPR/Cas9 system were designed as reported (He et al., 2017). The sgRNA target sequences of *DNG702, DNG701,* and *DNG704* are shown in Supplemental Table 6. *Agrobacterium tumefaciens* (strain EHA105) mediated transformation of rice was performed as described previously (Lin and Zhang, 2005; Zhang et al., 1997). Briefly, after infection by agrobacteria containing Crispr/Cas9 vector calli were put on selection media. The resistant calli were transferred to regeneration media to introduce seedlings. The transformants were genotyped with PCR (using primers sets corresponding to the sgRNAs) followed by DNA sequencing. Mutations in *DNG702, DNG701,* and *DNG704* were decoded by using the DSDecodeM tool (http://skl.scau.edu.cn/dsdecode/) (Liu et al., 2015; Xie et al., 2019). Lines with

mutations in both alleles were obtained at T0 generation. A homozygous *dng702* mutant line (with the same mutations in both alleles) was identified. Because of zygote abortion in *dng702* homozygous mutants, the homozygous mutant lines were maintained by ratooning after flowering (i.e. removing the aerial parts for shoot regeneration) (https://en.wikipedia.org/wiki/Ratooning). *The dng701* and *dng704* single mutants displayed no visible phenotype. Double null mutants of *DNG701* and *DNG704* were created by vectors containing the one guide RNA shared by the two genes. Crispr lines without exogenous DNA fragment insertion (CRISPR-negative lines) were used as control during methylation analysis.

Egg, Sperm, and unicellular zygote isolation

Rice plants were grown in either greenhouse or paddy field. Rice gametes and unicellular zygotes were isolated from mature anthers and ovaries of T3 generation transgenic plants as reported (Zhou et al., 2019),(Zhang et al., 2010). Briefly, mature anthers were soaked in 45% (w/v) sucrose and then transferred into 15% (w/v) sucrose to release sperm pairs. For isolation of eggs and unicellular zygotes, ovaries of non-pollinated and pollinated florets (about 6.5 h after pollination) were manually dissected under a dissection microscope. The dissociated ovules were then transferred into fresh 0.53 M mannitol and broken to release eggs or zygotes. All isolated cells were stained by FDA (Fluoresceinc diacetate, Sangon, 596-09-8) and collected by a micromanipulator system (Eppendorf, TransferMan® 4r). Despite three types of seeds produced by *dng701/4* mutation, we could not distinguish them at unicellular zygote stage the zygotes of *dng701/4* harvested were a mixture of the three different phenotypes.

RNA-seq and BS-seq library construction

For RNA-seq library construction, mRNA isolated from egg and unicellular zygote were reverse transcribed and amplified by utilizing SMART-Seq® v4 Ultra® Low Input RNA Kit (TAKARA, Cat. No.634889), cDNAs were sheared into 200-400 bp DNA fragments followed by purification using Agencourt AMPure Beads (Beckman Coulter, USA A63881). The rest steps were performed using the Truseq ChIP DNA library preparation kit (Illumina. IP-202-1012). BS-seq libraries were constructed using reported protocol (Clark et al., 2017) with modified primer adapter 2 oligo and iPCRtag primers (Supplemental Table 6). Fifty cells were pooled for each replicate for

bisulfite seq library construction. The globular embryo was isolated by laser microdissection. Briefly, seeds at 3 days after pollination were fixed in ethanol/acetic acid fixing solution (v 3:1) in 37°C for three times, then dehydrated with gradient ethanol and infiltrated by isopropanol followed by paraffin replacement and embedded. Globular embryos were isolated by dissection of the embedded seeds with ZEISS PALM MicroBeam system. DNA was isolated with DNA Micro Kit (Qiagen, 56304).

Paraffin section

Florets were marked by the flowering time and collected at 0, 1, 3, 4, and 5 days after fertilization. 50% FAA (50 ml absolute ethanol, 10ml 37% formaldehyde solution, 5 ml glacial acetic acid, add double distilled water to 100 ml) was used to fix samples. Paraffin embedding was performed as described(Itoh et al., 2005). Longitudinal slices were stained by toluidine blue (2% w/v) for two minutes at room temperature.

RNA-seq assay

Sequencing reads were mapped to genome (RGAP, version 7.0) by Hisat (0.1.6-beta) software, cufflinks (v2.2.1) suite was used to splice transcripts and calculate gene expression fold changes. Genes with fold change higher than 2 and Q value less than 0.01 were considered as differentially expressed genes. Expression matrix of all expressed genes for all samples was generated. Multidimensional scaling analysis (MDS) was performed to calculate the Euclidean distances of expression profiles of different samples, scatter plots of distances were drawn in R environment. Reported RNA-seq data were downloaded from SRA database (Anderson et al., 2017). The transcription profiles of reported rice egg, sperm, and vegetative cells were download from Gene Expression Omnibus (GEO) (accession number GSE50777)(Anderson et al., 2013).

BS-seq assay

The mapping, processing, and duplication filtering were all performed by using the standard procedure BS-seeker2 (v.2.1.8)(Guo et al., 2013). Clean reads were mapped to the rice genome

(RGAP, version 7.0). Duplication was removed, and uniquely mapped reads were retained for further analysis. Mean nuclear methylation levels were calculated by averaging methylation of individual cytosine in each context as described (Kim et al., 2019).

BS-seq data of mature embryo and seedlings were downloaded from Sequence Read Archive (SRA) database, the accession numbers were SRR771502 for mature embryo and SRR3503136, SRR3503137 for seedling (de Vega-Bartol et al., 2013; Tan et al., 2018).

Differentially methylated regions (DMRs) were identified within 50 bp bins (7,614,209 bins in total of rice genome) with at least 20 informative sequenced cytosines (i.e. the sum of the sequence depth of each cytosine multiplied by the number of cytosines within 50-bp bins at CG, CHG or CHH context) in both samples as described[18,](Kim et al., 2019). Numbers of bins that pass the requirement with at least 20 sequenced informative cytosines were provided in Supplemental Table 1. Methylation level in each bin was determined by dividing the total number of methylated cytosines with the total sequenced cytosine numbers. Bins with methylation differences greater than 0.5, 0.3, and 0.1 respectively at CG, CHG, and CHH contexts with P values <0.05 (Fisher's exactly test) were considered as differentially methylated regions. Considering replicate variance, only DMRs showing significant methylation in both replicates of the sample compared to the replicates of the other sample are retained (i.e. the CG, CHG, and CHH methylation difference between any replicate comparison combinations must be higher than 0.5, 0.3 and 0.1). To deduce tissue culture effect on DNA methylation, we isolated egg and zygote derived from callus-regenerated wild type (crWT) for BS-seq to compare with wild type. *dng* mutant DMRs that overlap with crWT-WT DMR were subtracted for further analysis. The detail information of the mutant DMRs in the replicates are provided in Supplemental Table 2.

Density plots were performed to show the frequency distribution of methylation differences of 50 bp bins between two samples, only the bins containing at least 20 informative sequenced cytosines in both samples and 0.7 CG, 0.3 CHG, or 0.1 CHH methylation ratios in either sample were retained as described (Kim et al., 2019). To show methylation dynamics during fertilization

(zygote versus egg) of DNG-demethylated regions, we firstly identified *dng* hyper-DMRs in egg or zygote versus wild type then calculated methylation differences of the regions between zygote and egg in wild type. The distribution of methylation differences was shown by density plots.

For the analysis in Supplemental Figure 15, more vigorous criteria were applied for DMR calling. Besides the above described criteria, the 50 bp-bin DMRs situated within 300 bp genomic sequences were merged and retained for further analysis. The remaining isolated DMRs within 100 bp of the genomic sequence were discarded.

For small RNA assay, the reported somatic data (SRR6420549, SRR6420550) was used (Tan et al., 2018). The mapped small RNA reads were counted for each DMR. RPKM was calculated to represent small RNA abundance. As control, random regions in the genome were chosen to show the background genomic small RNA enrichment.

Data availability

Genes sequence data from this article can be found in the Rice Genome Annotation Project website under the following accession numbers: DNG701, LOC_Os05g37350, DNG702, LOC_Os01g11900, DNG703, LOC_Os02g29230, DNG704, LOC_Os05g37410. All high throughput data in support of the finding of this study deposited to the Gene Expression Omnibus (GEO) under the accession number GSE143923. The globular embryo BS-seq data was under the accession number GSM4057604 and GSM4057605.

Author contribution

SZ produced and analyzed the data; XL produced the mutants, isolated the cells and participated in data production, QL participated in data analysis, WJ and YZ participated in material production, DXZ coordinated the project, analyzed the data and wrote the paper.

Acknowledgements

We thank Cheng Luo for help to isolate single cells. This work was supported by grants from the National Key Research and Development Program of China [2016YFD0100802], the National

Natural Science Foundation of China [31730049]; Huazhong Agricultural University Scientific & Technological Self-innovation Foundation [Program N° 2016RC003], the Fundamental Research Funds for the Central Universities [2662015PY228] and National Postdoctoral Program for Innovative Talents.

Supplemental Figures and tables

Supplemental Figure 1. Comparison of rice egg, sperm and zygote BS-seq data

Supplemental Figure 2. Differential methylation in egg, sperm, zygotes and embryos

Supplemental Figure 3. DNA methylation remodeling during post-zygotic stage

Supplemental Figure 4. *dng702* and *dng701/dng704* mutations affect seed development

Supplemental Figure 5. Effect of DNG mutations on gene DNA methylation in the gametes and zygote

Supplemental Figure 6**.** Comparison of *dng* mutant versus WT and crWT DMRs.

Supplemental Figure 7. Features of *dng702* and *dng701/4* DMRs.

Supplemental Figure 8. The heterozygous *dng702* mutants showed reduced effect on the egg methylome compared with that of the homozygous mutant

Supplemental Figure 9. Analysis of *dng702* sperm DMRs.

Supplemental Figure 10. DNG701/4-demethylated loci in egg are remethylated in zygotes

Supplemental Figure 11. DNG702-demethylated loci in sperm are remethylated in zygotes

Supplemental Figure 12. Hypermethylated regions in the sperm genome are demethylated by DNG701/4 in the zygotes.

Supplemental Figure 13. Analysis of replicates shown in Fig. 4 and Fig. 5.

Supplemental Figure 14. Rice egg and zygote RNA-seq data analysis

Supplemental Figure 15. Analysis of DMR called by more rigorous criteria.

Supplemental Table 1. BS-seq data analysis

Supplemental Table 2. Detail information of the mutant DMRs.

Supplemental Table 3. RNA-seq analysis data

Supplemental Table 4. Gene lists of Supplemental Figure 3H

Supplemental Table 5. Gene lists of Figure 6C

Supplemental Table 6. Primer sequences

References

Anderson, S.N., Johnson, C.S., Chesnut, J., Jones, D.S., Khanday, I., Woodhouse, M., Li, C., Conrad, L.J., Russell, S.D., and Sundaresan, V. (2017). The Zygotic Transition Is Initiated in Unicellular Plant Zygotes with Asymmetric Activation of Parental Genomes. Dev Cell 43:349-358 e344.

Anderson, S.N., Johnson, C.S., Jones, D.S., Conrad, L.J., Gou, X., Russell, S.D., and Sundaresan, V. (2013). Transcriptomes of isolated Oryza sativa gametes characterized by deep sequencing: evidence for distinct sex-dependent chromatin and epigenetic states before fertilization. Plant J 76:729-741.

Autran, D., Baroux, C., Raissig, M.T., Lenormand, T., Wittig, M., Grob, S., Steimer, A., Barann, M., Klostermeier, U.C., Leblanc, O., et al. (2011). Maternal epigenetic pathways control parental contributions to Arabidopsis early embryogenesis. Cell 145:707-719.

Baroux, C., Autran, D., Raissig, M.T., Grimanelli, D., and Grossniklaus, U. (2013). Parental contributions to the transcriptome of early plant embryos. Curr Opin Genet Dev 23:72-74.

Becker, C., Hagmann, J., Muller, J., Koenig, D., Stegle, O., Borgwardt, K., and Weigel, D. (2011). Spontaneous epigenetic variation in the Arabidopsis thaliana methylome. Nature 480:245-249.

Borg, M., and Berger, F. (2015). Chromatin remodelling during male gametophyte development. Plant J 83:177-188.

Borg, M., Jacob, Y., Susaki, D., LeBlanc, C., Buendia, D., Axelsson, E., Kawashima, T., Voigt, P., Boavida, L., Becker, J., et al. (2020). Targeted reprogramming of H3K27me3 resets epigenetic memory in plant paternal chromatin. Nat Cell Biol 22:621-629.

608 Borges, F., Donoghue, M.T.A., LeBlanc, C., Wear, E.E., Tanurdzic, M., Berube, B., Brooks, A., Thompson,
609 W.F., Hanley-Bowdoin, L., and Martienssen, R.A. (2021). Loss of Small-RNA-Directed DNA
610 Methylation in the Plant Cell Cycle Promotes Germline Reprogramming and Somaclonal
611 Variation. Curr Biol 31:591-600 e594.
612 Bouyer, D., Kramdi, A., Kassam, M., Heese, M., Schnittger, A., Roudier, F., and Colot, V. (2017). DNA
613 methylation dynamics during early plant life. Genome Biol 18:179.
614 Calarco, J.P., Borges, F., Donoghue, M.T., Van Ex, F., Jullien, P.E., Lopes, T., Gardner, R., Berger, F., Feijo,
615 J.A., Becker, J.D., et al. (2012). Reprogramming of DNA methylation in pollen guides epigenetic
616 inheritance via small RNA. Cell 151:194-205.
617 Choi, Y., Gehring, M., Johnson, L., Hannon, M., Harada, J.J., Goldberg, R.B., Jacobsen, S.E., and Fischer,
618 R.L. (2002). DEMETER, a DNA glycosylase domain protein, is required for endosperm gene
619 imprinting and seed viability in arabidopsis. Cell 110:33-42.
620 Clark, S.J., Smallwood, S.A., Lee, H.J., Krueger, F., Reik, W., and Kelsey, G. (2017). Genome-wide base-
621 resolution mapping of DNA methylation in single cells using single-cell bisulfite sequencing (scBS-
622 seq). Nat Protoc 12:534-547.
623 Cubas, P., Vincent, C., and Coen, E. (1999). An epigenetic mutation responsible for natural variation in
624 floral symmetry. Nature 401:157-161.
625 de Vega-Bartol, J.J., Simoes, M., Lorenz, W.W., Rodrigues, A.S., Alba, R., Dean, J.F., and Miguel, C.M.
626 (2013). Transcriptomic analysis highlights epigenetic and transcriptional regulation during zygotic
627 embryo development of Pinus pinaster. BMC Plant Biol 13:123.
628 Frost, J.M., Kim, M.Y., Park, G.T., Hsieh, P.H., Nakamura, M., Lin, S.J.H., Yoo, H., Choi, J., Ikeda, Y.,
629 Kinoshita, T., et al. (2018). FACT complex is required for DNA demethylation at heterochromatin
630 during reproduction in Arabidopsis. Proc Natl Acad Sci U S A 115:E4720-E4729.
631 Gehring, M. (2019). Epigenetic dynamics during flowering plant reproduction: evidence for
632 reprogramming? New Phytol 224:91-96.
633 Gehring, M., Huh, J.H., Hsieh, T.F., Penterman, J., Choi, Y., Harada, J.J., Goldberg, R.B., and Fischer, R.L.
634 (2006). DEMETER DNA glycosylase establishes MEDEA polycomb gene self-imprinting by allele-
635 specific demethylation. Cell 124:495-506.
636 Guo, W., Fiziev, P., Yan, W., Cokus, S., Sun, X., Zhang, M.Q., Chen, P.Y., and Pellegrini, M. (2013). BS-
637 Seeker2: a versatile aligning pipeline for bisulfite sequencing data. BMC Genomics 14:774.

638 Harris, C.J., Scheibe, M., Wongpalee, S.P., Liu, W., Cornett, E.M., Vaughan, R.M., Li, X., Chen, W., Xue, Y.,
639 Zhong, Z., et al. (2018). A DNA methylation reader complex that enhances gene transcription.
640 Science 362:1182-1186.
641 Hauser, M.T., Aufsatz, W., Jonak, C., and Luschnig, C. (2011). Transgenerational epigenetic inheritance in
642 plants. Biochimica et biophysica acta 1809:459-468.
643 He, X.J., Chen, T., and Zhu, J.K. (2011). Regulation and function of DNA methylation in plants and animals.
644 Cell Res 21:442-465.
645 He, Y., Zhang, T., Yang, N., Xu, M., Yan, L., Wang, L., Wang, R., and Zhao, Y. (2017). Self-cleaving
646 ribozymes enable the production of guide RNAs from unlimited choices of promoters for
647 CRISPR/Cas9 mediated genome editing. J Genet Genomics 44:469-472.
648 Hsieh, P.H., He, S., Buttress, T., Gao, H., Couchman, M., Fischer, R.L., Zilberman, D., and Feng, X. (2016).
649 Arabidopsis male sexual lineage exhibits more robust maintenance of CG methylation than
650 somatic tissues. Proc Natl Acad Sci U S A 113:15132-15137.
651 Hsieh, T.F., Ibarra, C.A., Silva, P., Zemach, A., Eshed-Williams, L., Fischer, R.L., and Zilberman, D. (2009).
652 Genome-wide demethylation of Arabidopsis endosperm. Science 324:1451-1454.
653 Ibarra, C.A., Feng, X., Schoft, V.K., Hsieh, T.F., Uzawa, R., Rodrigues, J.A., Zemach, A., Chumak, N.,
654 Machlicova, A., Nishimura, T., et al. (2012). Active DNA demethylation in plant companion cells
655 reinforces transposon methylation in gametes. Science 337:1360-1364.
656 Ingouff, M., Rademacher, S., Holec, S., Soljic, L., Xin, N., Readshaw, A., Foo, S.H., Lahouze, B., Sprunck, S.,
657 and Berger, F. (2010). Zygotic resetting of the HISTONE 3 variant repertoire participates in
658 epigenetic reprogramming in Arabidopsis. Curr Biol 20:2137-2143.
659 Ingouff, M., Selles, B., Michaud, C., Vu, T.M., Berger, F., Schorn, A.J., Autran, D., Van Durme, M., Nowack,
660 M.K., Martienssen, R.A., et al. (2017). Live-cell analysis of DNA methylation during sexual
661 reproduction in Arabidopsis reveals context and sex-specific dynamics controlled by
662 noncanonical RdDM. Genes Dev 31:72-83.
663 Itoh, J., Nonomura, K., Ikeda, K., Yamaki, S., Inukai, Y., Yamagishi, H., Kitano, H., and Nagato, Y. (2005).
664 Rice plant development: from zygote to spikelet. Plant Cell Physiol 46:23-47.
665 Kawakatsu, T., Nery, J.R., Castanon, R., and Ecker, J.R. (2017). Dynamic DNA methylation reconfiguration
666 during seed development and germination. Genome biology 18:171.
667 Kawakatsu, T., Stuart, T., Valdes, M., Breakfield, N., Schmitz, R.J., Nery, J.R., Urich, M.A., Han, X., Lister, R.,
668 Benfey, P.N., et al. (2016). Unique cell-type-specific patterns of DNA methylation in the root
669 meristem. Nat Plants 2:16058.

670 Kawashima, T., and Berger, F. (2014). Epigenetic reprogramming in plant sexual reproduction. Nature
671 reviews. Genetics 15:613-624.
672 Kim, M.Y., Ono, A., Scholten, S., Kinoshita, T., Zilberman, D., Okamoto, T., and Fischer, R.L. (2019). DNA
673 demethylation by ROS1a in rice vegetative cells promotes methylation in sperm. Proc Natl Acad
674 Sci U S A 116:9652-9657.
675 Kim, M.Y., and Zilberman, D. (2014). DNA methylation as a system of plant genomic immunity. Trends in
676 plant science 19:320-326.
677 La, H., Ding, B., Mishra, G.P., Zhou, B., Yang, H., Bellizzi Mdel, R., Chen, S., Meyers, B.C., Peng, Z., Zhu, J.K.,
678 et al. (2011). A 5-methylcytosine DNA glycosylase/lyase demethylates the retrotransposon Tos17
679 and promotes its transposition in rice. Proc Natl Acad Sci U S A 108:15498-15503.
680 Law, J.A., and Jacobsen, S.E. (2010). Establishing, maintaining and modifying DNA methylation patterns in
681 plants and animals. Nat Rev Genet 11:204-220.
682 Li, C., Xu, H., Fu, F.F., Russell, S.D., Sundaresan, V., and Gent, J.I. (2020a). Genome-wide redistribution of
683 24-nt siRNAs in rice gametes. Genome Res 30:173-184.
684 Lin, J.Y., Le, B.H., Chen, M., Henry, K.F., Hur, J., Hsieh, T.F., Chen, P.Y., Pelletier, J.M., Pellegrini, M.,
685 Fischer, R.L., et al. (2017). Similarity between soybean and Arabidopsis seed methylomes and
686 loss of non-CG methylation does not affect seed development. Proc Natl Acad Sci U S A
687 114:E9730-E9739.
688 Lin, Y.J., and Zhang, Q. (2005). Optimising the tissue culture conditions for high efficiency transformation
689 of indica rice. Plant cell reports 23:540-547.
690 Liu, J., Wu, X., Yao, X., Yu, R., Larkin, P.J., and Liu, C.M. (2018). Mutations in the DNA demethylase
691 OsROS1 result in a thickened aleurone and improved nutritional value in rice grains. Proc Natl
692 Acad Sci U S A 115:11327-11332.
693 Liu, W., Xie, X., Ma, X., Li, J., Chen, J., and Liu, Y.G. (2015). DSDecode: A Web-Based Tool for Decoding of
694 Sequencing Chromatograms for Genotyping of Targeted Mutations. Mol Plant 8:1431-1433.
695 Lu, Y., Song, Y., Liu, L., and Wang, T. (2021). DNA methylation dynamics of sperm cell lineage
696 development in tomato. Plant J 105:565-579.
697 Ma, X., Xing, F., Jia, Q., Zhang, Q., Hu, T., Wu, B., Shao, L., Zhao, Y., Zhang, Q., and Zhou, D.X. (2021).
698 Parental variation in CHG methylation is associated with allelic-specific expression in elite hybrid
699 rice**.** Plant Physiol, kiab088, https://doi.org/10.1093/plphys/kiab088

Narsai, R., Gouil, Q., Secco, D., Srivastava, A., Karpievitch, Y.V., Liew, L.C., Lister, R., Lewsey, M.G., and Whelan, J. (2017). Extensive transcriptomic and epigenomic remodelling occurs during Arabidopsis thaliana germination. Genome biology 18:172.

Niederhuth, C.E., Bewick, A.J., Ji, L., Alabady, M.S., Kim, K.D., Li, Q., Rohr, N.A., Rambani, A., Burke, J.M., Udall, J.A., et al. (2016). Widespread natural variation of DNA methylation within angiosperms. Genome biology 17:194.

Ono, A., Yamaguchi, K., Fukada-Tanaka, S., Terada, R., Mitsui, T., and Iida, S. (2012). A null mutation of ROS1a for DNA demethylation in rice is not transmittable to progeny. Plant J 71:564-574.

Park, J.S., Frost, J.M., Park, K., Ohr, H., Park, G.T., Kim, S., Eom, H., Lee, I., Brooks, J.S., Fischer, R.L., et al. (2017). Control of DEMETER DNA demethylase gene transcription in male and female gamete companion cells in Arabidopsis thaliana. Proc Natl Acad Sci U S A 114:2078-2083.

Park, K., Kim, M.Y., Vickers, M., Park, J.S., Hyun, Y., Okamoto, T., Zilberman, D., Fischer, R.L., Feng, X., Choi, Y., et al. (2016). DNA demethylation is initiated in the central cells of Arabidopsis and rice. Proc Natl Acad Sci U S A 113:15138-15143.

Penterman, J., Zilberman, D., Huh, J.H., Ballinger, T., Henikoff, S., and Fischer, R.L. (2007). DNA demethylation in the Arabidopsis genome. Proc Natl Acad Sci U S A 104:6752-6757.

Rodrigues, J.A., Ruan, R., Nishimura, T., Sharma, M.K., Sharma, R., Ronald, P.C., Fischer, R.L., and Zilberman, D. (2013). Imprinted expression of genes and small RNA is associated with localized hypomethylation of the maternal genome in rice endosperm. Proceedings of the National Academy of Sciences of the United States of America 110:7934-7939.

Schmitz, R.J., Schultz, M.D., Lewsey, M.G., O'Malley, R.C., Urich, M.A., Libiger, O., Schork, N.J., and Ecker, J.R. (2011). Transgenerational epigenetic instability is a source of novel methylation variants. Science 334:369-373.

Schoft, V.K., Chumak, N., Choi, Y., Hannon, M., Garcia-Aguilar, M., Machlicova, A., Slusarz, L., Mosiolek, M., Park, J.S., Park, G.T., et al. (2011). Function of the DEMETER DNA glycosylase in the Arabidopsis thaliana male gametophyte. Proceedings of the National Academy of Sciences of the United States of America 108:8042-8047.

She, W., Grimanelli, D., Rutowicz, K., Whitehead, M.W., Puzio, M., Kotlinski, M., Jerzmanowski, A., and Baroux, C. (2013). Chromatin reprogramming during the somatic-to-reproductive cell fate transition in plants. Development 140:4008-4019.

730 Stroud, H., Do, T., Du, J., Zhong, X., Feng, S., Johnson, L., Patel, D.J., and Jacobsen, S.E. (2014). Non-CG
731 methylation patterns shape the epigenetic landscape in Arabidopsis. Nat Struct Mol Biol 21:64-
732 72.
733 Tan, F., Lu, Y., Jiang, W., Wu, T., Zhang, R., Zhao, Y., and Zhou, D.X. (2018). DDM1 Represses Noncoding
734 RNA Expression and RNA-Directed DNA Methylation in Heterochromatin. Plant Physiol 177:1187-
735 1197.
736 Tan, F., Zhou, C., Zhou, Q., Zhou, S., Yang, W., Zhao, Y., Li, G., and Zhou, D.X. (2016). Analysis of
737 Chromatin Regulators Reveals Specific Features of Rice DNA Methylation Pathways. Plant Physiol
738 171:2041-2054.
739 Walker, J., Gao, H., Zhang, J., Aldridge, B., Vickers, M., Higgins, J.D., and Feng, X. (2018). Sexual-lineage-
740 specific DNA methylation regulates meiosis in Arabidopsis. Nat Genet 50:130-137.
741 Wang, G., and Kohler, C. (2017). Epigenetic processes in flowering plant reproduction. J Exp Bot 68:797-
742 807.
743 Xia, W., Xu, J., Yu, G., Yao, G., Xu, K., Ma, X., Zhang, N., Liu, B., Li, T., Lin, Z., et al. (2019). Resetting
744 histone modifications during human parental-to-zygotic transition. Science 365:353-360.
745 Xie, X., Ma, X., and Liu, Y.G. (2019). Decoding Sanger Sequencing Chromatograms from CRISPR-Induced
746 Mutations. Methods in molecular biology 1917:33-43.
747 Zemach, A., Kim, M.Y., Hsieh, P.H., Coleman-Derr, D., Eshed-Williams, L., Thao, K., Harmer, S.L., and
748 Zilberman, D. (2013). The Arabidopsis nucleosome remodeler DDM1 allows DNA
749 methyltransferases to access H1-containing heterochromatin. Cell 153:193-205.
750 Zemach, A., Kim, M.Y., Silva, P., Rodrigues, J.A., Dotson, B., Brooks, M.D., and Zilberman, D. (2010a).
751 Local DNA hypomethylation activates genes in rice endosperm. Proc Natl Acad Sci U S A
752 107:18729-18734.
753 Zemach, A., McDaniel, I.E., Silva, P., and Zilberman, D. (2010b). Genome-wide evolutionary analysis of
754 eukaryotic DNA methylation. Science 328:916-919.
755 Zhang, C., Hung, Y.H., Rim, H.J., Zhang, D., Frost, J.M., Shin, H., Jang, H., Liu, F., Xiao, W., Iyer, L.M., et al.
756 (2019). The catalytic core of DEMETER guides active DNA demethylation in Arabidopsis. Proc Natl
757 Acad Sci U S A 116:17563-17571.
758 Zhang, H., Lang, Z., and Zhu, J.K. (2018). Dynamics and function of DNA methylation in plants. Nat Rev
759 Mol Cell Biol 19:489-506.
760 Zhang, J., Xu, R.J., Elliott, M.C., and Chen, D.F. (1997). Agrobacterium-mediated transformation of elite
761 indica and japonica rice cultivars. Molecular biotechnology 8:223-231.

762 Zhang, Y.N., Wei, D.M., He, E.M., Miao, S., Tian, H.Q., and Russell, S.D. (2010). Isolation of Male and
763 Female Gametes of Rice. Crop Science 50:2457-2463.
764 Zhao, P., Zhou, X., Shen, K., Liu, Z., Cheng, T., Liu, D., Cheng, Y., Peng, X., and Sun, M.X. (2019). Two-Step
765 Maternal-to-Zygotic Transition with Two-Phase Parental Genome Contributions. Dev Cell 49:882-
766 893 e885.
767 Zhou, S., Jiang, W., Zhao, Y., and Zhou, D.X. (2019). Single-cell three-dimensional genome structures of
768 rice gametes and unicellular zygotes. Nat Plants 5:795-800.
769 Zhu, J.K. (2009). Active DNA demethylation mediated by DNA glycosylases. Annu Rev Genet 43:143-166.

770

Figure legends

Figure 1. DNA cytosine methylation in rice gametes, zygote, globular and mature embryos.

(**A**) Box plots showing overall cytosine methylation levels (mCG, mCHG, and mCHH) in gene and transposable element (TE) of rice Dongjin (DJ) egg (E), sperm (S), zygote (Z), globular embryo (GE, 3 day after pollination), mature embryo (ME, 7 day after pollination), and seedling (Se). Values of the methylomes are averages from the two replicates. The 25th and 75th percentiles (box), and median values are shown. The upper whisker extends from the hinge to the largest value no further than 1.5 x IQR from the hinge (where IQR is the inter-quartile range, or distance between the first and third quartiles). The lower whisker extends from the hinge to the smallest value at most 1.5 x IQR of the hinge. The significance of difference *P* values (Student *t*-test, one side, paired) between the indicated comparison are shown. (**B**) Density plot showing the frequency distribution of fractional methylation differences at 50 bp windows between the indicated samples. The significant difference *P* values (Student *t*-test, one side, paired) between the indicated comparison are shown. (**C**) Numbers of differentially methylated regions (DMR) of 50 bp between the indicated comparisons, distributed in TE, gene and intergenic regions. Enrichment P values (Fisher's exactly test, one side) of GE-Z hypo- and ME-GE hyper-CHH DMR in TE are 0.00144 and 2.1e-16, respectively. (**D**) Genome browser screenshots of mCG, mCHG, and mCHH in egg (E), sperm (S), and zygote (Z). Two replicates are shown. Grey illustrates DMRs.

Figure 2. Effects in DNA glycosylase mutations in seed and embryo development.

(A) Seed setting rates in *dng702* and *dng701/4* compared with wild type plants. Upper part: spikelet phenotypes. Lower part: box plots of seed setting ratio of three indicated genotype plants, n=10 panicles. The 25th and 75th percentiles (box), median and highest and lowest values are shown. **(B)** Percentages of the three indicated seed types observed in each genotype calculated from 1550, 1430, and 1600 seeds for wild type (WT), *dng702* and *dng701/4*, respectively. **(C, D)** Phenotypes of *dng702* (**C**) and *dng701/4* (**D**) ovaries before pollination (BP)

and at 1-5 days after pollination (DAP). Red arrows indicate the aborted embryos. Orange arrows indicate retarded embryo development. e/s: egg/synergid, cc: central cell, EGE: early globular embryo, GE: globular embryo, SAM: shoot apical meristem, ra: radicle. Bars = 50 μm.

Figure 3. DNG proteins demethylate DNA in rice gametes and unicellular zygote.

(**A**) Box plots of overall methylation levels in gene (upper part) and TE (lower part) of wild type (WT) and *dng702* and *dng701/4* mutant egg (E), sperm (S), and zygote (Z) and CRISPR-negative (crWT) and *dng702* heterozygous (*dng702* (+/-) egg. Two replicates per genotype are shown. The 25^{th} and 75^{th} percentiles (box), and median values are shown. The upper whisker extends from the hinge to the largest value no further than 1.5 x IQR from the hinge (where IQR is the inter-quartile range, or distance between the first and third quartiles). The lower whisker extends from the hinge to the smallest value at most 1.5 x IQR of the hinge. The values are averages of the two replicates per genotype. The significant difference (Student *T* test, one tailed) *P* values between the mutants and wild type are indicated. **, *P* value<0.01; *, *P* value <0.05. (**B**) Numbers of differentially methylated regions (DMR) in the indicated mutant cells relative to wild type after subtraction of the crWT-WT DMRs after subtraction of the crWT-WT DMRs. (**C**) Overlaps between *dng702* and *dng701/4* egg DMRs. Complementary numbers are indicated. (**D**) Overlap of *dng701/4* egg and zygote DMRs.

Figure 4. Genomic regions demethylated by DNG proteins in egg are remethylated in zygote.

(**A**) Density plots of frequency distributions of CG, CHG, and CHH methylation differences at 50 bp windows between *dng702* and wild type (WT) eggs. (**B**) Upper part: density plots of hypermethylated regions (pink to red) in the mutant shown in (**A**) with methylation differences between zygote and egg in wild type. Randomly selected regions (green) from A were plotted as control. Lower part: boxplots of zygote versus egg methylation differences of all regions (Z-E) and the colored regions in B. P values of difference significances (student's *t* test, two sides) are indicated. Y axis: zygote-egg methylation difference in wild type. The 25^{th} and 75^{th} percentiles (box), median and highest and lowest values are shown. (**C)** Numbers of *dng702* and *dng701/4* egg hyper DMRs that are hypermethylated in wild type zygotes versus eggs. **(D)** Screenshot example showing that a Z-E hypermethylated region hypermethylated in *dng702* egg.

Figure 5. Hypermethylated regions in the egg genome are demethylated by DNG701/4 in zygote.

(**A**) Density plot of frequency distributions of CG, CHG, and CHH methylation differences at 50-bp windows between *dng701/4* and wild type (WT) zygotes. (**B**) Upper part: density plots of hypermethylated regions (pink - red) identified in the mutant **A** with the methylation differences between zygote and egg in wild type. Randomly selected regions (green) from **A** were plotted as control. Lower part: boxplots of zygote-egg methylation differences of all regions (Z-E) and regions colored in **B**. P values of difference significances (student's *t* test, two sides) are indicated. Y axis: zygote-egg methylation difference in wild type. The 25^{th} and 75^{th} percentiles (box), median and highest and lowest values are shown. (**C**) Numbers of *dng701/4* egg hyper DMRs that are hypomethylated in wild type zygotes versus eggs. **(D)** Screenshot example showing that a Z-E hypomethylated region was hypermethylated in *dng701/4* zygote.

Figure 6. Function of DNG genes in egg and zygote gene expression.

(A) Scattering plotted of differential transcript levels in mutant eggs relative to wild type against $\log_2$FC. Red dots indicate differentially expressed genes (DEGs) (>2 fold, Q value<0.01). **(B)** Overlapping of *dng701* and *dng701/4* egg DEGs associated with hyper DMR. **(C)** Heat map clustering of 265 and 160 wild type egg-expressed (relative to zygote) genes that were downregulated and hypermethylated respectively in *dng702* and *dng701/4* eggs and 104 wild type zygote-expressed (relative to egg) genes that were upregulated and hypomethylated in *dng701/4* eggs. **(D)** IGV of two examples of the genes described in (**C**). **(E)** Scattering plot of transcript levels in *dng701/4* zygote versus wild type. **(F)** Overlapping of *dng701/4* egg and zygote DEGs associated with hyper DMR. The *P* values of overlapping significances (Fisher's exactly test) for up and down-regulated genes in egg and zygote are 1.23e-17 and 3.32e-17, respectively. **(G)** Density plots of wild type zygote-egg DEGs (>2 folds, p<0.01, Figure 1F) (left) in *dng701/4* background (right). Dashed lines indicate 2 fold changes. **(H)** Numbers of up- or downregulated genes in wild type zygote relative to egg that remained to be differentially (red and green) or no longer differentially (white and grey) expressed in *dng701/4* zygote. Grey represents DEGs associated with hyper DMRs in body or promoter in the mutant zygote. **(I)**

Numbers of cell cycle (CC) (GO: 0007049), DNA metabolism (DM) (GO: 0006259), and transcription factor (TF) (GO: 0003700) genes belonging to the grey box in (**D**). **(J)** Left panel, heat map of several zygote-activated (relative to the gametes) genes that were repressed and hyper-CHH methylated in *dng701/4* zygotes. Right panel, genome browser snapshots showing mCHH levels of three indicated genes in wild type and crWT egg (E) and zygote (Z) and *dng701/4* zygote (701/4-Z), sperm (701/4-S) and egg (701/4-E).

Figure 7. Model of DNA methylation remodeling before and after fertilization and function of the DNA demethylases DNG702 and DNG701/DNG704.

DNG702 and DNG701/704 are shown to target different euchromatin regions for DNA demethylation in gametes (egg and sperm). RNA-directed DNA methylation (RdDM) is suggested to be involved in the remodeling of CHH methylation in zygote and during postzygotic development. Dashed arrows indicate hypothetic actions.

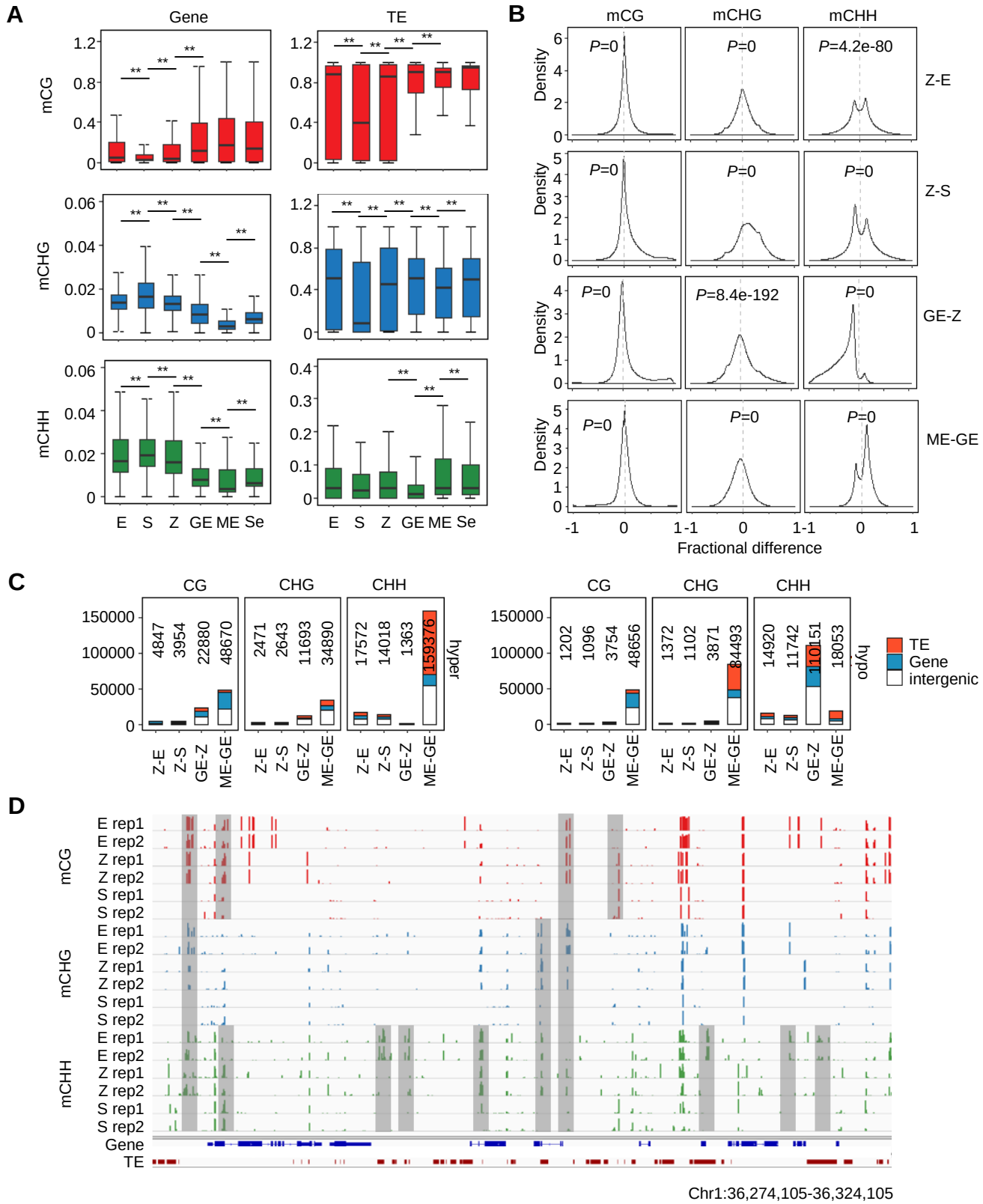
A
Gene
TE
mCG
mCHG
mCHH
E S Z GE ME Se
B
mCG
mCHG
mCHH
Density
P=0
P=4.2e-80
P=8.4e-192
Z-E
Z-S
GE-Z
ME-GE
Fractional difference
C
CG
CHG
CHH
hyper
hypo
TE
Gene
intergenic
Z-E
Z-S
GE-Z
ME-GE
D
E rep1
E rep2
Z rep1
Z rep2
S rep1
S rep2
Gene
TE
Chr1:36,274,105-36,324,105

Figure 1. DNA cytosine methylation in rice gametes, zygote, globular and mature embryos.
(**A**) Box plots showing overall cytosine methylation levels (mCG, mCHG, and mCHH) in gene and transposable element (TE) of rice Dongjin (DJ) egg (E), sperm (S), zygote (Z), globular embryo (GE, 3 day after pollination), mature embryo (ME, 7 day after pollination), and seedling (Se). Values of the methylomes are averages from the two replicates. The 25^{th} and 75^{th} percentiles (box), and median values are shown. The upper whisker extends from the hinge to the largest value no further than 1.5 x IQR from the hinge (where IQR is the inter-quartile range, or distance between the first and third quartiles). The lower whisker extends from the hinge to the smallest value at most 1.5 x IQR of the hinge. The significance of difference *P* values (Student *T* test, one side, paired) between the indicated comparison are shown. (**B**) Density plot showing the frequency distribution of fractional methylation differences at 50 bp windows between the indicated samples. The significant difference *P* values (Student *T* test, one side, paired) between the indicated comparison are shown. (**C**) Numbers of differentially methylated regions (DMR) of 50 bp between the indicated comparisons, distributed in TE, gene and intergenic regions. Enrichment P values (Fisher's exactly test, one side) of GE-Z hypo- and ME-GE hyper-CHH DMR in TE are 0.00144 and 2.1e-16, respectively. (**D**) Genome browser screenshots of mCG, mCHG, and mCHH in egg (E), sperm (S), and zygote (Z). Two replicates are shown. Grey illustrates differentially methylated regions.

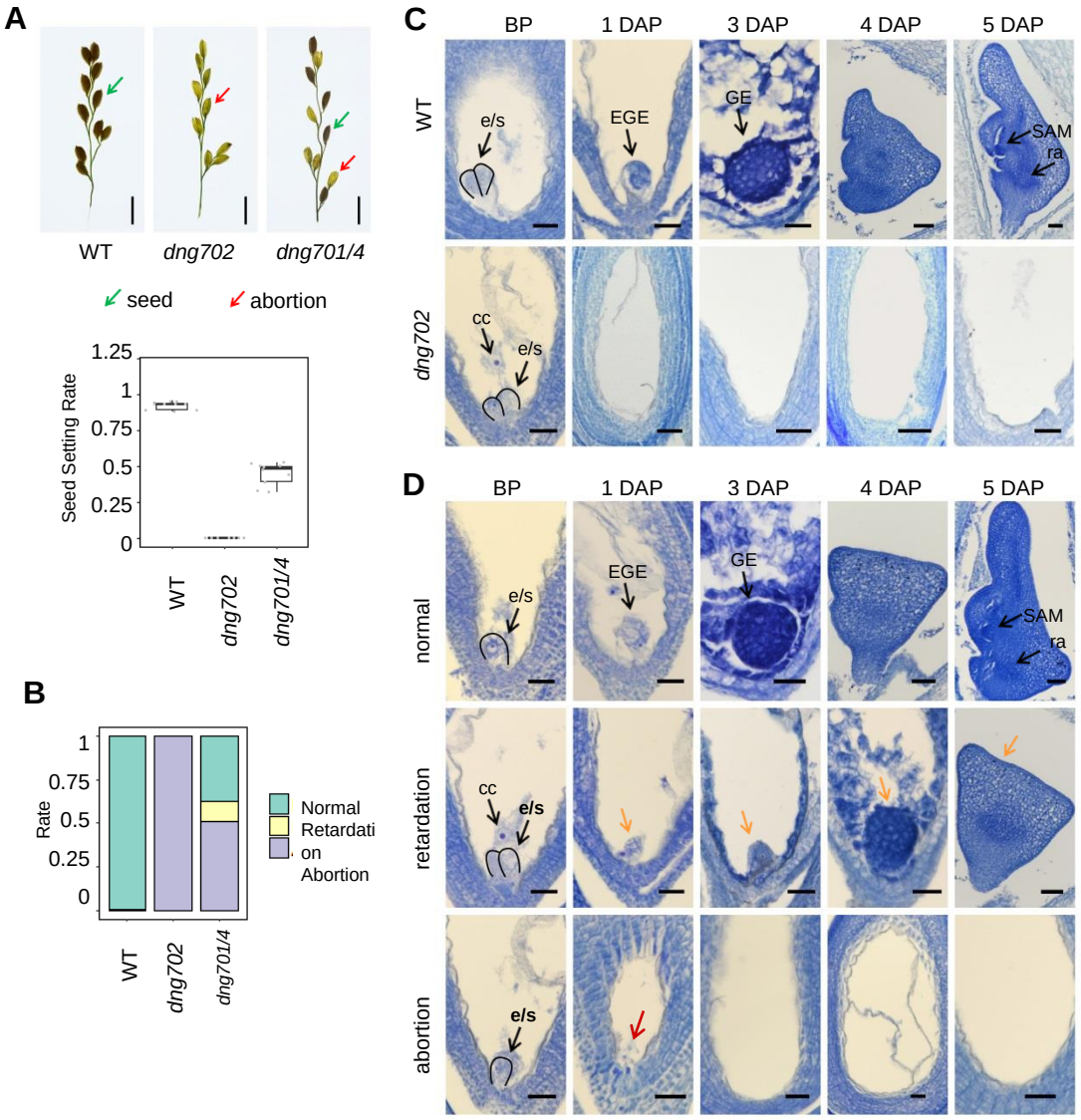

Figure 2. Effects in DNA glycosylase mutations in seed and embryo development.
(A) Seed setting rates in *dng702* and *dng701/4* compared with wild type plants. Upper part, spikelet phenotypes. Lower part, box plots of seed setting ratio of three indicated genotype plants, n=10 panicles. The 25th and 75th percentiles (box), median and highest and lowest values are shown. **(B)** Percentages of the three indicated seed types observed in each genotype calculated from 1550, 1430, and 1600 seeds for wild type (WT), *dng702* and *dng701/4*, respectively. **(C, D)** Phenotypes of *dng702* (**C**) and *dng701/4* (**D**) ovaries before pollination (BP) and at 1-5 days after pollination (DAP). Red arrows indicate the aborted embryos. Orange arrows indicate retarded embryo development. e/s: egg/synergid, cc: central cell, EGE: early globular embryo, GE: globular embryo, SAM: shoot apical meristem, ra: radicle. Bars = 50 μm.

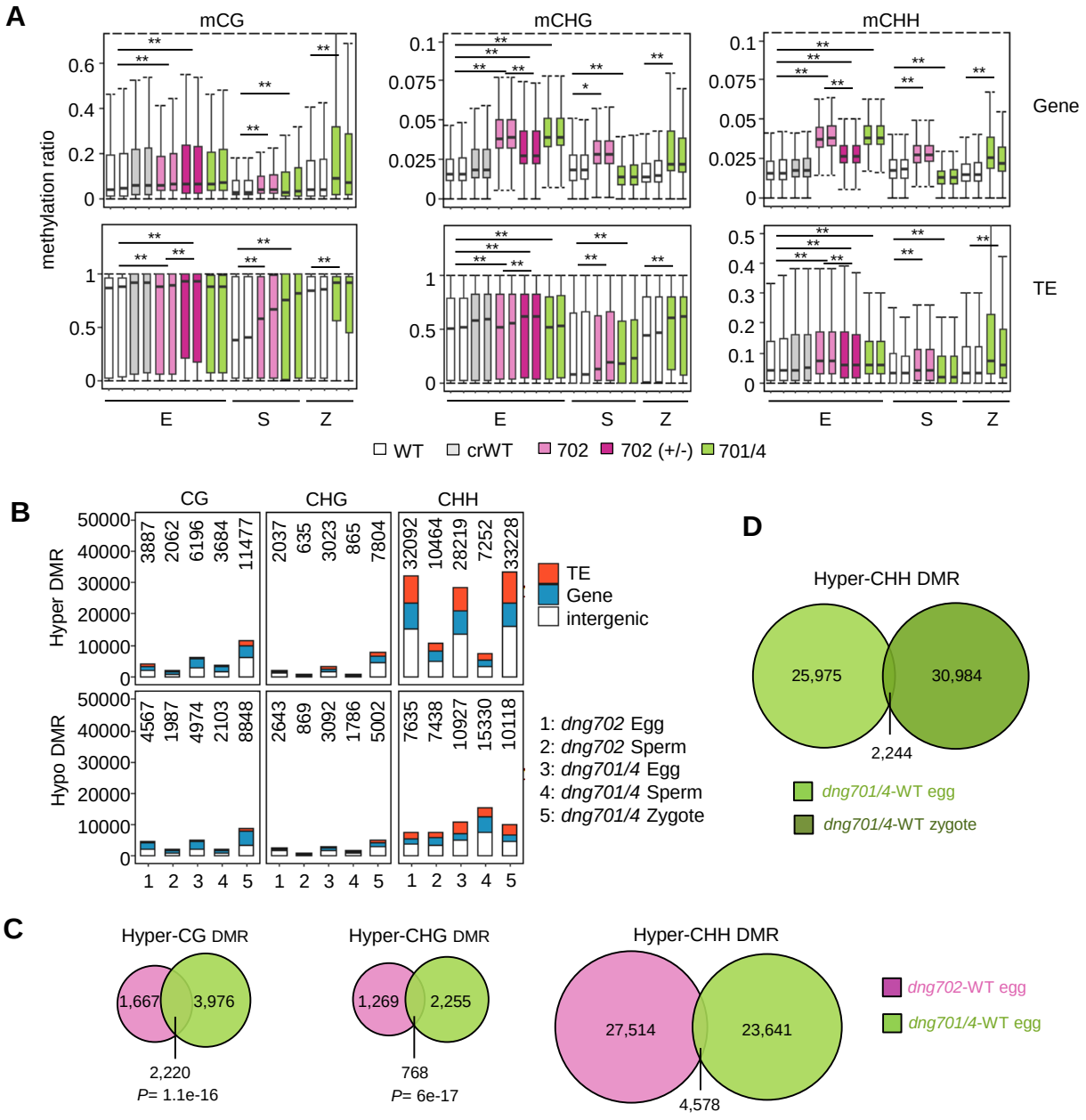

Figure 3. DNG proteins demethylate DNA in rice gametes and unicellular zygote.
(A) Box plots of overall methylation levels in gene (upper part) and TE (lower part) of wild type (WT) and *dng702* and *dng701/4* mutant egg (E), sperm (S), and zygote (Z) and CRISPR-negative (crWT) and *dng702* heterozygous (*dng702* (+/-) egg. Two replicates per genotype are shown. The 25th and 75th percentiles (box), and median values are shown. The upper whisker extends from the hinge to the largest value no further than 1.5 x IQR from the hinge (where IQR is the inter-quartile range, or distance between the first and third quartiles). The lower whisker extends from the hinge to the smallest value at most 1.5 x IQR of the hinge. The values are averages of the two replicates per genotype. The significant difference (Student *T* test, one tailed) *P* values between the mutants and wild type are indicated. **, *P* value<0.01; *, *P* value <0.05. **(B)** Numbers of differentially methylated regions (DMR) in the indicated mutant cells relative to wild type after subtraction of the crWT-WT DMRs. **(C)** Overlaps between *dng702* and *dng701/4* egg DMRs. Complementary numbers are indicated. **(D)** Overlap of *dng701/4* egg and zygote DMRs.

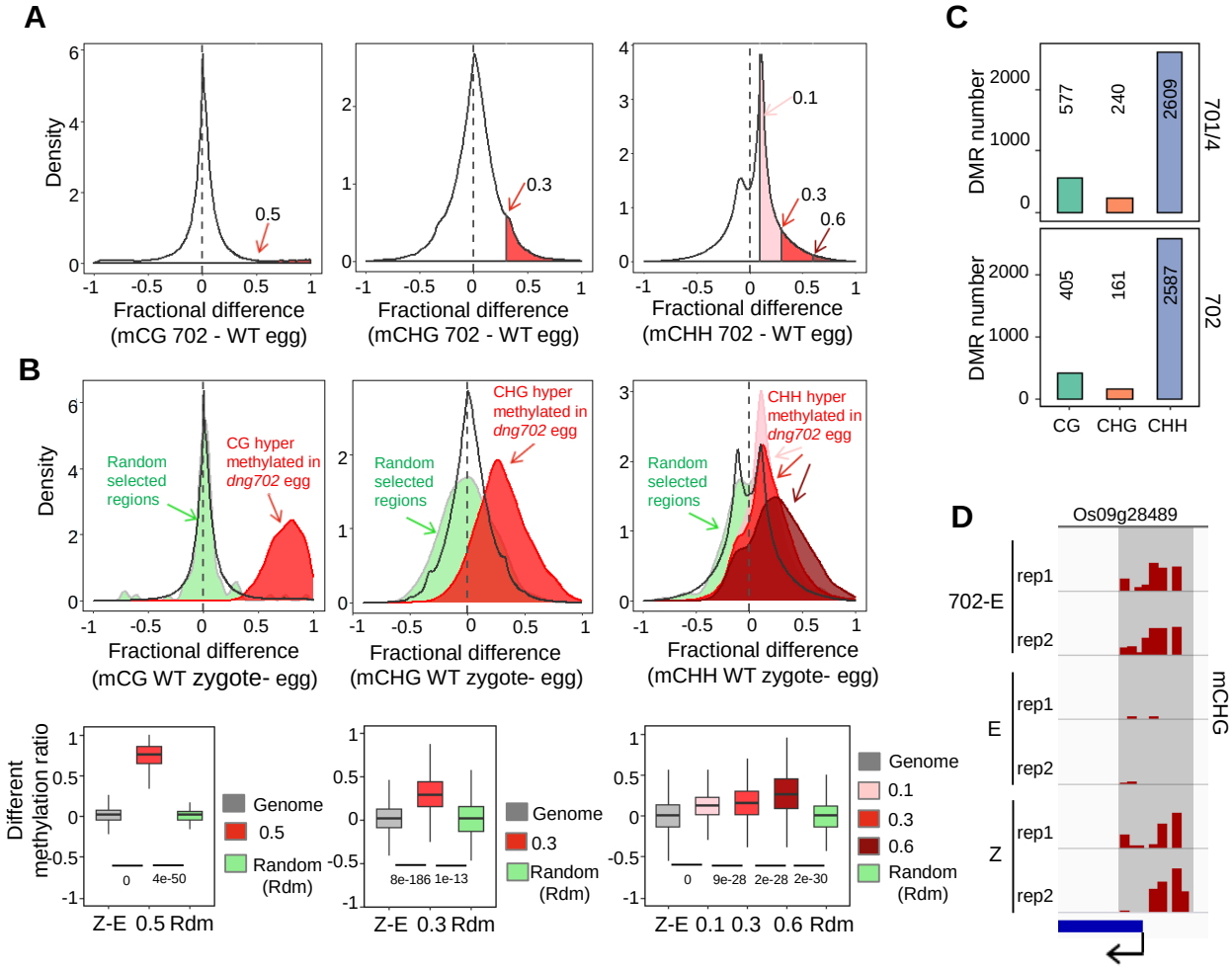

Figure 4. Genomic regions demethylated by DNG proteins in egg are remethylated in zygote.

(A) Density plots of frequency distributions of CG, CHG, and CHH methylation differences at 50 bp windows between *dng702* and wild type (WT) eggs. **(B)** Upper part: density plots of hypermethylated regions (pink to red) in the mutant shown in (A) with methylation differences between zygote and egg in wild type. Randomly selected regions (green) from **A** were plotted as control. Lower part: boxplots of zygote versus egg methylation differences of all regions (Z-E) and the colored regions in **B**. P values of difference significances (student's *t* test, two sides) are indicated. Y axis, zygote-egg methylation difference in wild type. The 25th and 75th percentiles (box), median and highest and lowest values are shown. **(C)** Numbers of *dng702* and *dng701/4* egg hyper DMRs that are hypermethylated in wild type zygotes versus eggs. **(D)** Screenshot example showing that a Z-E hypermethylated region was hypermethylated in *dng702* egg.

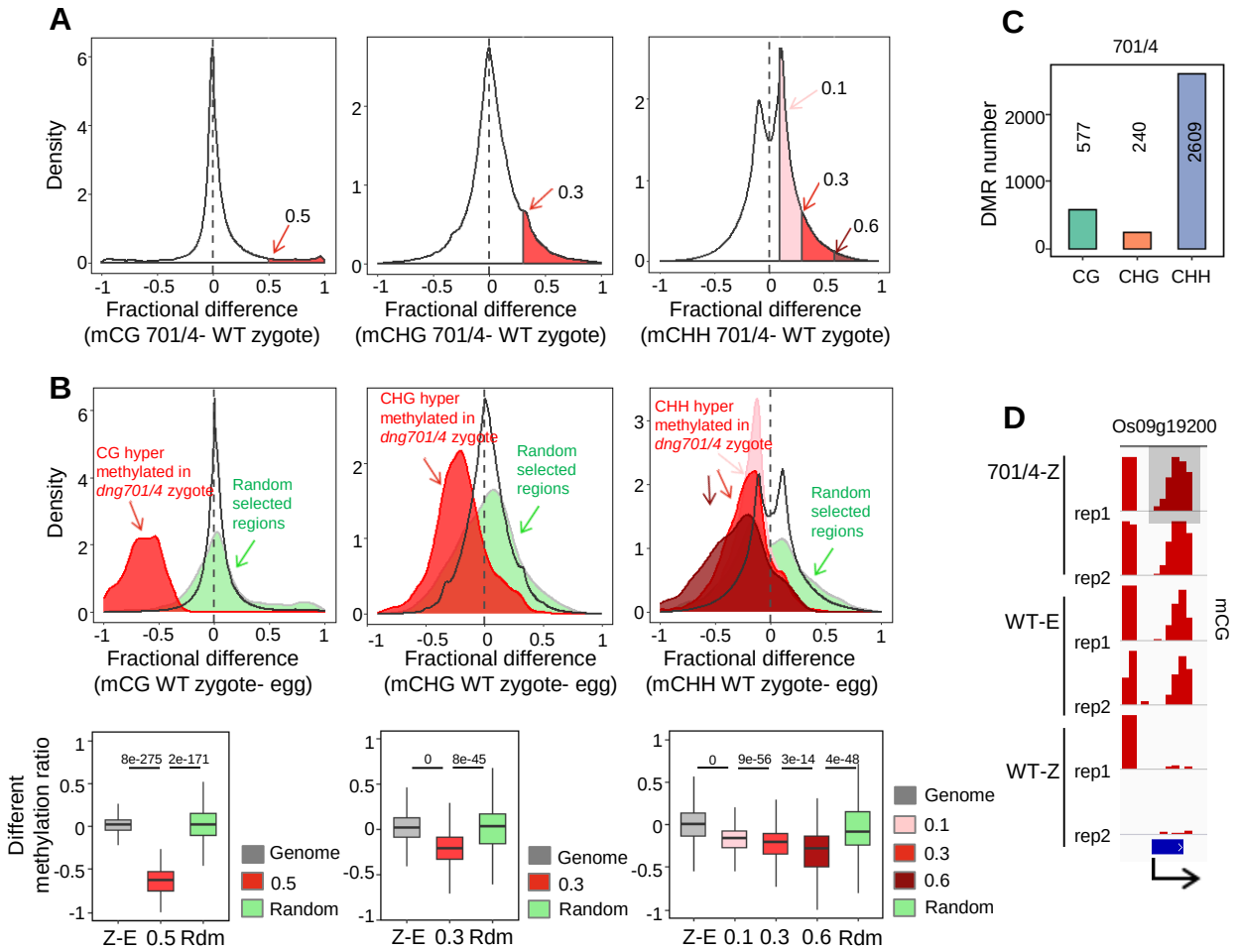

Figure 5. Hypermethylated regions in the egg genome are demethylated by DNG701/4 in zygote.
(A) Density plot of frequency distributions of CG, CHG, and CHH methylation differences at 50-bp windows between *dng701/4* and wild type (WT) zygotes. **(B)** Upper part: density plots of hypermethylated regions (pink - red) identified in the mutant **A** with the methylation differences between zygote and egg in wild type. Randomly selected regions (green) from **A** were plotted as control. Lower part: boxplots of zygote-egg methylation differences of all regions (Z-E) and regions colored in **B**. P values of difference significances (student's *t* test, two sides) are indicated. Y axis, zygote-egg methylation difference in wild type. The 25th and 75th percentiles (box), median and highest and lowest values are shown. **(C)** Numbers of *dng701/4* egg hyper DMRs that are hypomethylated in wild type zygotes versus eggs. **(D)** Screenshot example showing that a Z-E hypomethylated region was hypermethylated in dng701/4 zygote.

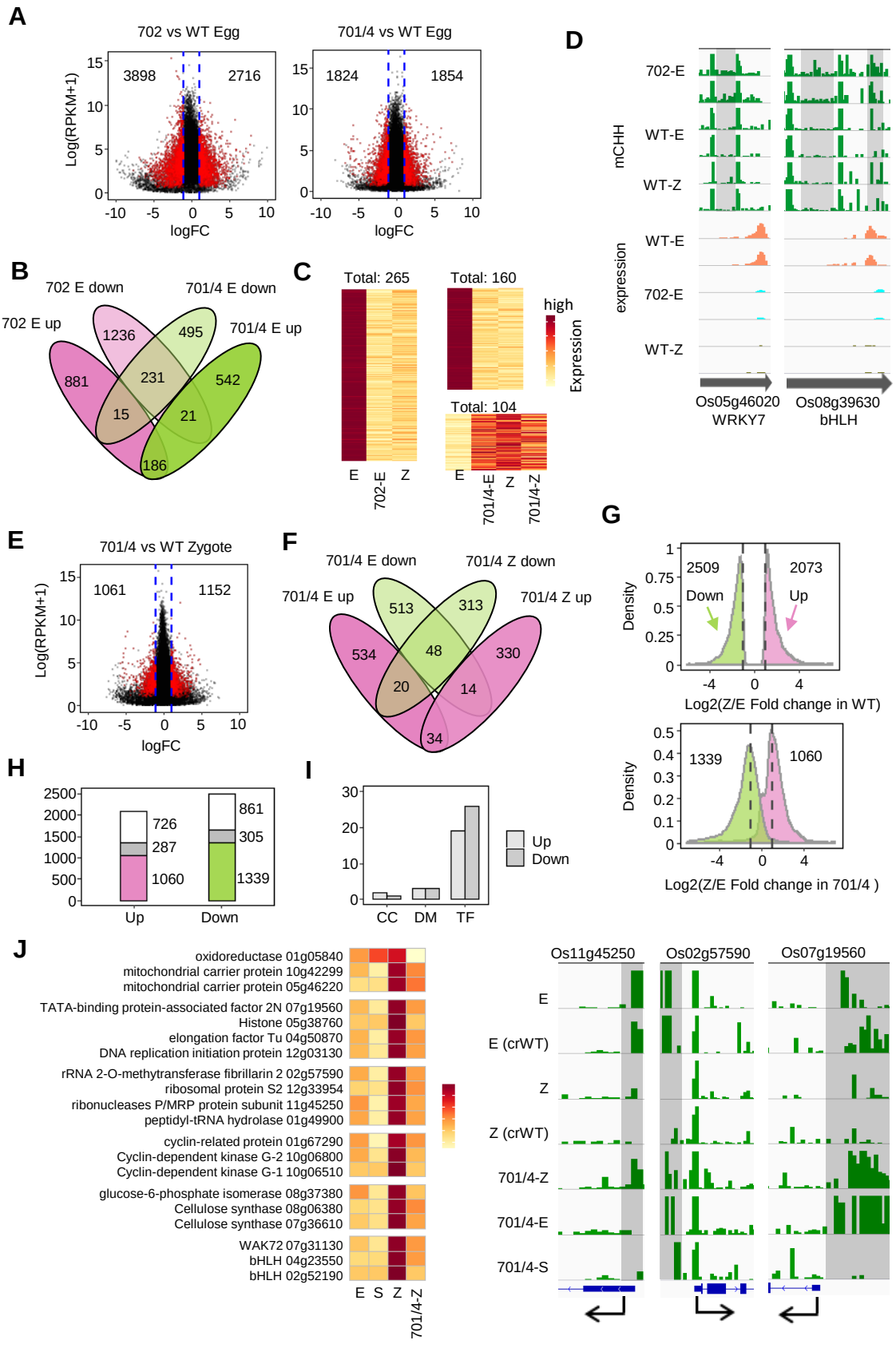
A
702 vs WT Egg
701/4 vs WT Egg
Log(RPKM+1)
logFC
3898
2716
1824
1854
B
702 E down
701/4 E down
702 E up
701/4 E up
1236
495
881
231
542
15
21
186
C
Total: 265
Total: 160
Total: 104
high
Expression
E
702-E
Z
701/4-E
701/4-Z
D
mCHH
expression
702-E
WT-E
WT-Z
Os05g46020
WRKY7
Os08g39630
bHLH
E
701/4 vs WT Zygote
1061
1152
F
701/4 E down
701/4 Z down
701/4 E up
701/4 Z up
513
313
534
48
330
20
14
34
G
Density
2509
Down
2073
Up
Log2(Z/E Fold change in WT)
1339
1060
Log2(Z/E Fold change in 701/4)
H
726
287
1060
861
305
1339
Up
Down
I
CC
DM
TF
Up
Down
J
oxidoreductase 01g05840
mitochondrial carrier protein 10g42299
mitochondrial carrier protein 05g46220
TATA-binding protein-associated factor 2N 07g19560
Histone 05g38760
elongation factor Tu 04g50870
DNA replication initiation protein 12g03130
rRNA 2-O-methytransferase fibrillarin 2 02g57590
ribosomal protein S2 12g33954
ribonucleases P/MRP protein subunit 11g45250
peptidyl-tRNA hydrolase 01g49900
cyclin-related protein 01g67290
Cyclin-dependent kinase G-2 10g06800
Cyclin-dependent kinase G-1 10g06510
glucose-6-phosphate isomerase 08g37380
Cellulose synthase 08g06380
Cellulose synthase 07g36610
WAK72 07g31130
bHLH 04g23550
bHLH 02g52190
E S Z 701/4-Z
Os11g45250
Os02g57590
Os07g19560
E
E (crWT)
Z
Z (crWT)
701/4-Z
701/4-E
701/4-S

Figure 6. Function of DNG genes in egg and zygote gene expression. (A) Scattering plotted of differential transcript levels in mutant eggs relative to wild type against $\log_2$FC. Red dots indicate differentially expressed genes (DEGs) (>2 fold, Q value<0.01). **(B)** Overlapping of *dng701* and *dng701/4* egg DEGs associated with hyper DMR. **(C)** Heat map clustering of 265 and 160 wild type egg-expressed (relative to zygote) genes that were downregulated and hypermethylated respectively in *dng702* and *dng701/4* eggs, and 268 and 104 wild type zygote-expressed (relative to egg) genes that were up-regulated and hypomethylated in *dng702* and *dng701/4* eggs, respectively. **(D)** IGV of two examples of the genes described in (**C**). **(E)** Scattering plot of transcript levels in *dng701/4* zygote versus wild type. **(F)** Overlapping of *dng701/4* egg and zygote DEGs associated with hyper DMR. The *P* values of overlapping significances (Fisher's exactly test) for up and down-regulated genes in egg and zygote are 1.23e-17 and 3.32e-17, respectively. **(G)** Density plots of wild type zygote-egg DEGs (>2 folds, p<0.01, Fig. 1F) (left) in *dng701/4* background (right). Dashed lines indicate 2 fold changes. **(H)** Numbers of up- or downregulated genes in wild type zygote relative to egg that remained to be differentially (red and green) or no longer differentially (white and grey) expressed in *dng701/4* zygote. Grey represents DEGs associated with hyper DMRs in body or promoter in the mutant zygote. **(I)** Numbers of cell cycle (CC) (GO:0007049), DNA metabolism (DM) (GO:0006259), and transcription factor (TF) (GO:0003700) genes belonging to the grey box in (**H**). **(J)** Left panel, heat map of several zygote-activated (relative to the gametes) genes that were repressed and hyper-CHH methylated in *dng701/4* zygotes. Right panel, genome browser snapshots showing mCHH levels of three indicated genes (A-D) in wild type and crWT egg (E), zygote (Z), and *dng701/4* zygote (701/4-Z), sperm (701/4-S) and egg (701/4-E).

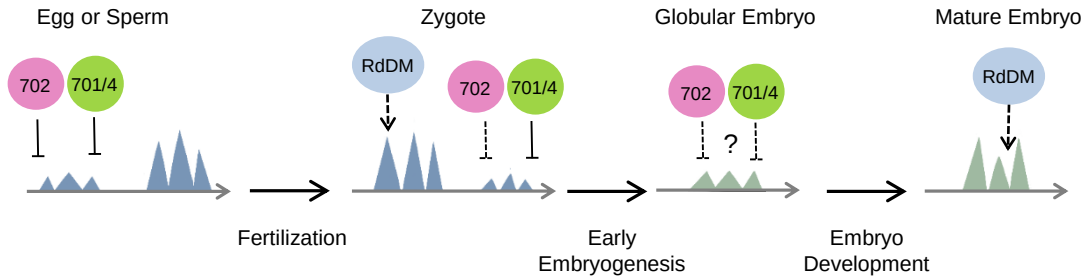

Figure 7. Model of DNA methylation remodeling before and after fertilization and function of the DNA demethylases DNG702 and DNG701/DNG704. DNG702 and DNG701/704 are shown to target different euchromatin regions for DNA demethylation in gametes (egg and sperm). RNA-directed DNA methylation (RdDM) is suggested to be involved in the remodeling of CHH methylation in zygote and during postzygotic development. Dashed arrows indicate hypothetic actions.