Dnmt1 binds and represses genomic retroelements via DNA methylation in mouse early embryos

Byungkuk Min †, Jung Sun Park †, Young Sun Jeong 1, Kyuheum Jeon 1,2 and Yong-Kook Kang 1,2,*

1 Development and Differentiation Research Center, Korea Research Institute of Bioscience Biotechnology (KRIIBB), 125 Gwahak-ro, Yuseong-gu, Daejeon 34141, South Korea and 2 Department of Functional Genomics, Korea University of Science and Technology, Daejeon 34113, South Korea

Received May 10, 2019; Revised June 10, 2020; Editorial Decision June 27, 2020; Accepted July 03, 2020

ABSTRACT

Genome-wide passive DNA demethylation in cleavage-stage mouse embryos is related to the cytoplasmic localization of the maintenance methyltransferase DNMT1. However, recent studies provided evidences of the nuclear localization of DNMT1 and its contribution to the maintenance of methylation levels of imprinted regions and other genomic loci in early embryos. Using the DNA adenine methylase identification method, we identified Dnmt1-binding regions in four- and eight-cell embryos. The unbiased distribution of Dnmt1 peaks in the genic regions (promoters and CpG islands) as well as the absence of a correlation between the Dnmt1 peaks and the expression levels of the peak-associated genes refutes the active participation of Dnmt1 in the transcriptional regulation of genes in the early developmental period. Instead, Dnmt1 was found to associate with genomic retroelements in a greatly biased fashion, particularly with the LINE1 (long interspersed nuclear elements) and ERVK (endogenous retrovirus type K) sequences. Transcriptomic analysis revealed that the transcripts of the Dnmt1-enriched retroelements were overrepresented in Dnmt1 knockdown embryos. Finally, methyl-CpG-binding domain sequencing proved that the Dnmt1-enriched retroelements, which were densely methylated in wild-type embryos, became demethylated in the Dnmt1-depleted embryos. Our results indicate that Dnmt1 is involved in the repression of retroelements through DNA methylation in early mouse development.

INTRODUCTION

In mammals, the first wave of global epigenetic reprogramming in the life cycle occurs during early embryogenesis and is important for the establishment of pluripotency (1). Initially, DNA methylation in the paternal chromosomes is drastically reduced shortly after fertilization (active demethylation). During cleavage, a progressive mode of demethylation affects both the maternal and paternal chromosomes, and this passive demethylation is thought to be achieved by the nuclear exclusion of DNA methyltransferase 1 (Dnmt1) in preimplantation-stage embryos (2–7).

Dnmt1 is considered as the main maintenance methyltransferase (8) because of its catalytic preference for hemimethylated DNA (9) and its localization on the replication fork during the S phase of the cell cycle (10,11). There is accumulating evidence for the role of Dnmt1 in maintaining methylation imprints in somatic cells (12) and preimplantation embryos (5,13). Genomic imprinting patterns play pivotal roles in development and growth; their deregulation greatly impacts the embryonic phenotype and may be the major cause of defects in methylation-deficient mutants (1). Thus, the observed cytoplasmic retention of Dnmt1 indicates that the maintenance of genomic imprints, and possibly other genomic repeats (5,14,15), is required during early embryonic development. In contrast to the early observations of the nuclear exclusion of Dnmt1, a somatic isoform of Dnmt1 (Dnmt1s) has been directly proven to be present in the nucleus of early embryos by means of oocyte-form Dnmt1 (Dnmt1o) knockout mice and a specific antibody discriminating the oocyte isoform from the somatic form (16). This finding has also been supported by other studies (5,13). Moreover, the complete knockout of both the Dnmt1o and Dnmt1s resulted in the complete abolishment of all methylation imprints (13,16,17). Therefore, although it is unclear how local methylation patterns are maintained over genome-wide demethylation, Dnmt1 is thought to contribute to maintaining DNA methylation patterns at certain genomic loci in cleavage-stage preimplantation embryos.

*To whom correspondence should be addressed. Tel: +82 42 860 4427; Fax: +82 42 860 4608; Email: ykkang@kribb.re.kr
†The authors wish it to be known that, in their opinion, the first two authors should be regarded as Joint First Authors.

© The Author(s) 2020. Published by Oxford University Press on behalf of Nucleic Acids Research.
This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com
Retrotransposable elements, or retroelements, involving long and short interspersed nuclear elements (SINEs and LINEs, respectively) and long terminal repeats (LTRs) are a significant component of mammalian genomes, occupying ~40% of mouse genomes (Mouse Genome Sequencing Consortium, 2002). The proven capacity for causing deleterious genomic rearrangements and the potent mechanisms for their silencing of neighboring genes in somatic tissues have led many to view retroelements as harmful parasites (18). In somatic cells, the expression of retroelements is attenuated by DNA methylation (19). However, in germ cells and early embryos where epigenetic reprogramming takes place, the resulting genome-wide demethylation lifts this chief repressive mechanism from transcription- and transposition-competent retroelements. In reality, retroelements are highly expressed in mature oocytes and cleaved embryos in mice, and different retroelements have specific, developmentally regulated expression patterns (20). This coordinated activation and repression of the same type of element with multiple genomic sites may provide robustness to the embryo, thereby allowing for temporal control of the newly formed zygotic genome after fertilization (21). For instance, endogenous retrovirus type L (ERV L) retroelements, which are specifically expressed at the two-cell stage (20), have been proposed to regulate a large number of genes during zygotic genome activation by promoting pervasive transcription into neighboring genes (22,23). In early human embryos, MLT2A1 and THE1A ERVs are expressed in an eight-cell-stage-specific fashion and LTR5_Hs is expressed in a morula-stage-specific fashion (24). In the case of LINE1 retroelements, they show the highest level of expression at the two-cell stage mouse embryo and function primarily at the chromatin level; in addition, LINE1 retroelements promote global chromatin decondensation and recondensation via transcriptional activity (21). Therefore, it appears persuasive that the transcriptional activation of retrotransposons during the preimplantation development stage may not be a nonspecific side effect of heterochromatin loss in the embryo and that the timely activation and repression of retroelement expression may have great significance for early developmental programs and embryo survival. Therefore, it is relevant to determine the regulatory role of Dnmt1 in this programmed development.

DNA adenine methylase identification (DamID) is one of the most comprehensive and versatile methods available for profiling protein–DNA interactions on a genomic scale (25–27). The Escherichia coli Dam protein methylates adenine (at the N6 position) in the GATC sequence; this protein does not exist in mammals. When fused to a chromatin protein of interest, Dam can be targeted in vivo to the binding sequences of the chromatin protein and can methylate the nearby GATC sequences. The adenine-methylated GATC sequences can be digested by the methylation-sensitive DpnI restriction enzyme, and the resulting DNA fragments can be used in microarrays or can be sequenced for mapping the binding sites. To assess chromatin binding on a genome-wide scale, the leading alternative to DamID is chromatin immunoprecipitation, for which the main limitation is the requirement for a highly specific antibody. In contrast, DamID does not require an antibody for the identification of protein-binding loci.

Because Dnmt1 may act on certain genomic loci to allow the loci to fulfill their roles in cleavage-stage mouse embryos, we attempted to identify genomic Dnmt1-binding loci, particularly focusing on a variety of retroelement families. We first verified the nuclear localization of GFP-Dnmt1 recombinant proteins after microinjection into the cytoplasm of a blastomere. Dnmt1-binding regions were identified using DamID and DpnI fragment enrichment by ligation-mediated PCR sequencing (DELP-seq). We found that Dnmt1 displayed a significant bias toward certain retroelement families that were increasingly transcribed when Dnmt1 was knocked down. Dnmt1 depletion in early mouse embryos via CRISPR/Cas9 technology revealed a considerable demethylation at these retroelement sequences. The identification of the Dnmt1-binding loci and the gross effect of Dnmt1 binding on the repression of retroelements via DNA methylation provide insight into the functions and mechanisms of actions of Dnmt1 during early development.

MATERIALS AND METHODS

Ethics statement

This study was carried out in strict accordance with the recommendations in the Guide for the Care and Use of Laboratory Animals of the National Livestock Research Institute of Korea. The protocol was approved by the Committee on the Ethics of Animal Experiments of the Korea Research Institute of Bioscience and Biotechnology.

Construction of Dnmt1 expression vectors

For cloning mouse Dnmt1s or Dnmt1o sequences, PCR was performed with cDNA obtained from HEK293 cells under the condition of 30 cycles of 94°C/30 s, 57°C/30 s and 72°C/3 m (see Supplementary File S1 for primer information). The resulting PCR products, which were devoid of the catalytic domains (Dnmt1s or Dnmt1o), were cloned into the EcoRI (NEB) and SalI (NEB) sites of pEGFP-C2 (Clontech) to make pEGFP-Dnmt1sc. Proper cloning was confirmed by sequencingsmadid clones. For construction of pIND-Dnmt1-V5-Dam, 3.4 kb of Dnmt1 fragment lacking the catalytic domain was amplified from mouse cDNA by PCR (see Supplementary File S1 for primer information) and cloned into pIND-V5-Dam plasmid (Addgene).

Microinjection and in vitro culture of mouse embryos

Mouse zygotes were collected from superovulated BCF1 (C57BL/6 × CBA/CA) females as described previously (28). Briefly, female BCF1 mice at 5 weeks of age were injected with 5 IU of pregnant mare serum gonadotrophin (PMSF), followed by 5 IU of human chorionic gonadotropins (hCG) 48 h apart, and mated with male mice. Successful mating was determined the following morning by detection of a vaginal plug. Mouse zygotes were transferred to M2 medium (Sigma) containing 0.1% (w/v) hyaluronidase to remove cumulus cells and cultured in M16 medium (Sigma) at 37°C, 5% CO2 in air (29).

Microinjection was performed 24 h after human chorionic gonadotropin injection (30). To visualize the pronuclei, cumulus-free zygotes were centrifuged at 13,000 rpm...
for 5 min. Linearized GFP-Dnmt1s−/− or GFP-Dnmt1o−− expression construct (4–20 ng/μl) was microinjected into the male pronuclei of mouse zygotes using an inverted microscope equipped with a micromanipulator (Leica). The injected zygotes were then cultured in M16 media for 48 h. The resulting four- or eight-cell stage embryos were collected, fixed using 4% formaldehyde and mounted with DAPI (Vector). The localization of Dnmt1 isoforms was evaluated by detecting GFP signals under a fluorescence microscope equipped with a microtome (Zeiss).

Production of recombinant GFP-Dnmt1 proteins and microinjection into mouse two-cell embryos

To produce recombinant GFP-Dnmt1 proteins, PCR was performed to obtain the GFP-Dnmt1s−/− part using the pEGFP-GFP-Dnmt1s−/− plasmid as a template under the condition of 20 cycles of 94°C/30 s, 46°C/45 s and 72°C/2.5 s (MJ gradient cycler). The primers used were 5′-GCGATATCAGTGGTAGAAGG-3′ and 5′-GGCATATCCGCGCTCGGTTG-3′. The PCR product was cloned into EcoRV (NEB) site of pET32 vector (Novagen) to make pET-32a(+) -GFP-Dnmt1s−/−. GFP-Dnmt1s−/− recombinant protein was produced after transformation into E. coli BL21 (DE3) strains and induction using 0.25 mM isopropyl-β-D-thiogalactopyranoside (IPTG; Sigma-Aldrich) for 2 or 4 h. The recombinant protein was purified using Ni2+-nitrilotriacetic acid column (Qiagen) and Sigma-Aldrich) for 2 or 4 h. The resulting four- or eight-cell embryos were harvested and subjected to Delp-seq. Genomic DNAs were isolated from the embryos by incubating them in an embryo lysis buffer (10 mM Tris–HCl, pH 8.0, 0.1 M EDTA, 0.5% SDS, 20 μg/ml RNase A and 0.4 mg/ml protease K) for 3 h at 56°C (31) and serially digested by DpnI (NEB) and FatI (NEB). The digested DNA fragments were ligated with 2.5 μM adapters specific to Dpni- or FatI-digested DNA ends (see Supplementary File S1 for adapter information) and used as a template for suppression PCR. Suppression PCR was performed using a cycling program of 18 cycles of 95°C/15 s, 45°C/1 m and 74°C/1 m (see Supplementary File S1 for primer information). Using 1 μl of suppression PCR products as templates, nested PCR was performed in a PCR program of 15 cycles of 95°C/10 s, 55°C/20 s and 72°C/1 m (see Supplementary File S1 for primer information). Each PCR was triplicated and pooled. For next-generation sequencing, adapters were removed and DNA fragments were end repaired and ligated with Illumina adapters. The NGS libraries were sequenced using HiSeq2000 (Illumina).

Dnmt1 knockdown using siRNAs

For microinjection of siRNAs, 50 μM of Dnmt1 or scrambled control siRNAs (see Supplementary File S1 for siRNA sequences) were injected into the cytoplasm of the zygote, and the embryos were cultured in M16 media until the four/eight-cell stages. The embryos at the desired stages were harvested and kept at −80°C for transcriptomic analysis. Knockdown efficiency was estimated by reverse transcriptase PCR. Pico-profiling products were diluted one hundredth, and 1 μl of the diluted product was used as a template. PCR was performed using AccuPower PCR Premix (Bioneer) along with Dnmt1 or Gapdh primers (see Supplementary File S1 for primer information) in the following PCR setting of 25 cycles of 97°C/30 s, 55°C/1 m and 72°C/1 m.
Generation of Dnmt1-depleted mouse embryos by CRISPR/Cas9 method and genotyping

A pair of single-guide RNAs (sgRNAs) was designed to target exon 32 of Dnmt1 gene and was synthesized (see Supplementary File S1 for sgRNA information; Macrogen, South Korea). Both sgRNAs (40 ng each) and Cas9 recombinant protein (40 ng; Toolgen, South Korea) were mixed and incubated in a microinjection buffer (0.1 mM EDTA, 10 mM Tris–HCl, pH 7.4) for 15 min at 37°C. Microinjection was performed under an inverted microscope equipped with a micromanipulator and a microinjector (Leica). The resulting four- and eight-cell stage embryos were treated with Tyrode’s solution (Sigma) for 2 min to remove zona pellucida, washed three times with 0.1% PVA–PBS and collected for either genotyping or MBD sequencing.

With a subset of the Dnmt1 CRISPR embryos, gene-editing efficiency was assessed. Individual embryos were placed in separate PCR tubes, and incubated in an embryo lysis buffer (20 mM Tris–HCl, pH 8.0, 0.9% Tween 20, 0.9% NP-40 and 2 mg/ml protease K) at 50°C for 30 min. For primary PCR, whole lysates were added to AccuPower PCR Premix (Bioneer, South Korea) with genotyping primers (Supplementary File S1), and subjected to PCR in 25 cycles of 95°C/30 s, 55°C/1 min and 72°C/30 s. Nested PCR was carried out with 0.1 μl of the primary PCR product as a template and internal primers in 35 cycles of 95°C/30 s, 55°C/1 min and 72°C/30 s. The products were visualized on agarose gels and sent for Sanger sequencing to detect any genomic alteration.

Pico-profiling and RNA sequencing library construction

Due to the limited starting amount of materials, we adopted the pico-profiling method as described in previous reports to enrich the transcriptomic materials (32). First, mRNAs were isolated from a pool of siDnmt1 or si-Control injected four/eight-cell embryos using Dynabeads mRNA DIRECT Kit (Invitrogen). Using the extracted mRNAs as templates, pico-profiling-specific adapter-tagged ds-cDNA fragments were generated by reacting mRNA and pico-profiling primers with SuperScript III (Invitrogen) followed by T4 DNA polymerase (NEB). Then, the adapter-tagged ds-cDNAs were amplified by 20 cycles of PCR using the anchor primers for the tagged adapters. The amplicons were digested by MlyI (NEB) to remove the pico-profiling adapters, and then the products were purified using AMPure XP beads (Beckman). Next, RNA-seq libraries were generated using the pico-profiling products according to the protocol suggested by NEBNext Ultra DNA Library Prep Kit for Illumina (NEB) and home-brew method previously reported (32). If not indicated otherwise, all materials were purchased from NEB. Two hundred nanograms of pico-profiling products were treated with T4 DNA polymerase, T4 polynucleotide kinase and Klenow enzymes to repair the ends of amplified DNA fragments. The reaction products were purified using AMPure XP beads (BeckmanCoulter) according to the manufacturer’s instructions. To adenylate 3’-ends of the end-repaired products, DNAs were reacted with Klenow fragment and 10 mM dATP for 2 h at 30°C. After purification with AMPure XP beads, NEBNext adapters with unique barcodes were ligated overnight at 16°C in the presence of diluted T4 DNA ligase. The next day, ligation mixtures were purified twice with AMPure XP beads and were enriched by PCR. Paired-end sequencing was performed using NextSeq 550 (Illumina) to generate raw sequencing data.

MBD sequencing and ChIP sequencing

Embryo sample (~300 blastomeres per sample group) was transferred to a Bioruptor microtube (Diagenode) with 20 μl lysis buffer (10 mM Tris–HCl, pH 8.0, 100 mM EDTA, pH 8.0, 0.5% SDS and 0.4 mg/ml protease K), and incubated sequentially at 56°C for 30 min and at 70°C for 15 min. The lysed chromatin were then sheared by a Bioruptor Pico sonicator (Diagenode). After 30 cycles of 30 s on and 90 s off shearing program, the sheared sample was purified using AMPure XP beads (Beckman) and then subjected to MBD-seq library production using MethylMiner Methylated DNA Enrichment Kit (Thermo) according to the manufacturer’s protocol. One microliter of streptavidin beads was precleared and incubated with 1 μl MBD-biotin proteins for 1 h at room temperature. After wash, the MBD-bead complexes were incubated with sheared DNA fragments and methyl/non-methyl spike-in DNAs overnight at 4°C. MBD-seq sensitivity and efficiency for each group were confirmed by estimating the relative amount of spike-in DNAs in elutes and flow-through. Illumina sequencing libraries were constructed using the captured DNA fragments as previously described (33), and sequencing data were generated by NextSeq 550 (Illumina).

For ChIP (chromatin immunoprecipitation) sequencing, 293T cells were plated on a six-well dish and, when they reached at 70–80% confluency, 2 μg of pEGFP-Dnmt1s, pEGFP-Dnmt1s+c or pEGFP-c2 (Clontech) plasmid was transfected into the cells using Lipofectamine LTX (Thermo) according to the manufacturer’s protocol. Cells were harvested 72 h after transfection and 1 × 10⁶ cells were used in ChIP as described previously (34). Briefly, for cross-linking of chromatin and proteins, cells were incubated in 1% formaldehyde solution for 10 min at room temperature, neutralized by glycine treatment and lysed to extract chromatin. The resulting chromatin was sheared to the size ranging from 200 to 500 bp, reacted with 2 μg α-GFP antibodies (Santa Cruz) or IgG control (Cell Signaling), and captured by Protein A magnetic beads (Bio-Rad). After reverse cross-linking and purification, the eluted DNA fragments were proceeded for Illumina sequencing library generation using NEBNext DNA Library Prep modules (NEB) and NEBNext Multiplex Oligos (NEB) according to the manufacturer’s protocol, and paired-end sequencing data were generated by NextSeq 550 (Illumina).

Bioinformatics analyses

Raw Illumina sequencing reads for all experiments were pre-processed to remove low-quality bases and adapter
sequences using 'TrimGalore' (https://www.bioinformatics.babraham.ac.uk), and aligned on the reference genome using BWA (4) for DELP-seq (mm9), MBD-seq (mm9) and ChIP-seq (hg38), while HISAT2 (mm9) (5) was used for mapping on the transcriptome.

Since we performed the suppression PCR for DELP-seq to obtain the fragments concurrently digested by DpnI/FatI, we only considered the DpnI/FatI-bound genomic loci for peak detection. Peaks were defined as DpnI/FatI loci with sizes of 100–500 bp (3,495,243 loci in mm9), and fold change >30 between Dnmt1Dam and Dnmt1 against each other. For peak detection, we first digested the reference genome (mm9) with DpnI and FatI restriction enzymes in silico, and then the fragments with sizes between 100 and 500 bp were extracted and saved as a bed file. Next, we counted the mapped reads of Dnmt1Dam or Dam embryos in the DpnI/FatI-bound loci using 'bedtools intersect', and the read enrichments were calculated by dividing the number of reads in Dnmt1Dam by that of Dam in each locus. Finally, a DpnI–FatI locus with >30-fold enrichment in Dnmt1Dam against Dam control was defined, for convenience, as a Dnmt1Dam ‘peak’, and vice versa for a Dam control ‘peak’. Genes (NCBI RefSeq) and retroelements (UCSC RepeatMasker) around peaks were annotated using 'bedtools closest' (https://bedtools.readthedocs.io/) to summarize the genomic distributions of peaks. To assess the read coverages over genes, CpG islands (CGIs), and retroelements using 'deepTools' (https://deeptools.readthedocs.io/), bed files containing peak intensities were converted to bigWig files using 'bedGraph-ToBigWig' (https://github.com/ENCODE-DCC/kentUtils) in order to input them into 'computeMatrix'. For retroelements and genes, 3 kb flanking regions from start and end positions were included in the computation, while CGIs were extended ±5 kb. Using the matrix files, heat maps and profile plots were generated using the following tools: 'plotHeatmap' and 'plotProfile'.

For analysis of RNA-seq data from s/Dnmt1 embryos, gene expression levels were calculated and normalized by HTSeq-DESeq2 pipeline (6,7). For quantification retroelement expression, 'bedtools intersect' was used to count only the uniquely mapped reads on specified retroelements (LINE1s >5 kb, ERVs >1 kb). To examine the association of Dnmt1 in coding gene expression, we obtained two publicly available transcriptome data, GSE18290 (35) and GSE66582 (36), and identified the genes possessing Dnmt1Dam or Dam peaks at promoters (TSS ±3 kb). Then, expression levels of the selected genes were profiled based on the number of peaks and presented as box plots.

For MBD-seq data analysis, genome was binned into 500-bp windows, and number of reads in each bin was counted and normalized by total read counts using 'bedtools intersect' and custom bash scripts. Additionally, we corrected the data for the MBD captured library sizes to preserve the original DNA methylation states in samples. Using the normalized data, epigenetic alteration on specified retroelements (LINE1s >5 kb, ERVs >1 kb) was analyzed between wild-type (WT) and Dnmt1-depleted embryos by 'bedtools intersect'. Using 'deepTools plotProfile', coverages of MBD-seq reads on coding genes were calculated. The normalized DNA methylation data of each group was converted to a bedGraph file to visualize them on UCSC Genome Browser.

The ChIP-seq data were quantified and normalized using HOMER (Hypergeometric Optimization of Motif EnRichment) (8). Relative coverage depths in 500-bp genomic bins over the GFP control were calculated to identify the full-length or truncated Dnmt1-enriched loci, and the number of overlapping bins between full-length or truncated Dnmt1 was counted to assess their resemblance in chromatin association. A bedGraph file for each sample and control was generated by HOMER.

Generation of all scatter/violin/box plots and statistical tests were executed in the R environment.

RESULTS

Nuclear occupancy of exogenous Dnmt1 in early mouse embryos

Dnmt1s and Dnmt1o are somatic and oocyte-specific forms of Dnmt1, respectively. We removed the C-terminal catalytic domain of Dnmt1s/o to yield truncated forms of GFP-Dnmt1s− and GFP-Dnmt1o−. These truncated forms possessed the N-terminal multidomain region that takes part in diverse regulatory functions (37) (Figure 1A). The exclusion of catalytic activity aimed to preclude potential harm due to overmethylation and to prevent other side effects that can interfere with the identification of the Dnmt1-binding regions. To prove a similarity in DNA-binding regions between the full-length GFP-Dnmt1s and the truncated GFP-Dnmt1s−, we performed ChIP analysis using α-GFP antibody and found that the genome-wide binding profile of the GFP-Dnmt1s− was very similar to that of the GFP-Dnmt1s in cultured cells (Supplementary Figure S1).

Both GFP-Dnmt1s− and GFP-Dnmt1o− showed a high level of similarity with the WT Dnmt1 in the cell cycle-dependent pattern of localization in the somatic cell nucleus (Supplementary Figure S2), agreeing with previous reports (10,38,39). To compare the locations of GFP-Dnmt1s− and GFP-Dnmt1o− proteins in mouse early embryos, we microinjected the expression plasmid into the zygote. The inspection of the resultant four- and eight-cell embryos revealed various GFP patterns that were not different between the GFP-Dnmt1s− and GFP-Dnmt1o− embryos (Figure 1B and Supplementary Figure S3). In most embryos, GFP signals were dispersed throughout the whole blastomere, including in the nucleus, while 25–50% of the embryos were marked by the nucleus-dense GFP signal (Figure 1B, d and g; also see the table in Supplementary Figure S3). The random incongruous localization of the GFP signal between the cytoplasm and the nucleus among embryos supports the temporary, cyclic localization of Dnmt1 to the nucleus in early embryos. The nuclear occupancy of exogenous Dnmt1 partially relied on the concentration of the injected vectors in that the fraction of nuclear GFP-positive embryos tended to increase with the concentration of microinjected DNA (data not shown), which agrees with the result of a previous study (3). To exclude the possibil-
Figure 1. Nuclear localization of Dnmt1 in cleavage-stage mouse embryos. (A) Structure of truncated GFP-Dnmt1s<sup>−</sup> and GFP-Dnmt1o<sup>−</sup> proteins. The carboxyl-terminal catalytic domain was removed (<sup>c</sup>−), and GFP was fused to the cDNA amino terminus of the human somatic variant of Dnmt1 (Dnmt1s) and to that of the oocyte-specific variant Dnmt1 (Dnmt1o). P, PCNA binding domain; N, nuclear localization signal; RF, replication foci targeting sequence; Zn, zinc finger-like (CXXC) motif; BAH, bromo-adjacent homology. (B) Variable expression of GFP-Dnmt1o<sup>−</sup> and GFP-Dnmt1s<sup>−</sup> among the different blastomeres of single embryos and between different embryos. As shown in the box (left), the transgene was microinjected into the male (♂) pronucleus of the mouse zygote; 2 days later, the cleaved embryos were examined for GFP expression (green) with DAPI (blue). GFP transgene alone injected embryos were used as a control. ♀, female pronucleus; pb, polar body. (C) Nuclear occupancy of recombinant Dnmt1 proteins. A 175 kDa His-tagged GFP-Dnmt1s<sup>−</sup> recombinant protein produced from <i>E. coli</i> after IPTG induction (left) was introduced into a single blastomere of each two-cell mouse embryo, as illustrated in Figure 1C. When observed 3 h later, GFP-Dnmt1s<sup>−</sup> proteins, injected into the cytoplasm, were detected in the nuclei in 50% of the microinjected embryos (Supplementary Figure S3). It demonstrated the active movement of GFP-Dnmt1s<sup>−</sup> from the cytoplasm into the nucleus without cell cycle progression (Figure 1C, c and d). The noninjected blastomere of the two-cell embryos did not exhibit GFP signal in the nuclei. Some GFP-Dnmt1s<sup>−</sup>-intense spots overlapped with DAPI-dense regions (Figure 1C, e and f), indicating that Dnmt1 is associated with specific chromosomal domains in the nucleus. In summary, these results suggest that the absence of a catalytic domain does not affect Dnmt1 localization in either the embryonic cells or the somatic cells and that the GFP-Dnmt1s<sup>−</sup> and GFP-Dnmt1o<sup>−</sup> proteins are comparable for the behaviors in the nucleus. Based on these results and the recent observations that this Dnmt1s somatic variant was present in the nucleus of early embryos (16,17), we decided to use the Dnmt1s<sup>−</sup> form in subsequent experiments.

Identification of Dnmt1-binding regions using DamID and Delp sequencing

The observation of ectopic Dnmt1 in the nucleus of early embryos prompted us to map the Dnmt1-binding regions. For this, we carried out a DamID experiment, which was used to identify the genomic targets of DNA-binding proteins (27). We constructed a vector expressing V5-Dam fused Dnmt1s<sup>−</sup> (DnmtDam; Figure 2A). The proper expression of the DnmtDam protein was observed in 293T cells (Supplementary Figure S4A). The expression mechanism of the inducible DamID system involving retinoic acid X receptor and VP16 transactivator is shown in Supplementary Figure S4B. The proper induction of DnmtDam by Ponasterone A treatment was confirmed by immunostaining (Supplementary Figure S4C) and western blotting (Supplementary Figure S4D). The Dnmt1 moiety guides the fusion proteins to the binding regions, and the Dam moiety, which does not exist in eukaryotes, subsequently methylates the surrounding adenines in the 5′-GATC-3′ sequence. The DpnI restriction enzyme specifically recognizes and cuts the Dam-methylated sequence (Supplementary Figure S4E).

We microinjected the DamID expression toolkit into mouse zygotes and cultured them in vitro until the
four/eight-cell stage (Figure 2B). The immunostaining of DamID-injected embryos using an anti-V5 antibody revealed a weak DnmtDam signal in the nuclei (Figure 2C). We used 71 DnmtDam and 61 Dam control embryos in library construction and sequencing (Figure 2D). Genomic DNA was isolated from the embryos and was subjected to DELP-seq, which began with the double digestion of genomic DNA with *DpnI* and *FatI* (5′-GATC-3′) and 3 kb to TSS) and distal promoters (from TSS/−10 kb to TSS/3 kb), gene bodies and intergenic regions. TSS and TES refer to transcription start and end sites, respectively. (G) Read coverages of DnmtDam (red) and Dam (blue) over genes and CGIs. The genes and CGIs of various lengths were scaled to fit into a fixed width surrounded by TSS/TES or S/E (start/end), respectively. For comprehensive analysis, the analytic regions were extended 5 kb from both the TSS and TES for genes and the S and E for CGIs. (H) Transcriptional profiles of genes having DnmtDam or Dam peaks at promoters. Consulting two publicly available transcriptional datasets (GSE18920 and GSE66582), the expression of genes with an increasing occurrence of DnmtDam or Dam peaks at the promoter regions (TSS/±3 kb) was profiled. The x-axes indicate the number of corresponding peaks at the promoters and the y-axes represent the expression levels of the associated genes. The number of genes with a specified number of peaks is denoted on each box plot. For statistical analysis, one-way ANOVA tests were performed.

Figure 2. Identification and characterization of the Dnmt1-binding regions in early mouse embryos by DamID and DELP sequencing. (A) Structure of Dnmt1–Dam (DnmtDam) fusion protein. V5 is an epitope for immunostaining. (B) DamID and DELP-seq (*DpnI* fragment enrichment by ligation-mediated PCR sequencing) procedures. Mouse yzogotes were injected with the DamID toolkit and were cultured until the four- and eight-cell stages. The embryos were treated with Ponasterone A for 24 h to induce the expression of DnmtDam for the methylation of nearby *DpnI* sites (5′-GATC-3′). Genomic DNA from the four/eight-cell embryo was subjected to DELP-seq. (C) Nuclear localization of DnmtDam in mouse embryos. Two- and four-cell nuclei were stained with the anti-V5 antibody (DnmtDam, green). The nuclei were stained for histone H3 lysine 9 acetylation (H3K9ac, red) and were counterstained with DAPI (blue). (D) Distribution of DnmtDam or Dam peaks for a nonspecific control. Of these 3.5 million reads, 40% are annotated as a DnmtDam ‘peak’ or a Dam ‘peak’. We detected the Dam peaks for a nonspecific control. Of these 3.5 million ages around the *Xist* X chromosome locus on the UCSC Genome Browser. However, we confess that there is a possibility that this Dnmt1-binding profile may not fully recapitulate that of full-length Dnmt1 as it was obtained from the truncated Dnmt1sc−−, even though the high similarity in the relative genomic coverage was shown between the full-length and the truncated Dnmt1 (Supplementary Figure S1).

No correlation between DnmtDam peaks at promoters and the regulation of gene expression

For ‘peak’ identification, we considered only those reads mapped to the *DpnI*- and *FatI*-bound genomic regions of length ranging from 100 to 500 bp in length and equivalent to the range of the sequencing library sizes in this study. In DELP-seq, differing from typical ChIP-seq, no conventional peak calling algorithms are applicable since genomic regions to be analyzed are predetermined by *DpnI*- *FatI*-restriction enzyme sites. For convenience, we defined a *DpnI*–*FatI*-bound locus where DnmtDam or Dam reads were predominantly enriched against each other (fold change >30) as a DnmtDam ‘peak’ or a Dam ‘peak’. We detected the Dam peaks for a nonspecific control. Of these 3.5 million
Preferential binding of DnmtDam to LINE1 and ERVK retroelements

The occurrence of the genome-wide events of epigenetic reprogramming during early development in mammals leads to the derepression of a variety of genomic retroelements and leads to genomic instability; in addition, the way that early embryos survive this instability is to quickly re-establish a set of defensive epigenetic mechanisms, including DNA methylation (21,42–44). We investigated whether Dnmt1 was involved in the regulation of retroelements in mouse early embryos. When the read coverages over various types of retroelements were plotted, the DnmtDam signal was denser in the LINE1 subfamilies (>5.0 kb in length; Figure 3A) and in IAPEz copies (>1.0 kb), the largest subfamily group in the ERVK (endogenous retrovirus type K) family (Figure 3B; for genomic copy numbers in individual subfamilies, see Supplementary Figure S6). In contrast, there was no DnmtDam signal enrichment in the ERV1 and ERVL families (>1.0 kb; Figure 3C). Based on this binding propensity of DnmtDam to certain retroelements, we attempted to make a list of the types of retroelements that DnmtDam binds by computing the read coverages of the whole genomic copies of individual retroelement subfamilies. Overall, the mean counts of DnmtDam reads over the LINE1 subfamilies were generally high and were significantly different (P = 0.016; Wilcoxon’s test) compared with those of the Dam reads (Figure 3D). However, there was no such difference (P > 0.38) among the other retroelement groups. We selected 10 subfamilies from each of the retroelement families based on their copy abundance in the mouse genome. The enrichment of DnmtDam reads was shown in all the selected LINE1 subfamilies at the P < 0.0001 level of significance (except for Lx, P < 0.01). Certain ERVK subfamilies, such as IAPEz, MMETn and RLTR10, were also shown to be predisposed by DnmtDam (P < 0.01). This result indicates that Dnmt1 binds to specific kinds of retroelements, such as LINE1 and ERVK sequences, in early mouse embryos.

Inverse relation between DnmtDam binding and the expression of LINE1 and IAPEz retroelements in mouse early embryos

We next addressed whether DnmtDam could transcriptionally repress retroelements in mouse early embryos. We microinjected a set of Dnmt1 (siDnmt1) or scrambled (Scr) siRNAs into the zygote and grew the zygotes in vitro. We obtained mRNAs from the resulting four/eight-cell embryos (n = 54 for siDnmt1 and n = 53 for Scr; Figure 4A), verified the reduction of Dnmt1 transcript in the knockdown embryos (Figure 4B) and further processed mRNA for pico-profiling (45,46), Illumina library construction and RNA sequencing. A similar number of reads were mapped at ~60% alignment efficiency. A comparison of transcriptomes using a scatter plot showed a similarity (R = 0.92) between the siDnmt1 and Scr embryos samples (Figure 4C). For reference, a total of 323 genes were upregulated (FC > 5, read count > 10), whereas 468 genes were downregulated in the siDnmt1 embryos. As for imprinted genes, 37 out of 134 genes were expressed (read count >0 in either siDnmt1 or Scr), and the number of genes with increased expression levels was twice that of the decreased (FC >2; 15 versus 8 genes) in the siDnmt1 embryos (Supplementary File S1). A variety of retroelements showed a significant increase in the expression level in siDnmt1 embryos that belong to either the LINE1 family (L1Md-A, L1Md-F, L1Md-F2, L1Md-T, L1Md-F3, L1Mus1 and L1Mus2) or the ERVK family (ETnERV2, IAPEz, MMERVK10C, MMETn and RLTR10; Figure 4D). As expected, these groups of retroelements with increased expression in the siDnmt1 embryos mostly corresponded to those preferred by Dnmt1 (see Figure 3E); this result exhibits the strong correlation between the two (Figure 4E). As shown from the direct comparison of DamID and the RNA-seq results using scatter plots for P-values, the L1Md-F2, L1Md-T, L1Md-A, L1Md-F3, L1Mus1, L1Md-F and IAPEz sequences were shown to have the highest statistical significance (Figure 4F). Meanwhile, the retroelements, including L1Md_Gf, L1Mus3, Lx and RMR16, appeared to display some disjunction between Dnmt1 binding and transcriptional repression, which may be due to their intrinsically low copy numbers and low expression levels in early embryos (Figure 4D); this renders the statistical evaluation of them quite tricky. Our result suggests that Dnmt1 specifically binds and transcriptionally represses a set of retroelements belonging to the LINE1 or ERVK family in early mouse embryos.

A profound loss of DNA methylation at retroelement sequences in Dnmt1-depleted mouse embryos

We next investigated whether Dnmt1 binding was associated with DNA methylation at the retroelement sequences. Using the CRISPR/Cas9 genome editing technology, we used the CRISPR/Cas9 genome editing technology, we...
Figure 3. Preferential binding of DnmtDam to LINE1 and ERVK retroelements in mouse embryos. (A) Enrichment of DnmtDam signals over LINE1 subfamilies. DnmtDam and Dam coverages over L1Md_T and L1_Mus1 (upper) and flanking regions (±3 kb); k-means (k = 2) clustered the DnmtDam reads over the indicated LINE1 subfamilies (lower). The elements of various lengths are scaled to fit into the fixed width surrounded by S (start) and E (end). (B) DamID signals over ERVK elements. The read coverages over IAPEz-int (upper) and the rest of the ERVK subfamilies (lower). (C) DamID signals over the ERV1 and ERVL elements. (D) Comparison of DnmtDam and Dam enrichment over the subfamilies of retroelements. Box and violin plots showing the distributions of the DnmtDam (pink) or Dam (blue) enrichment levels of individual retroelement subfamilies (black dots). Red diamonds indicate the trimmed means, and the corresponding subfamilies are linked by gray lines across DnmtDam and Dam charts for comparison. For statistical analysis, Wilcoxon signed-rank tests were performed. (E) DamID enrichment signals in each subfamily. In each of the LINE1, ERVK, ERV1 and ERVL retroelement families, the subfamilies of the top 10 highest copy numbers were selected to compare the DnmtDam (pink) and Dam (blue) enrichment over the individual genomic copies (dots). In each facet, the red dots in the middle of the violin plots indicate the trimmed means of enrichment levels, and they are linked by red lines to clarify the differences between the samples. In panels (A)–(E), only large-sized retroelement copies in the genome (>5 kb for LINE1s and >1 kb for ERVs) are included. Statistical significances were computed using Wilcoxon signed-rank tests. *P < 0.05; **P < 0.01; ***P < 0.001; ****P < 0.0001.

Figure 4. Effects of CRISPR/Cas9-mediated Dnmt1 disruption on DnmtDam enrichment of retroelements. (A) Schematic depicting the CRISPR/Cas9 mediated disruption of exon 32 of Dnmt1 gene encoding DNA methylase domain to remove its enzymatic activity (Supplementary Figure S7A). With a subset of four/eight-cell embryos obtained after microinjection of CRISPR/Cas9 toolkit, we first examined mutation frequency by PCR genotyping (47), and identified that all the CRISPR (CR) embryos had random deletions at the target regions (Supplementary Figure S7B). For MBD-seq, we used ~2 ng genomic DNA (~300 blastomeres per group) pooled from WT and CR embryos (n = 44 for WT embryos and n = 41 and 45 for CR1 and CR2 groups, respectively, as biological replicates; Supplementary Figure S7C), and performed MBD-based capture to enrich methylated DNA fragments before library construction for Illumina sequencing (Figure 5A). About 9.3 × 10^7 reads per group were mapped with ~91% alignment efficiency. The mean read counts per 500-bp genomic bin were very low in the CR groups compared to the control (Figure 5B). The histograms for the distribution of bins by FC values between the CR groups versus the control revealed a global loss of DNA methylation in the CR embryos (Figure 5C).

Visual inspection on the UCSC Genome Browser indicated that, in the WT embryos, methylated DNA fragments were largely enriched in retroelement-dense genomic stretches and usually spotted alongside DnmtDam peak clusters (Figure 5D). Those retroelements harboring DnmtDam signals were not evenly methylated; certain retroelement sequences were much less methylated or not methylated at all, depending on the size and the family they belonged to or possibly on the genomic context (see marked regions in Figure 5D). For instance, the full-length L1Md_A and L1Md_T sequences (>5.0 kb in length), LINE1 subfamilies that are considered as evolutionarily young and capable of independent mobilization (48), were relatively
strongly methylated, whereas a large fraction of full-length L1Md_E2 copies, depending on the genomic loci, were weakly methylated even though they were also closely occupied by DnmtDam signals, reflecting the notion that the recruitment of Dnmts did not necessarily lead to local methylation (49, 50) (Figure 5E). However, in the CR1 and CR2 embryos, heavy DNA methylation signals at these retroelements were, in general, markedly reduced (Figure 5D and E). IAPEz sequences, which also carried dense DnmtDam signals, were demethylated in the CR embryos but only partially (Figure 5C) as observed in Dnmt1-null mouse embryonic stem cells (51, 52), suggesting a joint participation of other Dnmt(s) in the methylation of IAPEz sequences. Figure 5F summarizes the degree of demethylation among the DnmtDam-rich retroelement families (Figure 3A and B) determined by calculating the mean read counts over the corresponding retroelement families. The results indicated that Dnmt1 was involved in the maintenance of DNA methylation at these retroelement sequences. Meanwhile, the profile of the mean DNA methylation levels of genic regions spanning between the transcription start and end sites also showed a demethylation of single-copy sequences in the CR embryos (Supplementary Figure S8A). With a special interest in imprinted genes, we looked into differentially methylated regions (DMRs) of imprinted loci. We found that both DnmtDam and methylation signals were co-localized at H19, Zrsr1, Zim3 and Kcqn1 gene DMRs (Supplementary Figure S8B) although the general weakness of both signals at unique genomic sequence leaves assignment of a detailed mechanistic study on Dnmt1 involve-
Figure 5. Dnmt1 is involved in DNA methylation of genomic retroelements in mouse early embryos. (A) Generation of Dnmt1-depleted embryos using CRISPR/Cas9 method and methyl-CpG capture for sequencing. Mouse zygotes were injected with Cas9 recombinant protein and a pair of sgRNAs targeting exon 33 of Dnmt1 gene and were cultured to the four- and eight-cell stages. Single WT embryo group and two Dnmt1 CRISPR embryo groups (CR1 and CR2) were separately subjected to methyl-CpG capture and library construction. MBD, methyl-CpG-binding protein. (B) Distribution of log$_2$-transformed normalized count in MBD-seq samples. Each gray dot represents a 500-bp genomic bin containing a number of reads normalized in counts per million. The mean count in each embryo group is shown in red dot with the actual mean values in counts per billion below. (C) Global loss of DNA methylation in CR embryos compared to the WT embryos. Histograms show the relative methylation levels in 500-bp genomic bins (x-axes) versus the frequency of bins (y-axes) in CR1 and CR2 embryos against the control (WT). Blue and red bars represent fractions of bins with methylation decrease and increase in CR embryos, respectively. (D) UCSC Genome Browser snapshots showing both DELP-seq (red) and MBD-seq signals (green for WT and deep red for CR groups) on a retroelement-dense or -sparse region chosen arbitrarily. Notably, retroelement loci with DnmtDam enrichment do not always show a dense methylation (see the loci marked with asterisk). (E) Dependence of DNA methylation levels on the type and the size of retroelement copy. UCSC Genome Browser snapshot shows only the large LINE1 (>1.0 or >5.0 kb in length as indicated) and ERV copies (>1.0 kb) extracted from the RepeatMasker (mm9) track. Arrows indicate the orientation of the retroelements. (F) Demethylation of retroelement copies in the CR1 and CR2 embryo groups. Box plots show difference in the DNA methylation levels between WT and CR embryos. Red dot represents the mean count (DNA methylation level) in each embryo group (see the explanation in B). Adjusted P-values ($P_{\text{adj}}$) are indicated.

DISCUSSION

Successful ChIP requires a good quality antibody and sufficient genomic material. If the number of cells is very limited, ChIP demands even higher quality antibody with extra specificity and sensitivity. ChIP experiments with early embryo samples are a challenge due to the above-mentioned reasons; at least tens of thousands of cells are demanded in recent low-cell-number ChIP-seq protocols (53) for typical transcription factors. We only had 500 diploid cells (blastosomes), an incomparably small scale for conventional ChIP with usual antibody. In these circumstances, DamID was a good alternative because of its enzyme-vouched specificity and fidelity in selecting global target loci free from uncertainty due to the use of different antibodies with variable performances. For your information, several studies have directly compared DamID with ChIP profiles, all of which found a strong similarity between the two methods (54–57). In this study, we identified and characterized the Dnmt1-binding genomic regions in mouse early embryos and used the DamID method as a tool for profiling the Dnmt1-binding sites. The distance between two DpnI sites is too long, the resulting DNA fragment is inadequate for sequencing. To overcome the size problem, we double digested the genome with DpnI and FatI enzymes; the latter cuts the genome 2-fold more frequently than the former. In addition, we performed suppression PCR to prevent the ligated sequences with the same DpnI or FatI adapter at both ends from being amplified and enriched those ligated sequences only with different adapters. Thus, we were able to reduce the average length of the amplicon library to 300–500 bp, which is appropriate for Illumina sequencing. This modified protocol is particularly useful for the analysis of samples, such as early mammalian...
embryos, because sequencing libraries prepared from such samples typically suffer from shallow read coverage and insufficient statistics, and this method produces an exhaustive collection of embryonic DNA libraries with high efficiency and minimal loss.

In typical somatic cells, Dnmt1 occupies replication forks to form replication machinery and continues to methylate newly synthesized DNA strands. Therefore, the Dnmt1-DamID experiment may have a high chance of producing nonspecific peaks over the whole chromosomes, which logically yields unintelligible results. In cleavage-stage mouse embryos, both global and passive modes of DNA demethylation occur, and the maintenance of DNA methylation is uncoupled from DNA replication, which greatly reduces the risk of generating nonspecific data from the Dnmt1-DamID experiment. However, the long-lasting scenario of the nuclear exclusion of early embryonic Dnmt1 is now being challenged. A fraction of Dnmt1 proteins are present in the nucleus and play a role in sustaining methylation imprints (5,13,16,17). Although Dnmt1 is nuclear, its full-time association with the replication fork during the S phase of the cell cycle is quite unlikely considering the global progressive demethylation in this period. Additionally, the nuclear Dnmt1 activity is limited and temporary based on our results (Figure 1B) and those of others (2–7). Therefore, the early embryo is a relevant and reasonable choice for Dnmt1-DamID experiments.

The transcriptomes of repetitive elements in preimplantation-stage embryos in humans (24) and mice (58) are characterized by dynamic changes over time that impart a stage-specific signature to the early embryo. The transcriptional activities of both the LINE1 and IAP retroelements of the ERVK family peaked at the two-cell stage and then continuously decreased as embryos advanced to the blastocyst stage in mice (21,58). This observation is consistent with the early study reporting that mouse oocytes and early embryos provide an environment unsuitable for the transcription of IAP and LINE1 retroelement families and provide an environment that is particularly suitable for the transcription of the MaLR and ERVL families (20). The unsuitable nuclear environment for LINE1 transcription notwithstanding, the requirement of LINE1 transcription for progression to the cleavage stage is interesting (21,59). The sharp decrease in the levels of LINE1 and IAP transcripts from the four-cell stage indicates the progressive repression of these retroelements during early development (58).

These early observations are consistent with our results of the enrichment of Dnmt1 at LINE1 and IAP retroelements in four/eight-cell embryos (Figure 3). Given the role of DNA methylation in maintaining genome stability by suppressing the transposition of retroelements (1), DNA methylation can serve as a critical mechanism of repression of retroelements in early embryos. Indeed, several Dnmts are involved in the suppression of retroelements in germ cells. In Dnmt3I-deficient germ cells, every type of retroelement is derepressed (60). IAP sequences are reactivated in Dnmt1-deficient fetal cells (61) and in Dnmt3a-deficient germ cells (62). LINE1s are demethylated in the absence of Dnmt3c in germ cells (63). In this study, we found that the Dnmt1-enriched LINE1 and IAP sequences were relatively highly methylated in WT embryos and were markedly demethylated in the Dnmt1-depleted embryos (Figure 5). Since these retroelements were the target loci of Dnmt1, and became demethylated and derepressed in the Dnmt1-depleted embryos, it indicated the involvement of Dnmt1 in the suppression of the LINE1 and IAP sequences via DNA methylation in early embryos. Meanwhile, several studies have reported a steady-state level of CpG methylation at the retroelement loci during the early developmental period (14,64,65), indicating that DNA methylation at these sequences is stably maintained against the genome-wide event of passive demethylation occurring in preimplantation-stage embryos (66,67). Thus, for now, there is little evidence for methylation change at these retroelements during this period, which leaves a question whether Dnmt1 has a role in the epigenetic regulation of retroelements that display a progressive repression from the four-cell stage on (58). To explain the potential role of Dnmt1 upon binding to retroelements, we noted that DNA methylation change occurred quite limitedly at a group of retroelement copies. For example, full-length LINE1 sequences, unlike most other 5′-truncated versions, have intact 5′-UTRs through which epigenetic regulators and transcription factors such as NANOG (24) control LINE1 expression. Consistently, DNA methylation was found highly concentrated on the large-sized LINE1 copies (Figure 5D and E). Despite the transcriptional competency of the full-length LINE1 copies, their extremely small fraction among the whole LINE1 population in the genome seldom allows methylation changes to be perceived with a global methylation assay even though DNA methylation changes at these LINE1 5′-UTRs. We indeed observed, from a separate study with mouse skeletal muscle tissues, an age-associated, size-dependent change of DNA methylation among the retroelement copies, which was otherwise not unearthed from the whole copy-based analysis (33).

Notably, the four-cell stage in mice is the time when retroelement expression begins to decline and the paternal genome of the zygote becomes densely modified with H3K9 methylation; the paternal genome, differing from the maternal genome carrying dense H3K9 methylation, stays seemingly unmethylated for H3K9 until the two-cell stage (29,68). It hints at a correlation between the establishment of global H3K9 methylation and the suppression of retroelements. In reality, in somatic and embryonic stem cells, SETDB1/Trim28/KRAB-ZNF-mediated H3K9 methylation, separately from or in combination with DNA methylation, represses certain retroelement families (69,70). Additionally, LINE1 sequences are repressed in a target-specific fashion in the experiment using the TALE-L1-KRAB construct in mouse embryos (21). Therefore, it is likely that DNA methylation cooperates with SETDB1/Trim28/KRAB-ZNF-mediated H3K9 methylation to transcriptionally regulate certain retroelement families in early mouse embryos. This may explain the uncomfortable gap between the extensive DNA demethylation at L1Md_A and L1Md_T families and their relatively small increase of transcription in the s/Dnmt1 embryos (Figures 4D and 5F).
DATA AVAILABILITY
Raw sequencing data for DELP-seq and RNA-seq were deposited under GSE111266: https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE111266
UCSC Genome Browser session link for DELP-seq and MBD-seq: http://genome.ucsc.edu/s/mbd00asis/DELPseq_MBDseq_Min_et_al_NAR

SUPPLEMENTARY DATA
Supplementary Data are available at NAR Online.

FUNDING
Korea Research Institute of Bioscience and Biotechnology Initiative Program; National Research Foundation of Korea [NRF-2019R1A2C1083916]. Funding for open access charge: Korea Research Institute of Bioscience and Biotechnology.

Conflict of interest statement. None declared.

REFERENCES
1. Messerschmidt,D.M., Knowles,B.B. and Solter,D. (2014) DNA methylation dynamics during epigenetic reprogramming in the germline and preimplantation embryos. Genes Dev., 28, 812–828.
2. Branco,M.R., Oda,M. and Reik,W. (2008) Safeguarding parental identity: Dnmt1 maintains imprints during epigenetic reprogramming in early embryogenesis. Genes Dev., 22, 1567–1571.
3. Cardoso,M.C. and Leonhardt,H. (1999) DNA methyltransferase is actively retained in the cytoplasm during early development. J. Cell Biol., 147, 25–32.
4. Carlson,L.L., Page,A.W. and Bestor,T.H. (1992) Properties and localization of DNA methyltransferase in preimplantation mouse embryos: implications for genomic imprinting. Genes Dev., 6, 2536–2541.
5. Howell,C.Y., Bestor,T.H., Ding,F., Latham,K.E., Mertinet,C., Trasler,J.M. and Chaillet,J.R. (2001) Genomic imprinting disrupted by a maternal effect mutation in the Dnmt1 gene. Cell, 104, 829–838.
6. Jeong,Y.S., Oh,K.B., Park,J.S., Kim,J.S. and Kang,Y.K. (2009) Cytoplasmic localization of oocyte-specific variant of porcine DNA methyltransferase-I during early development. Dev. Dyn., 238, 1666–1673.
7. Ratnam,S., Mertinet,C., Ding,F., Howell,C.Y., Clarke,H.J., Bestor,T.H., Chaillet,J.R. and Trasler,J.M. (2002) Dynamics of Dnmt1 methyltransferase expression and intracellular localization during oogenesis and preimplantation development. Dev. Biol., 245, 304–314.
8. Jones,P.A. and Liang,G. (2009) Rethinking how DNA methylation patterns are maintained. Nat. Rev. Genet., 10, 805–811.
9. Hermann,A., Goyal,R. and Jeltsch,A. (2004) The Dnmt1 DNA-(cytosine-5)-methyltransferase methylates DNA processively with high preference for hemimethylated target sites. J. Biol. Chem., 279, 48350–48359.
10. Leonhardt,H., Page,A.W., Weier,H.U. and Bestor,T.H. (1992) A targeting sequence directs DNA methyltransferase to sites of DNA replication in mammalian nuclei. Cell, 71, 855–873.
11. Vertino,P.M., Sekowski,J.A., Coll,J.M., Applegren,N., Han,S., Hickey,R.J. and Malkas,L.H. (2002) Dnmt1 is a component of a multiprotein DNA replication complex. Cell Cycle, 1, 416–423.
12. Li,E., Beard,C. and Jaenisch,R. (1993) Role for DNA methylation in genomic imprinting [see comments]. Nature, 366, 362–365.
13. Hiratsuka,R., Chiba,H., Kaneda,M., Tajima,S., Li,E., Minish,R. and Sasaki,H. (2008) Maternal and zygotic Dnmt1 are necessary and sufficient for the maintenance of DNA methylation imprints during preimplantation development. Genes Dev., 22, 1607–1616.
