The *H19* Imprinting Control Region Mediates Preimplantation Imprinted Methylation of Nearby Sequences in Yeast Artificial Chromosome Transgenic Mice

Eiichi Okamura,^a Hitomi Matsuzaki,^b Ryuuta Sakaguchi,^a Takuya Takahashi,^a Akiyoshi Fukamizu,^{b,c} Keiji Tanimoto^{b,c}

Graduate School of Life and Environmental Sciences, University of Tsukuba, Tsukuba, Ibaraki, Japan^a; Faculty of Life and Environmental Sciences, University of Tsukuba, Tsukuba, Ibaraki, Japan^b; Life Science Center of Tsukuba Advanced Research Alliance, University of Tsukuba, Tsukuba, Ibaraki, Japan^c

In the mouse *Igf2/H19* imprinted locus, differential methylation of the imprinting control region (*H19* ICR) is established during spermatogenesis and is maintained in offspring throughout development. Previously, however, we observed that the paternal *H19* ICR, when analyzed in yeast artificial chromosome transgenic mice (YAC-TgM), was preferentially methylated only after fertilization. To identify the DNA sequences that confer methylation imprinting, we divided the *H19* ICR into two fragments (1.7 and 1.2 kb), ligated them to both ends of a λ DNA fragment into which CTCF binding sites had been inserted, and analyzed this in YAC-TgM. The maternally inherited λ sequence, normally methylated after implantation in the absence of *H19* ICR sequences, became hypomethylated, demonstrating protective activity against methylation within the ICR. Meanwhile, the paternally inherited λ sequence was hypermethylated before implantation only when a 1.7-kb fragment was ligated. Consistently, when two subfragments of the *H19* ICR were individually investigated for their activities in YAC-TgM, only the 1.7-kb fragment was capable of introducing paternal allele-specific DNA methylation. These results show that postfertilization methylation imprinting is conferred by a paternal allele-specific methylation activity present in a 1.7-kb DNA fragment of the *H19* ICR, while maternal allele-specific activities protect the allele from *de novo* DNA methylation.

The mouse insulin-like growth factor 2 (*Igf2*) and noncoding *H19* genes, located about 90 kb apart on distal chromosome 7 (see Fig. 1A), are preferentially expressed from the paternal and maternal alleles, respectively. Monoallelic expression of these genes requires a differentially methylated region (DMR; also referred to as the imprinting control region, *H19* ICR) located from 2 to 4 kb 5' to the *H19* transcription start site (1, 2). The *H19* ICR is methylated in sperm but not in oocytes, and this allele-specific methylation pattern is maintained after fertilization (3). The CTCF insulator protein binds to the hypomethylated maternal *H19* ICR in somatic cells and blocks *Igf2* promoter activation by distant downstream enhancers. In contrast, the methylation of the paternal *H19* ICR silences *H19* gene transcription by inducing epigenetic changes at the *H19* gene promoter, while preventing CTCF binding to the *H19* ICR which then allows *Igf2* gene activation (4, 5). Deletion of the DNA methyltransferase (*Dnmt*) gene (6–8) or the *H19* ICR sequences from the mouse genome (1, 2) disrupts genomic imprinting at the *Igf2/H19* locus. For example, paternal deletion of the *H19* ICR from the mouse genome causes biallelic expression of the *H19* gene and fetal growth retardation, while its maternal deletion results in biallelic *Igf2* expression and fetal overgrowth.

DNA methylation and histone modifications are erased and reestablished during gametogenesis (9). Because some DMRs (*H19*, *Dlk1-Dio3*, *Rasgrf1*, and *Zdbf2*) acquire DNA methylation in the sperm and many others acquire DNA methylation in oocytes, and because they both are catalyzed by a common set of enzymes, *Dnmt3a*, *-3b*, and *-3L* (6–8, 10), DMRs like the *H19* ICR must harbor *cis* sequences to instruct their methylation only in the sperm, or their protection against *de novo* DNA methylation only in the eggs, although these mechanisms are not mutually exclusive. Despite intensive efforts to identify such sequences, none have ever been discovered (11–16).

Parental genomes undergo global epigenetic reprogramming during early embryogenesis. Immediately after fertilization, both paternal and maternal genomes are subject to genome-wide demethylation (17, 18). Along with this epigenetic change, hypermethylated CpGs flanking the paternal *H19* ICR become demethylated by the blastocyst stage (19). Thereafter, starting from around the blastocyst stage, both parental chromosomes are subject to *de novo* DNA methylation (20). Accordingly, hypomethylated CpGs flanking the *H19* ICR on both alleles become highly methylated in 12.5-day-postcoitum (dpc) embryos (19). Importantly, however, differential parental methylation of the *H19* ICR itself is maintained throughout embryogenesis. Therefore, it is postulated that mechanisms exist that allow the *H19* ICR to resist genome-wide reprogramming and to maintain differential methylation. While it is well established that CTCF binding to the maternal *H19* ICR is essential to maintain its hypomethylated status after implantation (21, 22), a molecular mechanism explaining how the paternal *H19* ICR is protected from genome-wide DNA demethylation after fertilization remains obscure. Furthermore, it is not known if the CTCF sites are sufficient to protect the fragment from *de novo* DNA methylation.

In order to reveal which *cis* DNA sequences are both required and sufficient for methylation imprinting at the *H19* ICR, we in-

Received 26 July 2012 Returned for modification 23 August 2012

Accepted 6 December 2012

Published ahead of print 10 December 2012

Address correspondence to Keiji Tanimoto, keiji@tara.tsukuba.ac.jp.

E.O. and H.M. contributed equally to this work.

Copyright © 2013, American Society for Microbiology. All Rights Reserved.

doi:10.1128/MCB.01003-12

serted a 2.9-kb DNA fragment encompassing the whole ICR into a 150-kb human β -globin yeast artificial chromosome (YAC), so that the transgene could be protected from the mouse genome environment, and then generated transgenic mice (TgM) (23). In these mice, the paternally inherited transgenic *H19* ICR was more heavily methylated than when it was maternally inherited, demonstrating that the fragment carries sufficient information to establish parent-of-origin differential methylation. Surprisingly, however, the transgenic *H19* ICR was not methylated in sperm. Essentially, the same paternal allele-specific methylation was also established postfertilization when a 2.9-kb *H19* ICR fragment was randomly inserted into the mouse genome (24) or a 2.4-kb *H19* ICR fragment was knocked in at the *alpha-fetoprotein* or *CD3* locus (25, 26). These results suggested that the *H19* ICR was primarily marked by an epigenetic modification other than DNA methylation in the germ cells and that, after fertilization, the mark would be translated into a differential methylation pattern. Meanwhile, methylation acquisition in sperm probably requires sequences lying beyond the 2.9-kb fragment, since the *H19* ICR became methylated in sperm when assayed in the context of 147-kb (*H19* gene locus) bacterial artificial chromosome (BAC) TgM (25).

In this study, as a first step to attempt to further define molecular mechanisms that lead to establishment of methylation imprinting, we set up experiments to examine whether (i) the *H19* ICR carries protective activity against methylation, (ii) the *H19* ICR harbors DNA sequences that instruct the initiation of DNA methylation, and (iii) the 2.9-kb *H19* ICR fragment can be further truncated without losing activity that is required to establish methylation imprinting after fertilization. Toward this end, a 2.3-kb λ DNA sequence, supplemented with four CTCF binding sites, was inserted into the middle of the *H19* ICR fragment, and this chimeric fragment was used to generate YAC-TgM; the λ DNA became differentially methylated depending on its parental origin. After the removal of *H19* ICR sequences by *in vivo* cre-loxP recombination, the λ DNA fragment alone was highly methylated irrespective of whether it was paternally or maternally inherited. Next, the *H19* ICR fragment was divided into two fragments, ICR43 (1.2 kb) and ICR21 (1.7 kb), and independent YAC-TgM lines carrying either of these two fragments were generated. DNA methylation analysis of the mice revealed that fragment ICR21, but not ICR43, harbored an activity permissive for paternal allele-specific DNA methylation after fertilization. These results strongly suggest that the *H19* ICR carries both allele-specific methylation and antimethylation protective activities in the paternal and maternal alleles, respectively.

MATERIALS AND METHODS

Generation of the λ +CTCF fragment. To facilitate plasmid construction, the following two oligonucleotides were phosphorylated, annealed, and ligated with SacI/KpnI-digested pBluescriptII/KS(+) vector, in which two BssHII sites were prospectively disrupted: 5'-CGGGATCCTA GGATCCGAGCT-3' and 3'-CATGGCCCTAGGATCCTAGGGC-5'. The resultant vector, pBSIIKS(+)-KBABS, carried KpnI-BamHI-AvrII-BamHI-SacI multicloning sites.

Five DNA fragments (L1 through L5) were PCR generated by using the primer sets indicated below and a 2,328-bp HindIII fragment from bacteriophage λ DNA as a template. Artificially introduced restriction enzyme sites and CTCF recognition motifs (*m1* to *m4*) with their flanking 4 nucleotides (nt) on both sides are underlined and italicized, respectively; the restriction enzymes and CTCF recognition motifs are

indicated in parentheses. The primers were as follows: L1, 5'-GCACC TAGGCGCCAAGCTTTGTGTGCCACCCA-3' (AvrII-KasI-HindIII) and 5'-TTCCTCGAGATGCCGCGTGGTGGCAGTACAATAACAGACA TTCCTAC-3' (XhoI-*m4*); L2, 5'-TAACCTCGAGGAAAAATGGCTAC GAAGT-3' (XhoI) and 5'-ACAGGCGCGCTACCGCGCGGTGGCAGCA TATAGTCATTCCAACCATCT-3' (BssHII-*m3*); L3, 5'-TTTGCGCGCC TGTCTTTGATGAGCATG-3' (BssHII) and 5'-GTATGTACAACAATT GCATGTCCAGAG-3' (BsrGI); L4, 5'-ATGCGCGCGCTGTACATTCA CTGCCGCGTGGCGCAACGTCGTTGGCGAAAAGTCATTAG-3' (BssHII-BsrGI-*m2*) and 5'-AATACGCGTGGCGATAATATTTATTG C-3' (MluI); and L5, 5'-TATTCCTCGAGACGCGTTTGTGCTGCCACCAC GCGGCAACTAGGTGTTTTAACTCGTG-3' (XhoI-MluI-*m1*) and 5'-G AGCCTAGGCGTACGAAGCTTTTCTAATTTAAACC-3' (AvrII-BsiWI-HindIII).

The AvrII/XhoI-digested L1 and XhoI/AvrII-digested L5 fragments were simultaneously ligated with the AvrII-digested pBSIIKS(+)-KBABS vector to generate pKBABS/L1+L5. Then, XhoI/BssHII-digested L2 and BssHII/MluI-digested L4 fragments were simultaneously ligated with the XhoI/MluI-digested pKBABS/L1+L5 vector to make pKBABS/L1+L2+L4+L5. Finally, the BssHII/BsrGI-digested L3 fragment was ligated with the BssHII/BsrGI-digested pKBABS/L1+L2+L4+L5 vector to make p λ +CTCF. In each cloning step, the correctness of DNA sequences was confirmed by DNA sequencing (DDBJ sequence accession no. AB691538).

Targeting vectors for homologous recombination in yeast. A backbone plasmid for the targeting construct, pHS1 (27), carried a human β -globin HS1 DNA fragment (from nt 13299 to 14250; HUMHBB; GenBank). The following double-stranded DNA (only the upper-strand sequences are shown) was subcloned into the HindIII site (at nt 13769 in HUMHBB) of pHS1 to generate pHS1/loxP5171-2272: 5'-AAGCTTATAACTTTCGTATAGTACACATTATACGAAGTTATGGATC CTAGGATCCATAAAGCTTTCGTATAGGATACTTTATACGAAGTTATA AGCTT-3' (HindIII-*loxP5171*-BamHI-AvrII-BamHI-*loxP2272*-HindIII; restriction enzyme sites are underlined, and *loxP* sequences are italicized). The 2.9-kb mouse ICR fragment (from nt 833 to 3696; AF049091 [23]) was subcloned into BamHI-digested pHS1/loxP5171-2272 in reverse orientation to generate pHS1/loxP5171-2272/ICR(-).

The pICR43(λ +CTCF(-))21 targeting vector was generated as follows. The following double-stranded DNA (only the upper-strand sequences are shown) was subcloned into the BspEI site (at nt 2519 in AF049091) of pHS1/loxP5171-2272/ICR(-) to generate pHS1/loxP5171-5171-2272-2272/ICR(-): 5'-TCCGGAAGCTTATAAAGCTTTCGTATAGTACACATTATACG AAGTTATGGATCCTAGGATCCATAAAGCTTTCGTATAGGATACTTTAT ACGAAGTTATAAGCTTCCGGA-3' (BspEI-HindIII-*loxP5171*-BamHI-AvrII-BamHI-*loxP2272*-HindIII-BspEI; restriction enzyme sites are underlined, and *loxP* sequences are italicized). Then, the AvrII DNA fragment excised from the p λ +CTCF was introduced into the AvrII site (between the *loxP5171* and *loxP2272* sequences) of pHS1/loxP5171-5171-2272-2272/ICR(-) in reverse orientation to generate the pICR43(λ +CTCF(-))21 targeting vector.

To generate the pICR43-21 targeting vector, the following double-stranded DNA (only the upper-strand sequences are shown) was subcloned into the blunt-ended BspEI site (at nt 2519 in AF049091) of pHS1/loxP5171-2272/ICR(-): 5'-AAATAAGCTTATAAAGCTTTCGTATA GGATACTTTATACGAAGTTATCCCGGGGATATCATAAAGCTTTCGTAT AGTACACATTATACGAAGTTATGTAACTAC-3' (HindIII-*loxP2272*-SmaI-EcoRV-*loxP5171*-HpaI; restriction enzyme sites are underlined, and *loxP* sequences are italicized).

In each cloning step, the correctness of DNA sequences was confirmed by DNA sequencing.

The targeting plasmid DNAs were linearized by digestion with SpeI (at nucleotide position 13670 in HUMHBB) and used for human β -globin YAC mutagenesis by homologous recombination.

TgM. Generation and structural analysis of human β -globin YAC-TgM have been described elsewhere (28). TgM ubiquitously expressing cre recombinase (29) were mated with ICR43(λ +CTCF)ICR21 or

ICR43-21 YAC-TgM to generate ICR43(λ +CTCF), (λ +CTCF)ICR21, and λ +CTCF TgM or ICR43 and ICR21 TgM, respectively. Successful cre-*loxP* recombination was confirmed by Southern blotting. For DNA methylation analysis, male or female TgM were crossed with female or male wild-type CD1 mice (ICR; Charles River Laboratories Japan, Yokohama, Japan), respectively, to obtain Tg offspring that inherited the transgene either paternally or maternally. Animal experiments were carried out in a humane manner and were approved by the Institutional Animal Experiment Committee of the University of Tsukuba. Experiments were conducted in accordance with the Regulation of Animal Experiments of the University of Tsukuba and the Fundamental Guidelines for the Proper Conduct of Animal Experiments and Related Activities in Academic Research Institutions under the jurisdiction of the Ministry of Education, Culture, Sports, Science and Technology of Japan.

Bisulfite sequencing. Genomic DNA from tail tips or whole testis was digested with XbaI and then treated with sodium bisulfite using the EZ DNA methylation kit according to the manufacturer's instructions (Zymo Research, Orange, CA). Unless otherwise noted in the figure legends, representative samples determined by Southern blotting methylation analyses were pooled and analyzed. Blastocysts obtained by natural mating were flushed from the uteri with M2 medium at 3.5 dpc and washed with M2 medium followed by phosphate-buffered saline (PBS). Oocytes were collected from the oviducts of superovulated hemizygous female TgM. The cumulus cells were removed by treatment with hyaluronidase and then washed repeatedly with M2 medium and PBS. Four to 14 blastocysts or 21 to 38 oocytes were embedded in agarose beads and treated with sodium bisulfite as described previously (24).

Transgene-specific nested PCR, cloning, and sequencing analysis were performed as described previously (23, 24). Subregions of the *H19* ICR (region I, nucleotides 1245 to 1995; region II, nucleotides 2845 to 3432; region II', nucleotides 3050 to 3432; GenBank accession no. [AF049091](#)) and λ +CTCF (region III, nucleotides 161 to 851; region IV, nucleotides 1804 to 2295; DDBJ accession no. [AB691538](#)) were amplified by nested PCR, primer sequences for which are listed below.

Potential PCR bias in bisulfite sequencing was assessed for sets of primers and corresponding target sequences, i.e., the methylated and unmethylated DNAs were mixed in defined proportions, bisulfite treated, and PCR amplified, and their nucleotide sequences after cloning were determined by sequencing (data not shown). Because recovery frequencies of methylated DNA were close to the expected value (50%), we concluded that a bisulfite sequencing bias in our experiment was negligible and did not affect the interpretation of results.

Southern blotting. Genomic DNA was prepared from tail tips of approximately 1- to 2-week-old TgM. For the DNA methylation analysis, the transgene sequence was first liberated by restriction enzymes and then subjected to methylation-sensitive enzymes, i.e., DNA from the ICR43(λ +CTCF)21 and λ +CTCF TgM was digested with BamHI and BstUI. After separation using agarose gel electrophoresis and Southern blot transfer to nylon membranes, the blots were hybridized with an α -³²P-labeled λ probe and subjected to X-ray film autoradiography.

Chromatin immunoprecipitation (ChIP) assay. The λ +CTCF and ICR43(λ +CTCF)21 YAC-TgM (2 to 4 months old) that inherited the transgene maternally were made anemic by phenylhydrazine treatment. Nucleated erythroid cells were collected from their spleens and fixed in PBS with 1% formaldehyde for 10 min at room temperature. Nuclei (3×10^7 cells) were digested with 0.01 unit/ μ l of micrococcal nuclease (MNase) at 37°C for 10 min to prepare primarily mononucleosome-sized chromatin. The chromatin was incubated with anti-CTCF antibody (06-917; Upstate Biotechnology, Lake Placid, NY) or normal rabbit IgG (sc-2027; Santa Cruz Biotechnology, Santa Cruz, CA) overnight at 4°C and was precipitated with preblocked Dynabeads protein G magnetic beads (Life Technologies, Carlsbad, CA). Immunoprecipitated materials were then washed extensively and reverse cross-linked. DNA was purified with the QIAquick PCR purification kit (Qiagen, Venlo, The Netherlands) and subjected to quantitative real-time PCR (qPCR) analysis. The qPCR was

performed with the Thermal Cycler Dice (TaKaRa Bio, Otsu, Japan) using SYBR Premix EX TaqII (TaKaRa Bio) with a cycling program as follows: an initial cycle for 10 s at 95°C, followed by 40 cycles of 5 s at 95°C and 30 s at 60°C. The *CD45* intron 6 (30) and the *Necdin* (31) regions were analyzed as negative controls. The *H19* ICR sequences (endogenous and transgenic loci) were analyzed as a positive control.

Primer sequences for nested and ChIP-quantitative PCRs. The first-round nested PCR primers are as follows: ICR-MA-5S1 (5'-GAATTGAGGATTATGTTTAGTGG-3') and BGLB-MA-3A1 (5'-TCTCGTCAAC CACCTTCATTAAC-3') for region I; ICR-MA-5S1 (5'-TATAGATGTT TTAGTTTAAATAAG-3') and ICR-MA-3A15 (5'-ACCAACCAATATAA CTCACATAAA-3') for region II; ICR-MA-5S1 and ICR-MA-3A6 (5'-A TATACACCTCTAAAATAATTCCC-3') for region II'; lambda-MA-3A4 (5'-ATCGTAAATAAATAACTAACCTAT-3') and lambda-MA-5S2 (5'-TTGTATGAGAGTAGATAGAAAAAG-3') for region III; lambda-MA-3A2 (5'-ATACCTTATTTTTTCTACTACAA-3') and lambda-MA-5S4 (5'-TTAAGTTTTGTGTGTTATTTATTA-3') for region IV.

The second-round nested PCR primers are as follows: ICR-MA-5S2 (5'-TTAAGGATTAGTATGAATTTTTGG-3') and ICR-MA-3A1 (5'-AA CATAACAATACTATAACCATAC-3') for region I; ICR-MA-5S4 (5'-G AATTTGGGGTATTAAAGTTTG-3') and ICR-MA-3A14 (5'-AAAA CATAAAAACATTATATACA-3') for region II; ICR-MA-5S4 and ICR-MA-3A5 (5'-AACCTAACTCATTCCTACACAAC-3') for region II'; lambda-MA-3A1 (5'-CTAACATTTATCTACATCATACCT-3') and lambda-MA-5S3 (5'-ATGGAATTTATTAGTTTTATTTTT-3') for region III; lambda-MA-3A3 (5'-CTAACTCCAACATATAATAACCC-3') and lambda-MA-5S1 (5'-ATTAGTAAGAAGATAGTAGTGATG-3') for region IV.

The ChIP-qPCR primers are as follows: CD45-In-6-5S (5'-CTTCTC TGACTCCAGATACTATTTTC-3') and CD45-In-6-3A (5'-ATAATAGT AAAATCACATTAATCCCCCAA-3') for *CD45* intron 6; *Necdin*-5S (5'-CATCGGTCTGCTCTGATCCGAA-3') and *Necdin*-3A (5'-CGCTGT CCTGCATCTCACAGTC-3') for *Necdin*; ICR-SEQ-5S15 (5'-CCTGAG GTACTGAACCTGGGTGA-3') and H19-ChIP-FWD (5'-GGGTTTATA CGCGGGAGTTG-3') for *H19* ICR; LC-ChIP-5S1 (5'-CGATGCACGCA ATGGTGTAG-3') and LC-ChIP-3A1 (5'-CAGCAAAACGCGTGCGGA TAA-3') for λ +CTCF.

Nucleotide sequence accession number. The λ +CTCF sequence (2,328 bp) was submitted to DDBJ and assigned accession no. [AB691538](#).

RESULTS

Generation of YAC-TgM bearing chimeric *H19* ICR sequences.

In our previous work, we inserted a bacteriophage λ DNA (HindIII-digested 2.3-kb) fragment into a human β -globin YAC at a position 3' to the locus control region (LCR) and generated YAC-TgM (32). As shown in Fig. 1A, this fragment bears a similar number and density of CpG motifs as does the 2.9-kb *H19* ICR fragment. In these TgM, the λ DNA sequences were substantially methylated in somatic cells regardless of their parental origin (data not shown), indicating that this fragment bears no allele-discriminating activity. To determine whether the *H19* ICR fragment possesses activities contributing to the DMR formation, we divided the fragment into two segments (at a BspEI site, Fig. 1A), each bearing either CTCF sites 4 and 3 (termed the ICR43 fragment; 1.2 kb) or sites 2 and 1 (ICR21; 1.7 kb). These subfragments were floxed by a set of *loxP5171* or *loxP2272* sites (33) and linked to either end of the 2.3-kb λ DNA (Fig. 1B, top). It is established that in the absence of CTCF binding sites the maternally inherited *H19* ICR fragment cannot maintain its hypomethylated state in somatic cells and becomes hypermethylated after implantation, both in the endogenous locus (22) and in transgenic contexts (34). Therefore, we speculated that an unmodified λ DNA fragment would be hypermethylated eventually, even if it were once hypom-

FIG 1 Generation and structural analysis of YAC-TgM. (A) (Top) Genomic structure of the *Igf2/H19* locus. The *Igf2* and *H19* genes (open boxes) are approximately 90 kb apart, and expression of both genes depends on the shared 3' enhancer in endodermal tissues (shaded box) (58). The *H19* ICR, located approximately at kb -2 to -4 relative to the transcription initiation site of the *H19* gene, is contained within a 2.9-kb *SacI* (Sa)-*BamHI* (B) fragment. (Bottom) Schematic representation of the *H19* ICR, the λ +CTCF, and the λ DNA fragments. The solid rectangles indicate the positions of the CTCF binding sites. The numbers of CpG sites in these fragments are shown on the right. (B) Schematic representation of the YAC transgenes. The positions of the β -like globin genes (open boxes) are shown relative to the LCR (gray box). *SfiI* restriction enzyme sites are located 5' to the LCR, within the LCR, and in the right arm of the YAC. Probes (solid rectangles) used for long-range structural analyses shown in panels C and D and expected restriction enzyme fragments with their sizes are shown. The enlarged map shows the detailed structure of the ICR43(λ +CTCF)21 and ICR43-21 fragments inserted between the LCR and the ϵ -globin gene. The λ +CTCF fragment is inserted at the *BspEI* site of the *H19* ICR fragment, and the positions of *loxP5171* and *loxP2272* are indicated as solid and open triangles, respectively. (C and D) Long-range structural analyses of the transgenes in the ICR43(λ +CTCF)21 (C) and ICR43-21 (D) YAC-TgM. DNA from thymus cells was digested with *SfiI* in agarose plugs and separated by pulsed-field gel electrophoresis, and Southern blots were hybridized separately to probes in panel B. Lines in panel C indicate that lanes were run on the same gel but were noncontiguous.

ethylated by the aid of *H19* ICR sequences. To avoid this situation, we supplemented the λ DNA sequences with four CTCF binding sites (termed λ +CTCF, Fig. 1A), so that the hypomethylated DNA status would be maintained. The chimeric fragment ICR43(λ +CTCF)21 was inserted into a human β -globin YAC, and TgM were generated (Fig. 1B). Two single-copy TgM lines (lines 2619 and 2653) were established, and the integrity of the YAC transgenes was confirmed by long-range Southern blot analysis (Fig. 1B and C). Then, the ICR43(λ +CTCF)21 TgM were crossed with cre-TgM to allow excision of the ICR43 and/or the ICR21 sequences. Offspring were analyzed for homologous recombination by Southern blotting (data not shown). After the cre transgene was removed by mating with non-Tg animals, the fixed genotypes of the ICR43(λ +CTCF) (see Fig. 6A), (λ +CTCF)ICR21 (see Fig. 6C), and λ +CTCF (see Fig. 3D) animals were confirmed by Southern blotting (data not shown).

DNA methylation analysis of the ICR43(λ +CTCF)21 fragment in YAC-TgM. We first examined the methylation status of the ICR43 and ICR21 regions in somatic cells of ICR43(λ +CTCF)21 YAC-TgM. Bisulfite sequencing (Fig. 2A to C) analysis revealed that both the ICR43 and ICR21 portions (regions II' and I, respectively) of the transgene were more heavily methylated when paternally inherited than when maternally inherited (P [region I] < 0.001 and < 0.001 or P [region II'] < 0.001 and < 0.001 for lines 2619 and 2653, respectively, according to the Mann-Whitney U test [http://quma.cdb.riken.jp/]). In the testis, these transgenic *H19* ICR sequences

were hardly methylated (Fig. 2D), while the endogenous *H19* ICR was heavily methylated (data not shown), indicating that paternal allele-specific methylation of these transgene fragments was established after fertilization. These phenotypes were essentially the same as those observed with the wild-type *H19* ICR sequences in our previous work (23). Therefore, the *H19* ICR could acquire differential methylation, even when it was split into two by inserting the λ +CTCF fragment into the middle.

Next, we examined the methylation status of the λ +CTCF portion of the ICR43(λ +CTCF)21 sequences in somatic cells of TgM. Bisulfite sequencing (Fig. 2A to C) analysis revealed that the λ +CTCF sequences (regions III and IV) were heavily methylated in somatic cells after paternal transmission (Fig. 2C), while they were only poorly methylated when maternally inherited (Fig. 2B) (P [region III] < 0.001 and < 0.001 or P [region IV] < 0.001 and < 0.001 for lines 2619 and 2653, respectively). Because the same sequences were only moderately methylated in the testis (Fig. 2D), the paternally inherited λ +CTCF region must have acquired most of its DNA methylation after fertilization. Samples were also subjected to Southern blot analysis, and consistent results were obtained (Fig. 3A to C).

DNA methylation analysis of the λ +CTCF fragment in YAC-TgM. To ask whether differential methylation observed in the λ +CTCF portion of the ICR43(λ +CTCF)21 transgene was in fact elicited by the flanking *H19* ICR sequences, we excised both ICR43 and ICR21 sequences from the ICR43(λ +CTCF)21 transgene by

A ICR43(λ +CTCF)21 transgene

B Maternal allele in somatic cells

C Paternal allele in somatic cells

D Testis cells

E Oocytes

in vivo cre-loxP recombination and generated λ +CTCF YAC-TgM (Fig. 3D and data not shown). In doing it this way, the activities of the distinct constructs could be compared at the same chromosomal integration site.

First, we examined the methylation status of the maternally inherited λ +CTCF sequences in somatic cells of the λ +CTCF TgM. Southern blot analysis (Fig. 3E) revealed that all 45 samples (22 and 23 from lines 2619 and 2653, respectively) exhibited partial methylation, indicating that addition of four CTCF sites alone to the λ DNA is not sufficient to fully protect the fragment from *de novo* DNA methylation. This methylation level in the λ +CTCF TgM was apparently higher than that of the maternally inherited λ +CTCF sequences in the ICR43(λ +CTCF)21 TgM (Fig. 3B), and bisulfite sequencing of regions III and IV of the λ +CTCF sequences (Fig. 2B and 4A and B) mostly confirmed the Southern blot results (P [region III] < 0.001 and < 0.005 or P [region IV] < 0.001 and = 0.212 for lines 2619 and 2653, respectively). Thus, by placing subfragments of the *H19* ICR on both sides, the maternally inherited λ +CTCF sequence became protected from methylation in somatic cells of the TgM. Bisulfite sequencing analysis of the λ +CTCF sequences in MII oocytes revealed that CpG sites were not methylated in either ICR43(λ +CTCF)21 (Fig. 2E) or λ +CTCF (Fig. 4E) TgM. Therefore, at least after fertilization, maternally inherited transgenic *H19* ICR carries an activity to protect bordering λ +CTCF sequences from *de novo* DNA methylation.

To examine if the inserted CTCF binding sites in the λ DNA fragment play any role in the protection against DNA methylation mechanism, we conducted chromatin immunoprecipitation (ChIP) assays with somatic cells of λ +CTCF and ICR43(λ +CTCF)21 YAC-TgM that inherited the transgene maternally (Fig. 5). The results clearly showed that CTCF was capable of binding to the λ DNA in the transgenic sequence context. Because CTCF was more enriched at the λ fragment in the ICR43(λ +CTCF)21 than in the λ +CTCF TgM (Fig. 5B), flanking *H19* ICR sequences may help to recruit the CTCF to the neighboring λ DNA fragment. It is also possible that the difference in CTCF binding between these two constructs may be a consequence, rather than a cause, of methylation levels of the fragment.

Next, we examined the methylation status of the paternally inherited λ +CTCF sequences in somatic cells of the λ +CTCF YAC-TgM. Southern blot analysis (Fig. 3F) revealed that, out of 60 paternal-inheritance TgM (27 and 33 from lines 2619 and 2653, respectively), 28 animals (47%) exhibited partial methylation while 32 (53%) showed hypermethylation. Bisulfite sequencing analysis of pooled DNA samples from each category of methylation pattern confirmed the Southern blot results (Fig. 4C). As indicated above, the λ +CTCF portion of the paternally inherited ICR43(λ +CTCF)21 transgene in somatic cells exhibited hypermethylation without exception (19 TgM) (Fig. 3C and data not shown). Therefore, by placing subfragments of the *H19* ICR on

both sides, paternally inherited λ +CTCF sequences became more highly methylated (higher penetration rate) in the somatic cells of YAC-TgM.

To further decipher the mode of differential methylation acquisition at the λ +CTCF sequence in greater detail, its methylation status in the testis of the λ +CTCF TgM was examined (Fig. 4D) and compared with that of the ICR43(λ +CTCF)21 TgM (Fig. 2D). Bisulfite sequencing analysis revealed that region IV (carrying the inserted CTCF site 4) of the λ +CTCF sequences was hardly methylated in either line of TgM, indicating that addition of the *H19* ICR sequences to the λ +CTCF fragment had no apparent effect on its methylation status in the testis, i.e., the methylation acquisition at region IV was enhanced after fertilization by the ICR43 and ICR21 sequences in the ICR43(λ +CTCF)21 transgene (Fig. 2C and 4C). On the other hand, methylation acquisition in region III (with the CTCF sites 1 and 2) of the ICR43(λ +CTCF)21 transgene had already been initiated in the testis (44 to 88%) (Fig. 2D), although the methylation level was still lower than that in somatic cells (96 to 97%) (Fig. 2C). Because the same sequences were poorly methylated in the λ +CTCF transgene at the same chromosomal sites (4.3 to 32%) (Fig. 4D), it is likely that the *H19* ICR sequences somehow instructed methylation acquisition in region III in the testis of the ICR43(λ +CTCF)21 TgM. Considering that the *H19* ICR portion of the transgene itself was not methylated in the testis of the ICR43(λ +CTCF)21 TgM (Fig. 2D), we speculate that a change in the surrounding environment of the λ +CTCF region (e.g., increased number of the CpG sites) rather than methylation imprinting activity of the *H19* ICR sequences generated this phenotype.

Collectively, these results demonstrate that the *H19* ICR carries both allele-specific methylation and antimethylation protective activities after fertilization and that these activities coordinately establish the differential methylation pattern in the ICR43(λ +CTCF)21 TgM.

DNA methylation analyses of the ICR43(λ +CTCF) and the (λ +CTCF)ICR21 fragments in YAC-TgM. To determine which part of the *H19* ICR sequences, ICR43 or ICR21, carries the activity that controls the methylation imprinting process, we generated ICR43(λ +CTCF) or (λ +CTCF)ICR21 TgM by performing and selecting for limited *in vivo cre-loxP* recombination in the ICR43(λ +CTCF)21 TgM (Fig. 6A and C and data not shown).

We first examined the methylation status of the maternally inherited transgenes in somatic cells of these YAC-TgM (Fig. 6). Bisulfite sequencing analysis revealed that both the ICR43 (region II' in Fig. 6B) and ICR21 (region I in Fig. 6D) sequences in the ICR43(λ +CTCF) and the (λ +CTCF)ICR21 TgM, respectively, exhibited low methylation status. At the same time, λ +CTCF portions in both YAC-TgM were moderately methylated (Fig. 6B and

FIG 2 DNA methylation status in the ICR43(λ +CTCF)21 YAC-TgM. (A) Map of the ICR43(λ +CTCF)21 fragment. Four regions (I, II', III, and IV) were analyzed by bisulfite sequencing in panels B to E and shown as gray bars beneath the map. (B and C) DNA methylation statuses of the ICR43 (II'), the λ +CTCF (III and IV), and the ICR21 (I) regions in somatic cells of the ICR43(λ +CTCF)21 YAC-TgM (lines 2619 and 2653) that inherited the transgenes either maternally (B) or paternally (C). Tail DNA was digested with XbaI and treated with sodium bisulfite. The four regions were amplified by nested PCR. PCR products were subcloned and then sequenced. Each horizontal row represents a single DNA template molecule. Methylated (filled circles) and unmethylated (open circles) CpG motifs are shown. Gray bars indicate the locations of the CTCF binding sites. The overall percentage of methylated CpGs is indicated beside each panel. (D) DNA methylation statuses of the ICR43, the λ +CTCF, and the ICR21 regions in testis germ cells of the ICR43(λ +CTCF)21 YAC-TgM (lines 2619 and 2653). Genomic DNA from the testis was analyzed by bisulfite sequencing, as described for panels B and C. (E) DNA methylation status of the λ +CTCF (III) region in MII oocytes of the ICR43(λ +CTCF)21 YAC-TgM (lines 2619 and 2653). Oocytes were embedded in agarose beads and treated with sodium bisulfite. The beads were separately and directly used to amplify region III of the transgene by nested PCR, and the resulting fragments were individually subcloned and sequenced.

FIG 3 DNA methylation status in the ICR43(λ +CTCF)21 and λ +CTCF YAC-TgM. (A and D) Partial restriction enzyme maps of the β -globin YAC transgene with the inserted ICR43(λ +CTCF)21 (A) or the λ +CTCF (D) fragment. Methylation-sensitive BstUI sites in the BamHI (B) fragment are displayed as vertical lines beneath each map of the transgene. A probe used for Southern blot analysis in panels B, C, E, and F is shown as a filled rectangle. (B and C) DNA methylation status of the λ +CTCF region in somatic cells of the ICR43(λ +CTCF)21 YAC-TgM (lines 2619 and 2653) that inherited the transgenes either maternally (B) or paternally (C). Tail DNA was digested with BamHI (-) and then BstUI (Bs), and the blot was hybridized with the λ probe. Tg, transgenic λ +CTCF. Here and in the following figures, asterisks indicate the positions of parental or methylated, undigested fragments. (E and F) DNA methylation status of the λ +CTCF region in somatic cells of the λ +CTCF YAC-TgM (lines 2619 and 2653) that inherited the transgenes either maternally (E) or paternally (F). Southern blot analysis of the tail DNA was conducted as described for panels B and C.

D), a result which was quantitatively similar to that seen in the λ +CTCF YAC-TgM (Fig. 4B) [P (region III) = 0.500 and = 0.9587 or P (region IV) = 0.087 and = 0.470 for lines 2619 and 2653, respectively, of the ICR43(λ +CTCF) TgM/ P (region III) = 0.645 and = 0.959 or P (region IV) = 0.507 and = 0.583 for lines 2619 and 2653, respectively, of the (λ +CTCF)ICR21 TgM]. These

results suggested that cooperativity between the ICR43 and ICR21 sequences was required to protect the maternally inherited λ +CTCF sequences from *de novo* DNA methylation in the YAC-TgM.

We next examined the methylation status of the paternally inherited transgenes in somatic cells of these YAC-TgM (Fig. 7).

A λ +CTCF transgene

B Maternal allele in somatic cells

C Paternal allele in somatic cells

D Testis cells

E Oocytes

FIG 4 DNA methylation status in the λ +CTCF YAC-TgM. (A) A map of the λ +CTCF fragment. Two regions (III and IV) were analyzed by bisulfite sequencing in panels B to E. (B and C) DNA methylation status of the λ +CTCF (III and IV) regions in somatic cells of the λ +CTCF YAC-TgM (lines 2619 and 2653) that inherited the transgenes either maternally (B) or paternally (C). Bisulfite sequencing analysis of the tail DNA was conducted as described in the legend to Fig. 2. For the paternally inherited transgene in panel C, representative samples determined by Southern blotting (no. 1 and 4 and no. 2 and 3 for partially and hypermethylated samples, respectively, in line 2619 or no. 2 to 5 and no. 1 and 6 for partially and hypermethylated samples, respectively, in line 2653 in Fig. 3F) were pooled and analyzed by bisulfite sequencing. (D) DNA methylation status of the λ +CTCF (III and IV) regions in testis germ cells of the λ +CTCF YAC-TgM (lines 2619 and 2653). Genomic DNA from the testis was analyzed by bisulfite sequencing, as described in the legend to Fig. 2. (E) DNA methylation status of the λ +CTCF (III) region in MII oocytes of the λ +CTCF YAC-TgM (lines 2619 and 2653). Genomic DNA from the oocytes was analyzed by bisulfite sequencing, as described in the legend to Fig. 2.

Both ICR43 and ICR21 fragments in the ICR43(λ +CTCF) and the (λ +CTCF)ICR21 TgM, respectively, were highly methylated. The λ +CTCF fragments in these TgM were also hypermethylated (Fig. 7B and D), and the methylation levels were equivalent to

those seen in the ICR43(λ +CTCF)21 TgM (Fig. 2C) [P (region III) = 0.951 and = 0.792 or P (region IV) = 0.548 and = 0.931 for lines 2619 and 2653, respectively, of the ICR43(λ +CTCF) TgM/ P (region III) = 0.549 and = 0.318 or P (region IV) = 0.500

FIG 5 ChIP experiments for CTCF binding in somatic cells. (A) Maps of the endogenous *H19*, transgenic λ +CTCF, and ICR43(λ +CTCF)21 loci. Target regions for qPCR amplification in panel B are shown as gray bars beneath each map. (B) ChIP was conducted for CTCF in somatic cells of the λ +CTCF and ICR43(λ +CTCF)21 YAC-TgM that inherited the transgenes maternally. The *CD45* intron 6 or *Necdin* gene and the *H19*ICR (shown in panel A) sequences were analyzed as negative and positive controls, respectively. Quantitative PCR was repeated at least three times for each sample. Fold enrichment of CTCF relative to IgG control (average values with standard deviations) was calculated and graphically depicted (average value of negative controls was set at 1.0). CTCF enrichment at the *H19*ICR seemed higher in the ICR43(λ +CTCF)21 than in the λ +CTCF TgM in both lines, presumably because the former carries two maternal *H19*ICR sequences (endogenous and transgenic) while the latter has only one.

and = 0.224 for lines 2619 and 2653, respectively, of the (λ +CTCF)ICR21 TgM]. According to these results, it seemed that, when paternally inherited, both ICR43 and ICR21 fragments contained activities sufficient to introduce DNA methylation into the adjacent λ +CTCF sequences.

DNA methylation in blastocysts. We previously reported that postfertilization methylation imprinting of the transgenic *H19* ICR in the β -globin locus commenced prior to implantation (34). Therefore, to ascertain whether the observed methylation deposition in the paternal ICR43 and ICR21 as well as the λ +CTCF fragments also occurred prior to implantation, we investigated the methylation status of the paternally inherited transgenes in blastocyst cells of the λ +CTCF, ICR43(λ +CTCF), and (λ +CTCF)ICR21 YAC-TgM (Fig. 8).

Bisulfite sequencing analysis revealed that the ICR43 region in the ICR43(λ +CTCF) TgM was hardly methylated (region II', Fig. 8C and D). Accordingly, the DNA methylation level in the λ +CTCF region was low (region IV, Fig. 8C and D), a result which was quantitatively similar to that seen in the λ +CTCF YAC-TgM (region IV, Fig. 8A and B) (P = 0.167 and 0.080 for lines 2619 and

FIG 6 DNA methylation status in the ICR43(λ +CTCF) and (λ +CTCF)ICR21 YAC-TgM (maternal allele). (A) A map of the ICR43(λ +CTCF) fragment. Three regions (II', III, and IV) were analyzed by bisulfite sequencing in panel B. (B) DNA methylation status of the ICR43 (II') and the λ +CTCF (III and IV) regions in somatic cells of the ICR43(λ +CTCF) TgM (lines 2619 and 2653) that inherited the transgenes maternally. Bisulfite sequencing analysis of the tail DNA was conducted as described in the legend to Fig. 2. (C) A map of the (λ +CTCF)ICR21 fragment. Three regions (I, III, and IV) were analyzed by bisulfite sequencing in panel D. (D) DNA methylation status of the λ +CTCF (III and IV) and the ICR21 (I) regions in somatic cells of the (λ +CTCF)ICR21 TgM (lines 2619 and 2653) that inherited the transgenes maternally. Bisulfite sequencing analysis of the tail DNA was conducted as described in the legend to Fig. 2.

2653, respectively). These results indicated that the ICR43 fragment was not capable of depositing DNA methylation by the blastocyst stage, indicating that both the ICR43 and the λ +CTCF sequences in the ICR43(λ +CTCF) TgM acquired DNA methylation after implantation (Fig. 7B). In contrast, the ICR21 sequences in the (λ +CTCF)ICR21 TgM acquired substantial methylation by the blastocyst stage (region I, Fig. 8E and F), and accordingly, the λ +CTCF region was also highly methylated (region IV, Fig. 8E and F) in comparison to the λ +CTCF TgM (region IV, Fig. 8A and B) (P = 0.045 and < 0.001 for lines 2619 and 2653, respectively). These results demonstrated that the ICR21, but not the ICR43, fragment contained an activity that was sufficient to introduce *de novo* DNA methylation into the adjacent λ +CTCF sequences on the paternal allele before implantation.

Generation of ICR43-21 YAC-TgM. Because the number of CpG and/or CTCF sites may be a functional component of the *H19* ICR, and because individual ICR43 or ICR21 subfragments

A ICR43(λ +CTCF) transgene **B** Paternal allele in somatic cells **C** (λ +CTCF)ICR21 transgene **D** Paternal allele in somatic cells

FIG 7 DNA methylation status in the ICR43(λ +CTCF) and (λ +CTCF)ICR21 YAC-TgM (paternal allele). (A) A map of the ICR43(λ +CTCF) fragment. Three regions (II', III, and IV) were analyzed by bisulfite sequencing in panel B. (B) DNA methylation status of the ICR43 (II') and the λ +CTCF (III and IV) regions in somatic cells of TgM (lines 2619 and 2653) that inherited the transgenes paternally. Bisulfite sequencing analysis of the tail DNA was conducted as described in the legend to Fig. 2. (C) A map of the (λ +CTCF)ICR21 fragment. Three regions (I, III, and IV) were analyzed by bisulfite sequencing in panel D. (D) DNA methylation status of the λ +CTCF (III and IV) and the ICR21 (I) regions in somatic cells of the (λ +CTCF)ICR21 TgM (lines 2619 and 2653) that inherited the transgene paternally. Bisulfite sequencing analysis of the tail DNA was conducted as described in the legend to Fig. 2.

carry lower numbers of these sites than do their parental fragment, they may not function in the absence of linked λ +CTCF sequences. To test if the ICR43 or ICR21 sequences alone are capable of establishing differential methylation status, each fragment was separately tested for its activity in YAC-TgM. To precisely compare the activities of these two regions and the parental *H19* ICR fragment, the ICR43 and ICR21 fragments were individually floxed by distinct sets of *loxP* sequences (*loxP5171* or *loxP2272*) and combined to employ a transgene coplacement strategy (ICR43-21, Fig. 1B) (35, 36). The floxed ICR43-21 fragment was inserted into the human β -globin YAC at a position 3' to the LCR (Fig. 1B), and two single-copy YAC-TgM lines (lines 41 and 191) were established. The integrity of the YAC transgenes in these animals was confirmed by long-range Southern blot analysis (Fig. 1B and D). In somatic cells of these mice, the ICR43-21 fragment

was differentially methylated in a parent-of-origin-dependent manner, as was expected from our previous results (data not shown) (23).

DNA methylation analyses of the ICR43 and ICR21 fragments in YAC-TgM. We next crossed ICR43-21 TgM with animals expressing cre recombinase to elicit cre-*loxP* recombination *in vivo*. cre recombinase randomly recognizes either a pair of *loxP2272* or a pair of *loxP5171* sites and excises the intervening DNA segments, which generates either ICR43 or ICR21 loci, respectively. Pups that carried either recombined transgene were determined by Southern blotting and then crossed with non-Tg animals to remove cre recombinase. Genotypes of the resultant ICR43 or ICR21 TgM sublines were confirmed by Southern blotting (Fig. 9A and data not shown).

To examine the methylation status of the ICR43 and ICR21 fragments in somatic cells of YAC-TgM, we analyzed their genomic DNA by bisulfite sequencing (Fig. 9) and Southern blot (data not shown) analyses. The ICR21 fragment was more heavily methylated when paternally inherited (Fig. 9B), which was identical to the methylation pattern seen in the 2.9-kb *H19* ICR (23) and the ICR43-21 fragments (data not shown), indicating that the ICR21 fragment alone is capable of recapitulating methylation imprinting in YAC-TgM. In contrast, the ICR43 fragment exhibited hypomethylation regardless of its parental origin (Fig. 9C) ($P = 0.544$ and 0.043 for lines 41 and 191, respectively), indicating that the paternal *H19* ICR is not able to acquire allele-specific methylation in the absence of the ICR21 sequences. In the testis of TgM, both ICR21 and ICR43 fragments were not methylated (data not shown), as was seen in the 2.9-kb *H19* ICR fragment (23), implying that methylation imprinting seen in the ICR21 fragment was established after fertilization. These results demonstrated that the ICR21, but not the ICR43, region contains sufficient *cis*-regulatory information to direct postfertilization, paternal-allele-specific DNA methylation. Importantly, these observations were highly consistent with the results of independently linking the ICR43 or ICR21 fragments to the λ +CTCF fragment and analyzing them at the blastocyst stage (Fig. 8C to F).

DISCUSSION

The majority of DMRs are methylated on the maternal allele. The combination of CpG spacing, histone modifications, and transcription spanning the DMR has been proposed to be important for maternal DMRs to acquire gametic methylation (37–40). Among a few paternal DMRs, *de novo* DNA methylation at the *Rasgrfl* DMR in prospermatogonia was established by small RNA (Piwi-interacting RNA [piRNA])-dependent mechanisms (41). Recently, Henckel et al. proposed a model in which transcription across the *H19* ICR might control methylation acquisition during spermatogenesis (9). However, this model seems unlikely as a mechanism to control postfertilization methylation imprinting as documented in our transgenic experiments, because the transcription that they described initiates from outside the 2.9-kb *H19* ICR, although it is possible that alternative promoters might drive transcription across the *H19* ICR in our transgenic experiments. Thus, the molecular mechanisms that dictate how allele-specific methylation takes place at the 2.9-kb *H19* ICR fragment and the identity of sequences responsible for the allele-discriminating process remain open questions.

We and others have reported that differential methylation of the *H19* ICR could be established even after fertilization at ectopic

FIG 8 DNA methylation status of the paternally inherited transgenes in blastocyst cells. Blastocysts that inherited the transgenes paternally were embedded in agarose beads and treated with sodium bisulfite. The beads were separately and directly used to amplify each region (region I, II', or IV) of the transgenes by nested PCR, and the resulting fragments were individually subcloned and sequenced. The results from single beads were presented together in a cluster. (A) A map of the λ +CTCF fragment. Region IV was analyzed by bisulfite sequencing in panel B. (B) DNA methylation status of region IV in the blastocyst cells of the λ +CTCF YAC-TgM (lines 2619 and 2653). (C) A map of the ICR43(λ +CTCF) fragment. Two regions (II' and IV) were analyzed by

FIG 9 DNA methylation status of the ICR21 and ICR43 fragments in somatic cells of YAC-TgM. (A) Maps of the endogenous *H19*, the ICR21, and the ICR43 fragments. Two regions (I and II) were analyzed by bisulfite sequencing and shown as gray bars above the transgene maps. (B and C) DNA methylation analysis of the ICR21 (B) and the ICR43 (C) regions in the TgM (lines 41 and 191) that inherited the transgenes either paternally (P) or maternally (M). Tail DNA was digested with *Xba*I and treated with sodium bisulfite. Region I of the ICR21 fragment (B) or region II of the ICR43 fragment (C) was amplified by nested PCR. PCR products were subcloned and then sequenced.

genomic sites (23–26). Subsequently, we found that the paternal transgenic *H19* ICR was already methylated at the blastocyst stage (34). Therefore, it is hypothesized that two plausible mechanisms are engaged in the establishment and maintenance of differential methylation of the transgenic *H19* ICR, though they are not mutually exclusive. First, as-yet-unknown epigenetic marks, set in the sperm, may be actively involved in introducing DNA methylation into nearby CpG sites from fertilization to the implantation period. Second, the mark set in the oocytes may somehow protect nearby CpG sites from allele-nonspecific *de novo* DNA methylation after fertilization, and even after implantation.

To test if these hypothetical activities indeed exist and, if they

bisulfite sequencing in panel D. (D) DNA methylation status of regions II' and IV in the blastocyst cells of the ICR43(λ +CTCF) YAC-TgM (lines 2619 and 2653). (E) A map of the (λ +CTCF)ICR21 fragment. Two regions (I and IV) were analyzed by bisulfite sequencing in panel F. (F) DNA methylation status of regions I and IV in the blastocyst cells of the (λ +CTCF)ICR21 YAC-TgM (lines 2619 and 2653).

do exist, to localize these activities within the 2.9-kb *H19* ICR fragment, we linked the *H19* ICR subfragments to the λ +CTCF fragment. The λ +CTCF fragment itself, even though it has four CTCF sites at the same position as that in the *H19* ICR, was substantially methylated upon maternal transmission (Fig. 3E and 4B). When the fragment was flanked by the ICR43 and ICR21 fragments [in the ICR43(λ +CTCF)21 TgM], the λ +CTCF sequences were methylated at low levels on the maternal allele (Fig. 2B and 3B), suggesting that the maternal *H19* ICR carries an activity that protects the adjacent heterologous CpG-rich sequences from *de novo* DNA methylation. Unidentified sequences in the *H19* ICR fragment in addition to the CTCF binding sites are probably required to exert this activity. Technically, we were unable to prove the existence of this activity at the preimplantation period, because the maternal λ +CTCF fragment remained unmethylated in blastocyst cells of TgM (data not shown). Therefore, a possible role of protective antimethylation activity in the “establishment” of postfertilization methylation imprinting at the *H19* ICR remains to be clarified.

When the ICR43 or ICR21 fragments were independently linked to the λ +CTCF fragment, they themselves exhibited low methylation levels on the maternal allele (Fig. 6B and D), which is consistent with the result of introducing the ICR43 or ICR21 sequences alone into the mouse genome (Fig. 9B and C). However, they did not protect the adjacent λ +CTCF fragment from *de novo* DNA methylation (Fig. 6B and D), i.e., they exhibited this activity only when the fragments flanked the λ +CTCF fragment. The underlying molecular mechanism of this phenomenon is currently uncertain. One possibility may be that interaction between multiple *cis*-regulatory elements scattered within the *H19* ICR fragment, mediated by DNA-binding proteins, may be essential for providing a robust activity that influences the methylation status of adjacent DNA either directly or indirectly by recruiting CTCF or by positioning nucleosomes (42). Alternatively, both ends of the λ +CTCF fragment may need to be flanked by the *H19* ICR sequences (ICR43 and/or ICR21). It is possible that the chromatin boundary activity of the CTCF (and additional) sites within the *H19* ICR protects the λ +CTCF fragment against the spread of DNA methylation from the outside.

Because protein (e.g., transcription factor) binding to the CpG sites protects them from *de novo* DNA methylation by DNMT3A (43), allele-specific protein binding may be involved in the protective antimethylation activity of the *H19* ICR after implantation. While CTCF binding sites in the maternal *H19* ICR are known to be essential to maintain its hypomethylation status after implantation in both endogenous (21, 22) and transgenic (34) contexts, they were not sufficient to protect the CpG-rich sequences, since the λ +CTCF fragment acquired methylation after implantation (Fig. 3E and 4B). It was reported that dyad Oct-binding sequences located in the ICR21 region mediated the maintenance of DNA hypomethylation at the *H19* ICR in cultured cells (44, 45). Curiously, the binding sequences are conserved in humans and were found to be mutated in Beckwith-Wiedemann syndrome (BWS) patients with ICR gain-of-methylation phenotypes, even though they have intact CTCF sites (46, 47). Therefore, the binding of CTCF as well as other factors, including Oct proteins, to the maternal *H19* ICR may be required to fully protect its CpG sites from *de novo* DNA methylation.

When the *H19* ICR subfragments (ICR43 and ICR21) were individually analyzed in YAC-TgM, only the ICR21 fragment ex-

hibited paternal allele-specific methylation, demonstrating that it carries sufficient information to establish postfertilization methylation imprinting (Fig. 9). This result is consistent with a previous observation showing that the 1.2-kb DNA element (exactly corresponding to the ICR43 region) can be removed from the *H19* ICR without disturbing its activity to establish differential DNA methylation at the endogenous locus (12). Therefore, the phenotype observed in the present study was unlikely to be an artifact of transgenic experimental investigation.

Based on our results, it is suggested that an unknown factor(s) in early embryos is actively involved in introducing DNA methylation into the ICR21, but not into the ICR43, fragments after paternal transmission. How such hypothetical factors distinguish the two fragments is currently unknown. Histone modifications could potentially regulate DNA methyltransferase (Dnmt) recruitment. For example, Dnmt3a and Dnmt3L preferentially interact with unmethylated lysine 4 of histone H3 (H3K4me0), and this interaction is inhibited by methylation at that residue (H3K4me/me2/me3) (39, 48, 49). Similarly, lysine 9 of histone H3 (H3K9) methylation has been shown to direct DNA methylation in several organisms (50–52). Therefore, histone modifications are capable of serving as primary epigenetic marks if they are set in germ cells. In support of this notion, it has been reported that, although histones are generally replaced with protamines during sperm maturation, some part of the genome, including the *Igf2/H19* allele, remained to be associated with histones (53, 54). Although H3K9me and H3K4me3 were reported to be enriched on the methylated and unmethylated DNA regions in somatic cells, respectively (39, 55), no difference in histone modification specific to either the germ line or a subregion of the *H19* ICR has ever been reported (9).

While knockout strategies are suitable for identifying “essential” *cis*-regulatory elements within the genome, transgenic studies allow a determination of whether a specific DNA sequence is sufficient for a specific function. We have previously shown by examining YAC-TgM that a 2.9-kb *H19* ICR fragment contained sufficient information to control imprinted methylation (23). In the current study, we were concerned about losing the imprinting activity by simply reducing the number of CpG sites upon dissecting the *H19* ICR fragment, which was one of the reasons why we linked the *H19* ICR subfragments to the CpG-rich λ DNA sequences. However, this concern proved to be unfounded, and we indeed demonstrated that the 1.7-kb ICR21 fragment was sufficient to establish postfertilization methylation imprinting in YAC-TgM. The activity also regulated methylation patterns in adjacent DNA regions in “heterologous” [$(\lambda$ +CTCF)ICR21 YAC-TgM], transgenic (ICR43-21 YAC-TgM), and probably endogenous loci, too, because DNA methylation in the downstream *H19* promoter region occurs after implantation (19). Although we cannot fully exclude the possibility that allele-specific protective antimethylation activity contributes to imprinted methylation patterns in the *H19* ICR before implantation, it seemed likely that allele-specific, methylation-introducing activity, rather than the protective activity, dominantly contributed to the formation of differential methylation in the ICR21 sequences, because the ICR43 fragment exhibited hypomethylation regardless of its parental origin (Fig. 9). By further narrowing the range of responsible sequences, we hope to identify molecules that act through these DNA sequences and thereby obtain insight into the methylation and genomic imprinting mechanisms. Furthermore, al-

though deregulation of methylation imprinting at the human *H19* ICR has been shown to be associated with Silver-Russell and Beckwith-Wiedemann syndromes (56, 57), causes of these epimutations are not fully understood. Uncovering the unknown epigenetic modifications that act as primary marks to establish imprinting will help elucidate the underlying molecular mechanisms of these epigenetic diseases.

ACKNOWLEDGMENTS

We thank Doug Engel (University of Michigan) and Jörg Bungert (University of Florida) for critically reading the manuscript and Y. Tanimoto for outstanding technical assistance.

This work was supported in part by a research grant from the Mochida Memorial Foundation for Medical and Pharmaceutical Research and a Grant-in-Aid for Young Scientists (S) from the Ministry of Education, Culture, Sports, Science and Technology of Japan (MEXT; KAKENHI grant number 20678002) to K.T. E.O. is a research fellow of the Japan Society for the Promotion of Science (JSPS; KAKENHI grant number 11J00587).

We declare that we have no conflict of interest.

REFERENCES

- Kaffer CR, Srivastava M, Park KY, Ives E, Hsieh S, Batlle J, Grinberg A, Huang SP, Pfeifer K. 2000. A transcriptional insulator at the imprinted *H19*/*Igf2* locus. *Genes Dev.* 14:1908–1919.
- Thorvaldsen JL, Duran KL, Bartolomei MS. 1998. Deletion of the *H19* differentially methylated domain results in loss of imprinted expression of *H19* and *Igf2*. *Genes Dev.* 12:3693–3702.
- Tremblay KD, Duran KL, Bartolomei MS. 1997. A 5' 2-kilobase-pair region of the imprinted mouse *H19* gene exhibits exclusive paternal methylation throughout development. *Mol. Cell. Biol.* 17:4322–4329.
- Bell AC, Felsenfeld G. 2000. Methylation of a CTCF-dependent boundary controls imprinted expression of the *Igf2* gene. *Nature* 405:482–485.
- Hark AT, Schoenherr CJ, Katz DJ, Ingram RS, Levorse JM, Tilghman SM. 2000. CTCF mediates methylation-sensitive enhancer-blocking activity at the *H19*/*Igf2* locus. *Nature* 405:486–489.
- Bourc'his D, Bestor TH. 2004. Meiotic catastrophe and retrotransposon reactivation in male germ cells lacking *Dnmt3L*. *Nature* 431:96–99.
- Kaneda M, Okano M, Hata K, Sado T, Tsujimoto N, Li E, Sasaki H. 2004. Essential role for de novo DNA methyltransferase *Dnmt3a* in paternal and maternal imprinting. *Nature* 429:900–903.
- Kato Y, Kaneda M, Hata K, Kumaki K, Hisano M, Kohara Y, Okano M, Li E, Nozaki M, Sasaki H. 2007. Role of the *Dnmt3* family in de novo methylation of imprinted and repetitive sequences during male germ cell development in the mouse. *Hum. Mol. Genet.* 16:2272–2280.
- Henckel A, Chebli K, Kota SK, Arnaud P, Feil R. 2012. Transcription and histone methylation changes correlate with imprint acquisition in male germ cells. *EMBO J.* 31:606–615.
- Hata K, Okano M, Lei H, Li E. 2002. *Dnmt3L* cooperates with the *Dnmt3* family of de novo DNA methyltransferases to establish maternal imprints in mice. *Development* 129:1983–1993.
- Bowman AB, Levorse JM, Ingram RS, Tilghman SM. 2003. Functional characterization of a testis-specific DNA binding activity at the *H19*/*Igf2* imprinting control region. *Mol. Cell. Biol.* 23:8345–8351.
- Drewell RA, Brenton JD, Ainscough JF, Barton SC, Hilton KJ, Arney KL, Dandolo L, Surani MA. 2000. Deletion of a silencer element disrupts *H19* imprinting independently of a DNA methylation epigenetic switch. *Development* 127:3419–3428.
- Ideraabdullah FY, Abramowitz LK, Thorvaldsen JL, Krapp C, Wen SC, Engel N, Bartolomei MS. 2011. Novel cis-regulatory function in ICR-mediated imprinted repression of *H19*. *Dev. Biol.* 355:349–357.
- Katz DJ, Beer MA, Levorse JM, Tilghman SM. 2005. Functional characterization of a novel *Ku70/80* pause site at the *H19*/*Igf2* imprinting control region. *Mol. Cell. Biol.* 25:3855–3863.
- Reed MR, Riggs AD, Mann JR. 2001. Deletion of a direct repeat element has no effect on *Igf2* and *H19* imprinting. *Mamm. Genome* 12:873–876.
- Szabo PE, Han L, Hyo-Jung J, Mann JR. 2006. Mutagenesis in mice of nuclear hormone receptor binding sites in the *Igf2*/*H19* imprinting control region. *Cytogenet. Genome Res.* 113:238–246.
- Mayer W, Niveleau A, Walter J, Fundele R, Haaf T. 2000. Demethylation of the zygotic paternal genome. *Nature* 403:501–502.
- Oswald J, Engemann S, Lane N, Mayer W, Olek A, Fundele R, Dean W, Reik W, Walter J. 2000. Active demethylation of the paternal genome in the mouse zygote. *Curr. Biol.* 10:475–478.
- Tomizawa S, Kobayashi H, Watanabe T, Andrews S, Hata K, Kelsey G, Sasaki H. 2011. Dynamic stage-specific changes in imprinted differentially methylated regions during early mammalian development and prevalence of non-CpG methylation in oocytes. *Development* 138:811–820.
- Hemberger M, Dean W, Reik W. 2009. Epigenetic dynamics of stem cells and cell lineage commitment: digging Waddington's canal. *Nat. Rev. Mol. Cell Biol.* 10:526–537.
- Engel N, West AG, Felsenfeld G, Bartolomei MS. 2004. Antagonism between DNA hypermethylation and enhancer-blocking activity at the *H19* DMD is uncovered by CpG mutations. *Nat. Genet.* 36:883–888.
- Schoenherr CJ, Levorse JM, Tilghman SM. 2003. CTCF maintains differential methylation at the *Igf2*/*H19* locus. *Nat. Genet.* 33:66–69.
- Tanimoto K, Shimotsu M, Matsuzaki H, Omori A, Bungert J, Engel JD, Fukamizu A. 2005. Genomic imprinting recapitulated in the human beta-globin locus. *Proc. Natl. Acad. Sci. U. S. A.* 102:10250–10255.
- Matsuzaki H, Okamura E, Shimotsu M, Fukamizu A, Tanimoto K. 2009. A randomly integrated transgenic *H19* imprinting control region acquires methylation imprinting independently of its establishment in germ cells. *Mol. Cell. Biol.* 29:4595–4603.
- Gebert C, Kunkel D, Grinberg A, Pfeifer K. 2010. *H19* imprinting control region methylation requires an imprinted environment only in the male germ line. *Mol. Cell. Biol.* 30:1108–1115.
- Park KY, Sellars EA, Grinberg A, Huang SP, Pfeifer K. 2004. The *H19* differentially methylated region marks the parental origin of a heterologous locus without gametic DNA methylation. *Mol. Cell. Biol.* 24:3588–3595.
- Tanimoto K, Sugiura A, Omori A, Felsenfeld G, Engel JD, Fukamizu A. 2003. Human beta-globin locus control region HS5 contains CTCF- and developmental stage-dependent enhancer-blocking activity in erythroid cells. *Mol. Cell. Biol.* 23:8946–8952.
- Tanimoto K, Liu Q, Grosveld F, Bungert J, Engel JD. 2000. Context-dependent EKLF responsiveness defines the developmental specificity of the human epsilon-globin gene in erythroid cells of YAC transgenic mice. *Genes Dev.* 14:2778–2794.
- Yu RN, Ito M, Saunders TL, Camper SA, Jameson JL. 1998. Role of *Ahch* in gonadal development and gametogenesis. *Nat. Genet.* 20:353–357.
- Shukla S, Kavak E, Gregory M, Imashimizu M, Shutinoski B, Kashlev M, Oberdoerffer P, Sandberg R, Oberdoerffer S. 2011. CTCF-promoted RNA polymerase II pausing links DNA methylation to splicing. *Nature* 479:74–79.
- Splinter E, Heath H, Kooren J, Palstra RJ, Klous P, Grosveld F, Galjart N, de Laat W. 2006. CTCF mediates long-range chromatin looping and local histone modification in the beta-globin locus. *Genes Dev.* 20:2349–2354.
- Shimotsu M, Matsuzaki H, Tanabe O, Campbell AD, Engel JD, Fukamizu A, Tanimoto K. 2007. Linear distance from the locus control region determines epsilon-globin transcriptional activity. *Mol. Cell. Biol.* 27:5664–5672.
- Lee G, Saito I. 1998. Role of nucleotide sequences of loxP spacer region in Cre-mediated recombination. *Gene* 216:55–65.
- Matsuzaki H, Okamura E, Fukamizu A, Tanimoto K. 2010. CTCF binding is not the epigenetic mark that establishes post-fertilization methylation imprinting in the transgenic *H19* ICR. *Hum. Mol. Genet.* 19:1190–1198.
- Shimizu T, Oishi T, Omori A, Sugiura A, Hirota K, Aoyama H, Saito T, Sugaya T, Kon Y, Engel JD, Fukamizu A, Tanimoto K. 2005. Identification of cis-regulatory sequences in the human angiotensinogen gene by transgene coplacement and site-specific recombination. *Mol. Cell. Biol.* 25:2938–2945.
- Tanimoto K, Sugiura A, Kanafusa S, Saito T, Masui N, Yanai K, Fukamizu A. 2008. A single nucleotide mutation in the mouse renin promoter disrupts blood pressure regulation. *J. Clin. Invest.* 118:1006–1016.
- Chotalia M, Smallwood SA, Ruf N, Dawson C, Lucifero D, Frontera M, James K, Dean W, Kelsey G. 2009. Transcription is required for establishment of germline methylation marks at imprinted genes. *Genes Dev.* 23:105–117.

38. Ciccone DN, Su H, Hevi S, Gay F, Lei H, Bajko J, Xu G, Li E, Chen T. 2009. KDM1B is a histone H3K4 demethylase required to establish maternal genomic imprints. *Nature* 461:415–418.
39. Henckel A, Nakabayashi K, Sanz LA, Feil R, Hata K, Arnaud P. 2009. Histone methylation is mechanistically linked to DNA methylation at imprinting control regions in mammals. *Hum. Mol. Genet.* 18:3375–3383.
40. Jia D, Jurkowska RZ, Zhang X, Jeltsch A, Cheng X. 2007. Structure of Dnmt3a bound to Dnmt3L suggests a model for de novo DNA methylation. *Nature* 449:248–251.
41. Watanabe T, Tomizawa S, Mitsuya K, Totoki Y, Yamamoto Y, Kuramochi-Miyagawa S, Iida N, Hoki Y, Murphy PJ, Toyoda A, Gotoh K, Hiura H, Arima T, Fujiyama A, Sado T, Shibata T, Nakano T, Lin H, Ichiyani K, Soloway PD, Sasaki H. 2011. Role for piRNAs and non-coding RNA in de novo DNA methylation of the imprinted mouse *Rasgrf1* locus. *Science* 332:848–852.
42. You JS, Kelly TK, De Carvalho DD, Taberlay PC, Liang G, Jones PA. 2011. OCT4 establishes and maintains nucleosome-depleted regions that provide additional layers of epigenetic regulation of its target genes. *Proc. Natl. Acad. Sci. U. S. A.* 108:14497–14502.
43. Han L, Lin IG, Hsieh CL. 2001. Protein binding protects sites on stable episomes and in the chromosome from de novo methylation. *Mol. Cell. Biol.* 21:3416–3424.
44. Hori N, Nakano H, Takeuchi T, Kato H, Hamaguchi S, Oshimura M, Sato K. 2002. A dyad oct-binding sequence functions as a maintenance sequence for the unmethylated state within the H19/Igf2-imprinted control region. *J. Biol. Chem.* 277:27960–27967.
45. Hori N, Yamane M, Kouno K, Sato K. 31 October 2012. Induction of DNA demethylation depending on two sets of Sox2 and adjacent Oct3/4 binding sites (Sox-Oct motifs) within the mouse H19/insulin-like growth factor 2 (Igf2) imprinted control region. *J. Biol. Chem.* [Epub ahead of print.] doi:10.1074/jbc.M112.424580.
46. Demars J, Shmela ME, Rossignol S, Okabe J, Netchine I, Azzi S, Cabrol S, Le Caignec C, David A, Le Bouc Y, El-Osta A, Gicquel C. 2010. Analysis of the IGF2/H19 imprinting control region uncovers new genetic defects, including mutations of OCT-binding sequences, in patients with 11p15 fetal growth disorders. *Hum. Mol. Genet.* 19:803–814.
47. Poole RL, Leith DJ, Docherty LE, Shmela ME, Gicquel C, Splitt M, Temple IK, Mackay DJ. 2012. Beckwith-Wiedemann syndrome caused by maternally inherited mutation of an OCT-binding motif in the IGF2/H19-imprinting control region, ICR1. *Eur. J. Hum. Genet.* 20:240–243.
48. Ooi SK, Qiu C, Bernstein E, Li K, Jia D, Yang Z, Erdjument-Bromage H, Tempst P, Lin SP, Allis CD, Cheng X, Bestor TH. 2007. DNMT3L connects unmethylated lysine 4 of histone H3 to de novo methylation of DNA. *Nature* 448:714–717.
49. Otani J, Nankumo T, Arita K, Inamoto S, Ariyoshi M, Shirakawa M. 2009. Structural basis for recognition of H3K4 methylation status by the DNA methyltransferase 3A ATRX-DNMT3-DNMT3L domain. *EMBO Rep.* 10:1235–1241.
50. Jackson JP, Lindroth AM, Cao X, Jacobsen SE. 2002. Control of CpNpG DNA methylation by the KRYPTONITE histone H3 methyltransferase. *Nature* 416:556–560.
51. Lehnertz B, Ueda Y, Derijck AA, Braunschweig U, Perez-Burgos L, Kubicek S, Chen T, Li E, Jenuwein T, Peters AH. 2003. Suv39h-mediated histone H3 lysine 9 methylation directs DNA methylation to major satellite repeats at pericentric heterochromatin. *Curr. Biol.* 13:1192–1200.
52. Tamaru H, Selker EU. 2001. A histone H3 methyltransferase controls DNA methylation in *Neurospora crassa*. *Nature* 414:277–283.
53. Arpanahi A, Brinkworth M, Iles D, Krawetz SA, Paradowska A, Platts AE, Saida M, Steger K, Tedder P, Miller D. 2009. Endonuclease-sensitive regions of human spermatozoal chromatin are highly enriched in promoter and CTCF binding sequences. *Genome Res.* 19:1338–1349.
54. Hammoud SS, Nix DA, Zhang H, Purwar J, Carrell DT, Cairns BR. 2009. Distinctive chromatin in human sperm packages genes for embryo development. *Nature* 460:473–478.
55. Delaval K, Govin J, Cerqueira F, Rousseaux S, Khochbin S, Feil R. 2007. Differential histone modifications mark mouse imprinting control regions during spermatogenesis. *EMBO J.* 26:720–729.
56. Choufani S, Shuman C, Weksberg R. 2010. Beckwith-Wiedemann syndrome. *Am. J. Med. Genet. C Semin. Med. Genet.* 154C:343–354.
57. Eggermann T. 2010. Russell-Silver syndrome. *Am. J. Med. Genet. C Semin. Med. Genet.* 154C:355–364.
58. Leighton PA, Saam JR, Ingram RS, Stewart CL, Tilghman SM. 1995. An enhancer deletion affects both H19 and Igf2 expression. *Genes Dev.* 9:2079–2089.