Non-CpG Methylation Occurs in the Regulatory Region of the Sry Gene

Koichiro NISHINO1–3), Naoko HATTORI1), Shun SATO1), Yoshikazu ARAI1), Satoshi TANAKA1), Andras NAGY3) and Kunio SHIOTA1)

1)Laboratory of Cellular Biochemistry, Animal Resource Sciences/Veterinary Medical Sciences, The University of Tokyo, Tokyo 113-8657, 2)JSPS Postdoctoral Fellowship for Research Abroad, Japan Society for Promotion of Science (JSPS), Tokyo 102-8472, Japan and 3)Samuel Lunenfeld Research Institute, Mount Sinai Hospital, Toronto, ON M5G 1X5, Canada

Abstract. The Sry (sex determining region on Y chromosome) gene is a master gene for sex determination. We previously reported that the Sry gene has tissue-dependent and differentially methylated regions (T-DMRs) by analyzing the DNA methylation states at CpG sites in the promoter regions. In this study, we found unique non-CpG methylation at the internal cytosine in the 5'-CCTGG-3' pentanucleotide sequence in the Sry T-DMR. This non-CpG methylation was detected in four mouse strains (ICR, BALB/c, DBA2 and C3H), but not in two strains (C57BL/6 and 129S1), suggesting that the CCTGG methylation is tentative and unstable. Interestingly, this CCTGG methylation was associated with demethylation of the CpG sites in the Sry T-DMR in the developmental process. A methylation-mediated promoter assay showed that the CCTGG methylation promotes gene expression. Our finding shows that non-CpG methylation has unique characteristic and is still conserved in mammals.

Key words: DNA methylation, Non-CpG methylation, Sry

DNA methylation plays an important role in normal differentiation, development and cloned animals [1–4]. Eukaryotic DNA contains 5-methylcytosine as the only methylated nucleotide in the genome. The methylation of cytosine occurs mainly at the CpG site in mammals [5–7]. Non-CpG methylation in mammals is quite rare, whereas non-CpG methylation as CNG (N can be any base) is common in plants [8], bacteria [9, 10] and yeast [11]. In mammals, DNA methylation of CpGs is involved in various phenomena from development through gene silencing, splicing, stabilizing and chromatin remodeling [1, 12–21]. In contrast, the biological role of non-CpG methylation in mammals is unknown.

Sry (sex determining region on Y chromosome) is the master gene for initiating the cascade leading to testicular differentiation in mammals [22]. Expression of the Sry gene is restricted to gonadal somatic cells at 10.5–12.5 days postcoitum (dpc) in the mouse [23–25]. The Sry gene produces two types of transcripts with different transcription start sites, linear and circular forms [24, 26]. The linear transcript is translated into protein and is expressed only from 10.5 to 12.5 dpc in the pre-Sertoli cells in the developing male gonad [23–25], whereas the circular form is thought to be untranslated [23, 27–29]. The Sry gene has tissue-dependent and differentially methylated regions (T-DMRs), and the CpG methylation kinetics coincide with the spatio-temporal expression of the Sry gene [19]. Thus, the Sry locus may be an interesting model for acquiring the epigenetic regulation and development of the gonad from the evolutionary aspect in mammals.

Materials and Methods

Mice and cell preparation

The experiments were carried out according to the guidelines for the care and use of laboratory animals, Graduate School of Agriculture and Life Sciences, The University of Tokyo. ICR (CD-1) mice, which have been kept as a closed colony in Japan and were designated as ICR-J in this experiment, were purchased from a Japanese dealer (Charles River Japan, Yokohama, Japan). Another ICR (CD-1) strain of mice, which has been kept as another closed colony in North America and was designated as ICR-A, was purchased from a North American dealer (Harlan, Indianapolis, IN, USA). C57BL/6, 129S1, BALB/c, DBA2 and C3H mice were purchased from a Canadian dealer (Charles River Canada, Quebec, Canada). Noon on the day when a vaginal plug was detected was designated as 0.5 dpc, and 8.5-dpc embryos, 11.5-dpc gonads separated from the mesonephros and livers and 15.5-dpc testes and livers were collected. Adult male (>3 months old) livers, kidneys and testes were also collected. To obtain 11.5-dpc gonadal somatic cells or adult testicular somatic cells, 11.5-dpc gonads or adult testes were pipetted to disperse single cells in a trypsin and collagenase solution. Separated cells were cultured in DMEM (Gibco BRL, Rockville, MD, USA) with 10% FBS at 37 C for 3 h.
Within this period, most of the somatic cells adhere to the dish surface, but the germ cells and blood cells do not. Then the floating cells were removed with the supernatant, and the remaining somatic cells were harvested after washing three times with PBS. To collect adult germ cells, adult testes were fixed with 4% (w/v) paraformaldehyde and embedded in paraffin. Five-micrometer-thick sections of adult testis were stained with Mayer’s hematoxylin and eosin, and then dried well. Spermatogonia, spermatocytes and round spermatids from the sections were isolated using a PALM MicroBeam system (PALM Microlaser Technologies A.G., Bernried, Germany) according to the manufacturer’s instructions. Germ cells from the sections were pooled (100–200 cells). Sperm was also obtained from the cauda epididymis. All samples were frozen in liquid nitrogen and stored at –80°C until use.

**Sodium bisulfite sequencing and MultiNA-COBRA**

A small amount of genomic DNA was extracted, treated with sodium bisulfite and subjected to restriction mapping and sequencing analyses as described previously [19]. The *Hind* III-digested genomic DNA were treated with 2.5 M bisulfite at 55°C for 16 h [19]. PCR amplification was performed using IMMOLASE™ DNA polymerase (Bioline, London, UK) and the primer set IBF/IBR for Region I or IIBF/IIBR1 or IIBR2 for Region II [19]. To determine the methylation status of individual CpG and CCTGG sites, the PCR product was gel extracted and subcloned into pGEM-T Easy vector (Promega, Madison, WI, USA), and then sequenced. The PCR reaction was performed at least three times, and 4–6 clones per PCR reaction were sequenced. The 3.8-kb DNA fragment (from nt4870 to nt8700; the numbering is based on the registered nucleotide sequence in the GenBank database; accession number X67204) including the 3.4-kb upstream region of the linear transcription start site (from nt6214 to nt8246) of the Sry gene was subcloned into pGL3-Basic vector (Promega). Cloning of unmethylated and CCTGG-methylated constructs was performed as described above. The gonadal cells of a 11.5-dpc fetus (pcgc11.5) from ICR-J were cultured on a 24-well plate (1.6 × 10^4 cells/well). The cells were then transfected with 0.165 pmol of reporter constructs with Effectene Transfection Reagent (Qiagen GmbH, Hilden, Germany). To normalize the firefly luciferase activity of the reporter constructs, an internal control plasmid (0.015 pmol) expressing Renilla luciferase (pRL-TK vector, Promega) was cotransfected into the cells. Promoter activity was measured 12 h after transfection into pcgc11.5. The activities of both luciferases were determined using a dual-luciferase reporter assay system (Promega) according to the manufacturer’s instructions. Experiments were performed at least three times independently.

**Results**

**The Sry T-DMRs contain CCTGG sequences**

There are two T-DMRs at the putative promoter region of circular and linear form transcripts, Region I and Region II, respectively (Fig. 1A), in the Sry locus [19]. Region I and Region II contain two CpG sites (a, nt5475; b, nt5688; the numbering is based on the registered nucleotide sequence in the GenBank database; accession number X67204) and five CpG sites (c, nt7411; d, nt7728; e, nt7741; f, nt7844; g, nt8004; Fig. 1A).

Bisulfite sequencing analysis revealed that CpGs in Region I were hypomethylated in the blastocyst, 11.5-dpc gonad, adult liver and adult kidney but were highly methylated in the 11.5-dpc liver in the ICR-J strain (Fig. 1B). The CpGs in Region II were hypomethylated in the 11.5-dpc gonad and the blastocyst, whereas they were hypermethylated in the 11.5-dpc liver, adult liver and kidney in the ICR-J strain (Fig. 1B). These results confirmed those of a previous study [19]. There is a CCTGG sequence in each of Region I (†, nt5174) and Region II (*, nt7476; Fig. 1A). An internal cytosine in the CCTGG was methylated in Region I in the blastocyst, 11.5-dpc gonad and adult liver and kidney, but not in the 11.5-dpc liver (Fig. 1B). In contrast, the CCTGG in Region II was not methylated in the 11.5-dpc liver and adult tissues (Fig. 1B). Interestingly, the methylation of CCTGG in Region II occurred in the blastocyst and gonad of 11.5-dpc, in which the CpGs were unmethylated. No methylation of symmetric and asymmetric non-CpG sites, including CGW, CCG, CA and CT sites, was observed in any sample examined.

To confirm further the presence of CCTGG methylation without
Fig. 1. A: CpG and CCWGG (W=A or T) sequences at the mouse Sry locus in ICR-J mice. Closed circles indicate the CpG dinucleotide, and open diamonds indicate the CCWGG pentanucleotide sequences. The Sry gene has two putative promoter regions of circular and linear transcripts (Region I and Region II). There are two CpGs (a, nt5475; b, nt5688) and one CCTGG site (nt5174) in Region I, and five CpGs (c, nt7411; d, nt7728; e, nt7741; f, nt7844; g, nt8004) and one CCTGG site (nt7476) in Region II. The numbering is based on the registered nucleotide sequence in the GenBank database (accession number X67204). Arrows indicate the primer positions for bisulfite PCR ( ). Open and closed diamonds indicate unmethylated and methylated CpGs, respectively. Crosses ( ) and asterisks (*) show the position of the CCTGG site. C: Detection of CCWGG methylation by methylation-sensitive restriction enzyme PspGI-mediated genomic PCR. After digestion with PspGI, the CCWGG methylated plasmid (CCWGG) and the CCWGG unmethylated plasmid (U), 11.5-dpc gonad and liver genomic DNAs were subjected to PCR (+lanes). PCR was also performed with undigested genomic DNA (−lanes). The 11.5-dpc gonad gave a PCR band, indicating that the CCTGG site was methylated.

DNA did not. When the CCWGG-methylated construct was used as a positive control, a band of the expected size was also detected, whereas the unmethylated construct, as a negative control, did not give an amplified product when the constructs were predigested by PspGI.

We performed the bisulfite reaction at 55 °C, which used to be one of the standard condition. This allows for the possibility that the CCTGG methylation detected at the Region II came from an experimental artifact because the cytosine to uracil conversion in the bisulfite reaction is often prevented by local folding. To address this issue, we confirmed conversion efficiency using three different methylated fragments (unmethylated, CpG-methylated and CCWGG-methylated constructs). Region II was amplified after the bisulfite reaction, and the conversion was determined at all cytosines by sequencing. Region II contains 94 cytosine residues, 6 of which are located at CpG sites and 1 of which is located at the CCWGG site (Fig. 2AB). Almost all unmethylated cytosines were converted to uracil in all the three different fragments (Fig. 2C). No CCTGG methylation was detected from the unmethylated construct, meaning that bisulfite conversion of the CCTGG site under the standard conditions, in so far as Region II was analyzed, was not influenced by a secondary structure. Thus, the existence of CCTGG methylation was confirmed by two different methods. The CCTGG of Sry T-DMR was differentially methylated in a tissue- and development-dependent manner. There is an inverse relationship between cytosine methylation in the CCTGG and CpG sites at the genomic locus.

Efficiency of CCTGG methylation detection by sodium bisulfite reaction

To evaluate the detection efficiency of CCWGG methylation, MultiNA-COBRA and bisulfite sequencing were performed using a mixture of differently methylated DNA fragments including Region II. The CCWGG-methylated and unmethylated fragments, unmethylated and CpG-methylated fragments, CCWGG-methylated and CpG/CCWGG-methylated fragments or CCWGG- and CpG-methylated fragments were then mixed at 3:1, 1:1 and 1:3 ratios. MultiNA-COBRA was performed using the methylation-sensitive restriction enzyme Hpy CH4 IV (New England Biolabs), and the level of DNA methylation at the CpG site was quantified based on the density of the PCR band. The CpG methylation levels of unmethylated, CCWGG-methylated, CpG-methylated and CpG/CCWGG-methylated fragments were 8.3 ± 0.2, 9.6 ± 0.1, 90.9 ± 0.3 and 85.3 ± 0.1%, respectively. The levels of CpG methylation at the HpyCH4IV site of the mixtures of unmethylated and CpG-methylated fragments were 21.8 ± 0.2% (3:1), 39.1 ± 0.2% (1:1) and 68.3 ± 0.4% (1:3), those of the mixtures of CCWGG-methylated and CpG/CCWGG-methylated fragments were 33.0 ± 0.2% (3:1), 72.7 ± 0.87% (1:1), and 87.0 ± 1.0% (1:3) and those of the mixtures of CCWGG-methylated and CpG-methylated fragments were 51.9 ± 0.9% (3:1), 72.0 ± 0.9% (1:1) and 84.6 ± 0.1% (1:3) (Fig. 3C). The CCWGG methylation was visibly lower than expected, although the mixtures of unmethylated and CpG-methylated fragments provided a methylation ratio close to the expected value. Sodium bisulfite sequencing also showed a marked tendency against detection of CCTGG methylation (Fig. 3EFGH).
The CCTGG methylation ratios estimated from bisulfite sequencing of the mixtures of CCWGG-methylated and unmethylated fragments were 64.7% (3:1), 17.6% (1:1) and 0.6% (1:3, Fig. 3E), and those of the mixtures of CCWGG-methylated and CpG-methylated fragments were 76.5% (3:1), 82.4% (1:1) and 94.1% (1:3, Fig. 3H). The CpG or CCTGG methylation ratio in other samples appeared almost equal to the expected ratio (Fig. 3FG). These data indicate that detection of CCTGG methylation, particularly in the combination of methylated and unmethylated CCWGG and of methylated CCWGG and CpG, is biased in the sodium bisulfite method.

**Inverse correlation between methylation of CCTGG and CpG sequences in Region II**

To examine the relationship between CCTGG and CpG methylation, the methylation status of Sry T-DMR was investigated in the developing gonad and other tissues in the ICR-J strain (Fig. 4). In Region II, hypermethylation of the CCTGG was detected in the gonad of 11.5 dpc. The 11.5-dpc fetal gonadal somatic cells were also hypermethylated. The CpGs were hypermethylated in the 8.5-dpc embryo and 11.5-dpc liver. In contrast, methylation of the CpGs was rare in the gonad and gonadal somatic cells at 11.5 dpc. The CCTGG was unmethylated in the 8.5-dpc embryo and in the liver at 11.5 and 15.5 dpc. In the adult testicular somatic cells, no CCTGG methylation was detected in Region II (Fig. 4B). Spermatogonia, spermatocytes and round spermatids showed hypermethylation of the CCTGG site and hypomethylation of the CpG sites (Fig. 4B). No methylation of symmetric and asymmetric non-CpG sites was observed in any sample examined except for the CCTGG site. Thus, the methylation status showed an inverse relationship between CpG and CCTGG sites in Region II, and this DNA methylation dynamic corresponded with the temporal expression pattern of the Sry gene.

**Variation of CCTGG methylation between mouse strains**

In an ICR mouse subcolony, CpGs in Region II were methylated partially in the fetal gonads of ICR-A mice (Fig. 5B). The partial methylation of CpGs also occurred in the C57BL/6, 129S1, BALB/c, DBA2 and C3H strains (Fig. 5B). COBRA analysis also indicated that the trend for the CpGs in Region II to become hypomethylated in the fetal gonad at 11.5 dpc relative to the fetal liver at 11.5 dpc was common among all mouse strains (data not shown). The CCTGG methylation was also observed in the ICR-A, BALB/c, DBA2 and C3H strains, whereas the C57BL/6 and 129S1 strains did not exhibit CCTGG methylation (Fig. 5B and 5C). No other DNA methylation of symmetric or asymmetric non-CpG sites...
was detected in the mice strains examined. These data are interesting in that they imply that methylation of the CCTGG site seems to depend on the genomic background.

**Effect of CpG or CCTGG methylation on Sry promoter activity**

We previously established a promoter assay system using pgc11.5 in vitro [19]. The reporter construct includes the 2.0-kb 5'-flanking region of the Sry gene, which is the putative linear promoter region and contains six CpG and one CCTGG sites. Interestingly, the CCWGG-methylated reporter construct showed increasing activity, up to 128%, compared with the unmethylated construct, although CpG methylation suppressed the activity to 43% in pgc11.5 (Fig. 6). The internal cytosine methylation of CCWGG seemed to promote the activity, even though the CpG methylation suppressed the activity (Fig. 6). These results indicate that the biological role of CCTGG methylation is distinct from that of CpG methylation.

**Discussion**

We found a unique non-CpG methylation of the internal cytosine within the sequence CCTGG at two T-DMRs in the Sry locus on the Y chromosome in the mouse. Sodium bisulfite sequencing is a conventional method for detecting DNA methylation in the genome, but here, we showed that CCWGG methylation is likely to be underestimated by sodium bisulfite sequencing because of a potential PCR and cloning bias. Therefore, CCWGG methylation in mammals may be greater than expected.

Molone et al. [30] found methylation at the CCWGG pentanucleotide sequence in an endogenous mammalian gene of primary effusion lymphoma and myeloma cell lines. Franchina et al. [31] and Agirre et al. [32] also reported that methylation of CCWGG can be found at the flanking region of the CpG island of the myf-3 gene and at the promoter of TP53. Imamura et al. [33] found non-CpG methylation at the Peg/Mest gene in the preimplantation embryo. We found a unique DNA methylation at the CCTGG site in the Sry promoter regions and that this methylation occurred...
CCWGG METHYLATION ON Sry GENE

Analyze the content of the text and create a natural representation of the information provided. The text discusses the reciprocal DNA methylation between CpG and CCTGG sites at the Region II in ICR-J. The text also mentions the genomic sequences of the 2.0-kb upstream region of mouse Sry, and the DNA methylation status of individual CpG and CCTGG sites in Region II in 11.5-dpc fetal gonads determined by bisulfite sequence analysis. The numbers in parentheses in (C) indicate the number of clones sequenced.

Male-to-female sex reversal has been reported after transfer of a certain Mus Domesticus-derived Y chromosome into the C57BL/6J inbred mouse strain [37–39]. Interestingly, the genomic sequence of the Sry promoter region including Region I and Region II in all mice strains examined (data not shown). The genomic sequences of the locus were identical in the 129S1, C57BL/6, BALB/c, DBA2 and C3H strains, supporting the interpretation that another mechanism besides the local sequence produces CCTGG methylation.

Male-to-female sex reversal has been reported after transfer of a certain Mus Domesticus-derived Y chromosome into the C57BL/6J inbred mouse strain [37–39]. Interestingly, the genomic sequence of the Sry promoter region including Region I and Region II in all mice strains examined (data not shown). The genomic sequences of the locus were identical in the 129S1, C57BL/6, BALB/c, DBA2 and C3H strains, supporting the interpretation that another mechanism besides the local sequence produces CCTGG methylation.

Male-to-female sex reversal has been reported after transfer of a certain Mus Domesticus-derived Y chromosome into the C57BL/6J inbred mouse strain [37–39]. Interestingly, the genomic sequence of the Sry promoter region including Region I and Region II in the ICR mice was the same as that in the Mus Domesticus strains. It will be valuable to examine whether the Mus Domesticus-derived Y chromosome can be affected by CCTGG methylation in the sex-determining period. The difference in the sensitivity to DNA methylation, such as at the CpG and/or CCWGG sites, between C57BL/6J and Mus Domesticus might be one reason why sex reversal occurs in B6-YDOM mice. If so, the strain-dependent phenomena may be reconciled by the epigenetic mechanism involving strain-dependent non-CpG methylation, although further study is needed using B6-YDOM mice. In addition, some studies have reported that CCWGG methylation causes transcriptional silencing [9, 30, 31, 40]. In contrast, CCWGG methylation at the Sry locus did not repress the gene expression in this study. CCWGG methylation seems to act as transcriptional promotion.

Recently, Lister et al. reported that human embryonic stem (ES) cells have abundant non-CpG methylation including CHG and

<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Fig. 4.</strong> Reciprocal DNA methylation between CpG and CCTGG sites at the Region II in ICR-J.</td>
<td><strong>Fig. 5.</strong> DNA methylation status depends on the mouse strain.</td>
</tr>
<tr>
<td><strong>A:</strong> DNA methylation status in Region II during development. Gon-S, gonadal somatic cells. <strong>B:</strong> DNA methylation status in the adult tissues. Tes-S, testicular somatic cells; Gonia, spermatogonia; Cyte, spermatocyte; R-tid, round spermatid. Open and closed circles indicate unmethylated and methylated CpGs, and open and closed diamonds indicate unmethylated and methylated CCTGGs, respectively. Asterisks (*) show the position of the CCTGG site. The letters c–f indicate the position of each CpG site, corresponding to those in Fig.1. Note that reciprocal methylation between CpG and CCTGG was highly consistent.</td>
<td><strong>A:</strong> Genomic structure of the 2.0-kb upstream region of mouse Sry. The circles and diamonds indicate CpG and CCTGG sites, respectively. Gray circles indicate CpG sites (h, nt7655; i, nt7767) present in the C57BL/6, 129S1, BALB/c DBA2, and C3H strains, but not in the ICR strains. The numbering is based on the registered nucleotide sequence in the GenBank database (accession number X67204). The letters c–f indicate the position of each CpG site corresponding to those in Fig.1. <strong>B:</strong> DNA methylation status of individual CpG and CCTGG sites in Region II in 11.5-dpc fetal gonads determined by bisulfite sequence analysis. <strong>C:</strong> Results of additional bisulfite sequencing of ICR-A, C57BL/6 and 129S1. Open and closed circles indicate unmethylated and methylated CpGs, and open and closed squares indicate unmethylated and methylated CCTGGs, respectively. Asterisks (*) show the position of the CCTGG site. The numbers in parentheses in (C) indicate the number of clones sequenced. Note that CCTGG methylation was not detected in the C57BL/6 and 129S1 strains.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Embryo</th>
<th>Gonad</th>
<th>Gon-S</th>
<th>Liver</th>
<th>Testis</th>
<th>Liver</th>
</tr>
</thead>
<tbody>
<tr>
<td>8.5 dpc</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>A</strong></td>
<td><strong>B</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adult</td>
<td>Tes-S</td>
<td>Gonia</td>
<td>Cyte</td>
<td>R-tid</td>
<td>Sperm</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
CHH sites (where H=A, C and T) [41]. Of the methylcytosines detected in the human ES cell line, 0.3% were in the CCWGG context [41]. Combined with our results, CCWGG methylation is conserved in mammals. The CCTGG methylation at the Sry locus is fragile, but further elucidation of the unique characteristics of non-CpG methylation should help in understanding mammalian epigenetic mechanisms.

Acknowledgments

We thank Dr MG Marinus (University of Massachusetts Medical School) for the GM2159 bacterial strain and Dr C Maeda for collecting the fetal gonads. This research was supported by the Japan Society for the Promotion of Science (JSPS, KN) and Grants-in-Aid for Scientific Research 21221008 (KS) and 20062003 (ST) from the Ministry of Education, Culture, Sports, Science and Technology of Japan.

References


