Nucleic Acids Research, 2003, Vol. 31, No. 16 ***4797–4804***
DOI: 10.1093/nar/gkg670

Methylation of CpG dinucleotides in the open reading frame of a testicular germ cell-specific intronless gene, *Tact1/Actl7b*, represses its expression in somatic cells

Mizue Hisano, Hiroshi Ohta, Yoshitake Nishimune and Masami Nozaki*

Department of Laboratory for Science Animal Experimentation, Research Institute for Microbial Diseases, Osaka University, 3-1, Yamadaoka, Suita, Osaka 565-0871, Japan

Received May 2, 2003; Revised and Accepted June 16, 2003

ABSTRACT

Methylation of CpG islands spanning promoter regions is associated with control of gene expression. However, it is considered that methylation of exonic CpG islands without promoter is not related to gene expression, because such exonic CpG islands are usually distant from the promoter. Whether methylation of exonic CpG islands near the promoter, as in the case of a CpG-rich intronless gene, causes repression of the promoter remains unknown. To gain insight into this issue, we investigated the distribution and methylation status of CpG dinucleotides in the mouse *Tact1/Actl7b* gene, which is intronless and expressed exclusively in testicular germ cells. The region upstream to the gene was poor in CpG, with CpG dinucleotides absent from the core promoter. However, a CpG island was found inside the open reading frame (ORF). Analysis of the methylation status of the *Tact1/Actl7b* gene including the 5′-flanking area demonstrated that all CpG sites were methylated in somatic cells, whereas these sites were unmethylated in the Tact1/Actl7b-positive testis. Transfection experiments with *in vitro*-methylated constructs indicated that methylation of the ORF but not 5′ upstream repressed *Tact1/Actl7b* promoter activity in somatic cells. Similar effects of ORF methylation on the promoter activity were observed in testicular germ cells. These are the first results indicating that methylation of the CpG island in the ORF represses its promoter in somatic cells and demethylation is necessary for gene expression in spermatogenic cells.

INTRODUCTION

In the genome of an adult vertebrate cell, 60–90% of the cytosines in CpG dinucleotides are methylated by DNA methyltransferase and the methylated cytosines tend to undergo deamination, resulting in a cytosine to thymine transition (1). This process leads to an overall reduction in the frequency of guanine and cytosine to ~40% of all nucleotides and a further reduction in the frequency of CpG dinucleotides to approximately a quarter of their expected frequency (2,3). The exception to CpG reduction in the genome is CpG islands, which have been estimated to constitute 1–2% of the mammalian genome (4).

Analyses of the positional relationship between CpG islands and mammalian genes have shown that CpG islands often include gene promoters and extend further downstream into transcribed regions (pro-CpG island) (5). Other CpG islands can also occur in downstream regions distant from the promoter (6) and some of these locate within exons (exonic CpG island). The persistence of CpG dinucleotides in the genome is dependent upon a lack of methylation of CpG. Pro-CpG islands of housekeeping genes are not methylated in any cell types and these genes are expressed ubiquitously. Conversely, some pro-CpG islands of tissue-specific genes are differentially methylated in non-expressing tissues. The methylation of pro-CpG islands may contribute to the repression of transcription noise and the control of tissue-specific gene expression (7). Additional studies show that *de novo* methylation of CpG islands residing within transcribed regions is permissive for transcription initiation (8–10) and that methylation of exonic CpG islands does not inhibit transcriptional elongation in mammalian cells (11). In contrast, methylation of the luciferase gene located immediately downstream of the promoter repressed RSV-LTR promoter activity ~10-fold (12) in mammalian cells. These data indicated that methylation of a transcribed portion of CpG-rich non-mammalian gene locating near a promoter can repress mammalian gene promoter activity. However, there are few studies concerning the methylation effects of a CpG-rich coding region of mammalian gene on its expression. To understand this issue, we examined the methylation effects of testicular germ cell-specific genes on their promoter activity.

Spermatogenesis is a sequential process by which undifferentiated germinal stem cells become mature spermatozoa, and the process consists of three well defined phases: the

*To whom correspondence should be addressed. Tel: +81 6 6879 8338; Fax: +81 6 6879 8339; Email: mnozaki@biken.osaka-u.ac.jp

mitosis of spermatogonia, the meiotic division of spermatocytes and spermiogenesis, in which haploid spermatids are transformed into morphologically and functionally differentiated spermatozoa (13). This complex and continuous process involves many germ cell-specific genes (14–16). These genes are strictly regulated in their expression in testicular germ cells. It is important to repress these genes in somatic cells and gene silencing is believed to be a major biological consequence of DNA methylation (17).

Several testis-specific genes are actually demethylated at CpG sites in testis, but methylated in non-expressing somatic tissues (18–23), indicating that further analyses of testicular genes will shed light on gene regulation by CpG methylation.

The *Tact1/Actl7b* gene (we call it *Tact1* hereafter) consists of a single exon and is exclusively expressed in haploid spermatid in testis. The specific expression is governed by an upstream sequence within the 2 kb region, as shown by transgenic analysis (24). Here we show that the *Tact1* gene contains a CpG island in its coding region, which is hypomethylated in testis and hypermethylated in somatic cells, and the *Tact1* promoter is repressed by methylation of the ORF. We further demonstrate that repression of the *Tact1* promoter by methylation within the gene occurs in testicular germ cells. These results indicated that internal methylation of the *Tact1* gene plays a part in repression in somatic cells in which the gene is not expressed and demethylation is necessary for expression in germ cells in testis.

MATERIALS AND METHODS

Generation of *Tact1* probes and Southern blotting

Tact1 probes for genomic Southern hybridization were generated by PCR using primers from the mouse genome sequence (24). For Probe 1, oligonucleotides corresponding to positions 304–323 (TMe1-3, 20mer oligonucleotide 5′-CAG GTT CAT TTG CAT GTC TG-3′) and 835–854 (TMe1-5, 20mer oligonucleotide 5′-TTT GTG TCC AAA GTA CCC CC-3′) downstream from the transcriptional terminal of the *Tact1* gene, and for Probe 2, corresponding to positions 1262–1281 (primer 2-5, 20mer oligonucleotide 5′-CAA GAA GAC CAG TGT GAC CT-3′) and 1570–1589 (primer 2-3, 20mer 5′-CTG AGC TGG CAA TAG TTG TG-3′) upstream from the transcriptional start site of the *Tact1* gene were synthesized. These primer sets (TMe1-3/TMe1-5 and 2-5/2-3) were used for PCR in a mixture containing 0.1 μg of mouse genomic DNA, 1 μM of each primer, 1× reaction buffer (TaKaRa), 0.4 mM dNTPs and 1 U of *TaKaRa Taq* DNA polymerase (TaKaRa). The reaction was run for 30 cycles with 1 min of denaturation at 94°C, 1 min of annealing at 60°C and 1.5 min of DNA synthesis at 72°C. A 10 min extension at 72°C was performed after completion of the 30 cycles. The PCR products of 551 bp generated with Probe 1 and 328 bp with Probe 2 were cloned into pT7 Blue T-vector (Novagen).

Genomic DNAs were prepared from an adult mouse (*129/Sv*) and 7-, 14-, 21- and 35-day-old mice (C57BL/6). These DNAs were digested with HindIII/EcoRV and a methyl-sensitive enzyme, HhaI or HpaII, or a methyl-insensitive enzyme, MspI. The resulting fragments were size fractionated, blotted on nitrocellulose filter and hybridized with labeled probes, as described previously (24). The signals were visualized using a Fuji Bio-image analyzer (BAS2000; Fuji Photo Film Co.).

Sodium bisulfite PCR sequencing

Bisulfite PCR sequencing was carried out as described (25) with minor modifications. Five micrograms of DNA digested with EcoRI was alkali-denatured in a total volume of 100 μl with 0.3 M NaOH at 42°C for 30 min. The denatured DNA solution was added to 1020 μl of 40.5% sodium bisulfite (pH 5), 60 μl of 10 mM hydroquinone and 10 μl of water, mixed gently, and incubated at 50°C for 16–18 h under mineral oil. This was followed by successive dialyses at 4°C in a large volume of (i) 5 mM sodium acetate/0.5 mM hydroquinone, pH 5.2, (ii) 0.5 mM sodium acetate (pH 5.2) and (iii) deionized water, to remove unreacted bisulfite. The dialyzed solution was dried using the freeze–dry method and resuspended in 100 μl of 10 mM Tris–HCl/0.1 mM EDTA, pH 7.5. NaOH was added to a final concentration of 0.3 M at room temperature for 10 min, followed by ammonium acetate (pH 7) to a final concentration of 3 M. The bisulfite-reacted DNA was ethanol precipitated and resuspended in 10 μl of 10 mM Tris–HCl/0.1 mM EDTA, pH 7.5. Primers for the bisulfite reaction were as follows: MCT2-5, 5′-GTA AGG AGG TTA GGA TGG AG-3′; MCT2-3, 5′-AAT CAT TAA CCT AAA TTC CC-3′; MCT3-5, 5′-AGA GAT GTT TTT ATG TTG AG-3′; MCT3-3, 5′-TCT TAA ACC TTA CCT AAC CC-3′; MCT10-5, 5′-TTG TTT TTT GGG TAG AGG AG-3′; MCT10-3, 5′-AAA AAA AAT AAA CCT AAA CC-3′; MCT11-5, 5′-GGT GAT GAT TTT TTT ATT TG-3′; MCT11-3, 5′-TCT ACT CAT ATA ACA AAA CC-3′; MCT12-5, 5′-TTA GAG ATG GGT ATT ATA TG-3′; MCT12-3B, 5′-TAA ACC ATA AAC TAC TCA AC-3′; MCT13-5, 5′-TGT TTT AGT GGG TTA ATT AG-3′; MCT13-3, 5′-ACA AAA ACA CTA CCA ACA CC-3′; MCT14-5, 5′-TTA GTT GAG TGG AAT TTG TG-3′; MCT14-3, 5′-AAC TCC CAT CTA AAT AAA CC-3′; MCT15-5, 5′-TGG GTT GTG ATG TTA TTA GG-3′; MCT15-3, 5′-CAA ACT CTA AAC TTT ATA AC-3′. The bisulfite-reacted genomic DNA was amplified by PCR. PCR amplifications were performed with nested primers in reaction mixtures containing 50 ng of DNA, 1 μM of each primer, 1× reaction buffer (TaKaRa), 0.4 mM dNTPs and 1 U of *TaKaRa Taq* DNA polymerase (TaKaRa). The reaction was run for 30 cycles with 1 min of denaturation at 94°C, 1 min of annealing at 55–60°C and 1.5 min of DNA synthesis at 72°C. A 1 min extension at 72°C was performed after completion of the 30 cycles. In the first instance, PCRs were performed with three primer pairs, MCT2 (MCT2-5/MCT2-3), MCT3 (MCT3-5/MCT3-3) or MCT12–14 (MCT12-5/MCT14-3). In the second, MCT10 (MCT10-5/MCT10-3), MCT11 (MCT11-5/MCT11-3) and MCT12 (MCT12-5/MCT12-3B) were amplified with the MCT2 PCR product, MCT13 (MCT13-5/MCT13-3) was amplified with the MCT12-14 PCR product, and MCT14 (MCT14-5/MCT14-3) and MCT15 (MCT15-5/MCT15-3) were amplified with the MCT3 PCR product. These second PCR products were ethanol precipitated and cloned into the vector pT7Blue. Individual clones were sequenced by the dideoxynucleotide chain-termination method using a ThermoSequenase kit (Amersham Pharmacia). The reaction products were analyzed using a LI-COR automatic sequencer Model 4000 (Li-COR).

Reporter plasmid constructs

The 5′-flanking fragment of the *Tact1* gene was used for generating reporter constructs to evaluate promoter activity. The fragment from –1938 to +44 was cloned into pEGFP-1 (T1G) or pGL3-Basic (T1L), as described previously (24). T1L was subjected to PCR with the 5′ primer TA-up2 (5′-CCC TTA AGT CAC CAA TGA GT-3′) and the 3′ primer RV4 (5′-GAC GAT AGT CAT GCC CCG CG-3′), corresponding to the luciferase upper region, to generate TSL (–130 to +44 into pGL3). A SmaI/HindIII 4.5 kb genomic fragment containing *Tact1* and 1.7 kb 5′- and 1.4 kb 3′-flanking sequences was cloned into pHSG399 to generate the TactS/H construct. The construct used for the standardization of luciferase activity by *in vivo* electroporation was an OAZ-t promoter (26) in pRL (Promega).

Cell culture and luciferase assay

NIH3T3 cells were spread at low density in 100 mm diameter plates and cultured in Dulbecco's modified Eagle's medium (DMEM; Life Technologies) containing 10% (v/v) fetal bovine serum (FBS) (Gibco BRL) for 3 days at 37°C in 5% CO_2 in air. Subsequently, they were seeded in 12-well tissue culture dishes at a density of 10^5 cells/well and grown for 24 h at 37°C in 5% CO_2 in DMEM with 10% (v/v) FBS. The culture dishes were washed with serum-free DMEM to remove old medium and the cells were transfected with Lipofectamine plus reagents (Gibco BRL) containing the constructed vectors and cultured in serum-free DMEM for 3 h at 37°C in 5% CO_2 and subsequently for 48 h in DMEM with 10% (v/v) FBS. The transfection efficiencies were normalized with respect to the co-transfection of 1.0 μg of pRL-SV40 vector (Promega). The harvesting of cells and assay of luciferase activity were performed using a PicaGene Dual Seapansy™ Luminescence kit (Nippon Gene) according to the manufacturer's recommendation.

Reverse transcription (RT)–PCR detection of mRNA

Total RNA was isolated from transfected NIH3T3 cells using Trizol reagent (Gibco BRL). Five micrograms of total RNA was subjected to first-strand cDNA synthesis with oligo-dT primer and Superscript™ II reverse transcriptase. PCR was performed in a mixture containing the RT product, 1 μM of each primer, 1× reaction buffer (TaKaRa), 0.4 mM dNTPs and 1 U of *TaKaRa Taq* DNA polymerase (TaKaRa) with the forward primer 1-5 (5′-AGA TGG CGA CAA AGA ACA GCC CTA GCC CTA A-3′) and the reverse primer 1-3 (5′-CAG AGA TGC CTC CAT GTT GAG CAG CTC ATG-3′) to detect a 338 bp fragment of the 5′ region of the *Tact1* transcript. The reaction was run for 30 cycles with 1 min of denaturation at 94°C, 1 min of annealing at 60°C and 1.5 min of DNA synthesis at 72°C, and then a 10 min extension at 72°C. Control detection of β-actin was performed with forward (MBA1, 5′-GTG GGC CGC CCT AGG CAC CA-3′) and reverse (MBA2, 5′-TGG CCT TAG GGT TCA GGG GG-3′) primers under the same conditions.

In vitro methylation of DNAs

Reporter constructs and DNA fragments were methylated *in vitro* with 1 U of SssI methylase (New England Biolabs) for each microgram of DNA in the presence of 0.16 mM *S*-adenosylmethionine at 37°C for 3 h. It was confirmed by digestion with the restriction enzyme HhaI followed by electrophoresis that the methylation at all CpG sites was completed (data not shown).

In vivo electroporation

Reporter constructs containing the *Tact1* upstream or OAZt promoter, reporter gene and SV40 poly(A) additional signal, were amplified by PCR with primers and purified with a SUPREC™-PCR column (TaKaRa). For luciferase: 5′ primer (RV primer 3, 5′-CTA GCA AAA TAG GCT GTC CC-3′) and 3′ primer (RV primer 4, 5′-GAC GAT AGT CAT GCC CCG CG-3′). For EGFP: 5′ primer (Green5′, 5′-TTC TTT CCT GCG TTA TCC C-3′) and 3′ primer (Green3′, 5′-CGA TTT CGG CCT ATT GGT TA-3′). For pRL: 5′ primer (RL5, 5′-TTT TGC TCA CAT GGC TCG AC-3′) and 3′ primer (RL3, 5′-CGC ACA TTT CCC CGA AAA GT-3′). A DNA solution containing 0.5 μg/ml of each DNA fragment, 100 μM of general caspase inhibitor (Z-VAD-FMK; R&D Systems) and 0.1 mg/ml trypan blue (Nacalai tesque) was injected into the seminiferous tubules. The injection was performed according to the method of Ogawa *et al.* via the efferent ducts (27). Approximately 12 μl of DNA solution was injected into the seminiferous tubules. Electron pulses were charged with an electric pulse generator (Electroporator CUY-21; Tokiwa-Science). The testes were held between a pair of tweezer-type electrodes (Tokiwa-Science) and square electric pulses were applied eight times in four different directions at 40 V, for 50 ms durations with 950 ms intervals. After 16–20 h, transfected testes were collected for luciferase assay and histological examination. For histology, testes were fixed in 4% paraformaldehyde for 16 h. After the removal of extra paraformaldehyde by washing with PBS, the testes were dehydrated with 100% acetone for 1 h and embedded in glycol methacrylate (Thechnovit 8100; Heraeus Kulger). Serial sections of the testes were prepared at a thickness of 5 μm. The sections were examined under a fluorescent microscope for the expression of EGFP. After green fluorescent was photographed, the sections were stained with hematoxylin and observed under a photomicroscope.

RESULTS

The mouse *Tact1* gene contains a CpG island

The *Tact1* gene consists of a single exon of 1407 bp (24) containing 56 CpGs, of which 55 are located in the ORF (Fig. 1). Dense clustering of CpG dinucleotides is found at the 5′ ends of many genes, some of which are called CpG islands. CpG islands are defined as sequences longer than 200 or 500 bp, with a ratio of the observed to expected number of CpG (CpGo/e) >0.6 and a G+C frequency (%GC) >0.5 (3,28). The overall sequence and the ORF of the *Tact1* gene are rich in CpG (CpGo/e = 0.46; %GC = 0.586 and CpGo/e = 0.49; %GC = 0.593; 1284 bp) and 575 bp within the ORF corresponded to the definition of a CpG island (size = 575 bp; CpGo/e = 0.605; %GC = 0.578) (Fig. 1). The *Tact1* CpG island is located 200 bp downstream from the transcription initiation site. On the other hand, the 5′ upstream region of the *Tact1* gene is not CpG-rich (CpGo/e = 0.180; %GC = 0.472). Most CpG islands are distributed 5′ upstream to the first exon and first intron

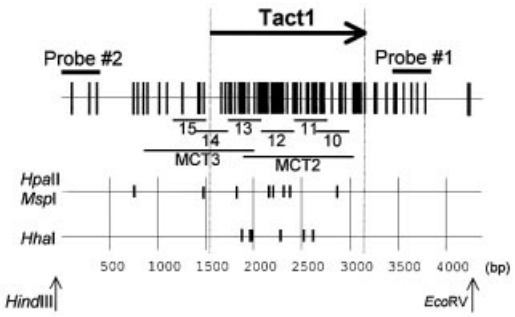

Figure 1. Distribution of CpGs and restriction sites in the *Tact1* gene, and primer design. At the top of the figure, vertical bars indicate all CpGs in the 4.4 kb HindIII/EcoRV (vertical arrows) fragment. A horizontal arrow shows the *Tact1* transcriptional site, and bold and thin lines show probes used for genomic Southern and PCR-synthesized regions, respectively. At the bottom of the figure, methyl-sensitive HhaI and HpaII, and methyl-insensitive MspI sites in the 4.4 kb HindIII/EcoRV fragment are shown.

(29). However, some CpG islands present within the ORF tend to have a lower %GC on average than the 5′ CpG islands, tend to have a slightly decreased mean for the occurrence of CpGo/e and tend to be shorter (3,6). In this context, the CpG island within the ORF of *Tact1* is unique and similar in features to some repetitive sequences including *Alu* (30).

CpG sites of the *Tact1* gene were hypermethylated in somatic cells and hypomethylated in adult testis

CpG islands of housekeeping genes are always unmethylated (2). CpG islands of some tissue-specific genes change their methylation status in a tissue- or development-specific manner (31). The *Tact1* gene is expressed in haploid spermatids in testis but not in any other tissues (16,24). To elucidate whether the CpG island within the *Tact1* gene is methylated in various tissues, HpaII/MspI or HhaI digestion and Southern blot analysis were performed using genomic DNA prepared from mouse tissues. Figure 2A shows the profiles of Southern blots of the digested DNAs from testis, kidney and liver of adult mice hybridized with the upstream (2) and downstream (1) probes as shown in Figure 1. Digestion of genomic DNA with EcoRV/HindIII generated a 4.4 kb fragment containing six HhaI and eight MspI/HpaII sites. Neither HhaI nor HpaII could digest the EcoRV/HindIII fragment from liver and kidney, showing that all sites are completely methylated in these tissues. The genomic DNA from testis was almost completely digested with either HpaII or HhaI within the *Tact1* gene, although the HpaII site, 0.9 kb upstream from the gene, was not digested at all and the HpaII site, 0.24 kb upstream was only partially digested (Fig. 2A). These findings indicated the possibility that internal CpG sites of the *Tact1* gene are demethylated in the adult testis and completely methylated in kidney and liver. We found that genomic DNA from 7-day-old testis was not digested (Fig. 2B). A 4.4 kb fragment from the 14-, 21- and 35-day-old testis DNA gradually disappeared in an age-dependent manner and was almost digested in 35-day-old as in adult testis (Fig. 2). Spermatogenesis initiates a few days after birth and 7-day-old testis contains mainly somatic cells and fewer spermatogonia (32). The 35-day-old testis contains spermatogonia, spermatocytes, spermatids and spermatozoa, and a very small number of somatic cells. These results indicated that somatic cells in

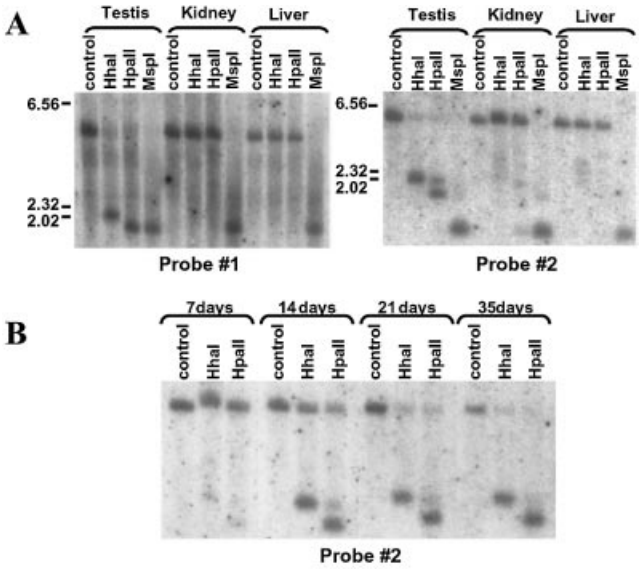

Figure 2. Testis DNAs were not methylated but somatic DNAs were methylated at the *Tact1* locus. (**A**) DNAs were isolated from adult mouse testis, kidney and liver, and digested with EcoRV and HindIII. They were then digested with HhaI, HpaII or MspI, and subjected to Southern blot analysis for the downstream (probe 1) or upstream (probe 2) region as shown in Figure 1. The control fragment was digested with EcoRV and HindIII (4.4 kb). The bars in the left margin indicate the size of the marker (λ/HindIII). (**B**) DNAs were isolated from day 7, 14, 21 and 35 mouse testes, and subjected to Southern blot analysis using probe 2.

testis are methylated, and during the differentiation of spermatogenic cells, methylated CpG sites might be demethylated.

In the testicular germ cells that express the *Tact1* gene, upstream and internal CpG sites appeared to be demethylated, but the method did not allow us to evaluate the precise extent of demethylation at each site. We resorted to the sodium bisulfite modification procedure to assess the methylation status of CpG sites in the *Tact1* gene, of which 49 are located between +60 and +1123 and five between –382 and –130 (Fig. 1). The analyzed *Tact1* fragments were amplified by PCR applied to genomic DNA treated with sodium bisulfite. The PCR products were then cloned and several clones were sequenced. We observed that the CpG dinucleotides were hardly methylated at all not only in the CpG island but also at other CpG sites including 5′ upstream of the *Tact1* gene in the adult testis (Fig. 3). In adult testis, there are small numbers of somatic cells, although most cells are spermatogenic. As each sequence represents the methylation profile of a single DNA molecule, it is possible to detect methylated sequences in somatic cells in testis. In contrast to the adult testis, liver that did not express the gene contained a fully methylated *Tact1* sequence. All of the CpG sites were heavily methylated in liver, kidney and prepuberal testis suggesting that the *Tact1* gene is methylated in all somatic cells and methylation might be required for the arrest of *Tact1* transcription.

Methylation of CpG sites within the *Tact1* gene represses expression

In a previous study, data from our transgenic studies indicated that the upstream sequence of *Tact1* functions as a testis germ cell-specific promoter. However, *in vitro*, the *Tact1* upstream

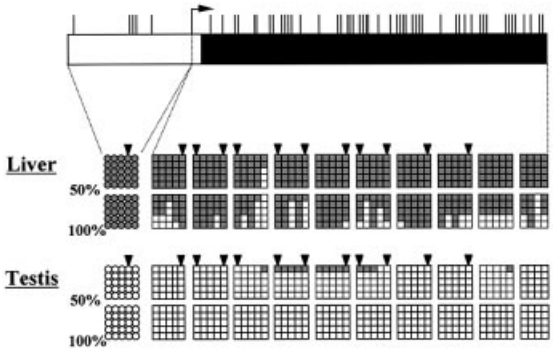

Figure 3. Methylation profile of CpGs at the *Tact1* locus in the testis and liver analyzed by the bisulfite PCR method. The analyzed region of the gene is represented at the top with a broken arrow at the transcription initiation site, vertical bars at the locations of CpGs, and a closed square representing the ORF. Genomic DNA was treated with bisulfite and the *Tact1* region was amplified by PCR, cloned and sequenced. Diagrams show the DNA methylation levels of liver (top) and testis (bottom). Boxes and circles indicate downstream from the transcriptional initiation site and 5′ flank, respectively. The CpG methylation status of each site is shown by a percentage based on sequence analysis. Closed squares or circles correspond to methylated cytosines, and open squares or circles correspond to unmethylated cytosines. Vertical numbers of squares or circles represent the percentage of methylated clones. Arrowheads indicate the cytosine of methylation-sensitive restriction sites.

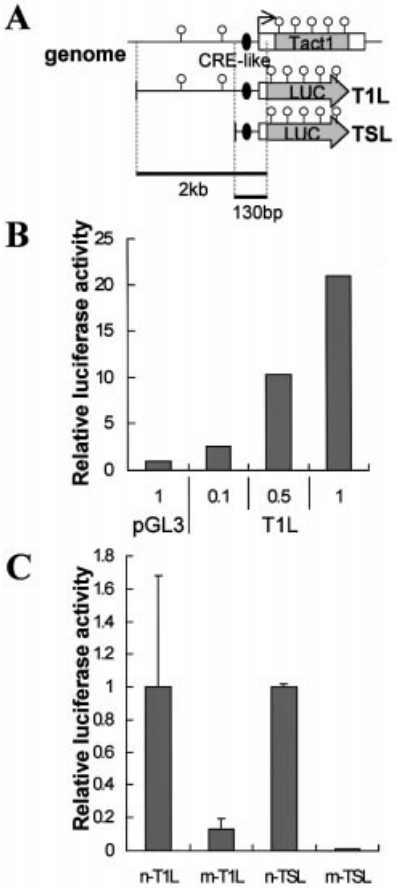

Figure 4. Effects of methylation of the reporter gene on *Tact1* promoter activity determined by transient assay with NIH3T3. (**A**) Diagrams of T1L and TSL constructs. Construct of T1L or TSL including the 2 kb or 130 bp *Tact1* upstream region linked to the luciferase gene. Open and gray squares indicate the untranslated region and ORF of *Tact1*, respectively. Black ovals indicate the positions of the *CRE*-like motif and open circles show the region that includes CpGs. (**B**) T1L construct (0.1, 0.5 or 1 μg) was transiently transfected into NIH3T3 cells, and luciferase activity was analyzed. pGL3 is a negative control vector without a promoter. Fold luciferase activity was normalized to that obtained with pGL3 as 1. (**C**) Reporter constructs were methylated by SssI methylase (m-T1L or m-TSL) or not methylated (n-T1L or n-TSL) and transiently transfected into NIH3T3 cells. The luciferase activity of the non-methylated construct was set to 1.0.

sequence was found to be active when T1L was transiently transfected in the NIH3T3 somatic cell line (Fig. 4B). A possible explanation is that regulatory constraints, which govern the transcriptional repression of the endogenous gene in somatic cells, do not operate on the construct when introduced into the cells by transient transfection. Other studies have shown some germ cell-specific genes to be regulated by methylation at CpG dinucleotides (20,23,33). Since five CpG sites located upstream of the *Tact1* gene were methylated in somatic cells and demethylated in testis (Fig. 3), it is possible that *Tact1* is also regulated by CpG methylation of the upstream sequence. To obtain more direct evidence that this modification affects the expression, T1L and TSL were methylated with the SssI methylase and tested for activity. T1L contained the 2 kb *Tact1* upstream region (–1938 to +44) with 16 CpG sites and TSL contained the 130 bp (–130 to +44) core promoter with no CpG sites linked to luciferase (Fig. 4A). We transiently transfected these constructs into NIH3T3 cells and assayed the effect of methylation on the direct expression of the luciferase gene (Fig. 4C). Both constructs showed a significant reduction (>80%) in activity following SssI methylase treatment compared with untreated controls. These results showed the repression of expression of these constructs due to CpG methylation except for the upstream region. Overall, the luciferase gene is CpG-rich (size of ORF = 1650 bp; CpGo/e = 1.075; %GC = 0.468), although these values do not correspond to the definition of a CpG island. These results indicated that the methylation of CpG sites in the ORF downstream from the promoter may repress *Tact1* promoter activity.

To confirm this possibility in the *Tact1* gene, we prepared *Tact1* genomic DNA containing 2 kb of the upstream region and the *Tact1* transcription unit and assayed the effects of methylation on the expression by RT–PCR. We could clearly detect the band of RT–PCR products from the transfectant using non-methylated *Tact1* genomic DNA, whereas RT–PCR products from methylated DNA could only be observed faintly, indicating that methylation of CpG sites in the ORF of the *Tact1* gene repressed *Tact1* promoter activity (Fig. 5).

Methylation of CpG sites within the ORF inhibits *Tact1* promoter activity in testicular germ cells

From our data, methylation of either the reporter or the *Tact1* ORF exerted a strong negative effect on the *Tact1* promoter in NIH3T3 cells suggesting that methylation inactivates the endogenous *Tact1* promoter in somatic cells. *Tact1* is hypomethylated and expressed in testicular germ cells. If *Tact1* is methylated in testis, *Tact1* expression would be repressed. To test this point, we examined the effects of methylation on expression in testicular germ cells *in vivo*. We

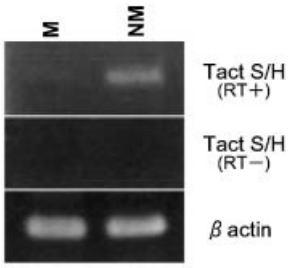

Figure 5. Tact1 promoter activity was repressed by methylated CpGs in the *Tact1* ORF. A HindIII/SmaI fragment (4.5 kb) containing the *Tact1* gene was cloned into pHSG399 vector (Tact S/H), which was methylated (M) by SssI methylase or not (NM) and transfected into NIH3T3 cells. Transcription of the *Tact*1 gene was analyzed by RT–PCR and compared with that of β-actin.

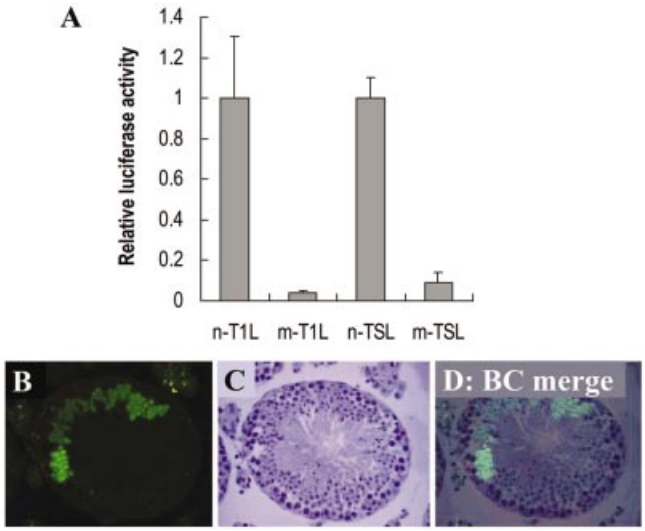

Figure 6. *In vivo* electroporation assay showed that *Tact1* promoter activity was repressed by methylated CpG in its ORF. (**A**) T1L and TSL were methylated by SssI methylase, injected into seminiferous tubules and electric pulses were applied. Luciferase activities of methylated or non-methylated T1L (m-T1L or n-T1L) and TSL (m-TSL or n-TSL) were standardized with each of the non-methylated constructs. (**B**–**D**) T1G is expressed only in haploid germ cells. The T1G fragment including the 2 kb Tact1 upstream region (–1938 to +44 bp) linked to EGFP, was transfected into seminiferous tubules. (B) EGFP expression of testis cross-sections. (C) Hematoxylin staining of (B). (D) A merged image of (B) and (C).

introduced the T1G DNA fragment into testicular germ cells by a combination of DNA injection into seminiferous tubules and subsequent *in vivo* electroporation, and detected EGFP signals specifically in haploid germ cells (Fig. 6B–D). These results showed that the 2 kb upstream of *Tact1* drives the EGFP gene in spermatids but not in Sertoli cells, although DNA should be introduced efficiently into Sertoli cells (34,35). To examine the effects of methylation quantitatively, we used T1L and TSL, and assayed their luciferase activity. These DNA fragments with or without methylation were injected into adult mice (5 weeks old), and the luciferase activities were determined 16–20 h after *in vivo* electroporation. These reporting fragments contain the *Tact1* promoter, reporter luciferase gene and SV40 poly(A) addition signal, but not the vector sequence. Both constructs exhibited a significant reduction (80–90%) in activity following methylation compared with the unmethylated controls (Fig. 6A), showing that methylation of the reporter gene represses *Tact1* promoter activity and the mechanism of repression using DNA methylation functions in testicular germ cells. In addition, these experiments can exclude the possibility that methylation of the vector sequence affects promoter activity. The findings strongly suggested that demethylation of the ORF during spermatogenesis is necessary to express the *Tact1* gene.

DISCUSSION

In this report, we showed that the testicular germ cell-specific gene *Tact1* consisting of a single exon contains a CpG island in its coding region. CpG sites in *Tact1* are hypermethylated in somatic cells and hypomethylated in testis. The methylation of CpG sites in the coding region of *Tact1* represses the promoter activity in somatic cells. In addition, this mechanism of repression via methylation of CpG also occurs in testicular germ cells.

The ORF of the *Tact1* gene is rich in CpG, with a 575 bp stretch defined as a CpG island. On the other hand, the 5′-flanking region of the gene has no CpG islands, although CpG islands are believed to be associated with the 5′ region of genes including promoters (pro-CpG islands). Recently, however, many CpG islands have also been found within exons but not over the transcription start site of genes (exonic CpG islands) (3). Exonic CpG islands are shorter, with less CpGo/e and less %GC than pro-CpG islands. They are dominantly found in repeated transposable elements including *Alu* (36), as well as in genes of limited expression (9) and intronless pseudogenes (37). These transposable CpG-rich sequences were integrated into the genome randomly and retain their CpG content. The *Tact1* gene is intronless and believed to be generated by retroposition, but with no short direct repeats on both sides of the gene and poly(A) remnant in the 3′ downstream (24). The present study showed that a CpG island exists in the exon and CpG is rare in the flank, supporting the notion that the *Tact1* gene is generated by retroposition. In general for retroposons, because mRNAs lack a promoter in the 5′-flanking sequence of the genomic DNA, the inserted cDNA reverse-transcribed from the mRNA cannot be expressed in any tissues. As a result, most retroposons evolve into non-functional processed pseudogenes. Nevertheless, expressing and functional intronless genes have been frequently reported in testis, implying that testicular germ cells provide an environment for the transcription of retroposons, which facilitates the expression from promoter-like sequences. The acquisition of promoter activity leads to the generation of functional retroposons, which may play important roles in germ cell development in mammals.

We found that CpG sites in the *Tact1* gene including an internal CpG island and the 5′ upstream region were methylated in somatic tissues, but not in testicular germ cells. Conventional wisdom holds that CpG islands of housekeeping genes are unmethylated in almost all tissues, with the exception of X inactivation, imprinting and tumors (29). Sp1 sites within the core sequence of CpG islands were believed to be required to prevent methylation (38,39), but there are no Sp1 motifs around the *Tact1* gene. Therefore,

methylation of the CpG island of the *Tact1* gene in somatic cells might be due to the absence of a target element which can prevent methylation by methyltransferases. Why is the *Tact1* gene demethylated in testis? The data in Figure 2B showed that demethylation of *Tact1* in testis occurs between 7 and 14 days postpartum, during which spermatogonia actively proliferate, indicating that the *Tact1* gene is passively demethylated during the spermatogonia expansion process. Such a demethylation mechanism has been postulated by *in situ* alterations of methylated DNA during germ cell development in testis using anti-5-methylcytosine antibodies. However, it should be noted that demethylated chromatids retain mildly methylated regions, indicating that a part of the newly synthesized chromosomes maintain a methylated status (40). Methylation imprints of H19 begin to be acquired in prenatal gonocytes, are completed in the pachytene stage, which occurs after birth, and are maintained up to spermatozoa (41), indicating that restricted sites within the genome are actually methylated during spermatogenesis. This is consistent with the accumulation of Dnmt1 protein in spermatogonia and preleptotene spermatocytes (42). *Alu* repeats also containing a CpG island in the ORF are hypermethylated in somatic cells and hypomethylated in the male germ line (43). A specific *Alu* binding protein has been isolated from spermatozoa, which selectively protects *Alu* elements from methylation and may be responsible for the undermethylated state of *Alu*s in the male germ line (44). A similar situation in *Alu* implies that demethylation of the *Tact1* gene in spermatogenic cells is also due to some methylation-protecting mechanism.

The present results showed that methylation of the firefly (*Photinus pyralis*) luciferase coding region, which is rich in CpG (CpGo/e = 1.08; %GC = 0.468), represses *Tact1* promoter activity. Repression of transcription by CpG methylation within the luciferase gene has been reported using an RSV LTR-luciferase construct (12). In addition, transgenes utilizing the bacterial ORF including the *lac* repressor, which is also CpG-rich, are hypermethylated and transcriptionally silent in all somatic tissues, but expressed in testis of transgenic mice (45). These repressive effects can be explained by the presence of a mechanism that might have evolved originally to allow for the repression of the expression of parasitic DNA sequences that spread through the mammalian genome. In contrast, despite being a single copy functional gene, *Tact1* is shown to be regulated by CpG methylation in its ORF. So far, several testicular germ cell-specific genes repressed by CpG methylation have been reported and they are regulated by methylation within the promoter region (18,20–23). Therefore, this is the first report of regulation by CpG methylation in the ORF of a testicular germ cell-specific functional gene. It is intriguing that the regulation of genes involved in germ cell function is likely to be analogous to the intrinsic defense mechanism against parasitic DNAs.

The internal CpG island is located 200 bp downstream of the transcription initiation site of the *Tact1* gene and there are only four CpG sites upstream of this CpG island. The methylation of dense CpG sites in the ORF may be significant for repression of the promoter activity. The significant repression of the *Tact1* promoter activity by CpG methylation in the exon of the gene could be affected at either the transcription elongation or the initiation level. Some mammalian genes contain exonic CpG islands, located between exon 2 and exon 11, a considerable distance downstream of the promoter (46,47). Methylation of these exonic CpG islands could not prevent transcription indicating that transcriptional elongation is not inhibited by methylation itself and that the methylation-induced local chromatin structure is not sufficient to eliminate elongation. Accordingly, it seems unlikely that DNA methylation would function to prevent transcription elongation in *Tact1*. Alternatively, the transcriptional inhibition by CpG methylation of the ORF near the promoter could be to prevent promoter activity at the level of the chromatin structure. The local chromatin structure is changed by the binding of methyl-CpG binding domain (MBD) proteins to a short patch of methylated DNA and these proteins subsequently recruit other proteins including sin3A and histone deacetylase (48). The local chromatin structure might spread through unmethylated regions. Depending on the sizes of the methylated and unmethylated regions, the distance of the methylated region from the promoter, and strength of the promoter activity, this chromatin structure could silence adjacent unmethylated regions. The *Tact1* exonic CpG island, which is 575 bp long, is located 200 bp from the promoter and hypermethylated in somatic cells. This situation may be significant in preventing the promoter activity.

We showed that the mechanism of repression by methylation is also present in testicular germ cells. We suggest that MBD protein plays some role in the repression of transcription. This is consistent with the fact that some MBD proteins are present in testis. The overall level of methylation in chromosomes has been observed to increase after meiosis (49). DNA methylation and MBD proteins may play important roles in spermiogenesis. In conclusion, although transcriptional repression of the *Tact1* gene may require methylation in somatic cells, reactivation appears to require the complete demethylation of the ORF.

ACKNOWLEDGEMENTS

This work was supported by a Grant-in-Aid for Scientific Research on Priority Areas (C) Genome Science from the Ministry of Education, Culture, Science and Technology of Japan.

REFERENCES

1. Kass,S.U., Pruss,D. and Wolffe,A.P. (1997) How does DNA methylation repress transcription? *Trends Genet.*, **13**, 444–449.
2. Bird,A.P. (1986) CpG-rich islands and the function of DNA methylation. *Nature*, **321**, 209–213.
3. Takai,D. and Jones,P.A. (2002) Comprehensive analysis of CpG islands in human chromosomes 21 and 22. *Proc. Natl Acad. Sci. USA*, **99**, 3740–3745.
4. Antequera,F. and Bird,A. (1993) Number of CpG islands and genes in human and mouse. *Proc. Natl Acad. Sci. USA*, **90**, 11995–11999.
5. Bird,A.P. (1987) CpG islands as gene markers in the vertebrate nucleus. *Trends Genet.*, **3**, 342–347.
6. Larsen,F., Gundersen,G., Lopez,R. and Prydz,H. (1992) CpG islands as gene markers in the human genome. *Genomics*, **13**, 1095–1107.
7. Siegfried,Z., Eden,S., Mendelsohn,M., Feng,X., Tsuberi,B.Z. and Cedar,H. (1999) DNA methylation represses transcription *in vivo*. *Nature Genet.*, **22**, 203–206.
8. Issa,J.P., Vertino,P.M., Boehm,C.D., Newsham,I.F. and Baylin,S.B. (1996) Switch from monoallelic to biallelic human IGF2 promoter

methylation during aging and carcinogenesis. *Proc. Natl Acad. Sci. USA*, **93**, 11757–11762.
9. Jones,P.A. (1999) The DNA methylation paradox. *Trends Genet.*, **15**, 34–37.
10. Zion,M., Ben-Yehuda,D., Avraham,A., Cohen,O., Wetzler,M., Melloul,D. and Ben-Neriah,Y. (1994) Progressive *de novo* DNA methylation at the bcr-abl locus in the course of chronic myelogenous leukemia. *Proc. Natl Acad. Sci. USA*, **91**, 10722–10726.
11. Gonzalgo,M.L., Hayashida,T., Bender,C.M., Pao,M.M., Tsai,Y.C., Gonzales,F.A., Nguyen,H.D., Nguyen,T.T. and Jones,P.A. (1998) The role of DNA methylation in expression of the p19/p16 locus in human bladder cancer cell lines. *Cancer Res.*, **58**, 1245–1252.
12. Hsieh,C.L. (1997) Stability of patch methylation and its impact in regions of transcriptional initiation and elongation. *Mol. Cell. Biol.*, **17**, 5897–5904.
13. Hecht,N. (1986) Regulation of gene expression during mammalian spermatogenesis. In Pedersen,R.A. (ed.), *Experimental Approaches to Mammalian Embryonic Development*. Cambridge University Press, Cambridge, pp. 153–193.
14. Wolgemuth,D.J. and Watrin,F. (1991) List of cloned mouse genes with unique expression patterns during spermatogenesis. *Mamm. Genome*, **1**, 283–288.
15. Penttila,T.L., Yuan,L., Mali,P., Hoog,C. and Parvinen,M. (1995) Haploid gene expression: temporal onset and storage patterns of 13 novel transcripts during rat and mouse spermiogenesis. *Biol. Reprod.*, **53**, 499–510.
16. Fujii,T., Tamura,K., Masai,K., Tanaka,H., Nishimune,Y. and Nojima,H. (2002) Use of stepwise subtraction to comprehensively isolate mouse genes whose transcription is up-regulated during spermiogenesis. *EMBO Rep.*, **3**, 367–372.
17. Colot,V. and Rossignol,J.L. (1999) Eukaryotic DNA methylation as an evolutionary device. *Bioessays*, **21**, 402–411.
18. Iannello,R.C., Young,J., Sumarsono,S., Tymms,M.J., Dahl,H.H., Gould,J., Hedger,M. and Kola,I. (1997) Regulation of Pdha-2 expression is mediated by proximal promoter sequences and CpG methylation. *Mol. Cell. Biol.*, **17**, 612–619.
19. Zhang,L.P., Stroud,J.C., Walter,C.A., Adrian,G.S. and McCarrey,J.R. (1998) A gene-specific promoter in transgenic mice directs testis-specific demethylation prior to transcriptional activation *in vivo*. *Biol. Reprod.*, **59**, 284–292.
20. De Smet,C., Lurquin,C., Lethe,B., Martelange,V. and Boon,T. (1999) DNA methylation is the primary silencing mechanism for a set of germ line- and tumor-specific genes with a CpG-rich promoter. *Mol. Cell. Biol.*, **19**, 7327–7335.
21. Singal,R., vanWert,J., Bashambu,M., Wolfe,S.A., Wilkerson,D.C. and Grimes,S.R. (2000) Testis-specific histone H1t gene is hypermethylated in nongerminal cells in the mouse. *Biol. Reprod.*, **63**, 1237–1244.
22. Kroft,T.L., Jethanandani,P., McLean,D.J. and Goldberg,E. (2001) Methylation of CpG dinucleotides alters binding and silences testis-specific transcription directed by the mouse lactate dehydrogenase C promoter. *Biol. Reprod.*, **65**, 1522–1527.
23. Xie,W., Han,S., Khan,M. and DeJong,J. (2002) Regulation of ALF gene expression in somatic and male germ line tissues involves partial and site-specific patterns of methylation. *J. Biol. Chem.*, **277**, 17765–17774.
24. Hisano,M., Yamada,S., Tanaka,H., Nishimune,Y. and Nozaki,M. (2003) Genomic structure and promoter activity of the testis haploid germ cell-specific intronless genes, Tact1 and Tact2. *Mol. Reprod. Dev.*, **65**, 148–156.
25. Hajkova,P., El-Maarri,O., Engemann,S., Oswald,J., Olek,A. and Walter,J. (2002) DNA-methylation analysis by the bisulfite-assisted genomic sequencing method. *Methods Mol. Biol.*, **200**, 143–154.
26. Ike,A., Yamada,S., Tanaka,H., Nishimune,Y. and Nozaki,M. (2002) Structure and promoter activity of the gene encoding ornithine decarboxylase antizyme expressed exclusively in haploid germ cells in testis (OAZt/Oaz3). *Gene*, **298**, 183–193.
27. Ogawa,T., Arechaga,J.M., Avarbock,M.R. and Brinster,R.L. (1997) Transplantation of testis germinal cells into mouse seminiferous tubules. *Int. J. Dev. Biol.*, **41**, 111–122.
28. Gardiner-Garden,M. and Frommer,M. (1987) CpG islands in vertebrate genomes. *J. Mol. Biol.*, **196**, 261–282.
29. Cross,S.H. and Bird,A.P. (1995) CpG islands and genes. *Curr. Opin. Genet. Dev.*, **5**, 309–314.
30. Schmid,C.W. (1998) Does SINE evolution preclude Alu function? *Nucleic Acids Res.*, **26**, 4541–4550.
31. Thomassin,H., Flavin,M., Espinas,M.L. and Grange,T. (2001) Glucocorticoid-induced DNA demethylation and gene memory during development. *EMBO J.*, **20**, 1974–1983.
32. Bellve,A.R., Cavicchia,J.C., Millette,C.F., O'Brien,D.A., Bhatnagar,Y.M. and Dym,M. (1977) Spermatogenic cells of the prepuberal mouse. Isolation and morphological characterization. *J. Cell Biol.*, **74**, 68–85.
33. Iannello,R.C., Gould,J.A., Young,J.C., Giudice,A., Medcalf,R. and Kola,I. (2000) Methylation-dependent silencing of the testis-specific Pdha-2 basal promoter occurs through selective targeting of an activating transcription factor/cAMP-responsive element-binding site. *J. Biol. Chem.*, **275**, 19603–19608.
34. Yamazaki,Y., Fujimoto,H., Ando,H., Ohyama,T., Hirota,Y. and Noce,T. (1998) In vivo gene transfer to mouse spermatogenic cells by deoxyribonucleic acid injection into seminiferous tubules and subsequent electroporation. *Biol. Reprod.*, **59**, 1439–1444.
35. Yomogida,K., Yagura,Y. and Nishimune,Y. (2002) Electroporated transgene-rescued spermatogenesis in infertile mutant mice with a sertoli cell defect. *Biol. Reprod.*, **67**, 712–717.
36. Ponger,L., Duret,L. and Mouchiroud,D. (2001) Determinants of CpG islands: expression in early embryo and isochore structure. *Genome Res.*, **11**, 1854–1860.
37. Glusman,G., Sosinsky,A., Ben-Asher,E., Avidan,N., Sonkin,D., Bahar,A., Rosenthal,A., Clifton,S., Roe,B., Ferraz,C. *et al.* (2000) Sequence, structure and evolution of a complete human olfactory receptor gene cluster. *Genomics*, **63**, 227–245.
38. Macleod,D., Charlton,J., Mullins,J. and Bird,A.P. (1994) Sp1 sites in the mouse aprt gene promoter are required to prevent methylation of the CpG island. *Genes Dev.*, **8**, 2282–2292.
39. Brandeis,M., Frank,D., Keshet,I., Siegfried,Z., Mendelsohn,M., Nemes,A., Temper,V., Razin,A. and Cedar,H. (1994) Sp1 elements protect a CpG island from *de novo* methylation. *Nature*, **371**, 435–438.
40. Coffigny,H., Bourgeois,C., Ricoul,M., Bernardino,J., Vilain,A., Niveleau,A., Malfoy,B. and Dutrillaux,B. (1999) Alterations of DNA methylation patterns in germ cells and Sertoli cells from developing mouse testis. *Cytogenet. Cell Genet.*, **87**, 175–181.
41. Lucifero,D., Mertineit,C., Clarke,H.J., Bestor,T.H. and Trasler,J.M. (2002) Methylation dynamics of imprinted genes in mouse germ cells. *Genomics*, **79**, 530–538.
42. Mertineit,C., Yoder,J.A., Taketo,T., Laird,D.W., Trasler,J.M. and Bestor,T.H. (1998) Sex-specific exons control DNA methyltransferase in mammalian germ cells. *Development*, **125**, 889–897.
43. Hellmann-Blumberg,U., Hintz,M.F., Gatewood,J.M. and Schmid,C.W. (1993) Developmental differences in methylation of human Alu repeats. *Mol. Cell. Biol.*, **13**, 4523–4530.
44. Chesnokov,I.N. and Schmid,C.W. (1995) Specific Alu binding protein from human sperm chromatin prevents DNA methylation. *J. Biol. Chem.*, **270**, 18539–18542.
45. Scrable,H. and Stambrook,P.J. (1997) Activation of the lac repressor in the transgenic mouse. *Genetics*, **147**, 297–304.
46. Macleod,D., Ali,R.R. and Bird,A. (1998) An alternative promoter in the mouse major histocompatibility complex class II I-Abeta gene: implications for the origin of CpG islands. *Mol. Cell. Biol.*, **18**, 4433–4443.
47. Bender,C.M., Gonzalgo,M.L., Gonzales,F.A., Nguyen,C.T., Robertson,K.D. and Jones,P.A. (1999) Roles of cell division and gene transcription in the methylation of CpG islands. *Mol. Cell. Biol.*, **19**, 6690–6698.
48. Bird,A.P. and Wolffe,A.P. (1999) Methylation-induced repression–belts, braces and chromatin. *Cell*, **99**, 451–454.
49. Loukinov,D.I., Pugacheva,E., Vatolin,S., Pack,S.D., Moon,H., Chernukhin,I., Mannan,P., Larsson,E., Kanduri,C., Vostrov,A.A. *et al.* (2002) BORIS, a novel male germ-line-specific protein associated with epigenetic reprogramming events, shares the same 11-zinc-finger domain with CTCF, the insulator protein involved in reading imprinting marks in the soma. *Proc. Natl Acad. Sci. USA*, **99**, 6806–6811.