Large scale mapping of methylcytosines in CTCF-binding sites in the human H19 promoter and aberrant hypomethylation in human bladder cancer

Daiya Takai, Felicidad A. Gonzales, Yvonne C. Tsai, Mathew J. Thayer and Peter A. Jones*

Department of Biochemistry and Molecular Biology and 1Department of Urology, USC/Norris Comprehensive Cancer Center, Keck School of Medicine of the University of Southern California, 1441 Eastlake Avenue, Los Angeles, CA 90033, USA and 2Vollum Institute and Department of Molecular and Medical Genetics, Oregon Health Sciences University, 3181 S.W. Sam Jackson Park Road, Portland, OR 97201, USA

Received May 11, 2001; Revised and Accepted September 10, 2001

The methylation status of binding sites of the insulator protein, CTCF, in the H19 promoter has been suggested as being critical to the regulation of imprinting of the H19/IGF2 locus located in chromosome 11p15. In this study, we have analyzed the methylation of all of seven potential CTCF-binding sites in the human H19 promoter since the methylation status of these sites has not been reported. We found that all the binding sites except the sixth were hypermethylated whereas only the sixth binding site showed allele-specific methylation in normal human embryonic ureteral tissue. We also analyzed the methylation status of these sites in human–mouse somatic-cell-hybrid clones containing a single copy of human chromosome 11 and which were treated with 5-aza-2'-deoxycytidine (5-aza-CdR) to yield clones which expressed human IGF2 and H19 mutually exclusively of each other. In most of the clones, a correlation between methylation of the sixth CTCF-binding site and expression of IGF2 was observed. Therefore, we analyzed the methylation status of this site in human bladder cancer and found hypomethylation of the paternal allele in two of six informative cases. These results demonstrate that only the sixth CTCF-binding site acts as a key regulatory domain for switching between H19 or IGF2 expression, whereas the other sites are not subject to allele-specific methylation. Loss of methylation imprinting of H19 is linked to hypomethylation of the paternal allele in human bladder cancer, unlike the situation in Wilms’ tumor and colon cancer where the maternal allele becomes hypermethylated.

INTRODUCTION

Imprinted genes are specifically expressed from either the maternal or paternal allele in mammalian development. Several imprinted genes have been identified (1,2) and a particularly well documented example is the H19/IGF2 located in chromosome 11p15, in which the two genes are closely linked and expressed mutually exclusively of each other. IGF2 is expressed only from the paternal allele, whereas H19 is transcribed only from the maternal allele. The position of endodermal enhancers, which lie downstream of H19, are essential for the imprinting of the IGF2 and H19 genes (3). The promoter region of H19 is methylated only on the paternal allele in both mouse (4,5) and human (6), and this region is regarded as a key domain, in the control of imprinting of this locus (7).

Recently, the upstream region of the mouse H19 gene was shown to contain four binding sites for a vertebrate enhancer-blocking protein CTCF (8–10). The promoter of the human H19 contains seven potential CTCF-binding sites all upstream of the transcriptional initiation site and methylation of these sites can preclude binding by CTCF (8–10). Each CTCF-binding site of the human H19 promoter contains three to five CpG dinucleotides (CpGs) within an ∼50 bp stretch and could potentially play a role in imprinting. To analyze the methylation status of these CpGs, we utilized bisulfite sequencing (11,12) which is capable of mapping all methylated and unmethylated cytosines in a particular sequence. Frevel et al. (13) and Vu et al. (14) reported allele-specific methylation of the region containing the sixth CTCF-binding site in Wilms’ tumor and in human embryonic tissue, respectively, but methylation mapping of some of the other potential sites has not been reported. Thus, we were interested in the biological significance of the seven CTCF sites and their methylation status in the H19 promoter.

Using the bisulfite-sequencing method, we report the methylation status of all of the seven CTCF-binding sites in human embryonic tissue, human bladder cancer and adjacent normal bladder tissue since loss of imprinting (LOI) in various types of cancer has been reported (1). We show that only the sixth CTCF-binding site shows allele-specific methylation and show a correlation between the methylation status of this site with expression of either IGF2 or H19 in human–mouse somatic-cell-hybrid clones containing a single copy of the human chromosome 11. In contrast to the situation in Wilms’ tumor (13) and colorectal cancer (15) where hypermethylation

*To whom correspondence should be addressed. Tel: +1 323 865 0816; Fax: +1 323 865 0102; Email: jones_p@ccnt.hsc.usc.edu
of the sixth CTCF-binding site has been reported, we found hypomethylation of this region in two of six bladder cancer cases.

RESULTS

The sixth CTCF-binding site shows allele-specific methylation whereas other CTCF-binding sites are methylated in normal human embryonic ureteral tissue

Bisulfite genomic sequencing analysis revealed the methylation status of 30 CpGs contained within all the potential CTCF-binding sites. Sites 1, 2, 3, 4, 5, and 7 were methylated in normal embryonic ureteral tissues (average percent methylated cytosines: 84, 79, 93, 78, 90 and 80%, respectively), suggesting that they were not utilized for control of imprinting (Fig. 1A). Only the sixth CTCF-binding site (average percent methylated cytosines: 64%) and two CpGs close to the H19 transcriptional start site (average percent methylated cytosines: 69%), which were previously revealed to be differentially methylated (14) (DMR; Fig. 1A), were found to be ~50% methylated. CpGs upstream from the fifth CTCF-binding site were highly methylated and a possible imprinting and/or methylation border occurred between the fifth and the sixth CTCF-binding sites.

Previously, the region containing the sixth CTCF-binding site, which is 2 kb upstream of the transcription start of H19, was reported to show allele-specific methylation in human fetal tissues (14). This region contains several well characterized single nucleotide polymorphisms which allows distinction of the paternal and maternal alleles. Bisulfite genomic sequencing of cloned individual molecules revealed allele-specific methylation in this site, especially of the five CpG dinucleotides in the sixth CTCF-binding site (Fig. 1B). Embryos 1–4 were heterozygous in this region. Sporadic methylation of the maternal allele was common but the paternal allele was almost completely methylated in all cases suggesting that the completely methylated state was better maintained than the unmethylated state.

Correlation of methylation status of the sixth CTCF-binding site and expression switch between IGF2 or H19 in human–mouse somatic-cell hybrids

To determine the relationship between the methylation of the CTCF-binding sites and switching of expression between H19 or IGF2, we next analyzed human–mouse somatic-cell-hybrid clones which contained a single copy of human chromosome 11. The hybrids were exposed to 5-aza-2′-deoxycytidine (5-aza-CdR), a demethylating agent that inhibits DNA methyltransferase (16), and clones established that had various expression patterns for the two genes (Fig. 2A). Expression of H19 and IGF2 were mutually exclusive in the cloned lines (Fig. 2B). We then analyzed the methylation status of the CTCF-binding sites (Fig. 2C and D). The CpGs in the sixth CTCF-binding site and the two CpGs close to the H19 transcriptional start were hypomethylated in clones expressing H19 (clones A and B, Fig. 2C). On the other hand, the seventh CTCF-binding site showed significantly higher methylation levels than the sixth CTCF-binding site, suggesting that it was not utilized for control of imprinting either. Other CpGs were relatively hypermethylated. Interestingly, clone C, which expressed neither IGF2 nor H19 (Fig. 2B), showed hypermethylation in the sixth CTCF-binding site but hypomethylation at the two CpGs close to the H19 transcriptional start and the first CTCF-binding site. A possible explanation for this result was that interaction of CTCF with the wrong CTCF-binding site resulted in the region functioning as an insulator but prevented activation of the H19 promoter. Bisulfite genomic sequencing of cloned individual molecules revealed that clone A had one completely methylated molecule and clone B had two completely methylated molecules (Fig. 2D) suggesting that re-methylation spread throughout this region once it had been initiated past a certain threshold level.

Aberrant hypomethylation of the sixth CTCF-binding site in human bladder cancer

As described above, the variable methylation of the sixth CTCF-binding site could potentially play a significant role in the regulation of expression of either IGF2 or H19. We also analyzed the methylation status of the sixth CTCF-binding site in human bladder cancer cases and corresponding normal bladder tissues. From preliminary studies of adjacent normal bladder samples, four of seven cases were demonstrated to have different polymorphic alleles at this site; thus, bisulfite sequencing was performed on these four cases (Fig. 3A). In case 1, almost complete hypomethylation of some of the paternal alleles was observed in the cancer sample. In case 4, patchy hypomethylation in the vicinity of the sixth CTCF-binding site of the paternal allele was observed. In all of the cases, the maternal alleles showed sporadically methylated CpGs; however, in the maternal allele, no apparent difference between cancer and adjacent normal tissue was detected with respect to the methylation of the maternal allele.

These results suggested that aberrant hypomethylation might play a role in loss of imprinting especially breakdown of maintenance of mono-allelic expression of H19. We analyzed further bladder cancer samples, containing different polymorphic alleles both in the promoter of H19 and the RsaI site in H19 itself. An additional two out of eight cases contained different polymorphic alleles (cases 5 and 6, Fig. 3A and C). In these two cases, the allele-specific methylation status of the sixth CTCF-binding site was maintained and no obvious changes were observed in the other CTCF-binding sites in the upstream region of H19. Allele-specific expression analysis of H19, which was determined by RT–PCR and restriction digestion of RsaI (Fig. 3D and E), revealed mono-allelic expression of H19 in all four samples including the bladder cancer with aberrant hypomethylation in the sixth CTCF-binding site. Although we had no clinical cases for directly supporting the hypothesis that the sixth CTCF-binding site might play a role in the regulation of expression of either IGF2 or H19, it is concluded that demethylation in the paternal allele in bladder cancer, which was quite rare in normal tissue, could be more predominant than methylation in the maternal allele in this locus and might play a role in overexpression of H19 in advanced stage bladder cancer previously reported by Cooper et al. (17).
DISCUSSION

The human H19 promoter has seven CTCF-binding sites, any one or a combination of which might be methylated in an allele-specific manner. However, from our study of these sites in normal embryos and hybrid cells, only the sixth CTCF-binding site showed allele-specific methylation. Thus, the 7-fold repeat of the CTCF-binding site in the human H19 promoter might not be a ‘fail-safe’ mechanism for the insulating mechanism via binding by CTCF, but might be the result of duplication as in the case of pseudogenes. From the results of our hybrid cell analysis, we speculate that methylation or demethylation of the wrong CTCF-binding sites in the H19 promoter also affects the regulation of expression either IGF2 or H19.

In human bladder cancer, we demonstrated aberrant hypomethylation in the sixth CTCF-binding site in the paternal allele. Until now, no report has analyzed the imprinting status of H19 and IGF2 in human bladder cancer, although overexpression of H19 in advanced stage bladder...
cancer (17), and a correlation between expression of \textit{H19} and poor prognosis has been reported (18). Hypomethylation of the paternal allele in bladder cancer could cause LOI and overexpression of \textit{H19}. Unfortunately, we were only able to find two cases which contained the correct combinations of polymorphisms in the promoter and exon of \textit{H19} to check whether the methylation change resulted in bi-allelic expression. None of these samples showed bi-allelic expression.

Figure 2. (A) Schematic of the establishment of human–mouse somatic hybrids cell lines containing a single human chromosome. The human chromosome 11 with full methylation and therefore presumably of paternal origin was transferred to the 10T1/2 mouse cell line from HDm-18 by microcell-mediated chromosome transfer. The hybrid cells were treated with 5-aza-CdR twice. Cloned lines were passaged for DNA and RNA analyses. (B) Expression of \textit{IGF2} and \textit{H19} is mutually exclusive from a single human chromosome. RT–PCR analysis of human \textit{IGF2} and \textit{H19} expression levels in control cell lines and in subclones derived from the 10T1/2 cells containing human chromosome 11 treated with 5-aza-CdR. Lane 1 is a control containing no cDNA. Lane 2 is a control which contained cDNA prepared from confluent 10T1/2 cells to demonstrate that the primers used resulted in specific amplification of human \textit{IGF2} and \textit{H19} sequences. Expression levels of \textit{IGF2} and \textit{H19} in subclones isolated from untreated 10T1/2 hybrid cells are shown in lane 8 and subclones isolated from 5-aza-CdR-treated 10T1/2 hybrid cells in lanes 3–7. RT–PCR was also performed with mouse GAPDH-specific primers to demonstrate similar levels of the RNA template in the various cDNA pools used. PCR products were analyzed by Southern blot analysis, hybridized with radiolabeled internal oligomers and exposed to X-ray films. (C) A physical map of human \textit{H19} promoter and seven potential CTCF-binding sites and methylation status of seven CTCF-binding sites in mouse–human somatic hybrids cells. The sixth CTCF-binding site and DMR showed a lower level of methylation in the hybrid cell clones, A and B (expressing only human \textit{H19}). Other sites showed various levels of methylation. (D) The methylation status of individual cloned molecules containing the sixth CTCF-binding site was determined by bisulfite sequencing.
so that we do not know whether the IFG2/H19 imprint is lost in a significant number of bladder cancers.

We have analyzed the methylation status of a 500 bp sequence containing the sixth CTCF-binding site utilizing difference plots (Fig. 4). In the normal samples, the methylation status of the first 19 of the total of 27 CpGs was relatively well maintained on both the paternal and maternal genomes. However, in the last eight CpGs variable levels of
hypermethylation were observed on the maternal genome suggesting a border of methylation and/or imprinting at approximately position 19. In the tumor samples, the maternal pattern appeared similar to that of the normal samples; however, considerable variability and hypomethylation of the paternal allele was apparent. Also surprising was the difference in the direction of methylation change in different types of cancer. In Wilms’ tumor (13) and colon cancer (15), hypermethylation of the maternal allele was predominant whereas bladder cancer showed hypomethylation in the paternal allele. In lung cancer, hypomethylation of the \textit{H19} promoter and biallelic expression of \textit{H19} (19) has been reported. Therefore, the direction of methylation change might be cancer-type specific and be reflective of the molecular defect associated with each cancer.

In association with this study, we found two new single nucleotide polymorphisms in the vicinity of the sixth CTCF-binding site, thus, the total number of known single nucleotide polymorphisms in this 500 bp region is seven. Three of these are probably due to conversion of methylated Cpg to TpG since these Cpg dinucleotides are methylated in sperm and methylated Cpg is less stable than unmethylated Cpg (20). However, the number of polymorphisms is unusually large considering that the average rate for single nucleotide polymorphisms is approximately one per 1250–1900 bp (21,22).

These results demonstrate that the sixth CTCF-binding site acts as a key regulatory domain for switching of \textit{H19}/\textit{IGF2} expression and that LOI of \textit{H19} is linked to hypomethylation of paternal allele in human bladder cancer.

MATERIALS AND METHODS

Cell culture

Cells from the C3H/10T1/2 Cl 8 line (10T1/2), at passage 16, were grown in Dulbecco’s modified Eagle medium (DMEM) containing 10% heat-inactivated fetal calf serum, penicillin (100 U/ml) and streptomycin (100 µg/ml; Life Technologies, Rockville, MD). Microcell hybrid cells containing a neo-marked human fibroblast chromosome 11 were established as described (23,24). Microcell hybrid cells were treated with 1 \times 10^{-6} M 5-aza-CdR (Sigma, St Louis, MO) \sim 24 h after seeding. A stock solution of 5-aza-CdR was made in PBS and filter sterilized. The medium containing 5-aza-CdR was removed 24 h after treatment and replaced with fresh medium. Afterwards, cells were reseeded and treated again with 1 \times 10^{-6} M 5-aza-CdR. Seven days after the second treatment with 5-aza-CdR, cells were reseeded at 500 cells/60 mm dish. Individual hybrid colonies were picked by using glass cloning rings and passaged in medium containing 250 µg/ml of G418.

DNA/RNA isolation from tissues and cell lines

Normal fetal ureteral tissues of 22–23 weeks of gestation were obtained at the Los Angeles County-University of Southern California Medical Center. Matched pairs of normal and tumor bladder specimens were obtained from patients treated at the Los Angeles County-University of Southern California Medical Center and the University of Southern California/ Norris Comprehensive Cancer Center (Los Angeles, CA). DNA was isolated using standard procedures (25) by treatment with proteinase K and phenol–chloroform extraction. RNA was isolated using standard procedures (26) by acid phenol–guanidium thiocyanate–chloroform extraction.

Treatment of DNA with sodium bisulfite

A modified protocol by Rein (27) was used for this analysis. Briefly, 1 µg of DNA was incubated with 0.6 M NaOH (freshly prepared) in a final volume of 20 µl for 15 min at 37°C. Cytosines were sulfonated by adding 120 µl of a solution containing 3.6 M sodium bisulfite (freshly prepared; Sigma)
and 0.6 mM hydroquinone (freshly prepared; Sigma) and incubating the sample in a Perkin-Elmer model 480 thermocycler for a period of 5 h by the following cycling protocol: 30 s at 95°C and then 15 min at 50°C. The DNA sample was the desalted using the Wizard DNA Clean-Up System and desulfonation achieved by incubation with 0.6 M NaOH for 5 min at room temperature. The converted DNA was precipitated and re-suspended in 10 µl of H2O; 1 µl was used for PCR amplification.

PCR sequencing

PCRs were performed in a 25 µl volume containing 2.5 µl of buffer (100 mM Tris–HCl pH 8.3 at 25°C, 500 mM KCl, 15 mM MgCl2, 0.01% gelatin; Sigma), 2 µl of dNTPs (2.5 mM), 20 pM primers, 0.2 U of Taqstart antibody (Clontech, Palo Alto, CA), 1 U of Taq DNA polymerase (Sigma) and 1 µl of DNA. Amplifications were performed in a Robocycler (Stratagene, La Jolla, CA) for 38 cycles of 95°C for 30 s, annealing temperature for 60 s and 72°C for 60 s. The PCR primers used were as follows: (the first, 232 bp, 50°C) sense 5'-GTA TTT TTG GAG GTT TTT TAT TAA G-3', antisense 5'-ACA CCT AAC CTA AAA AAC CTA AAA C-3'; (the second, 320 bp, 50°C) sense 5'-AGG TGT TTT AGT TTT TGG GAT GAT A-3', antisense 5'-CCA TAA ATA TTC TAT CCA TCA TTA-3'; (the third, 219 bp, 50°C) sense 5'-GGT TTT TGG TAG GTA TAG AAA TTG-3', antisense 5'-CAC CTA ACT TAA ATA ACC CAA AAC-3', sequencing primer 5'-GTA GTA TAT TGG TAT TGG-3'; (the fourth, 257 bp, 54°C) sense 5'-GTT TTT GGT AGG TTT AAG AG-3', antisense 5'-TAA ATA TCC TAT CCC TAA TAA C-3'; (the fifth, 272 bp, 55°C) sense 5'-TTT TGT AGG TTT GTT GGT AG-3', antisense 5'-TCC CAT AAA TAT CCT ATA CCT C-3'; (the sixth, 503 bp, 48°C) sense 5'-GTA GGG TTT TGG GTA GAT T-3', antisense 5'-CAC TAA AAA AAC AAT TAT CCA TTA C-3'; (the seventh, 254 bp, 55°C) sense 5'-GAG TAT TTT TAT CCG AGT AT-3', antisense 5'-AAA AAT TCT CAA ACT TTT CCA TAA A-3'. For subcloning of individual molecules, the PCR products were cloned into pGem-T Easy vector (Promega, Madison, WI). The PCR products were purified by MicroSpin S-300 HR columns (Amersham Pharmacia Biotech, Uppsala, Sweden) and individual plasmid molecules were sequenced by an automated DNA sequencer at the Norris Cancer Center microchemical core laboratory. The reliability of direct sequencing of PCR products (28) by the fluorescence dye terminator and thin-layer gel electrophoresis sequencing system was confirmed by comparison of the direct sequencing data and the population of molecules sequenced following individual cloning (data not shown).

Reverse transcription–polymerase chain reaction (RT–PCR)

RT–PCR was performed as described previously (29) with the following conditions for the various reactions. A volume of cDNA corresponding to 100 ng of template was used for all reactions. The PCR primers used are: (human H19, 26 cycles, 60°C) sense 5'-GGG CTC GGG GAG TGT GGG-3', antisense 5'-GGC AAG GTG CTC AAC ACT CA-3', probe 5'-CGG TGG AGG AGC TCA GCT CT-3'; (mouse GAPDH, 21 cycles, 73°C) sense 5'-CAG CCT CGT CCC GTC GTA GAC AAA ATG G-3', antisense 5'-TTC TGG GTG GCA GTG ATG GCA TGG A-3', probe 5'-CGG TGC TGA GTA GG G-3'. The PCR products were analyzed by Southern blot analysis (29), hybridized with radiolabeled internal oligomers and exposed to X-ray film.

Analysis of the H19 RsaI polymorphism

To assess allele-specific expression of H19, we performed RT–PCR and restriction digestion by Rsal. Total RNA was treated with DNase I (Life Technologies) before the reverse transcription reaction to avoid genomic DNA contamination. The absence of genomic DNA in DNase I-treated RNA was confirmed by over-cycling the PCR reaction (40 cycles) using the same primer set. A volume of cDNA corresponding to 100 ng of template was used for reactions using primer: sense 5'-ACC CCC TGC GGT GGA CGG TT-3', antisense 5'-TGG AAT GCT TGA AGG CTG CT-3'. Amplifications were performed in a PTC-100 (MJ Research, Waltham, MA) for 10 cycles of 95°C for 30 s, 55°C for 60 s and 72°C for 60 s following 25 cycles of 95°C for 30 s, 57°C for 60 s and 72°C for 60 s. PCR products were digested by Rsal (New England Biolabs, Beverly, MA) and electrophoresed on a 2% agarose gel.

Percentage of the methylation status of the sixth potential CTCF-binding site and its vicinity

The percentage of methylation levels in normal embryo, normal bladder and bladder cancer was calculated on each allele of all samples. The paternal and maternal alleles were identified by single nucleotide polymorphisms, and the percentage of ‘hypermethylation in maternal allele’ and ‘hypomethylation in paternal allele’ was determined by: (no. of hypermethylated molecules derived from maternal allele) / (no. of total maternal alleles) and (no. of hypomethylated molecules derived from paternal allele) / (no. of total paternal alleles), respectively. Embryo 5 had identical polymorphisms in this region, thus the allele was assumed according to the methylation pattern.

ACKNOWLEDGEMENTS

We thank the staff at the USC/Norris Comprehensive Cancer Center microchemical core laboratory for the sequencing. This work was supported by grant CA83867 from the National Cancer Institute.

REFERENCES


