MOLECULAR CARCINOGENESIS 49:999–1006 (2010)

Unveiling the Methylation Status of CpG Dinucleotides in the Substituted Segment of the Human p53 Knock-In (*Hupki*) Mouse Genome

Sang-In Kim,[1] Monica Hollstein,[2] Gerd P. Pfeifer,[1] and Ahmad Besaratinia[1]*

[1]*Department of Cancer Biology, Beckman Research Institute of the City of Hope National Medical Center, Duarte, California*
[2]*LIGHT Laboratories, University of Leeds, Leeds, UK*

Methylated cytosines within CpG dinucleotides (mCpGs) along the DNA-binding domain of the *TP53* tumor suppressor gene (exons ~5–8) are the single most significant mutational target in human cancers. The *hu*man *p*53 *k*nock-*in* (*Hupki*) mouse model was constructed using gene-targeting technology to create a mouse strain that harbors human wild-type *TP53* DNA sequences spanning exons 4–9 in both copies of the mouse *p53* gene. To date, however, the methylation status of cytosines within CpGs in the substituted segment of the *Hupki* mouse genome has not been determined. This lack of information deserves special attention because DNA methylation in mammals, which occurs almost exclusively within CpG dinucleotides, is a dynamic process throughout developmental stages and may vary among different species. Here, we have investigated the status of CpG methylation in the substituted segment of the *Hupki* mouse genome, and compared it to the methylation profile of the corresponding segment in the human genome using the combined bisulfite-restriction analysis and sodium bisulfite genomic sequencing. We found that all cytosines within CpGs of the *TP53* DNA-binding domain, on both the coding and noncoding strands, were heavily methylated in *Hupki* fibroblasts, as they were in human fibroblasts. This is in keeping with the fully methylated status of *TP53* CpGs that is known to prevail in adult human tissues. The remarkably similar patterns of cytosine methylation within CpG dinucleotides in *Hupki* cells and human cells further validates the suitability of mutagenesis assays in *Hupki* cells for experimental induction of *TP53* mutations that have been observed in human tumors.

Key words: cancer; DNA methylation; mouse model; mutation; *TP53* tumor suppressor gene

INTRODUCTION

Tumor-driving mutations in the *TP53* gene are frequent events in human cancers [1–4], and *TP53* generally stands at the top of the list of the most frequently mutated genes even when all coding sequences of the human genome are analyzed [5]. The vast majority of human *TP53* mutations arise from a single point mutation in the segment encoding the DNA-binding domain of the TP53 protein (roughly exons 5–8) [1,2]. The methylated CpGs (mCpGs) in this genomic segment constitute the most prominent mutational target in the *TP53* gene in a variety of human cancers [1–4]. The significance of mCpGs in human *TP53* mutagenesis is borne out by the observation that *TP53* mutational hotspots in certain types of human cancer localize almost exclusively to mCpG-containing codons [6]. For example, both lung and colon cancer mutational hotspots contain predominantly mCpGs in their sequence contexts, and nonmelanoma skin cancer mutational hotspots cluster at pyrimidine-mCpG sequence contexts [5].

Molecular mechanisms contributing to inherent hypermutability of mCpGs in the human genome have not been fully elucidated [6]. Depending on cancer type and etiology, however, a number of factors leading to the high frequency of tumor-driving mutations at CpG sites of *TP53* have been proposed [5]. For example, the propensity of methylated cytosines within CpG dinucleotides to undergo spontaneous hydrolytic deamination to thymine, or inflammation-derived nitrosative/oxidative stress that gives rise to promutagenic metabolites, for example, nitric oxide, may partially explain the prevalence of C to T transition mutations at mCpG sites in the *TP53* gene in colon and other internal cancers [7]. Or, enhanced reactivity of mCpGs with electrophilic compounds present in tobacco smoke, for example, polycyclic aromatic hydrocarbons, may account for the preponderance of G to T transversions within CpG sequences in the *TP53* gene in

Additional supporting information may be found in the online version of this article.

Abbreviations: mCpGs, methylated CpGs; *Hupki*, human p53 knock-in; CpGs, 5′-CpG-3′ dinucleotides; COBRA, combined bisulfite-restriction analysis; FBS, fetal bovine serum; PCR, polymerase chain reaction.

*Correspondence to: Department of Cancer Biology, Beckman Research Institute of the City of Hope National Medical Center, Duarte, California.

Received 15 June 2010; Revised 27 July 2010; Accepted 20 August 2010

DOI 10.1002/mc.20683

Published online 13 October 1010 in Wiley Online Library (wileyonlinelibrary.com).

smoking-related lung cancer [8]. Further, elevated energy absorption of methylated cytosines, especially in the long wavelength range of ultraviolet radiation, increases photodimer formation and targeted mutagenesis, which could explain the predominance of single- and tandem C to T transitions within pyrimidine-mCpG sequences in the *TP53* gene in sunlight-associated nonmelanoma skin cancer [9]. Importantly, mutations in CpG dinucleotides have a particularly deleterious impact on TP53 tumor suppressor functions, and are strongly selected for in human tumors [6].

The *hu*man *p*53 *k*nock-*i*n (*Hupki*) mouse model was constructed using gene-targeting technology to create a mouse strain that harbors human wild-type *TP53* DNA sequences in both copies of the mouse *p53* gene [10]. Replacement of exons 4–9 of the endogenous mouse *p53* alleles in the *Hupki* mouse with the homologous normal human *TP53* gene sequences has offered a humanized version of the *TP53* gene in a murine cell context [10]. The *Hupki* mouse model system is a unique research tool for experimental recapitulation of human *TP53* mutagenesis [10]. The utility of the *Hupki* mouse model system for studying the underlying mechanism of human *TP53* mutagenesis has been demonstrated in a number of in vivo animal experiments [11,12] and in vitro cell culture experiments [13–17]. Given the prominence of mCpGs in the human *TP53* mutagenesis [5], and the significant dynamics of DNA methylation during embryonic development and across species [18–20], the status of CpG methylation in the substituted segment of *Hupki* mouse genome deserves special attention. Investigating the methylation status of 5′-CpG-3′ dinucleotides (CpGs) in the human segment of the Hupki mouse genome is particularly important as primary fibroblasts from explanted embryos of *Hupki* mice are being increasingly used for generating carcinogen-induced *TP53* mutations [21,22]. Thus, establishing the methylation status of cytosines within CpG dinucleotides in the *Hupki* *TP53* gene can further validate the accuracy of this experimental system for modeling human *TP53* mutagenesis and carcinogenesis. In the present study, we have determined the *TP53* CpG methylation profile in the substituted segment of *Hupki* embryonic fibroblasts and compared it to the methylation profile of the same segment when present in normal human neonatal fibroblasts, using two independent assays, including the combined bisulfite-restriction analysis (COBRA) [23], and sodium bisulfite genomic sequencing [24].

MATERIALS AND METHODS

Cell Culture

Primary mouse embryonic fibroblasts, prepared from 13.5-d-old embryos of *Hupki* mice [10], were grown in Dulbecco's modified Eagle's medium (DMEM) (Irvine Scientific, Santa Ana, CA) supplemented with 10% fetal bovine serum (FBS). For comparison purposes, early passage normal human fibroblasts, prepared from neonatal foreskin, were cultivated under the same conditions. When reaching nearly confluency, all cultures were harvested by trypsinization, and subjected to genomic DNA isolation using a standard phenol and chloroform extraction and ethanol precipitation protocol [25]. The DNA was dissolved in TE buffer (10 mM Tris–HCl, 1 mM EDTA, pH 7.5), and kept at −80°C until further analysis.

COBRA and Bisulfite Genomic Sequencing

To establish the status of cytosine methylation within CpGs in the human segment of the *Hupki* mouse genome, we used both the COBRA [23], and bisulfite genomic sequencing techniques [24] to find methylated CpGs in both coding and noncoding strands of exons 5–8 of the *TP53* gene. For comparison, we performed parallel analysis on the genomic DNA of normal human fibroblasts. Briefly, 2 μg of total genomic DNA were subjected to sodium bisulfite treatment using the Qiagen EpiTect kit according to the manufacturer's instructions (Qiagen, Valencia, CA). The purified bisulfite-treated DNA was subsequently analyzed by standard COBRA and genomic sequencing assays, as described previously [26]. Methodologically, both assays rely on the principle that sodium bisulfite treatment of DNA can selectively deaminate cytosine, but not 5-methylcytosine, to uracil, thus, resulting in a primary sequence change in the DNA [23,24]. This sequence conversion is exploited to differentiate between unmethylated and methylated cytosine upon subsequent restriction enzyme digestion [23] and direct sequencing [24] in the COBRA and genomic sequencing assays, respectively. In COBRA, bisulfite-treated genomic DNA serves as template in a polymerase chain reaction (PCR) to amplify the region of interest for an ensuing DNA methylation analysis. The successive bisulfite treatment and PCR amplification convert unmethylated cytosine to thymine ($C \rightarrow U \rightarrow T$) and methylated cytosine to cytosine ($mC \rightarrow mC \rightarrow C$). These sequence conversions lead to methylation-dependent retention of preexisting restriction enzyme sites if the CpG dinucleotides were methylated prior to bisulfite treatment, or of creation of new restriction sites if the dinucleotides were unmethylated. Thus, restriction enzyme digestion can reveal methylation-dependent sequence differences at specific restriction sites in the PCR product of a locus of interest in bisulfite-treated DNA [23]. Direct genomic sequencing of bisulfite-treated and PCR-amplified DNA, on the other hand, can provide information on the status of cytosine methylation throughout the locus of interest [24].

For both COBRA and genomic sequencing analysis, we designed specific primers to PCR amplify

exons 5–8 of the *TP53* gene on both coding and noncoding strands, individually. The primer sequences used for PCR amplification of all analyzed target sequences are listed in Supplementary Table 1. The PCR primers were designed to be complementary to the bisulfite-converted DNA sequences with no CpG dinucleotide in the corresponding region of original unconverted DNA. The avoidance of CpG dinucleotides within the primer sequences ensures that the amplification step of PCR does not discriminate between DNA templates according to their original methylation status. In other words, the lack of CpGs in the sequence of PCR primers results in amplification of the sequence of interest in between the two primers, regardless of the DNA methylation status of that sequence in the original genomic DNA. The restriction enzymes used for digestion of PCR products in all COBRA analysis are also identified in Supplementary Table 1 and Figures 1 and 2. For genomic sequencing, the PCR products obtained after bisulfite conversion of genomic DNA were cloned into the TOPO-TA cloning vector (Invitrogen, Inc., Carlsbad, CA) according to the manufacturer's instructions. Randomly selected clones from the genomic DNA of *Hupki* fibroblasts versus counterpart human cells were sequenced using an ABI-3730 DNA Sequencer (ABI Prism, PE Applied BioSystems, Foster City, CA).

Although both the COBRA and genomic sequencing assays are routine and highly validated molecular biology techniques [27,28], we have made standard controls by preparing unmethylated DNA controls (negative controls), which were the human *TP53*-specific PCR products of each of exons 5–8. We have subsequently methylated these unmethylated DNA controls by in vitro treatment with M. SssI CpG methyltransferase (New England Biolabs, Ipswich, MA) to prepare methylated DNA controls (positive controls). The above-mentioned standard controls served as negative and positive controls, respectively, in all runs of the assay. We have processed both the negative and positive controls in parallel to all experimental samples, and included them throughout all assay runs.

RESULTS AND DISCUSSION

In the present study, we have scanned the genomic DNA of *Hupki* mouse fibroblasts and of human

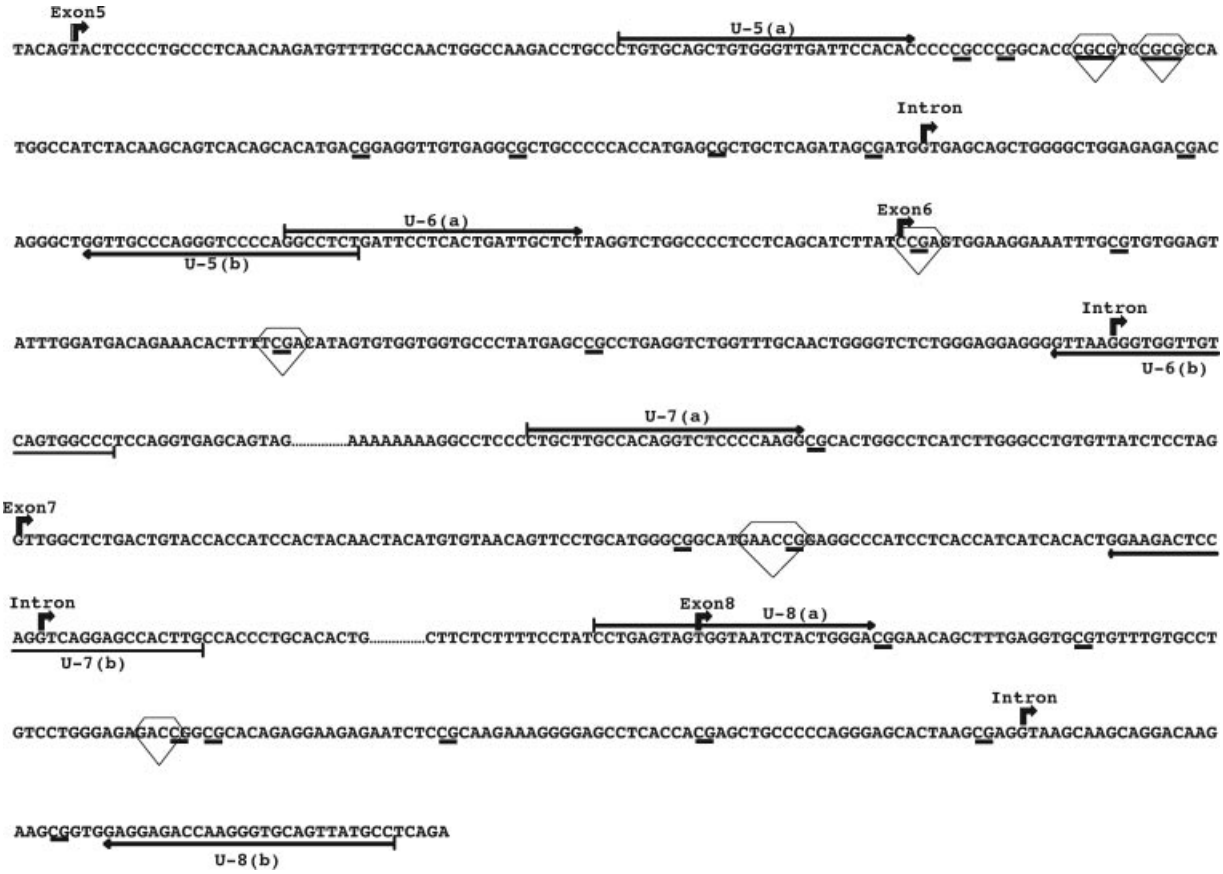

Figure 1. Schematic representation of the analyzed target sequences in the *TP53* gene (coding strand). The location and designation of all primers are indicated along the target sequences (see also Supplementary Table 1). The start positions of exons and introns are indicated by arrows. CpG dinucleotides are underlined. Restriction-enzyme recognition sites are shown within diamond boxes.

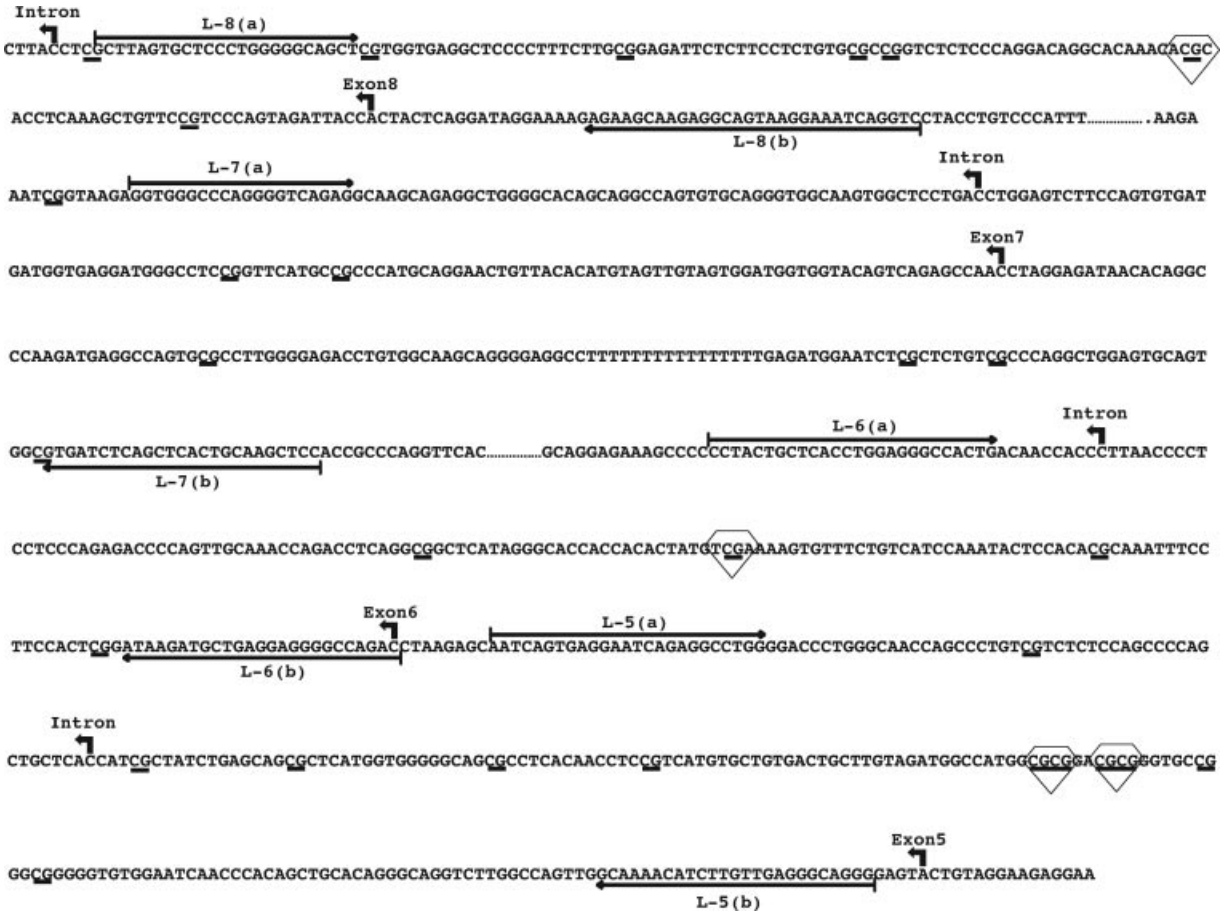

Figure 2. Schematic representation of the analyzed target sequences in the *TP53* gene (noncoding strand) (see legend of Figure 1 and Supplementary Table 1).

fibroblasts for methylated CpGs at specific restriction sites in exons 5–8 of the *TP53* gene on both coding and noncoding strands using standard COBRA analysis [23]. For confirmation, we have used genomic sequencing to establish, at the level of nucleotide resolution, the methylation status of each cytosine within every individual CpG throughout the human *TP53* DNA-binding domain in both the *Hupki* and human cells. As shown in Figures 3–5 (a) (upper panels), there were remarkably similar patterns of DNA methylation in the substituted segment of the *Hupki* mouse genome versus human genome, as demonstrated by our COBRA analysis of all the examined target sequences. Specifically, all cytosines within CpG dinucleotides at all restriction sites examined by the COBRA analysis showed extensive methylation in both the *Hupki* and human genomes. The striking similarities between methylation status of cytosines within CpG dinucleotides at specific restriction sites in exons 5, 6, 7, and 8 of the *TP53* gene on both coding and noncoding strands in the *Hupki* mouse genome versus human genome are illustrated in the upper panels of Figures 3–5. It is worth mentioning that the background signals of the undigested original products in some positive controls (i.e., in vitro methylated DNA prepared from the human *TP53*-specific PCR products of each of exons 5–8 that were subjected to restriction enzyme digestion: see Figures 3–5 (a); [+] lanes in 'Pos') are most likely due to less than 100% efficiency of methylating DNA in vitro with M. SssI CpG methyltransferase. Likewise, the background signals of undigested products in some of the human and *Hupki* samples imply that the analyzed targets in the respective samples are less than 100% methylated (see Figures 3–5 (a); [+] lanes in 'Human' and '*Hupki*').

In agreement with our COBRA analysis, genomic sequencing of the substituted segment of the *Hupki* mouse DNA and the corresponding sequences in human DNA showed fully methylated cytosines within virtually every individual CpG throughout exons 5–8 of the *TP53* gene on both coding and noncoding strands. Representative DNA methylation profiles in exons 5, 7, and 8 of the *TP53* gene on both DNA strands in the *Hupki* mouse genome versus human genome are illustrated in the lower panels of Figures 3–5. These three exons comprise the most frequently mutated region in the DNA-binding domain of the *TP53* gene in all cancer types

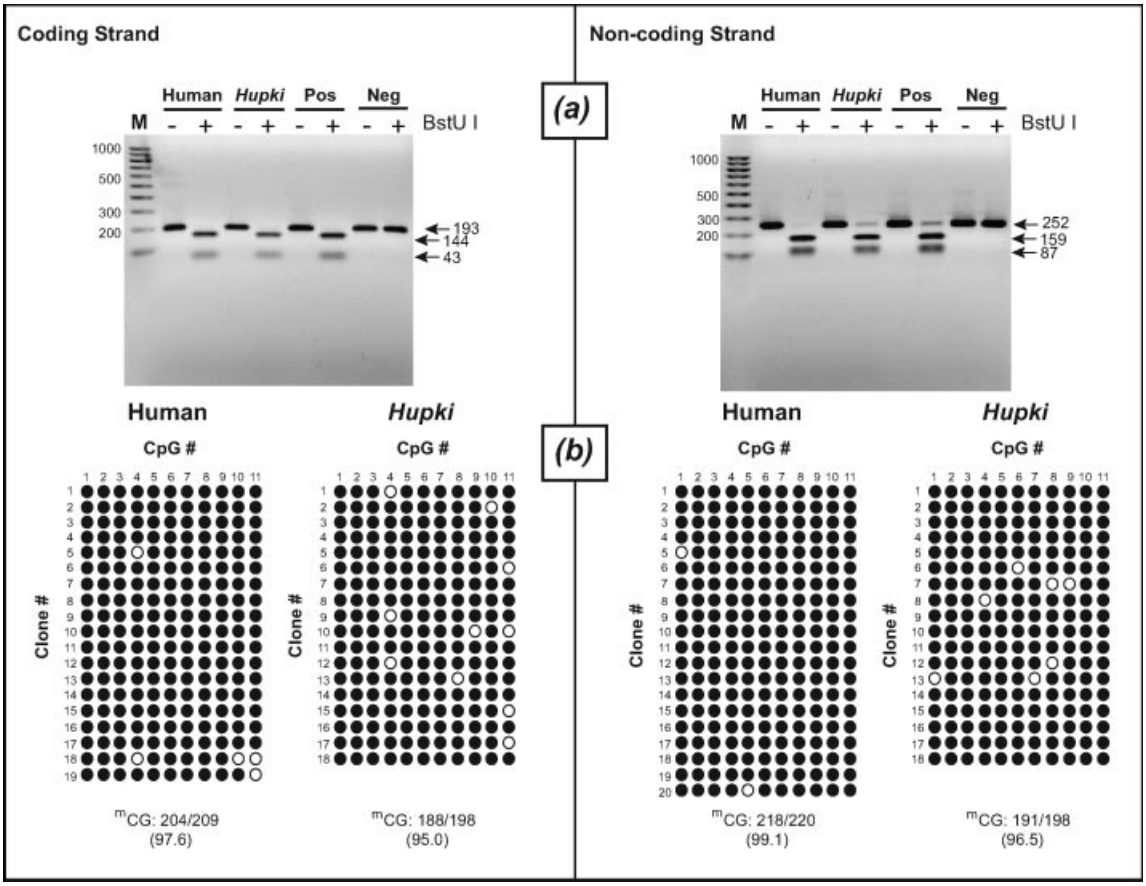

Figure 3. Comparison of DNA methylation profiles in exon 5 of the *TP53* gene between the *Hupki* mouse and human genomes. (a) Genomic DNA samples of the *Hupki* mouse fibroblasts and human fibroblasts were scanned for methylated CpGs at specific restriction sites in exon 5 of the *TP53* gene on both coding and noncoding strands using standard COBRA analysis [23]. ''BstU I'' is the restriction enzyme used for digestion of bisulfite-treated and PCR-amplified DNA on both coding and noncoding strands. ''+'' and ''−'' refer to the presence and absence, respectively, of the restriction enzyme in the reaction mix. All restriction enzyme digested and undigested products are indicated by arrows. We have made standard controls by preparing unmethylated DNA controls (negative controls), which were the human *TP53*-specific PCR products of exon 5 for coding and noncoding strands, individually. We have subsequently methylated these unmethylated DNA controls by in vitro treatment with M. SssI CpG methyltransferase (New England Biolabs) to prepare methylated DNA controls (positive controls). The above-mentioned standard controls served as negative and positive controls, respectively, in all runs of the assay. Pos = positive control; Neg = negative control; M = size marker. (b) Genomic sequencing was performed to establish the methylation status of each cytosine within every individual CpG dinucleotide throughout the target sequence in the genomic DNA of *Hupki* cells and human cells. There are 11 CpGs in the PCR product of exon 5 on both coding and noncoding DNA strands. (●) = Methylated CpG; (○) = unmethylated CpG; ^{m}CG: absolute number of methylated CpGs/total CpGs (% methylated CpGs); Minor differences in ^{m}CG% found between the *Hupki* mouse genome and human genome were not statistically significant (Fisher's exact test).

combined [6]. The importance of exons 5, 7, and 8 of the *TP53* gene in carcinogenesis is highlighted by the finding that the most prominent mutational hotspots in this tumor suppressor gene in human cancers localize to several codons that are part of these three exons [6]. The overall percentage of methylated cytosines within all CpGs on the coding DNA strand of the *Hupki* mouse genome versus human genome, as estimated from the CpG methylation status at each site in each of 11–20 clones per exon examined, were 95.0 and 97.6 in exon 5, 100.0 and 100.0 in exon 7 plus part of its corresponding intron, and 93.4 and 94.3 in exon 8, respectively. The overall percentage of methylated cytosines within all CpGs on the noncoding DNA strand of the *Hupki* mouse genome versus human genome were 96.5 and 99.1 in exon 5, 95.6 and 97.9 in exon 7 plus part of its corresponding intron, and 98.6 and 97.5 in exon 8, respectively. None of the differences in the percentage of methylated cytosines within CpGs found between the *Hupki* mouse genome and human genome was statistically significant (Fisher's exact test). We have further analyzed these data to compare the methylation status of CpGs in every individual target locus within exons 5–8 of the *Hupki* mouse genome and the counterpart human genome. In all cases, there was no statistically significant difference between the methylation status of CpGs in specific target loci, individually, within exons 5–8 of the *Hupki* mouse genome and the corresponding methylation status in human genome (χ^2 test or Fisher's exact test, as appropriate). Moreover, we

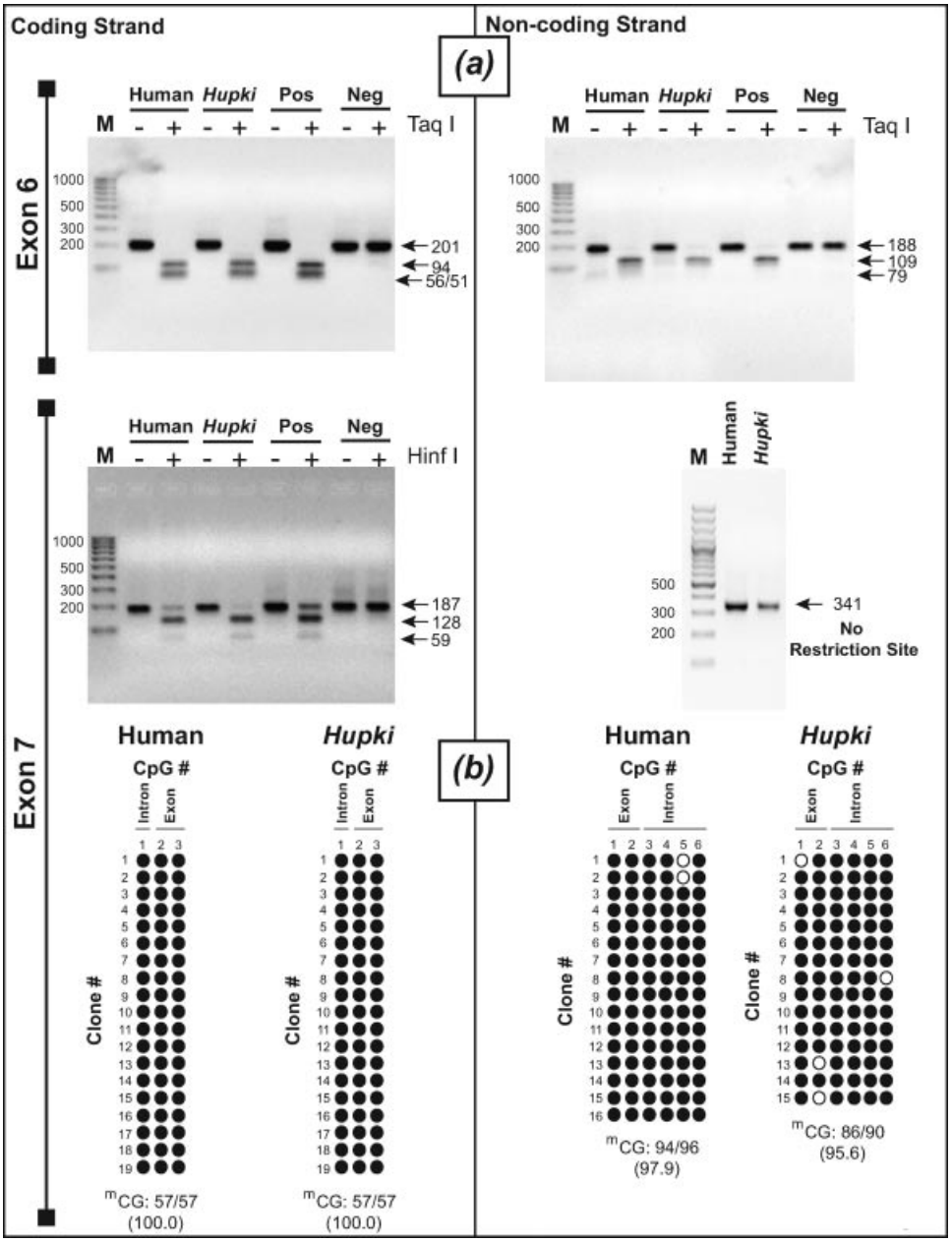

Figure 4. Comparison of DNA methylation profiles in exon 6 and exon 7 of the *TP53* gene between the *Hupki* mouse and human genomes. (a) Genomic DNA samples of the *Hupki* mouse fibroblasts and human fibroblasts were scanned for methylated CpGs at specific restriction sites in exon 6 and exon 7 of the *TP53* gene on both coding and noncoding strands using standard COBRA analysis [23]. ''Taq I'' is the restriction enzyme used for digestion of bisulfite-treated and PCR-amplified DNA on both coding and noncoding strands of exon 6. ''Hinf I'' is the restriction enzyme used for digestion of bisulfite-treated and PCR-amplified DNA on the coding strand of exon 7; no known restriction site was found in the PCR product of the noncoding strand of exon 7, however (as such, we have shown the PCR products of the noncoding strand, while emphasizing that the absence of restriction sites in these PCR products precluded restriction enzyme digestion of the COBRA analysis). ''+'' and ''−'' refer to the presence and absence, respectively, of the restriction enzyme in the reaction mix. All restriction enzyme digested and undigested products are indicated by arrows. Pos = positive control; Neg = negative control (see legend of Figure 3); M = size marker. (b) Genomic sequencing was performed to establish the methylation status of each cytosine within every individual CpG dinucleotide throughout the target sequence in the genomic DNA of *Hupki* cells and human cells. The methylation status of all cytosines within CpG dinucleotides in exon 7 plus part of its corresponding intron is shown on both DNA strands. There are three and six CpGs, respectively, in the PCR product of exon 7 plus part of its corresponding intron on coding and noncoding DNA strands. (●) = Methylated CpG; (○) = unmethylated CpG; ^{m}CG: absolute number of methylated CpGs/total CpGs (% methylated CpGs); minor differences in ^{m}CG% found between the *Hupki* mouse genome and human genome were not statistically significant (Fisher's exact test).

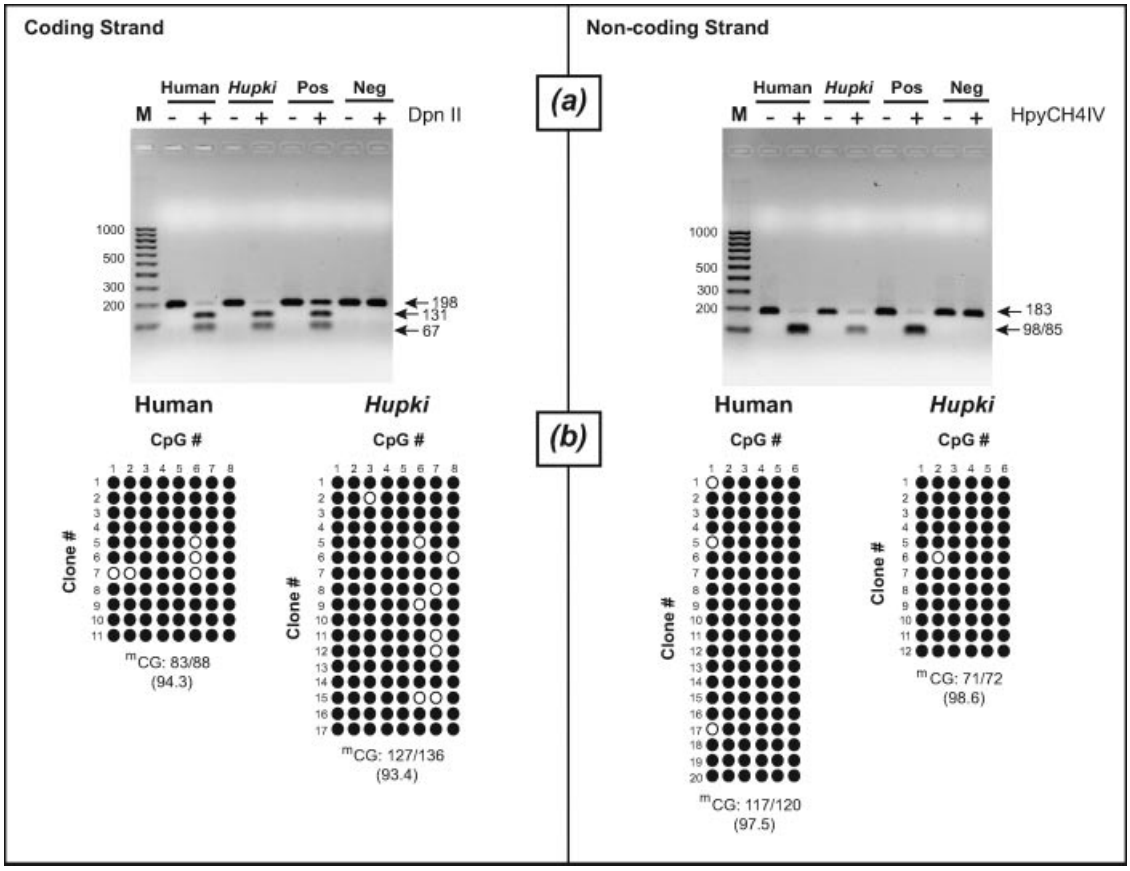

Figure 5. Comparison of DNA methylation profiles in exon 8 of the *TP53* gene between the *Hupki* mouse and human genomes. (a) Genomic DNA samples of the *Hupki* mouse fibroblasts and human fibroblasts were scanned for methylated CpGs at specific restriction sites in exon 8 of the *TP53* gene on both coding and noncoding strands using standard COBRA analysis [23]. ''Dpn II'' and ''HpyCH4IV are the restriction enzymes used for digestion of bisulfite-treated and PCR-amplified DNA on the coding and noncoding strands, respectively. ''+'' and ''−'' refer to the presence and absence, respectively, of the restriction enzyme in the reaction mix. All restriction enzyme digested and undigested products are indicated by arrows. Pos = positive control; Neg = negative control (see legend of Figure 3); M = size marker. (b) Genomic sequencing was performed to establish the methylation status of each cytosine within every individual CpG dinucleotide throughout the target sequence in the genomic DNA of *Hupki* cells and human cells. There are eight and six CpGs, respectively, in the PCR product of exon 8 on coding and noncoding DNA strands. (●) = Methylated CpG; (○) = unmethylated CpG; ^{m}CG: absolute number of methylated CpGs/total CpGs (% methylated CpGs); minor differences in ^{m}CG% found between the *Hupki* mouse genome and human genome were not statistically significant (Fisher's exact test).

examined the data to determine whether the minor methylation differences observed at certain CpG sites within individual exons in the *Hupki* mouse genome and human genome could result in statistically significant differences in DNA methylation patterns in the exon of interest between the respective genomes. In all cases, the overall CpG methylation patterns were similar between the *Hupki* mouse genome and the human genome for every individual exon examined (Hypergeometric test [29]). Altogether, these findings are consistent with our previous report in which we demonstrated that the *TP53* sequences along exons 5–8 are highly methylated at every CpG site on both DNA strands in all human tissues and cell lines tested [30]. Of note, we had previously established an extensive methylation of CpG dinucleotides in the mouse *p53* gene, as well ([31], and unpublished data).

Lastly, we have recently studied the effect of a known carcinogen, benzo[*a*]pyrene diol epoxide (B[*a*]PDE), on DNA methylation in human cells, and demonstrated that the *TP53* gene is not targeted by this chemical treatment [32]. We have also performed similar investigations on other known or suspect chemical/physical carcinogens to determine whether *TP53* DNA methylation profile can be modulated by carcinogen treatment (manuscript in preparation). In all cases, however, we did not detect any significant changes in the *TP53* gene methylation profile consequent to treatment with various carcinogens. Obviously, it would be interesting to see the response of the *Hupki* and human cells to substances that target *TP53*; however, to the best of our knowledge, no agent has ever been identified as a substance that targets *TP53*, and modifies its methylation pattern.

In conclusion, we have observed extensive cytosine methylation within CpG dinucleotides of human *TP53* sequences in embryonic fibroblasts from *Hupki* mice. The profile of cytosine methylation at CpG sites in the human segment of the *Hupki* mouse genome is indistinguishable from that found in the same DNA segment in human fibroblasts. In both cases, all cytosines within these dinucleotides are highly methylated throughout exons 5–8 of the *TP53* gene on both coding and noncoding DNA strands. The concordance with human cell data of the CpG methylation status in the substituted segment of the *Hupki* mouse genome, which is the most frequently hit target in human cancers [5,6], lends further support to the suitability of the *Hupki* model for investigating human *TP53* mutagenesis and carcinogenesis [21,22]. Our findings are particularly significant for scientific disciplines with a general interest in the field of cancer research, specifically cancer biology. The verified utility of the *Hupki* model system makes it appealing to investigators who share a common interest in the field of experimental and molecular cancer research.

ACKNOWLEDGMENTS

This work was supported by grants from the University of California Tobacco Related Disease Research Program (18KT-0040 to A.B.), and National Cancer Institute (R01CA084469 to G.P.P.).

REFERENCES

1. Hofseth LJ, Hussain SP, Harris CC. p53: 25 years after its discovery. Trends Pharmacol Sci 2004;25:177–181.
2. Olivier M, Petitjean A, Marcel V, et al. Recent advances in p53 research: An interdisciplinary perspective. Cancer Gene Ther 2009;16:1–12.
3. Petitjean A, Achatz MI, Borresen-Dale AL, Hainaut P, Olivier M. TP53 mutations in human cancers: Functional selection and impact on cancer prognosis and outcomes. Oncogene 2007;26:2157–2165.
4. Vogelstein B, Lane D, Levine AJ. Surfing the p53 network. Nature 2000;408:307–310.
5. Pfeifer GP, Besaratinia A. Mutational spectra of human cancer. Hum Genet 2009;125:493–506.
6. Pfeifer GP. p53 mutational spectra and the role of methylated CpG sequences. Mutat Res 2000;450:155–166.
7. Goodman JE, Hofseth LJ, Hussain SP, Harris CC. Nitric oxide and p53 in cancer-prone chronic inflammation and oxyradical overload disease. Environ Mol Mutagen 2004;44: 3–9.
8. Besaratinia A, Pfeifer GP. Second-hand smoke and human lung cancer. Lancet Oncol 2008;9:657–666.
9. Pfeifer GP, You YH, Besaratinia A. Mutations induced by ultraviolet light. Mutat Res 2005;571:19–31.
10. Luo JL, Yang Q, Tong WM, Hergenhahn M, Wang ZQ, Hollstein M. Knock-in mice with a chimeric human/murine p53 gene develop normally and show wild-type p53 responses to DNA damaging agents: A new biomedical research tool. Oncogene 2001;20:320–328.
11. Luo JL, Tong WM, Yoon JH, et al. UV-induced DNA damage and mutations in Hupki (human p53 knock-in) mice recapitulate p53 hotspot alterations in sun-exposed human skin. Cancer Res 2001;61:8158–8163.
12. Tong WM, Lee MK, Galendo D, Wang ZQ, Sabapathy K. Aflatoxin-B exposure does not lead to p53 mutations but results in enhanced liver cancer of Hupki (human p53 knock-in) mice. Int J Cancer 2006;119:745–749.
13. Liu Z, Muehlbauer KR, Schmeiser HH, Hergenhahn M, Belharazem D, Hollstein MC. p53 mutations in benzo(a)pyrene-exposed human p53 knock-in murine fibroblasts correlate with p53 mutations in human lung tumors. Cancer Res 2005;65:2583–2587.
14. Reinbold M, Luo JL, Nedelko T, et al. Common tumour p53 mutations in immortalized cells from Hupki mice heterozygous at codon 72. Oncogene 2008;27:2788–2794.
15. Liu Z, Hergenhahn M, Schmeiser HH, Wogan GN, Hong A, Hollstein M. Human tumor p53 mutations are selected for in mouse embryonic fibroblasts harboring a humanized p53 gene. Proc Natl Acad Sci USA 2004;101:2963–2968.
16. vom Brocke J, Krais A, Whibley C, Hollstein MC, Schmeiser HH. The carcinogenic air pollutant 3-nitrobenzanthrone induces GC to TA transversion mutations in human p53 sequences. Mutagenesis 2009;24:17–23.
17. Whibley C, Odell AF, Nedelko T, et al. Wild-type and HUPKI (human P53 knock-in)murine embryonic fibroblasts: P53/ARF pathway disruption in spontaneous escape from senescence. J Biol Chem 2010;285:11326–11335.
18. Morgan HD, Santos F, Green K, Dean W, Reik W. Epigenetic reprogramming in mammals. Hum Mol Genet 2005;14: R47–R58.
19. Suzuki MM, Bird A. DNA methylation landscapes: Provocative insights from epigenomics. Nat Rev Genet 2008;9:465–476.
20. Jones PA, Baylin SB. The epigenomics of cancer. Cell 2007; 128:683–692.
21. Besaratinia A, Pfeifer GP. Applications of the human p53 knock-in (Hupki) mouse model for human carcinogen testing. FASEB J 2010;24:2612–2619.
22. Kucab JE, Phillips DH, Arlt VM. Linking environmental carcinogen exposure to TP53 mutations in human tumours using the human TP53 knock-in (Hupki) mouse model. FEBS J 2010;277:2567–2583.
23. Xiong Z, Laird PW. COBRA: A sensitive and quantitative DNA methylation assay. Nucleic Acids Res 1997;25:2532–2534.
24. Frommer M, McDonald LE, Millar DS, et al. A genomic sequencing protocol that yields a positive display of 5-methylcytosine residues in individual DNA strands. Proc Natl Acad Sci USA 1992;89:1827–1831.
25. Pfeifer GP, Chen HH, Komura J, Riggs AD. Chromatin structure analysis by ligation-mediated and terminal transferase-mediated polymerase chain reaction. Methods Enzymol 1999;304:548–571.
26. Tommasi S, Karm DL, Wu X, Yen Y, Pfeifer GP. Methylation of homeobox genes is a frequent and early epigenetic event in breast cancer. Breast Cancer Res 2009;11:R14.
27. Eads CA, Laird PW. Combined bisulfite restriction analysis (COBRA). Methods Mol Biol 2002;200:71–85.
28. Hajkova P, el-Maarri O, Engemann S, Oswald J, Olek A, Walter J. DNA-methylation analysis by the bisulfite-assisted genomic sequencing method. Methods Mol Biol 2002;200: 143–154.
29. Adams WT, Skopek TR. Statistical test for the comparison of samples from mutational spectra. J Mol Biol 1987;194:391–396.
30. Tornaletti S, Pfeifer GP. Complete and tissue-independent methylation of CpG sites in the p53 gene: Implications for mutations in human cancers. Oncogene 1995;10:1493–1499.
31. You YH, Szabo PE, Pfeifer GP. Cyclobutane pyrimidine dimers form preferentially at the major p53 mutational hotspot in UVB-induced mouse skin tumors. Carcinogenesis 2000;21:2113–2117.
32. Tommasi S, Kim SI, Zhong X, Wu X, Pfeifer GP, Besaratinia A. Investigating the epigenetic effects of a prototype smoke-derived carcinogen in human cells. PLoS ONE 2010;5: e10594.

TABLE 1: Detailed information on genomic sequences, primer designs, and restriction enzymes used for bisulfite-based CpG methylation analysis of the human *TP53* gene.

Nucleotide positions +13104 - +13296 **"Exon 5 plus intron" Coding strand**				
Original DNA sequence	**5'CTGTGCAGCTGTGGGTTGATTCCACACCCCCGCCCGGCACCCGCGTCCGCGCCATGGCCATCTACAA GCAGTCACAGCACATGACGGAGGTTGTGAGGCGCTGCCCCCACCATGAGCGCTGCTCAGATAGCGATG GTGAGCAGCTGGGGCTGGAGAGACGACAGGGCTGGTTGCCCAGGGTCCCCAGGCCTCT**			
DNA sequence after bisulfite treatment	**5'TTGTGTAGTTGTGGGTTGATTTTATATTTTCGTTCGGTATTCGCGTTCGCGTTATGGTTATTTATAAGT AGTTATAGTATATGACGGAGGTTGTGAGGCGTTGTTTTTATTATGAGCGTTGTTTAGATAGCGATGGTG AGTAGTTGGGGTTGGAGAGACGATAGGGTTGGTTGTTTAGGGTTTTTAGGTTTTT**			
Forward primer	**Designation: U-5 (a)**	**t_m (°C): 53.2**	**# Bases: 27**	**Sequence: 5'-TTGTGTAGTTGTGGGTTGATTTTATAT-3'**
Reverse primer	**Designation: U-5 (b)**	**t_m (°C): 52.6**	**# Bases: 25**	**Sequence: 5'-AAAAACCTAAAAACCCTAAACAACC-3'**
PCR product size	**193 bp**			
# CpGs in PCR product	**11**			
Restriction enzyme recognition site	**CG▼CG**			
# Restriction sites in PCR product	**2**			
Restriction enzyme	**BstU I**			
Product size after restriction enzyme digestion	**Digested: 6 bp, 43 bp, 144 bp**			**Non-digested: 193 bp**

Nucleotide positions +13290 - +13490 "Exon 6 plus intron" Coding strand				
Original DNA sequence	5'GGCCTCTGATTCCTCACTGATTGCTCTTAGGTCTGGCCCCTCCTCAGCATCTTATCCGAGTGGAAGGA AATTTGCGTGTGGAGTATTTGGATGACAGAAACACTTTTCGACATAGTGTGGTGGTGCCCTATGAGCCG CCTGAGGTCTGGTTTGCAACTGGGGTCTCTGGGAGGAGGGGTTAAGGGTGGTTGTCAGTGGCCC			
DNA sequence after bisulfite treatment	5'GGTTTTTGATTTTTTATTGATTGTTTTTAGGTTTGGTTTTTTTTTAGTATTTTATTCGAGTGGAAGGAA ATTTGCGTGTGGAGTATTTGGATGATAGAAATATTTTTCGATATAGTGTGGTGGTGTTTTATGAGTCGT TTGAGGTTTGGTTTGTAATTGGGGTTTTTGGGAGGAGGGGTTAAGGGTGGTTGTTAGTGGTTT			
Forward primer	**Designation:** **U-6 (a)**	**t_m (°C):** **59.16**	**# Bases:** **27**	**Sequence:** **5'-GGTTTTTGATTTTTTATTGATTGTTTT-3'**
Reverse primer	**Designation:** **U-6 (b)**	**t_m (°C):** **59.64**	**# Bases:** **24**	**Sequence:** **5'-AAACCACTAACAACCACCCTTAAC-3'**
PCR product size	**201 bp**			
# CpGs in PCR product	**4**			
Restriction enzyme recognition site	**T▼CGA**			
# Restriction sites in PCR product	**2**			
Restriction enzyme	**Taq I**			
Product size after restriction enzyme digestion	**Digested:** **51 bp, 56 bp, 94 bp**		**Non-digested:** **201 bp**	

Nucleotide positions +13938 - +14124 "Exon 7 plus intron" Coding strand				
Original DNA sequence	5'CTGCTTGCCACAGGTCTCCCCAAGGCGCACTGGCCTCATCTTGGGCCTGTGTTATCTCCTAGGTTGG CTCTGACTGTACCACCATCCACTACAACTACATGTGTAACAGTTCCTGCATGGGCGGCATGAACCGGAG GCCCATCCTCACCATCATCACACTGGAAGACTCCAGGTCAGGAGCCACTTG			
DNA sequence after bisulfite treatment	5'TTGTTTGTTATAGGTTTTTTTAAGGCGTATTGGTTTTATTTTGGGTTTGTGTTATTTTTTAGGTTGGTT TTGATTGTATTATTATTTATTATAATTATATGTGTAATAGTTTTTGTATGGGCGGTATGAATCGGAGGTT TATTTTTATTATTATTATATTGGAAGATTTTAGGTTAGGAGTTATTTG			
Forward primer	**Designation:** **U-7 (a)**	**t_m (°C):** **48.6**	**# Bases:** **25**	**Sequence:** **5'-TTGTTTGTTATAGGTTTTTTTAAGG-3'**
Reverse primer	**Designation:** **U-7 (b)**	**t_m (°C):** **52.1**	**# Bases:** **27**	**Sequence:** **5'-CAAATAACTCCTAACCTAAAATCTTCC-3'**
PCR product size	**187 bp**			
# CpGs in PCR product	**3**			
Restriction enzyme recognition site	**G▼AATCG**			
# Restriction sites in PCR product	**1**			
Restriction enzyme	**Hinf I**			
Product size after restriction enzyme digestion	**Digested:** **59 bp, 128 bp**			**Non-digested:** **187 bp**

Nucleotide positions +14443 - +14640 "Exon 8 plus intron" Coding strand				
Original DNA sequence	5'CCTGAGTAGTGGTAATCTACTGGGACGGAACAGCTTTGAGGTGCGTGTTTGTGCCTGTCCTGGGAGA GACCGGCGCACAGAGGAAGAGAATCTCCGCAAGAAAGGGGAGCCTCACCACGAGCTGCCCCCAGGGA GCACTAAGCGAGGTAAGCAAGCAGGACAAGAAGCGGTGGAGGAGACCAAGGGTGCAGTTATGCC			
DNA sequence after bisulfite treatment	5'TTTGAGTAGTGGTAATTTATTGGGACGGAATAGTTTTGAGGTGCGTGTTTGTGTTTGTTTTGGGAGAG ATCGGCGTATAGAGGAAGAGAATTTTCGTAAGAAAGGGGAGTTTTATTACGAGTTGTTTTTAGGGAGTA TTAAGCGAGGTAAGTAAGTAGGATAAGAAGCGGTGGAGGAGATTAAGGGTGTAGTTATGTT			
Forward primer	**Designation:** **U-8 (a)**	**t_m (°C):** **52.2**	**# Bases:** **25**	**Sequence:** **5'-TTTGAGTAGTGGTAATTTATTGGGA-3'**
Reverse primer	**Designation:** **U-8 (b)**	**t_m (°C):** **54.1**	**# Bases:** **26**	**Sequence:** **5'-AACATAACTACACCCTTAATCTCCTC-3'**
PCR product size	**198 bp**			
# CpGs in PCR product	**8**			
Restriction enzyme recognition site	**▼GATC**			
# Restriction sites in PCR product	**1**			
Restriction enzyme	**Dpn II**			
Product size after restriction enzyme digestion	**Digested:** **67 bp, 131bp**		**Non-digested:** **198 bp**	

Nucleotide positions +13060 - +13311 "Exon 5 plus intron" Non-coding strand				
Original DNA sequence	5'AATCAGTGAGGAATCAGAGGCCTGGGGACCCTGGGCAACCAGCCCTGTCGTCTCTCCAGCCCCAGCT GCTCACCATCGCTATCTGAGCAGCGCTCATGGTGGGGGGCAGCGCCTCACAACCTCCGTCATGTGCTGT GACTGCTTGTAGATGGCCATGGCGCGGACGCGGGTGCCGGGCGGGGGGTGTGGAATCAACCCACAGCT GCACAGGGCAGGTCTTGGCCAGTTGGCAAAACATCTTGTTGAGGGCAGGG			
DNA sequence after bisulfite treatment	5'AATTAGTGAGGAATTAGAGGTTTGGGGATTTTGGGTAATTAGTTTTGTCGTTTTTTTAGTTTTAGTTG TTTATTATCGTTATTTGAGTAGCGTTTATGGTGGGGGTAGCGTTTTATAATTTTCGTTATGTGTTGTGAT TGTTTGTAGATGGTTATGGCGCGGACGCGGGTGTCGGGCGGGGGGTGTGGAATTAATTTATAGTTGTAT AGGGTAGGTTTTGGTTAGTTGGTAAAATATTTTGTTGAGGGTAGGG			
Forward primer	**Designation:** **L-5 (a)**	**t_m (°C):** **52.8**	**# Bases:** **25**	**Sequence:** **5'-AATTAGTGAGGAATTAGAGGTTTGG-3'**
Reverse primer	**Designation:** **L-5 (b)**	**t_m (°C):** **52.0**	**# Bases:** **25**	**Sequence:** **5'-CCCTACCCTCAACAAAATATTTTAC-3'**
PCR product size	**252 bp**			
# CpGs in PCR product	**11**			
Restriction enzyme recognition site	**CG▼CG**			
# Restriction sites in PCR product	**2**			
Restriction enzyme	**BstU I**			
Product size after restriction enzyme digestion	**Digested:** **6 bp, 87 bp, 159 bp**			**Non-digested:** **252 bp**

<table>
<tr><th colspan="5">Nucleotide positions +13320 - +13507
“Exon 6 plus intron” <u>Non-coding strand</u></th></tr>
<tr><td>Original DNA sequence</td><td colspan="4">5’CCTACTGCTCACCTGGAGGGCCACTGACAACCACCCTTAACCCCTCCTCCCAGAGACCCCAGTTGCA
AACCAGACCTCAGG<u>CG</u>GCTCATAGGGCACCACCACACTATGT<u>CG</u>AAAAGTGTTTCTGTCATCCAAATAC
TCCACA<u>CG</u>CAAATTTCCTTCCACT<u>CG</u>GATAAGATGCTGAGGAGGGGCCAGAC</td></tr>
<tr><td>DNA sequence after bisulfite treatment</td><td colspan="4">5’<u>TTTATTGTTTATTTGGAGGGTTATTG</u>ATAATTATTTTTAATTTTTTTTTTTAGAGATTTTAGTTGTAAAT
TAGATTTTAGG<u>CG</u>GTTTATAGGGTATTATTATATTATGT<u>CGA</u>AAAGTGTTTTTGTTATTTAAATATTTTA
TA<u>CG</u>TAAATTTTTTTTTATT<u>CG</u>GATAAGATGTTGAGGAGGGGTTAGAT</td></tr>
<tr><td>Forward primer</td><td>Designation:
L-6 (a)</td><td>t_m (°C):
58.81</td><td># Bases:
26</td><td>Sequence:
5’-TTTATTGTTTATTTGGAGGGTTATTG-3’</td></tr>
<tr><td>Reverse primer</td><td>Designation:
L-6 (b)</td><td>t_m (°C):
59.68</td><td># Bases:
25</td><td>Sequence:
5’-ATCTAACCCCTCCTCAACATCTTAT-3’</td></tr>
<tr><td>PCR product size</td><td colspan="4">188 bp</td></tr>
<tr><td># CpGs in PCR product</td><td colspan="4">4</td></tr>
<tr><td>Restriction enzyme recognition site</td><td colspan="4">T▼CGA</td></tr>
<tr><td># Restriction sites in PCR product</td><td colspan="4">1</td></tr>
<tr><td>Restriction enzyme</td><td colspan="4">Taq I</td></tr>
<tr><td>Product size after restriction enzyme digestion</td><td colspan="2">Digested:
79 bp, 109 bp</td><td colspan="2">Non-digested:
188 bp</td></tr>
</table>

Nucleotide positions +13846 - +14186 "Exon 7 plus intron" Non-coding strand				
Original DNA sequence	5'GGTGGGCCCAGGGGTCAGAGGCAAGCAGAGGCTGGGGCACAGCAGGCCAGTGTGCAGGGTGGCAAGTGGCTCCTGACCTGGA GTCTTCCAGTGTGATGATGGTGAGGATGGGCCTCCGGTTCATGCCGCCCATGCAGGAACTGTTACACATGTAGTTGTAGTGGATG GTGGTACAGTCAGAGCCAACCTAGGAGATAACACAGGCCCAAGATGAGGCCAGTGCGCCTTGGGGAGACCTGTGGCAAGCAGGG GAGGCCTTTTTTTTTTTTTTTGAGATGGAATCTCGCTCTGTCGCCCAGGCTGGAGTGCAGTGGCGTGATCTCAGCTCACTGCAAG CTCC			
DNA sequence after bisulfite treatment	5'GGTGGGTTTAGGGGTTAGAGGTAAGTAGAGGTTGGGGTATAGTAGGTTAGTGTGTAGGGTGGTAAGTGGTTTTTGATTTGGAGT TTTTTAGTGTGATGATGGTGAGGATGGGTTTTCGGTTTATGTCGTTTATGTAGGAATTGTTATATATGTAGTTGTAGTGGATGGTG GTATAGTTAGAGTTAATTTAGGAGATAATATAGGTTTAAGATGAGGTTAGTGCGTTTTGGGGAGATTTGTGGTAAGTAGGGGAGGT TTTTTTTTTTTTTTTTTGAGATGGAATTTCGTTTTGTCGTTTAGGTTGGAGTGTAGTGGCGTGATTTTAGTTTATTGTAAGTTTT			
Forward primer	**Designation:** **L-7 (a)**	**t_m (°C):** **54.6**	**# Bases:** **20**	**Sequence:** **5'-GGTGGGTTTAGGGGTTAGAG-3'**
Reverse primer	**Designation:** **L-7 (b)**	**t_m (°C):** **46.7**	**# Bases:** **25**	**Sequence:** **5'-AAAACTTACAATAAACTAAAATCAC-3'**
PCR product size	**341 bp**			
# CpGs in PCR product	**6**			
Restriction enzyme recognition site	**None**			
# Restriction sites in PCR product	**None**			
Restriction enzyme	**None**			
Product size after restriction enzyme digestion	**Digested:** **None**		**Non-digested:** **341 bp**	

Nucleotide positions +14403 - +14585 "Exon 8 plus intron" Non-coding strand				
Original DNA sequence	5'GCTTAGTGCTCCCTGGGGGCAGCTCGTGGTGAGGCTCCCCTTTCTTGCGGAGATTCTCTTCCTCTGT GCGCCGGTCTCTCCCAGGACAGGCACAAACACGCACCTCAAAGCTGTTCCGTCCCAGTAGATTACCAC TACTCAGGATAGGAAAAGAGAAGCAAGAGGCAGTAAGGAAATCAGGTC			
DNA sequence after bisulfite treatment	5'GTTTAGTGTTTTTTGGGGGTAGTTCGTGGTGAGGTTTTTTTTTTTTGCGGAGATTTTTTTTTTTTGTGC GTCGGTTTTTTTTAGGATAGGTATAAATACGTATTTTAAAGTTGTTTCGTTTTAGTAGATTATTATTATT TAGGATAGGAAAAGAGAAGTAAGAGGTAGTAAGGAAATTAGGTT			
Forward primer	**Designation:** L-8 (a)	**t_m (°C):** 53.8	**# Bases:** 24	**Sequence:** 5'-GTTTAGTGTTTTTTGGGGGTAGTT-3'
Reverse primer	**Designation:** L-8 (b)	**t_m (°C):** 55.3	**# Bases:** 30	**Sequence:** 5'-AACCTAATTTCCTTACTACCTCTTACTTCT-3'
PCR product size	183 bp			
# CpGs in PCR product	6			
Restriction enzyme recognition site	A▼CGT			
# Restriction sites in PCR product	1			
Restriction enzyme	HpyCH4IV			
Product size after restriction enzyme digestion	**Digested:** 85 bp, 98 bp		**Non-digested:** 183 bp	

- CpG sites are underlined.
- Restriction-enzyme recognition sites are double-underlined.
- Primer directions are indicated by arrow in the DNA sequence.