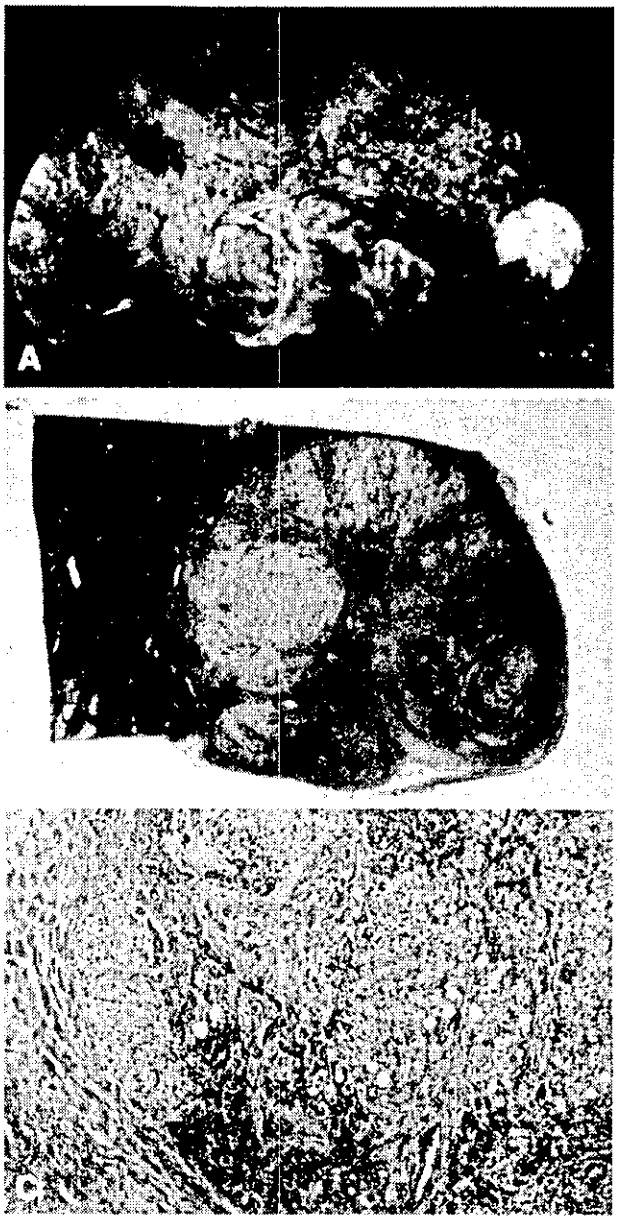



Figure 2. Photographs showing the gross (A), loupe (B) appearance, and histology (C) of the tumor. The tumor is located on the right lobe of the liver and is well circumscribed; the nontumoral liver shows marked enlargement with fatty degeneration (A). The multinodular masses vary in size and are separated by fibrous tissue (B). Microscopically, extensive bile duct proliferation is seen at the margin of the pseudolobules, with hepatocellular bile-ductular transformation (C). The connective tissue in the septa is mature and compact. The hepatocytes are small and polygonal, resembling embryonal hepatocytes.

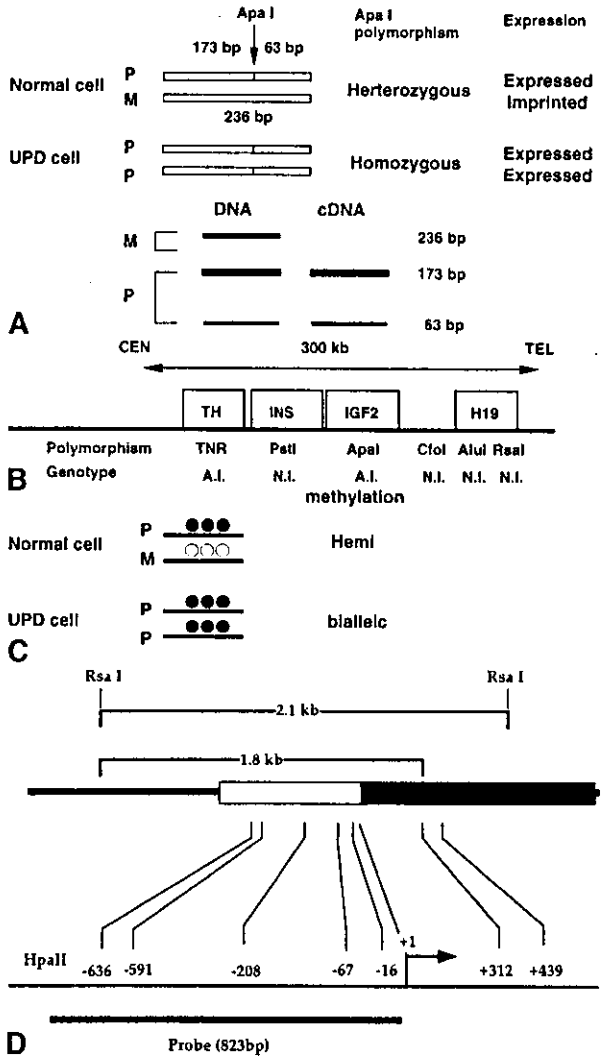



Figure 3. Schematic diagram summarizing the studies used for analysis of genetic and epigenetic mosaicism. A: Insulin-like growth factor 2 (*IGF2*) allelic imbalance and allele-specific expression. M, maternal allele; P, paternal allele; UPD, uniparental disomy. B: Genotypes for tyrosine hydroxylase (*TH*), insulin (*INS*), *IGF2*, and *H19* polymorphic markers. A.I., allelic imbalance; CEN, centromere; N.I., not informative; TEL, telomere; TNR, tetranucleotide repeat. *TH* is mapped about 300 kb centromeric to *H19*. The genotype was determined from results for the lung, normal liver, mixed hamartoma of the liver, and nesidioblastosis specimens. C: *H19* methylation and epigenetic mosaicism. Open circles, unmethylated allele; closed circles, methylated allele; M, maternal allele; P, paternal allele. D: Schematic illustration of *H19* promoter. *Rsa*I and *Hpa*II sites from −636 to +439 are shown by the vertical bar. The transcription start site is indicated by an arrow and shown as +1. The probe used for Southern blot analysis is shown.

ylation-insensitive isozyme *Msp*I, resolved on a 0.8% agarose gel, and transferred to HybondN+ (Amersham, Buckinghamshire, UK) in 1.5 M NaCl/0.25 N NaOH. The membranes were prehybridized at 65°C for 3 h in 3× SSPE (1× SSPE = 0.15 M NaCl, 1 mM EDTA, and 0.01 M sodium phosphate buffer), 5× Denhardt's solution, 1% SDS, and 10 mg poly A/ml, and then hybridized with the H19

probe labeled with [α-$^{32}$P]-dCTP by the random-primer method. After overnight hybridization, blots were washed twice with 1× SSPE, 1% SDS at 65°C and twice with 1× SSPE, 1% SDS at 65°C and then exposed to X-ray film. The H19 probe was an 823-bp fragment hybridizing to the promoter region of the *H19* gene [13].

## RESULTS

Previous studies have demonstrated the presence of somatic mosaicism for paternal isodisomy by documenting two grossly disproportionate alleles on 11p15.5 using polymorphic markers, and by analyzing *H19* promoter hypermethylation [3–8]. In this study, to investigate the somatic mosaicism in a BWS patient with paternal isodisomy, a series of molecular studies was performed in some organs or tissues obtained at autopsy or surgery. To characterize a genetic mosaicism, analyses of *IGF2* polymorphism and other markers on 11p15.5 to determine allelic imbalance were performed. It is conceivable that the allelic imbalance was the result of mosaicism for two types of cells: normal cells heterozygous for *IGF2 Apa*I polymorphism and paternal isodisomic cells homozygous for *IGF2* (Fig. 3A). Other polymorphic markers, including *H19* (*Rsa*I, *Alu*I, *Cfo*I) and the insulin gene (*Pst*I), were not informative in all of the tissues (Fig. 3B). To investigate whether the lost allele is of maternal origin, allele-specific *IGF2* expression was determined (Figs. 3A,4C). The paternal expression of *IGF2* is consistent with the evidence of allelic imbalance, with the paternal allele being more intense than the maternal allele in BWS with UPD. Furthermore, to confirm paternal isodisomy, allele-specific methylation of *H19* promoter was also determined. Epigenetic mosaicism is also characterized by two populations of cells. In normal cells, *H19* is methylated on the paternal allele, but is unmethylated on the maternal allele. In UPD cells, *H19* is biallelically methylated (Figs. 3C, 5) [6,7,12]. Thus, an increased degree of methylation of *H19* promoter indicates an increase in the proportion of the *H19* paternal allele (biallelically methylated cells).

### Heterozygosity of *TH* and *IGF2*

All showed an allelic imbalance, because the upper bands were weaker in density than the lower bands (Fig. 4A). In nontumoral liver and nesidioblastosis, the upper bands were much weaker than the lower ones. In MHL and lung tissue, the upper bands were slightly weaker than the lower bands. Both alleles were clearly retained, however, in the normal control liver.

*IGF2* also showed an allelic imbalance for *Apa*I polymorphism (Fig. 4B). The upper bands (236 bp) in normal liver and nesidioblastosis were weaker in density than the lower bands (173 bp). Reduction of the upper allele of *IGF2* in nesidioblastosis was the greatest of all the specimens. The upper and lower bands were almost equal in density in MHL. The upper band was slightly denser than the lower band in the lung tissue.

Genomic DNA in this case was identified as having an allelic imbalance for the *TH* tetranucleotide repeat polymorphism and *IGF2 Apa*I polymorphism with evidence of disproportionate alleles in nesidioblastosis, MHL, and nontumoral liver and lung tissue (Fig. 4). The ratios of paternal-to-maternal allele at the *TH* and *IGF2* loci in this case were different for each tissue examined (somatic mosaicism). A particularly gross disproportion in alleles for *IGF2* and *TH* was seen in the DNA of the nesidioblastosis tissue and nontumoral liver.

### *IGF2* allelic expression

PCR and reverse transcription (RT)-PCR restriction fragment length polymorphism (RFLP) were performed to determine allele-specific *IGF2* expression (Fig. 4C). A specimen from nesidioblastosis showed monoallelic expression. Since *IGF2* is expressed from the paternal allele, the 173-bp and 63-bp bands were most likely derived from the paternal allele. Thus, allele-specific *IGF2* expression of paternal origin was found in the nesidioblastosis specimen, consistent with the normal imprinting pattern. The paternal expression of *IGF2* was consistent with evidence of allelic imbalance in the nesidioblastosis specimen, with the paternal allele being more intense than the maternal allele in the syndrome.

### *H19* methylation

Control pancreatic tissue and lung tissue showed a significant reduction in signal intensity of the 2.1-kb fragment and an increase in the 1.8-kb fragment and low-molecular-weight signals, indicating

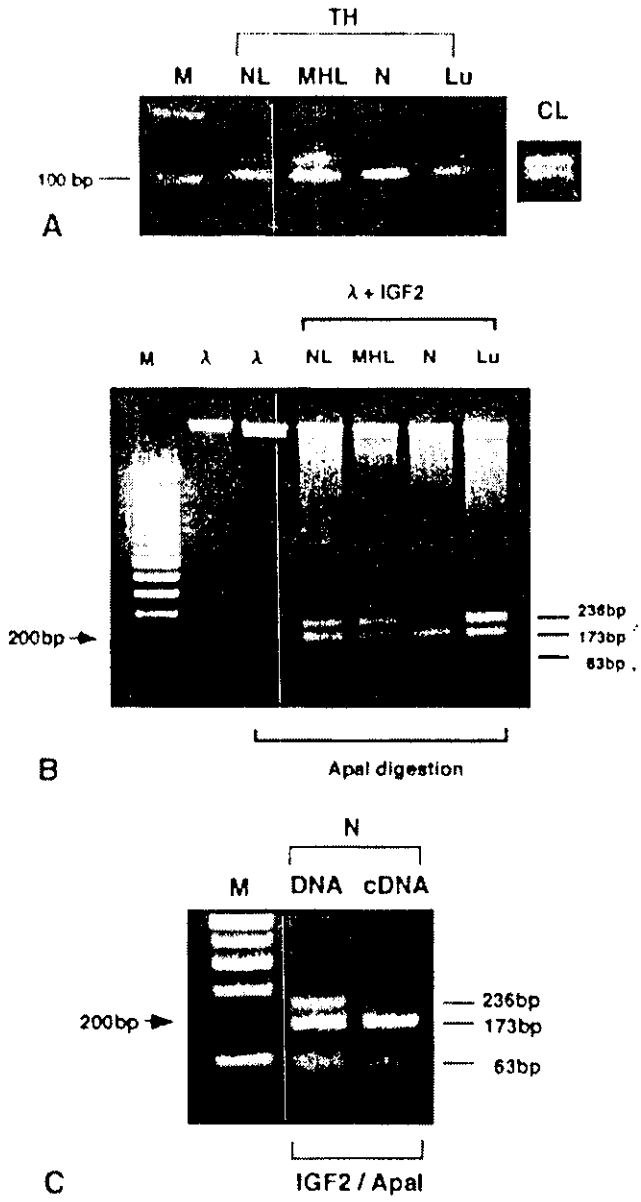



**Figure 4.** Allelic imbalance for *TH* (A) and *IGF2* (B), and allele-specific expression of *IGF2* in nesidioblastosis (C). Tetranucleotide repeat polymorphism of *TH* was analyzed in specimens from nontumoral liver (NL), lung tissue (Lu), nesidioblastosis (N), mixed hamartoma of the liver (MHL), and control normal liver (CL). To analyze allelic imbalance for *IGF2*, polymerase chain reaction (PCR) products of *IGF2* were digested with *Apa*I. The internal control (λDNA) was used to exclude the possibility of incomplete digestion. Digestion of λDNA with *Apa*I cleaved the 48,502 base pair (bp) λDNA fragment into 38,416-bp and 10,086-bp fragments. The second lane from the left is λDNA (λ), and the third lane from the left is λDNA with *Apa*I digestion. Both PCR products of *IGF2* and λDNA in all specimens were completely digested. An *Apa*I polymorphism in *IGF2* (Fig. 3A) was used to analyze allele-specific expression. A specimen from nesidioblastosis (N) was analyzed for the *Apa*I polymorphism by PCR amplification of genomic DNA (DNA) and RT-PCR of mRNA (cDNA), followed by digestion with *Apa*I. M, 100-bp ladder.

hemimethylation and allele-specific methylation (Fig. 3D, Fig. 5). Generation of 1.8-kb fragment is often observed with the fully methylated 2.1-kb fragment, because the methylation of the *Hpa*II site in exon 1 (+312) of the *H19* paternal allele is not always methylated. Hypermethylation of the maternal allele in specimens from the nontumoral liver tissue, MHL, lung tissue, and nesidioblastosis was demonstrated by generation of 2.1-kb and 1.8-kb fragments and a loss of low-molecular-weight unmethylated DNA signals (indicated as LMW) by *Rsa*I and *Hpa*II digestion. Nontumoral liver tissue (hepatomegaly) and pancreatic tissue (diffuse nesidioblastosis) showed hypermethylation at the *H19* promoter. Both normal lung tissue and MHL also showed increased methylation, but not as completely as the liver and pancreatic tissues in BWS (Fig. 5). Lung tissue showed the weakest methylation among the tissues examined in BWS. A slight decrease in signal intensity of the 2.1-kb and 1.8-kb fragments and an increase in the low-molecular-weight signal were seen in the lung and MHL specimens, suggesting an increase in the proportion of hemimethylated cells in the tissues examined in BWS.

In addition, decreased intensity of the maternal allele of *IGF2* was shown to correlate with reduction of the *H19* hypomethylated allele, also consistent with duplication of the paternal allele. Thus, there appears to be allelic imbalance that resulted in mosaicism of two types of cells: normal cells with biparental 11p15 DNA and the paternal isodisomic cells (Fig. 3A,C). These results are highly suggestive of mosaic paternal uniparental isodisomy, based on the finding of excess homozygosity and allelic imbalance at 11p15.5 and high methylation of the *H19* promoter region [6–8,15]. The somatic mosaicism we observed in this study was detectable not only genetically but also epigenetically; we therefore use the term "genetic/epigenetic mosaicism."

## DISCUSSION

Numerous cases of an association between BWS and benign tumors of a hamartomatous nature have been reported, including nephroblastomatosis, nesidioblastosis, and adrenal cytomegaly. It is not known why BWS patients have several associated hamartomatous conditions: It has been pos-

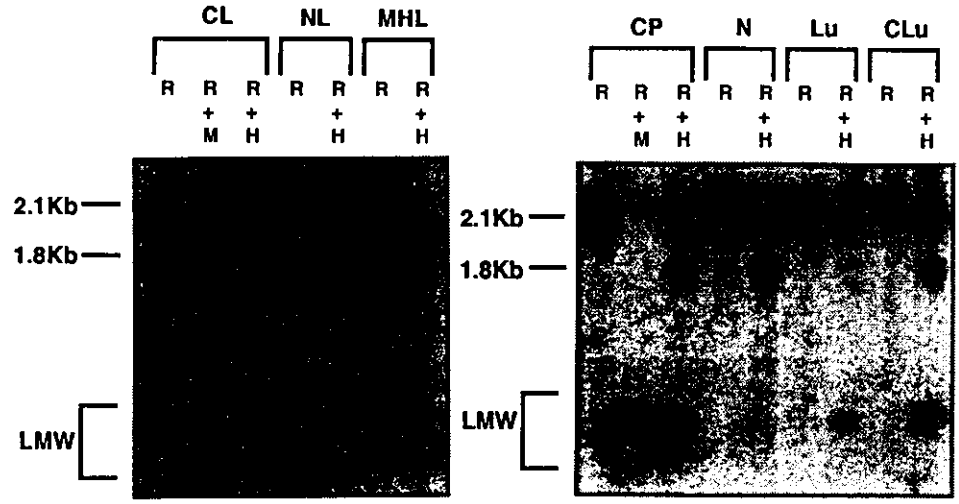



**Figure 5.** Southern blot analysis of methylation of *H19* gene. CL, control liver tissue; NL, nontumoral liver; MHL, mixed hamartoma of the liver; CP, control normal pancreatic tissue; N, nesidioblastosis; Lu, lung tissue; CLu, control normal lung tissue; R, *Rsa*I; M, *Msp*I; H, *Hpa*II; LMW, low-molecular-weight unmethylated DNA.

tulated that the hyperplastic growth of the organs may delay maturation, but recent progress in research on the molecular basis for the genesis of BWS has clarified the etiology of these associated hamartomatous conditions. Several imprinted genes are localized to 11p15.5. IGF2 is essential for normal fetal growth and *IGF2* is one of the candidate genes for BWS. Its activity is regulated by genomic inactivation of the maternal allele. The organs hypertrophied in BWS correlate with the cell types that prominently express *IGF2* during normal development [16]. *H19* is a paternally imprinted growth regulatory gene of unknown function. On the paternal allele, the *H19* promoter region is heavily methylated at CpG residues, leading to repression of *H19* expression. Loss of imprinting of the *IGF2* gene has occasionally been linked to biallelic inactivation of *H19*, resulting in a bipaternal epigenotype of *IGF2* and *H19*, leading to somatic overgrowth and predisposition to Wilms' tumor [17–19].

From studies in disomic patients, a genetic/epigenetic mosaicism has been proposed in which two populations of cells arise as a consequence of a postzygotic recombination between the maternal and paternal homologue: one with *IGF2* expression from both alleles and *H19* biallelic methylation (paternal isodisomic cell), and the other with *IGF2* expression from only the paternal allele and *H19* allele-specific methylation (normal cell). The extent and location of genetic/epigenetic changes may be responsible for the variable clinical features of BWS [5–8,12]. The presence of allelic loss or various degrees of disproportional alleles on 11p15.5 has been observed in patients with UPD, not only in embryonal tumors but in nontumoral tissues (blood, tongue, kidney), and these findings provide further support for the hypothesis that somatic mosaicism results in overgrowth and/or asymmetrical growth, such as hemihypertrophy, macroglossia, nephromegaly, and other organomegalies [3,4,20]. We identified our BWS patient as having a mosaic paternal uniparental isodisomy, and we demonstrated genetic/epigenetic mosaicism in some of the patient's organs. Although the tissues we investigated were all of endodermal origin (lung, liver, pancreas), each tissue appeared to have a different ratio of somatic mosaicism. These findings also support the evidence that this case of BWS with UPD originated from a postzygotic event [5,6].

UPD is thought to be a potential cause of nesidioblastosis. De Lonaly et al. reported that specific loss of maternal alleles of the imprinted chromosome region 11p15 occurs only in the focal form of the nesidioblastosis (focal adenomatous hyperplasia), not in the diffuse form [21,22]. However, our case of BWS showed a diffuse form of nesidioblastosis and exhibited a similar alteration of 11p15.5 due to constitutional UPD mosaicism with an excessive proportion of paternal isodisomic cells, which is equivalent to maternal loss of

heterozygosity (LOH). During normal embryologic development of the pancreas, islets form by budding from the ductular epithelium. *IGF2* is expressed in pancreatic islets and plays a role in their proliferation. The continuous formation of islets in nesidioblastosis may be due to biallelic expression of *IGF2*. The insulin gene is located between the *TH* and the *IGF2* genes (Fig. 3B). It is not known whether the insulin gene is imprinted in humans, but if it is biallelically expressed as a result of UPD, it may contribute to hyperinsulinism.

MHL includes embryonic-type cells with biliary channel formation in a fashion reminiscent of normal hepatic embryogenesis [10]. These findings suggest that MHL may be hamartomatous. MHL has never been reported in association with any anomalies. The MHL described by Toda et al. is of particular interest [23], because the tumor was congenital and the patient was born at full term with a high birth weight of 4.4 kg (>+1.5 SD). However, other signs of BWS were not seen in this patient, thus the patient may have had the incomplete form of BWS. Hemihypertrophy has been found to be the most common feature of BWS with UPD [7] and also has the potential for developing into embryonal tumors [12]. Two examples of focal nodular hyperplasia have been reported in patients with hemihypertrophy [24]. It is possible, however, that in these reports, focal nodular hyperplasia was confused with MHL. Thus, MHL may in some way be related to BWS, although the data for a UPD-related etiology of MHL are less clear than for nesidioblastosis. MHL may be a lesion with a decrease in the proportion of paternal isodisomic cells surrounded by a marked increase in the proportion of disomic cells. The hepatocytes within MHL resemble embryonal hepatocytes and may have monoallelic expression of *IGF2*, while the larger hepatocytes in the nontumoral liver may have biallelic expression of *IGF2*. This may result in slower growth and immaturity within the MHL than in the surrounding hepatocytes that exhibit hyperplastic growth. The above findings suggest that genetic/epigenetic mosaicism contributes to growth imbalance, which may result in the formation of some hamartomatous lesions associated with BWS. Recently, *IGF2/H19* epigenetic mosaicism has been demonstrated in nephroblastomatosis [25].

This is the first report of nesidioblastosis in which the histopathological findings at surgery were compared with those at autopsy a year later, and the differentiation and development of the islets and acinar tissue were more conspicuous at 13 months. Nesidioblastosis associated with BWS may correct itself during development. This finding may be supported by the fact that the growth rate and bone age of BWS patients are higher during the first few years, but normalize thereafter [2]. The macroglossia in these patients has also been found to resolve during development. These corrections observed in BWS may be related to the stage-specific *IGF2* expression seen in the human liver [26]. Some epigenetic changes may correct the abnormal development of nesidioblastosis, however, unfortunately, we were not able to investigate the DNA of the nesidioblastosis specimen obtained at autopsy.

## ACKNOWLEDGMENTS

This work was supported by Grants-in-Aid for Scientific Research from the Ministry of Education, Science and Culture of Japan (10152256, 10307004) and Cancer Research Grants from the Ministry of Health and Welfare (9-14, 10D-1).

## REFERENCES

1. Sotelo-Avila C, Gonzalez-Crussi F, Fowler J. Complete and imcomplete forms of Beckwith-Wiedemann syndrome: their oncogenic potential. Pediatr 1980;96:47–50.
2. Cohen M. A comprehensive and critical assessment of overgrowth and overgrowth syndrome. Adv Hum Genet 1989;18:209–212.
3. Schneid H, Seurin D, Vazquez M, Gourmelen M, Cabrol S, Le Bouc Y. Parental allele specific methylation of the human insulin-like growth factor II gene and Beckwith-Wiedemann syndrome. J Med Genet 1993;30:353–362.
4. Schneid H, Vazquez M, Seurin D, Bouc Y. Loss of heterozygosity in non-tumoral tissue in two children with Beckwith-Wiedemann syndrome. Growth Regul 1991;1: 168–170.
5. Henry I, Puech A, Riesewijk A, et al. Somatic mosaicism for partial paternal isodisomy in Wiedemann-Beckwith syndrome; a post-fertilization event. Eur J Hum Genet 1993;1:19–29.
6. Catchpoole D, Lam W, Valler D, et al. Epigenetic modification and uniparental inheritance of H19 in Beckwith-Wiedemann syndrome. J Med Genet 1997;34:353–359.
7. Slatter R, Elliot M, Welham K, et al. Mosaic uniparental disomy in Beckwith-Wiedemann syndrome. J Med Genet 1994;31:749–753.
8. Reik W, Brown K, Slatter R, Satori P, Elliot M, Maher E. Allelic methylation of H19 and IGF2 in the Beckwith-Wiedemann syndrome. Hum Mol Genet 1994;3:1297–1301.
9. Slocia E, Capella C, Kloppel G. Tumor-like lesions of the

endocrine pancreas. In: Rosai J, LHS, eds. Atlas of Tumor Pathology. Washington, DC: Armed Forces Institute of Pathology, 1997;237–246.
10. Rhodes R, Marchildon M, Luebke D, Edmondson H, Mikity V. A mixed hamartoma of the liver: light and electron microscopy. Hum Pathol 1978;9:211–221.
11. Goossens A, Gepts W, Saundubray J. Diffuse and focal nesidioblastosis. A clinicopathological study of 24 patients with persistent neonatal hyperinsulinemic hypoglycemia. Am J Surg Pathol 1989;13:766–775.
12. Morison I, Becroft D, Taniguchi T, Woods C, Reeve AE. Somatic overgrowth associated with overexpression of insulin-like growth factor II. Nat Med 1996;2:311–316.
13. Fukuzawa R, Umezawa A, Ochi K, Ikeda H, Hata J-I. High frequency of inactivation of the imprinted *H19* gene in "sporadic" hepatoblastoma. Int J Cancer 1999;82:490–497.
14. Steenman M, Rainier S, Dobry C, Grundy P, Horon I, Feinberg A. Loss of imprinting of IGF2 is linked to reduced expression and abnormal methylation of H19 in Wilms' tumor. Nat Genet 1994;7:433–439.
15. Reik W, Brown K, Schneid H, Bouc Y, Bickmore W, Maher E. Imprinting mutations in the Beckwith-Wiedemann syndrome suggested by an altered imprinting pattern in the IGF2-H19 domain. Hum Mol Genet 1995;4:2379–2385.
16. Hedborg F, Holmgren L, Sandstedt B, Ohlsson R. The cell type-specific IGF2 expression during early human development correlates to pattern of overgrowth and neoplasia in the Beckwith-Wiedemann syndrome. Am J Pathol 1994; 145:802–817.
17. Ogawa O, Becroft D, Morisin I, et al. Constitutional relaxation of insulin-like growth factor II gene imprinting association with Wilm's tumor and gigantism. Nat Genet 1993;5:408–412.
18. Weksberg R, Shen D, Fei Y, Song Q, Squire J. Disruption of insulin-like growth factor 2 imprinting in Beckwith-Wiedemann syndrome. Nat Genet 1993;5:143–150.
19. Moulton T, Crenshaw T, Hao Y, et al. Epigenetic lesions at the H19 locus in Wilms' tumor patients. Nat Genet 1994; 7:440–447.
20. Chao L, Huff V, Tomlinson G, Riccard V, Strong L, Saunders G. Genetic mosaicism in normal tissue of Wilms' tumour patients. 1993;3:127–131.
21. De Lonlay P, Fournet J-C, Rahier J, et al. Somatic deletion of the imprinted 11p15 region in sporadic persistent hyperinsulinemic hypoglycemia of infancy is specific of focal adenomatous hyperplasia and endorses partial pancreatectomy. J Clin Invest 1997;100:802–807.
22. Verarre V, Fournet J-C, De Lonlay P, et al. Paternal mutation of the sulfonylurea receptor (*SUR1*) gene and maternal loss of imprinted 11p15 imprinted gene lead to persistent hyperinsulinism in focal adenomatous. J Clin Invest 1998;102:1286–1291.
23. Toda T, Yamada M, Nakama B, Yamazato M, Hokama A, Muto Y. Immunohistochmical and electron microscopic findings in a case of mixed hamartoma of the liver. Jpn J Surg 1990;20:101–106.
24. Arthur C. Primary hepatic tumor in children. In: Fingold M, ed. Pathology of Neoplasia in Children and Adolescents. Philadelphia WB Saunders, 1986;352–354.
25. Okamoto K, Morison I, Taniguchi T, Reeve AE. Epigenetic changes at the insulin-like growth factor II/H19 locus in developing kidney is an early event in Wilms tumorigenesis. Proc Natl Acad Sci USA 1997;94:5367–5371.
26. Ekström TJ, Cui H, Li X, Ohlsson R. Promoter-specific IGF2 imprinting status and its plasticity during human liver development. Development 1995;121:309–316.

ELSEVIER

Biochimica et Biophysica Acta 1518 (2001) 137–144

BIOCHIMICA ET BIOPHYSICA ACTA
BBA

www.bba-direct.com

Promoter paper

# Epigenetic mark sequence of the H19 gene in human sperm

Toshio Hamatani [a], Hiroyuki Sasaki [c], Ko Ishihara [c], Naoko Hida [a], Tetsuo Maruyama [a], Yasunori Yoshimura [a], Jun-ichi Hata [b], Akihiro Umezawa [b,*]

[a] *Department of Obstetrics and Gynecology, Keio University School of Medicine, Tokyo 160-8582, Japan*
[b] *Department of Pathology, Keio University School of Medicine, 35 Shinanomachi, Shinjuku-ku, Tokyo 160-8582, Japan*
[c] *Division of Human Genetics, Department of Integrated Genetics, National Institute of Genetics and Department of Genetics, Graduate University for Advanced Studies, Mishima, Shizuoka 411-8540, Japan*

Received 20 November 2000; received in revised form 1 February 2001; accepted 1 February 2001

**Abstract**

We have investigated the epigenetic mark in the human H19 gene. The H19 promoter is methylation-free in human sperm, but it is methylated in the paternally derived allele of most adult tissues. Consequently, the H19 gene is exclusively transcribed from the maternal allele. It was demonstrated that the differentially methylated region (DMR) located 2 kb upstream from mouse *H19* is essential for the imprinting of *H19*. A 39 bp sequence in DMR has a high degree of similarity between humans, mice and rats. The highly conserved 15 bp core region of the consensus sequence contains four methylatable sites, and thus has been proposed as a potential imprinting mark region. In this study, fine epigenetic sequencing analysis was performed on the sperm DNA in comparison with other adult organs. Interestingly, the conserved sequence of the potential mark region was methylated in almost all the sperm genomes analyzed. Furthermore, the single dinucleotide CpG, whose methylation affects the accessibility of the element to CTCF, was methylated in the conserved core in the human sperm. These results suggest that the human core sequences may act as an imprinting center in the reciprocal monoallelic expression of *H19*. 

*Keywords:* H19; Imprinting; Bisulfite; Methylation; Sperm; Human

## 1. Introduction

Genomic imprinting describes the phenomenon of heritable parent-of-origin-specific expression of genes. Two contiguous genes, the gene for insulin-like growth factor II (IGF2) and H19 gene, are imprinted in both humans and mice: the paternal allele of IGF2 and the maternal allele of H19 are expressed in most tissues [1–4]. This reciprocal imprinting has been intensively studied to understand the mechanisms of genomic imprinting. However, the molecular mechanisms that determine the reciprocal imprinting of the human Igf2 and H19 genes have never been fully understood.

The methylation of the H19 promoter is strictly regulated during development. The H19 promoter is methylation-free in human sperm, but the paternally derived H19 promoter is methylated at a later stage of development and in adult organs [5]. Epigenetic marks, such as methylation and/or the chromatin of certain elements, have been hypothesized to the control methylation status of the H19 promoter. DNA methylation analyses of the mouse H19 gene have identified a differentially methylated region (DMR) located 2 kb upstream from H19. The H19 DMR plays a pivotal role in the imprinting of both Igf2 and H19 in the mouse [6] and harbors a maternal-specific chromatin conformation in the form of nuclease-hypersensitive sites and unique dimethyl sulfate reactivity that may confer a new maternal epigenetic imprint [7,8]. The DMR is unmethylated when maternally derived, but methylated when paternally derived. This allele-specific methylation is maintained throughout murine development and may therefore function as an imprinting mark [9,10].

The maternal chromatin conformation represents a DNA boundary element that limits access of Igf2 to the 3′ H19 enhancers [8], therefore silencing the maternal Igf2. Furthermore, several lines of evidence from independent experiments have suggested that the DMR plays dual roles in imprinting: a methylated DMR on the paternal chro-

* Corresponding author. Fax: +81-3-3353-3290;
E-mail: au@med.keio.ac.jp

0167-4781/01/$ – see front matter 
PII: S0167-4781(01)00190-7

mosome appears to act as an inactivation center that methylates and silences the neighboring H19, whereas an unmethylated DMR on the maternal chromosome may serve as a boundary element or a chromatin insulator that blocks the interaction between Igf2 and the downstream enhancers [11–13].

A more recent, closer examination of the DMR by homology plot analysis has identified a 39 bp sequence with a high degree of similarity in humans, mice, and rats [14]. This sequence element is repeated six times in humans (hcs1–hcs6) (Fig. 1A) [5,15] and five times in mice (mcs1–mcs5) and rats (rcs1–rcs5) [14,16,17]. The highly conserved 15 bp core region of the consensus sequence contains four methylatable CG dinucleotide (CpG) sites (Fig. 1D–B). Moreover, four (mcs1–mcs3 and mcs5) of the five mouse elements are associated with a major maternal-specific DNase I hypersensitive site [8,18], suggesting that they might bind a regulatory protein(s). In humans, a number of *Hpa*II and *Hha*I sites in the H19 DMR have also been shown to be methylated in sperm and on the paternal alleles of somatic cells, and this methylation was found to be preserved in pooled 8–32 cell stage embryos [15].

An evolutionarily conserved sequence hcs within these potential imprinting control regions (ICRs) of humans (Fig. 1B) may be involved in the control of H19 and IGF2 imprinting as well as the mouse conserved sequence mcs. The ICR has boundary or insulator activity, but apparently no silencing activity [19–21]. The core sequences of mcs bind a protein called CTCF [19,20,22], which is required for boundary function in the chicken β-globin genetic locus [20]. Binding of CTCF to its recognition site is abolished when the ICR is methylated, even when only a single CpG dinucleotide is methylated [23], and mutated binding sites no longer show enhancer-blocking activity. These results establish a strong correlation between methylation and an open or unblocked boundary.

Conventional DNA methylation studies using genomic Southern and methylation-sensitive restriction enzymes such as *Hpa*II and *Hha*I offer a convenient method of probing the DNA methylation status of the CG nucleotide sequences (CpGs) recognized by the restriction enzymes. The methylation status of other CpGs, which constitute the majority of such sites, is not detected by this methodology. By contrast, bisulfite genomic sequencing allows investigation of the methylation status of every cytosine nucleotide in the selected genomic DNA fragment. Although Frevel et al. demonstrated the methylation status of cytosines in hcs6 to determine the downstream boundary of the human DMR [5], the methylation of cytosines in the core sequences of other human conserved sequences has not been well characterized. The results of the present study suggest that the epigenetic status of cytosines in the core sequences of hcs may act as an imprinting mark in the reciprocal monoallelic expression of the H19 gene.

## 2. Materials and methods

### 2.1. Sequence data

The upstream H19 sequence used in this study is available from GenBank, and its accession number is AF125183 [5].

### 2.2. Human semen, kidney, and brain samples

Human semen was obtained from volunteers after obtaining their informed written consent. Human kidney and brain samples were also obtained after obtaining the informed written consent of bereaved families. The investigations were conducted according to the guidelines approved by the Ethics Committee of Keio University School of Medicine.

### 2.3. Preparation of human semen

Semen samples were obtained from volunteers whose ejaculate was considered to be normal based on WHO criteria [24]. After liquefaction, the semen was filtered through a 50 μm pore nylon membrane filter (Nippon Rikagaku Kikai, Tokyo, Japan), and centrifuged by discontinuous Percoll density gradients (20, 40, 60, and 80% Percoll) with human tubal fluid (HTF; Life Technologies, MD, USA) at 1000×*g* for 25 min [25]. The sperm pellet was resuspended in HTF containing 3.5% human serum albumin (HSA), and after gentle layering on 200 μl of 80% Percoll, it was centrifugated at 500×*g* for 10 min. The sperm layer on 80% Percoll was separated and gently layered beneath the HTF with 3.5% HSA, and motile sperm were allowed to swim up during incubation at 37°C in 5% $CO_2$ in air for 60 min. Approx. 80% of the upper layer containing motile sperm was withdrawn [26]. The sperm suspension was washed again with PBS and centrifuged again as described above. The precipitated sperm was utilized for DNA extraction. These procedures were performed to eliminate contamination by lymphocytes and degraded sperm DNA fragments in the semen plasma.

### 2.4. DNA extraction and bisulfite modification

DNAs were extracted and purified from human sperm, kidney, and brain by standard methods [27]. Prior to bisulfite treatment, 8 μg of genomic DNA was digested for 16 h at 37°C with *Eco*RI. The digested DNAs were ethanol-precipitated and resuspended in 80 μl of $H_2O$. The bisulfite treatment to convert unmethylated cytosine to uracil was carried out as described elsewhere [28] with the following modifications. The bisulfite reaction under mineral oil was performed at 55°C for 16 h in a total volume of 947.6 μl containing 3.4 M sodium bisulfite (Sigma) and 0.5 mM hydroquinone (Sigma), and was followed

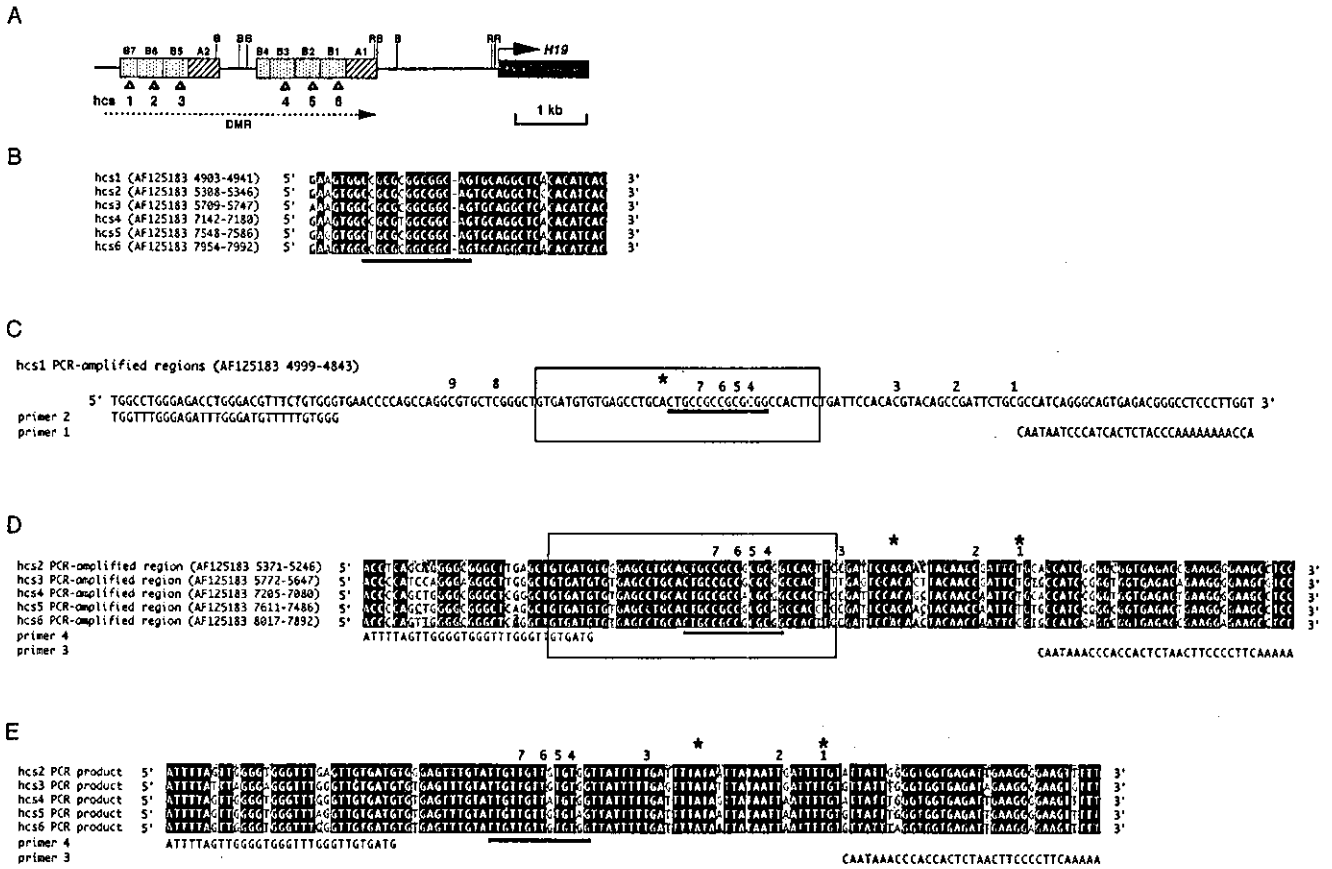



Fig. 1. Map of the upstream DMR of human *H19*. (A) Hatched boxes (A1 and 2) and stippled boxes (B1–7) represent two types of the human 400 bp repeat sequences; open arrowheads, locations of the conserved 39 bp sequences; B, *Bam*HI; R, *Eco*RI. (B) Human sequence alignment (hcs1–hcs6) of the evolutionarily conserved 39 bp sequence elements in the upstream H19 DMR of mammals. The nucleotide positions (in parentheses) are according to the numbering schemes for AF125183. The highly homologous bases in hcs1–hcs6 are shaded. The highly conserved 15 bp core (CCGCGNGGNGGC-AG) of the evolutionarily conserved sequence is underlined. (C) The PCR-amplified sequence (including hcs1) from the bottom strand with primers 1 and 2. The CG dinucleotides (CpGs) between the primers are numbered on top. The framed sequence shows the evolutionarily conserved sequence hcs1. The highly conserved 15 bp core of the consensus sequence is underlined. The asterisks at the top indicate the location of single base pair polymorphisms: A/G (in the bottom strand, nucleotide 4924, AF125183). (D) The possible PCR-amplified sequences (including hcs2–5, or 6) from bottom strand with primers 3 and 4. The CpGs between the primers are numbered on top. The framed sequences show the evolutionarily conserved sequences (hcs2–6). The highly homologous bases among hcs2–6 are shaded. The highly conserved core of the consensus sequence is underlined. The asterisks at the top indicate the location of single base pair polymorphisms (in the bottom strand), in 5′ to 3′ order: hcs3, A/G (nucleotide 5701, AF125183); hcs5, T/C (nucleotide 7523, AF125183). (E) The possible PCR products from the bisulfited bottom strand including the regions of hcs2–6.

by denaturation with alkali. Reactions were desalted using the Wizard DNA cleanup system (Promega). The DNA was eluted in 75 μl of $H_2O$, incubated with 0.3 M NaOH for 15 min at 37°C, neutralized with 3 M ammonium acetate, and precipitated in ethanol. The bisulfite-treated DNA was then resuspended in 50 μl of $H_2O$ and stored at −20°C.

### 2.5. PCR amplification, cloning, and sequencing

The PCR primer sequences were chosen on the basis of containing several guanines and as few cytosines as possible and no simple sequence repeats or homopolymers. After bisulfite modification, the two strands of the duplex are no longer complementary and require a different set of primers for each strand. This primer design also avoids selective amplification of strands based on presumptions about methylation status [29]. The strand being targeted in this study is the 'bottom' strand (Fig. 1C–E). A 156 or 126 bp region from the 'bottom' strand of the human H19 gene was amplified with Taq Gold polymerase (Applied Biosystems, Japan). The primer sequences were: primer 1, 5′-ACCAAAAAAAACCCATCTCACTACCCTAA-TAAC-3′; primer 2, 5′-TGGTTTGGGAGATTTGGGA-TGTTTTTGTGGG-3′; primer 3, 5′-AAAAACTTCCC-CTTCAATCTCACCACCCAAATAAC-3′; primer 4, 5′-ATTTTAGTTGGGGTGGGTTTGGGTTGTGATG-3′.

The PCR primers 3 and 4 were designed to amplify across the core sequences of hcs2–6 (Fig. 1D,E). As a result of the bisulfite PCR with primers 3 and 4, 126 bp regions including the core sequences of hcs3 and hcs5 were amplified, but the regions of the core sequences of hcs2, hcs4, and hcs6 were never amplified.

All reactions contained 0.6 μM primers, 0.2 mM

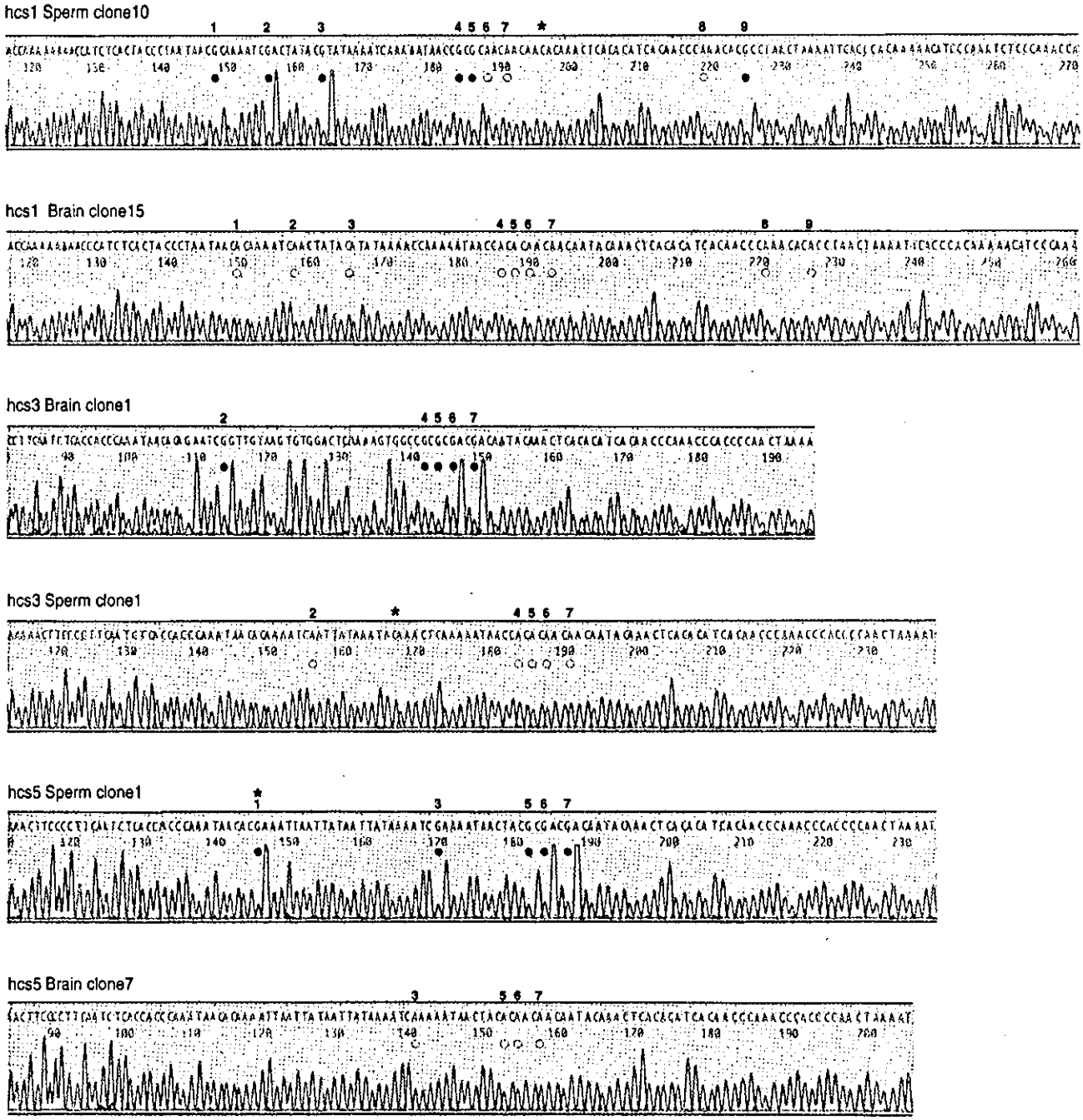



Fig. 2. DNA sequence of a single clone derived from the PCR-amplified hcs regions. Circles represent the cytosines of CpGs: open circles, unmethylated cytosines; filled circles, methylated cytosines. The numbers at the top correspond to the numbers of the CpGs shown in Fig. 1C–E. The asterisks at the top indicate the location of single base pair polymorphisms: hcs1, T/C (nucleotide 4924, AF125183); hcs3, T/C (nucleotide 5701, AF125183); hcs5, A/G (nucleotide 7523, AF125183).

dNTPs, 50 mM KCl, 10 mM Tris-HCl (pH 8.3), and 1.25 units of AmpliTaq Gold DNA polymerase (Applied Biosystems). The PCRs were performed on a GeneAmp Thermal Cycler-9700 (Applied Biosystems) using the following programs: one cycle at 95°C for 13 min, five cycles at 94°C for 0.5 min, annealing temperature for 1 min, 72°C for 2 min, and one cycle at 72°C for 15 min. The PCR products obtained were gel-purified by electroelution with a DNA trapping kit (DNA CELL, Daiichi-Kagaku). Purified fragments were ligated into pGEM-easy T vector (Promega) and cloned into highly competent *Escherichia coli* host cells (JM109, Toyobo). Random white colonies were screened for the correct-sized insert in the plasmid (*Eco*RI digestion). Typically, 30–90% of white colonies contained the insert of interest. These DNAs were then sequenced on an ABI310 automatic sequencer by dye ter-

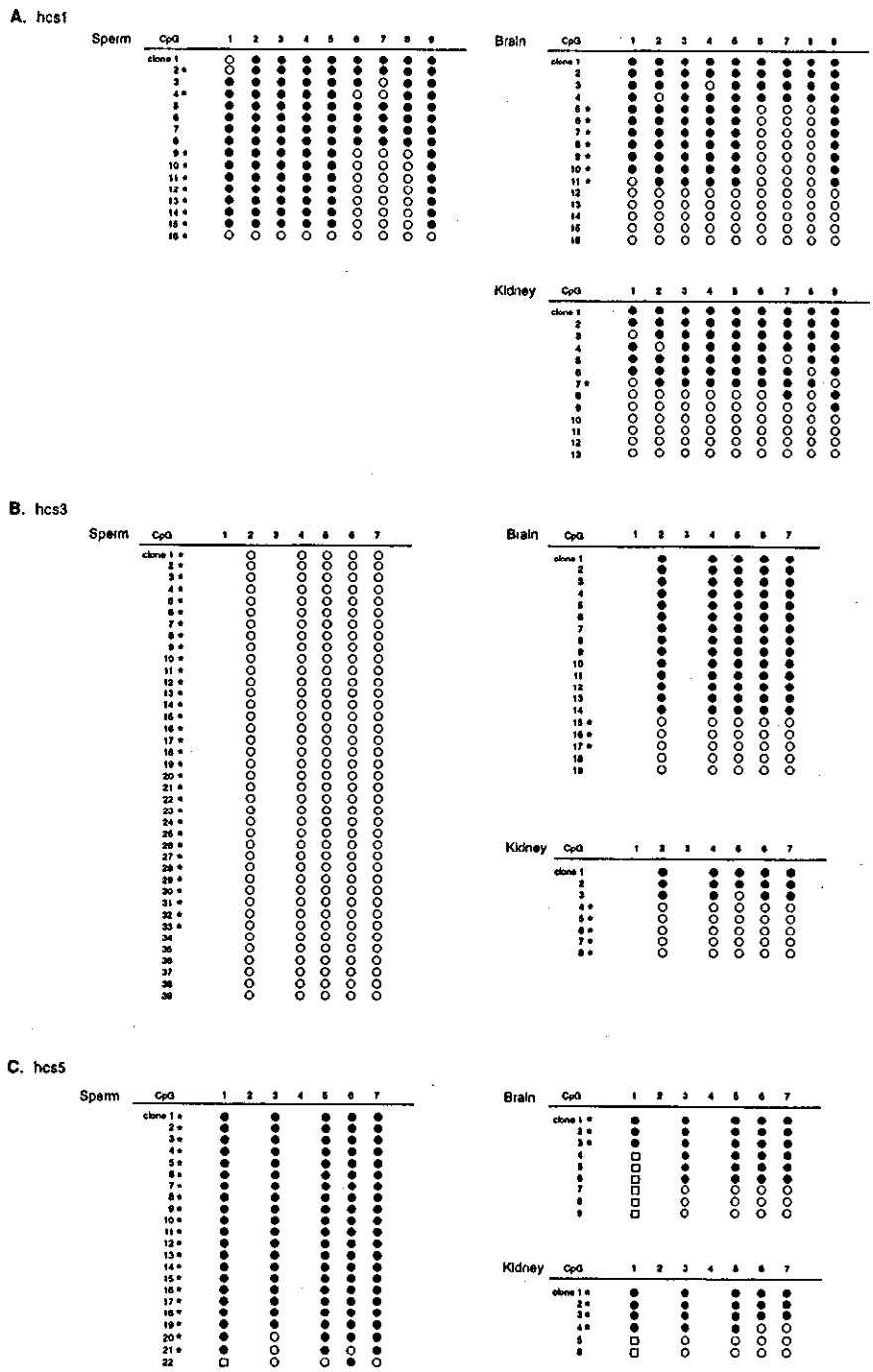



Fig. 3. Bisulfite genomic sequencing analysis of hcs1, hcs3 and hcs5. Genomic DNAs from sperm, brain and kidney were bisulfite-treated and PCR-amplified and single clones were sequenced to show the methylation pattern of CpGs (numbered in Fig. 1C,D) located in the H19 upstream DMR. These data are summarized in the bar graphs in Fig. 4. The heterogeneity of the methylation status of the individual strands is evident in this representation of the primary data. Each clone corresponds to an individual strand of DNA. CpGs are represented by circles: filled circles, methylated cytosines; open circles, unmethylated cytosines. There was a single base pair polymorphism (T/C) at the location of the first CpG of the hcs5 region (AF125183 7523). The open square stands for a thymine or an unmethylated cytosine due to the polymorphism.

minator chemistry (Applied Biosystems) with the M13 reverse universal primers. The data from sperm DNAs were obtained from two independent sets of DNA modification, amplification, and cloning. The full set of DNA sequence data is available from the corresponding author (A.U.).

## 3. Results

The core sequences of hcs1 were successfully amplified in spite of the difficulties of PCR amplification in the repetitive sequence region. There were nine CpGs between primers 1 and 2 (Fig. 1C). Representative detailed meth-

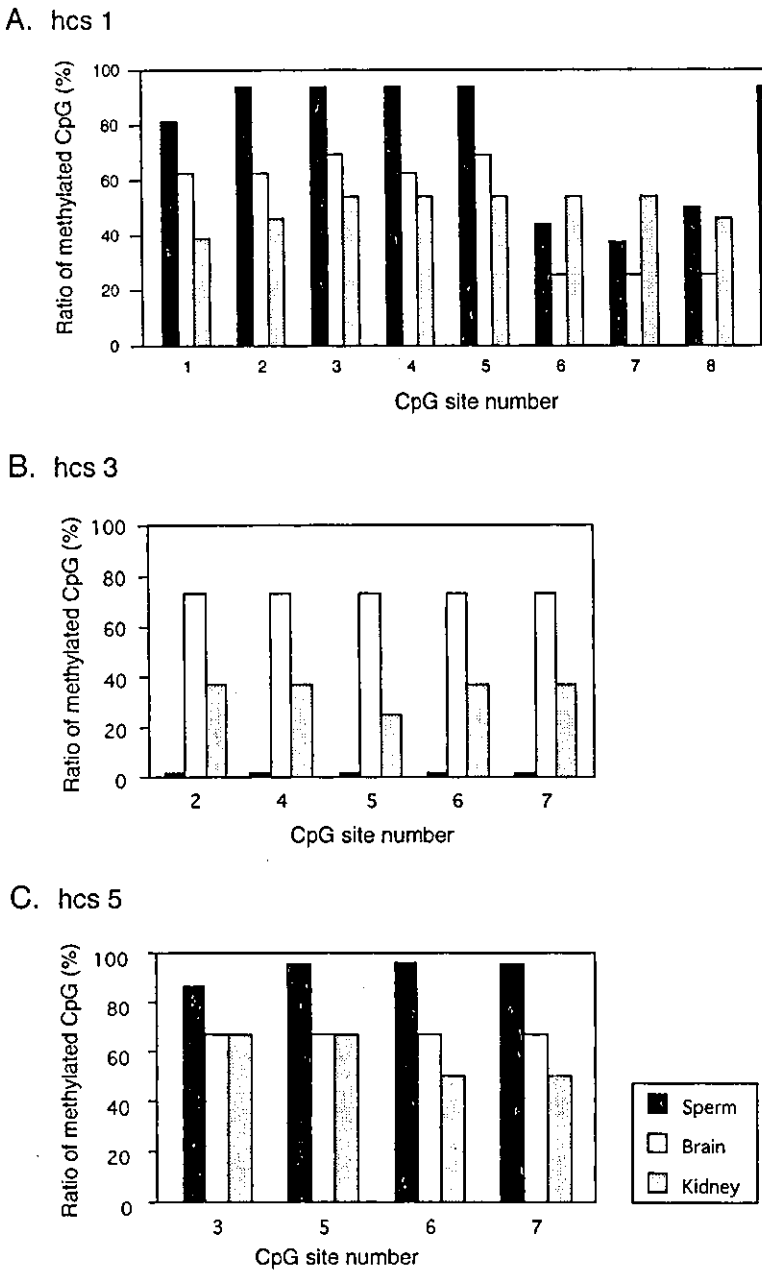



Fig. 4. Methylation frequency of CpGs in hcs1, hcs3 and hcs5. The numbers beneath the bars indicate the positions of the CpGs shown in Fig. 1. The hcs1 is 156 bp in length and spans from nucleotide 4843 to 4999 (AF125183), including nine assayable CpGs (Fig. 1C). The hcs3 is 126 bp in length and spans from nucleotide 5647 to 5772 (AF125183), including five assayable CpGs (Fig. 1D). The hcs5 is 126 bp in length and spans from nucleotide 7486 to 7611 (AF125183), including four assayable CpGs (Fig. 1D). Each bar (filled bar, sperm DNA; open bar, brain DNA; stippled bar, kidney DNA) indicates the percentage of clones that were methylated.

ylation patterns from single clones are shown in Fig. 2. The methylation pattern and frequency of residual cytosines of hcs1 CpGs are shown in Fig. 3A and 4A, respectively. Sperm DNA was highly methylated in CpGs #1, 2, 3, 4 and 5 in hcs1. By contrast, only approx. 50% of the same CpG sites, #1–5, were shown to be methylated in kidney and brain DNAs.

Since Frevel et al. demonstrated that the A1 repeat (Fig. 1A) may constitute the 3′ boundary of the imprint mark, we investigated the epigenetic sequences to determine the extent of the 3′ boundary of the imprint mark or DMR. The core sequences of hcs3 and hcs5 were successfully amplified when primers 3 and 4 were used. The amplified region includes 5 and 4 of the CpGs in hcs3 and hcs5, respectively (Fig. 1D).

The detailed methylation patterns from single clones of paternal and maternal alleles are shown in Fig. 3. The proportion of clones derived from bisulfite-modified DNAs with residual cytosines of hcs3 CpGs or hcs5 CpGs are shown in Fig. 4. Sperm DNA lacks methylation

at CpGs in hcs3 with CpG sites #2, 4, 5, 6 and 7, although kidney and brain DNAs have been shown to be approx. 50% methylated at every CpGs site (Figs. 3 and 4). Sperm DNA was highly methylated in the CpGs #3, 5, 6 and 7 of hcs5, whereas only approx. 50% of the same CpGs sites were shown to be methylated in kidney and brain DNAs, implying that the maternal and paternal alleles are differentially methylated. Single base pair polymorphisms can be detected in the hcs3 and hcs5 region (Fig. 1D) in 5′ to 3′ order (AF125183): A/G (nucleotide 5701 in hcs3), A/G (nucleotide 5692 in hcs3), and T/C (nucleotide 7523 in hcs5). Thus, hcs3 and hcs5 are informative for parental alleles. The methylated allele is considered to be of paternal origin, because the H19 gene is paternally imprinted.

## 4. Discussion

This study focused on analyzing methylation of the evolutionarily conserved sequence within the potential imprinting mark regions neighboring the H19 gene in human sperm in comparison with certain organs. The methylation profile of sperm DNA revealed that the core sequences in hcs1 and hcs5 may be highly methylated. The methylation profile of CpGs in the human H19 DMR of human sperm reinforced the hypothesis that the hcs are the site of an ICR. While the H19 promoter is free of methylation in human and murine sperm (our unpublished observations), it is methylated in the paternally derived genome of most adult tissues [5,30]. Thus, methylation of hcs 1 and hcs 5 in the direct 400 bp repeat in human sperm may play an important role in determining the methylation of the H19 promoter and thus establish the imprinting of this locus.

This evolutionarily conserved sequence within these potential imprinting mark regions has a recognition site for CTCF. The binding site of CTCF contains several CpGs, and binding of CTCF to the recognition site is methylation-sensitive. Methylation of a single dinucleotide CpG #4 is essential for the accessibility of the element to CTCF, the same as with other transcription factors. This analyzed region consists of a 400 bp sequence of six direct repeats and thus is difficult to differentially amplify by PCR. Despite this difficulty, our fine methylation analysis by a bisulfite-PCR sequencing technique revealed that CpG #4 of hcs1 was methylated in almost all the sperm genomes analyzed (15 out of 16 genomes) and in approx. 50% of the genomes of adult organs, supporting the hypothesis that the evolutionarily conserved sequence functions as the imprinting mark of the human H19 gene through binding of CTCF to this site.

The methylation frequency of the same repeat sequence has recently been reported to increase mainly during the developmental stage of gonocytes in both alleles of the male germ line in mice [6]. Cytosines of nucleotide 3109 in mcs4 and nucleotide 3343 in mcs5 (GenBank accession No. AF049041) correspond to CpG #4 of hcs1–6 (Fig. 1C–E in this study and Fig. 4A in ref. [6]) and are the essential CpG sites for CTCF binding. Interestingly, both CpG sites become methylated during the development of primordial germ cells, and the corresponding sites are specifically methylated in the human sperm. However, this is not true of female primordial germ cells. The evidence that the particular 39 bp repeats of DMR are methylated in sperm strengthens the possibility of the existence of the predicted imprinting center region, which is similarly modulated in both human and mice.

## Acknowledgements

The authors wish to thank T. Yoshikawa, M. Okazaki, R. Sakurai, Y. Yamamoto, and N. Sakai (Department of Obstetrics and Gynecology, Keio University School of Medicine) for valuable assistance. Helpful suggestions from M. Fukuma and M. Kishi are also appreciated. This work was supported by a grant from the Ministry of Education, Science and Culture to J.H. and A.U., by Keio University Special Grant-in-Aid for Innovative Collaborative Research Project to J.H. and A.U., by a National Grant-in-Aid for the Establishment of a High-Tech Research Center at Private Universities, by Keio University Grant-in-Aid for Encouragement of Young Medical Scientists to T.H. and by Keio Health Counseling Center to T.H.

## References

[1] T.M. DeChiara, E.J. Robertson, A. Efstratiadis, Parental imprinting of the mouse insulin-like growth factor II gene, Cell 64 (1991) 849–859.
[2] N. Giannoukakis, C. Deal, J. Paquette, C.G. Goodyer, C. Polychronakos, Parental genomic imprinting of the human IGF2 gene, Nat. Genet. 4 (1993) 98–101.
[3] M.S. Bartolomei, S. Zemel, S.M. Tilghman, Parental imprinting of the mouse H19 gene, Nature 351 (1991) 153–155.
[4] Y. Zhang, T. Shields, T. Crenshaw, Y. Hao, T. Moulton, B. Tycko, Imprinting of human H19: allele-specific CpG methylation, loss of the active allele in Wilms tumor and potential for somatic allele switching, Am. J. Hum. Genet. 53 (1993) 113–124.
[5] M.A. Frevel, S.J. Sowerby, G.B. Petersen, A.E. Reeve, Methylation sequencing analysis refines the region of H19 epimutation in Wilms tumor, J. Biol. Chem. 274 (1999) 29331–29340.
[6] T. Ueda, K. Abe, A. Miura, M. Yuzuriha, M. Zubair, M. Noguchi, K. Niwa, Y. Kawase, T. Kono, Y. Matsuda, H. Fujimoto, H. Shibata, Y. Hayashizaki, H. Sasaki, The paternal methylation imprint of the mouse H19 locus is acquired in the gonocyte stage during foetal testis development, Genes Cells 5 (2000) 649–659.
[7] P.E. Szabo, G.P. Pfeifer, J.R. Mann, Characterization of novel parent-specific epigenetic modifications upstream of the imprinted mouse H19 gene, Mol. Cell. Biol. 18 (1998) 6767–6776.
[8] A.T. Hark, S.M. Tilghman, Chromatin conformation of the H19 epigenetic mark, Hum. Mol. Genet. 7 (1998) 1979–1985.
[9] K.D. Tremblay, J.R. Saam, R.S. Ingram, S.M. Tilghman, M.S. Bartolomei, A paternal-specific methylation imprint marks the alleles of the mouse H19 gene, Nat. Genet. 9 (1995) 407–413.

[10] K.D. Tremblay, K.L. Duran, M.S. Bartolomei, A 5′ 2-kilobase-pair region of the imprinted mouse H19 gene exhibits exclusive paternal methylation throughout development, Mol. Cell. Biol. 17 (1997) 4322–4329.
[11] J.L. Thorvaldsen, K.L. Duran, M.S. Bartolomei, Deletion of the H19 differentially methylated domain results in loss of imprinted expression of H19 and Igf2, Genes Dev. 12 (1998) 3693–3702.
[12] A.L. Webber, R.S. Ingram, J.M. Levorse, S.M. Tilghman, Location of enhancers is essential for the imprinting of H19 and Igf2 genes, Nature 391 (1998) 711–715.
[13] J.V. Schmidt, J.M. Levorse, S.M. Tilghman, Enhancer competition between H19 and Igf2 does not mediate their imprinting, Proc. Natl. Acad. Sci. USA 96 (1999) 9733–9738.
[14] M.P. Stadnick, F.M. Pieracci, M.J. Cranston, E. Taksel, J.L. Thorvaldsen, M.S. Bartolomei, Role of a 461-bp G-rich repetitive element in H19 transgene imprinting, Dev. Genes Evol. 209 (1999) 239–248.
[15] Y. Jinno, K. Sengoku, M. Nakao, K. Tamate, T. Miyamoto, T. Matsuzaka, J.S. Sutcliffe, T. Anan, N. Takuma, K. Nishiwaki, Y. Ikeda, T. Ishimaru, M. Ishikawa, N. Niikawa, Mouse/human sequence divergence in a region with a paternal-specific methylation imprint at the human H19 locus, Hum. Mol. Genet. 5 (1996) 1155–1161.
[16] M.A. Frevel, J.J. Hornberg, A.E. Reeve, A potential imprint control element: identification of a conserved 42 bp sequence upstream of H19, Trends Genet. 15 (1999) 216–218.
[17] K. Ishihara, N. Hatano, H. Furuumi, R. Kato, T. Iwaki, K. Miura, Y. Jinno, H. Sasaki, Comparative genomic sequencing identifies novel tissue-specific enhancers and sequence elements for methylation-sensitive factors implicated in Igf2/H19 imprinting, Genome Res. 10 (2000) 664–671.
[18] S. Khosla, A. Aitchison, R. Gregory, N.D. Allen, R. Feil, Parental allele-specific chromatin configuration in a boundary-imprinting-control element upstream of the mouse H19 gene, Mol. Cell. Biol. 19 (1999) 2556–2566.
[19] A.T. Hark, C.J. Schoenherr, D.J. Katz, R.S. Ingram, J.M. Levorse, S.M. Tilghman, CTCF mediates methylation-sensitive enhancer-blocking activity at the H19/Igf2 locus, Nature 405 (2000) 486–489.
[20] A.C. Bell, G. Felsenfeld, Methylation of a CTCF-dependent boundary controls imprinted expression of the Igf2 gene, Nature 405 (2000) 482–485.
[21] C. Kanduri, C. Holmgren, M. Pilartz, G. Franklin, M. Kanduri, L. Liu, V. Ginjala, E. Ulleras, R. Mattsson, R. Ohlsson, The 5′ flank of mouse H19 in an unusual chromatin conformation unidirectionally blocks enhancer-promoter communication, Curr. Biol. 10 (2000) 449–457.
[22] P. Szabo, S.H. Tang, A. Rentsendorj, G.P. Pfeifer, J.R. Mann, Maternal-specific footprints at putative CTCF sites in the H19 imprinting control region give evidence for insulator function, Curr. Biol. 10 (2000) 607–610.
[23] A.C. Bell, A.G. West, G. Felsenfeld, The protein CTCF is required for the enhancer blocking activity of vertebrate insulators, Cell 98 (1999) 387–396.
[24] WHO, WHO Laboratory Manual for the Examination of Human Semen and Sperm Cervical-Mucus Interaction, Cambridge University Press, Cambridge, 1992, p. 44.
[25] S. Kaneko, S. Oshio, K. Kobanawa, T. Kobayashi, H. Mohri, R. Iizuka, Purification of human sperm by a discontinuous Percoll density gradient with an inner column, Biol. Reprod. 35 (1986) 1059–1063.
[26] A. Lopata, M.J. Patullo, A. Chang, B. James, A method for collecting motile spermatozoa from human semen, Fertil. Steril. 27 (1976) 677–684.
[27] A. Umezawa, H. Yamamoto, K. Rhodes, M.J. Klemsz, R.A. Maki, R.G. Oshima, Methylation of an ETS site in the intron enhancer of the keratin 18 gene participates in tissue-specific repression, Mol. Cell. Biol. 17 (1997) 4885–4894.
[28] S.J. Clark, J. Harrison, C.L. Paul, M. Frommer, High sensitivity mapping of methylated cytosines, Nucleic Acids Res. 22 (1994) 2990–2997.
[29] D.M. Woodcock, C.B. Lawler, M.E. Linsenmeyer, J.P. Doherty, W.D. Warren, Asymmetric methylation in the hypermethylated CpG promoter region of the human L1 retrotransposon, J. Biol. Chem. 272 (1997) 7810–7816.
[30] H. Sasaki, A.C. Ferguson-Smith, A.S. Shum, S.C. Barton, M.A. Surani, Temporal and spatial regulation of H19 imprinting in normal and uniparental mouse embryos, Development 121 (1995) 4195–4202.

# SPECIFIC BISULFITE MODIFICATION OF CTG TRIPLET REPEATS OF THE ANDROGEN RECEPTOR GENE, A GENE ASSOCIATED WITH THE TRIPLET REPEAT DISEASE X-LINKED SPINAL AND BULBAR MUSCULAR ATROPHY (KENNEDY DISEASE)

Kensuke Ochi[1, 2], Hiroyuki Nozaki[3], Fumiaki Tanaka[5], Shingo Kato[4], Ryuji Fukuzawa[1], Gen Sobue[5], Yasuo Fukuuchi[3], Yoshiaki Toyama[2], Jun-ichi Hata[1], and Akihiro Umezawa *[1]

Department of [1]Pathology , [2]Orthopaedics, [3]Neurology, and [4]Microbiology, Keio University School of Medicine, 35 Shinanomachi, Shinjuku-ku Tokyo 160-8582, Japan
[5]Department of Neurology, Nagoya University School of Medicine, 65 Tsurumai-cho, Showa-ku Nagoya 466, Japan

* Correspondence: au@med.keio.ac.jp

*(Accepted November 11, 2000)*

## SUMMARY

Expansion of triplet (CAG)n repeats has been associated with triplet-repeat diseases such as X-linked spinal and bulbar muscular atrophy (SBMA). To elucidate the molecular mechanism of the down-regulated expression of the human androgen receptor gene (hAR) in SBMA patients, we speculated that a certain percentage of the CAG triplets are methylated. We employed the bisulfite method to determine methylation of CAG triplets. Although the bisulfite reaction modified the cytosines in the plus strand CAG repeats, a certain percentage of the minus strand CTG repeats were not modified both in genomic DNA from patients and in bacterial plasmids. This unexpected modification was not seen in unexpanded triplets. Thus, we conclude that a peculiar secondary structure around the expanded CAG or CTG triplets protects itself against bisulfite chemical modification. This specific DNA structural property may account for the pathogenesis of CAG triplet repeat diseases.

KEY WORDS: androgen receptor; triplet repeat; bisulfite reaction; Kennedy (SBMA)

## INTRODUCTION

Expansion of CAG triplet repeats at codon 58 in exon 1 of the X-linked human androgen receptor gene (hAR) leads to cause X-linked spinal and bulbar muscular atrophy (SBMA) (1-3). SBMA is an X-linked recessive disorder, characterized by adult onset, and slow progressive proximal spinal and bulbar muscular weakness and atrophy, often with mild androgen insensitivity (4). It has been reported to be a lower motor neuron disease, but no other known hAR mutations, even those that cause complete androgen insensitivity, show motor neuron symptoms. The levels of both hAR mRNA and protein in spinal cords are decreased in SMBA (5, 6). Thus, we compared the methylation state of spinal cord and lymphocyte DNA in SBMA patients with control subjects, using the bisulfite method. We incidentally observed strand-specific bisulfite modification of the hAR gene which is a responsible for the CAG triplet repeat disease SBMA. This unique bisulfite modification of the hAR gene results from specific DNA structural property around CAG or CTG triplets of the hAR gene.

## PATIENTS, MATERIALS AND METHODS

### Patients

Description of patients are summarized in table 1. In brief, lymphocytes were obtained from seven patients with SBMA (cases 1-7) at department of neurology, Keio university school of medicine, and spinal cords were obtained from two autopsied patients (cases 8 and 9) at department of neurology, Nagoya university school of medicine. For control, lymphocytes from two unaffected males and spinal cord from one autopsied unaffected male who died of acute myocardial infarction were analyzed. They were autopsied at department of pathology, Keio university school of medicine.

### Bisulfite chemical modifications

Genomic DNA purified by standard methods was subjected to modification by bisulfite as described previously (7-10). For the plus strand of hAR gene, the sense primer was TGAGTGTAGTATTTTTTGGTGTTAGTTTG (bp +134 - +162), and the antisense primer was CAACCTCTCTCAAAATAACACTCCAAAACC (bp +386 - +357). For the minus strand of hAR gene, the sense primer was ATCTACCCTCAACCACCATCC AAAACCTACC (bp +28 - +58), and the antisense primer was GGATGTAATTTTTT TTGGGGTGGTATTTTAGGG (bp +391 - +359). For amplification of CAG repeats in the bacterial plasmid containing the hAR gene, a new antisense primer, ACTTAAA AAAAACCATCCTCATCACCCTACTAC (+290 - +261), was prepared for the plus strand.

The amplified DNA was cloned into pGEM-T (Promega), and multiple independent isolates were sequenced with an A.L.F. DNA sequencer (Pharmacia). In this reaction, all bisulfite-modified cytosine residues are converted to thymine, while bisulfite-resistant

Table 1. Description of patients and source of genome DNA

| | sex | age of onset | primary symptom | tissue analyzed |
|---|---|---|---|---|
| case 1 | male | 40 | dysbasia | lymphocytes |
| 2 | male | 54 | rhinolalia | lymphocytes |
| 3 | male | 30 | tremor | lymphocytes |
| 4 | male | 30 | tremor | lymphocytes |
| 5 | male | 45 | dysbasia | lymphocytes |
| 6 | male | 35 | muscular atrophy | lymphocytes |
| 7 | male | 39 | dysbasia | lymphocytes |
| 8 | male | 47 | unknown, died at an age of 64 | spinal cord |
| 9 | male | 61 | unknown, died at an age of 74 | spinal cord |
| Normal lymphocytes | | | | |
| 1 | male | | blood sample was taken at an age of 26 | lymphocytes |
| 2 | male | | blood sample was taken at an age of 33 | lymphocytes |
| Normal spinal cord | | | | |
| | male | | died of myocardial infarction at an age of 77 | spinal cord |

cytosine residues remain as cytosines. After PCR, several clones were obtained from the genomic DNA of each patient.

## RESULTS

We prepared lymphocyte DNAs from 7 SBMA patients and a normal control male, and spinal cord DNA from two SBMA patients and spinal cord DNAs from normal male as controls (11). Analysis of lymphocyte genomic DNA and spinal cord genomic DNA from SBMA patients by bisulfite chemical modifications showed that the minus strands were protected from bisulfite modification although almost all the plus strand cytosine bases were modified (Fig. 1A, 1C and 2A). This phenomenon was seen only on expanded triplets, not on $(CTG)_3$, $(CAG)_6$ and $(CAG)_3$, $(CTG)_6$ on same strands.

The same result was obtained also from genomic DNA obtained from autopsy cases of SBMA from affected spinal cord. Again, the minus strands were not modified by bisulfite, while the plus strands were completely modified by bisulfite (Fig. 1B, 1D and 2B). Control lymphocyte and spinal cord DNA from unaffected subjects was modified on the plus and minus strands by bisulfite (Fig. 3).

**Spinal cord, minus strand**

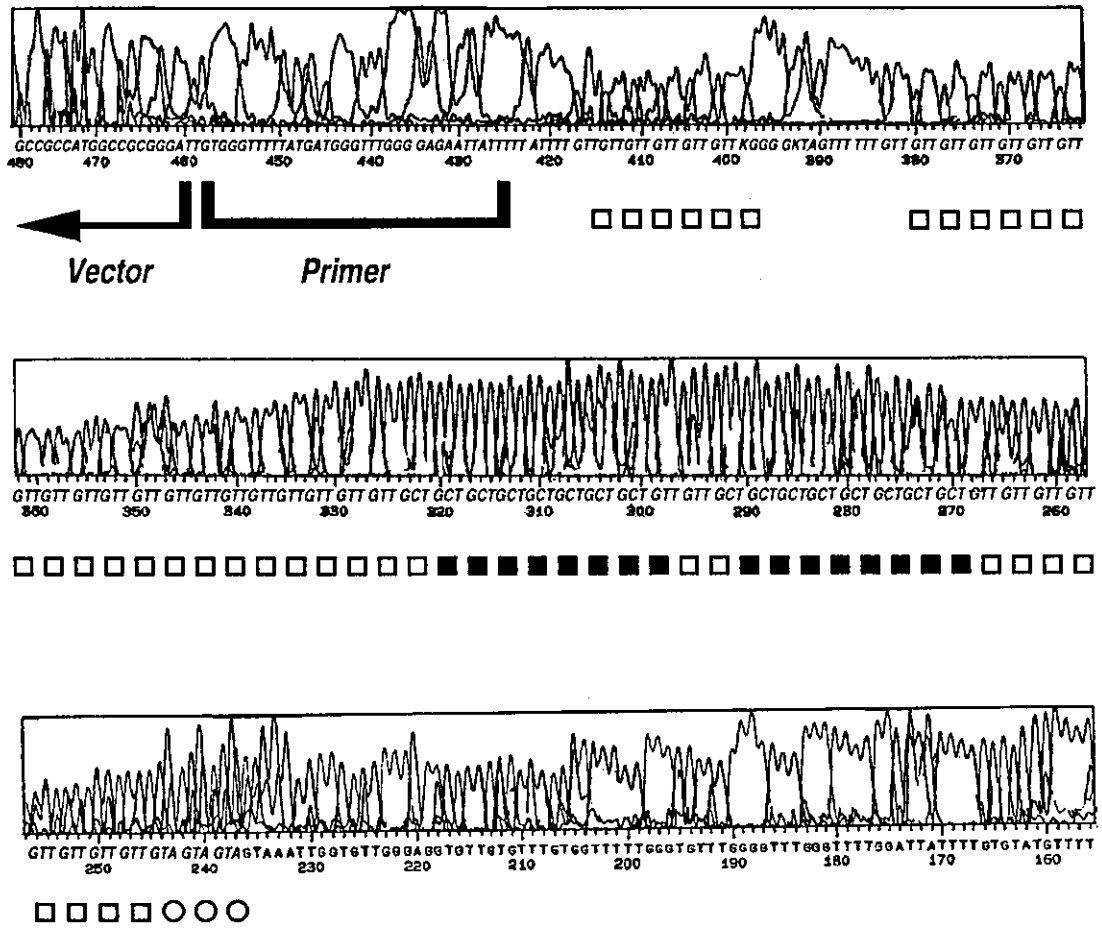
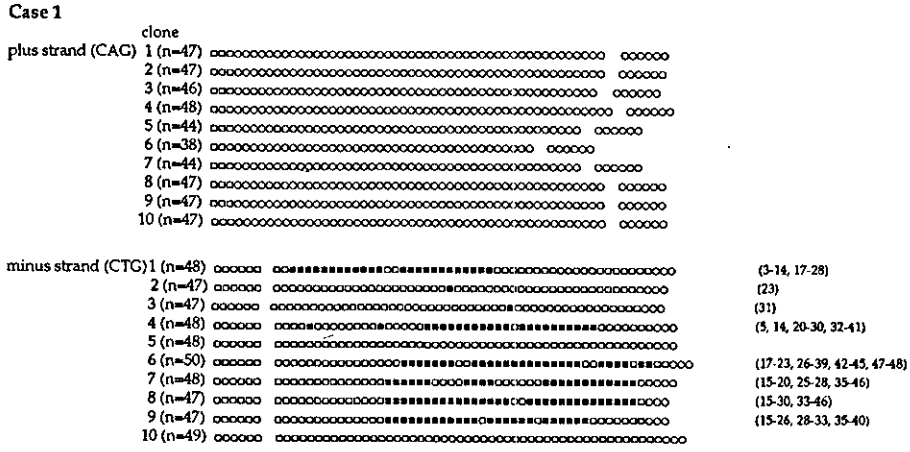
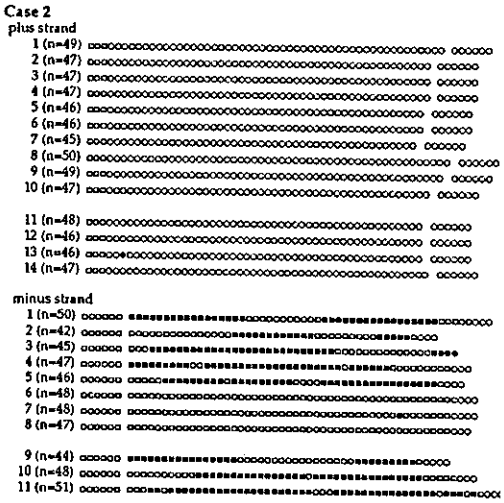
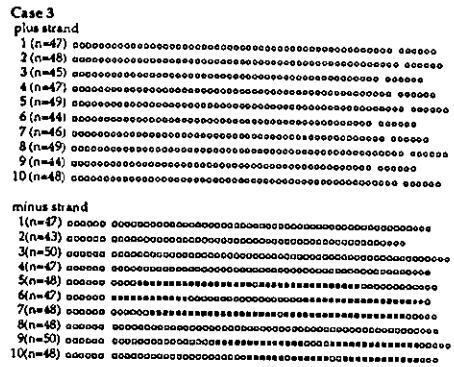
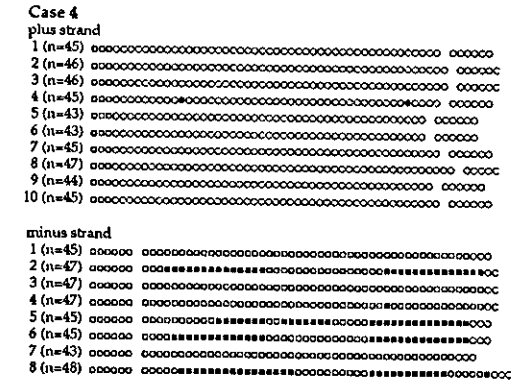


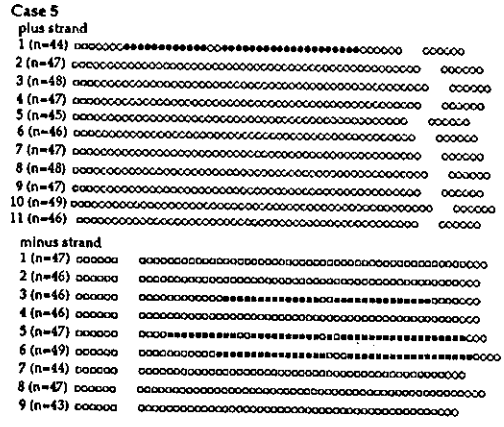

Fig. 1. Representative sequence pattern of bisulfite-modified genome from the patients and normal males. Genomic DNAs were obtained from the spinal cords of the patient. For the plus strand and the minus strand, modified genomic DNAs were amplified as described in the patients and methods. In this reaction, all bisulfite-modified cytosine residues are converted to thymine, while bisulfite-resistant cytosine residues remain as cytosine.


Case 1
clone
plus strand (CAG) 1 (n=47)
2 (n=47)
3 (n=46)
4 (n=48)
5 (n=44)
6 (n=38)
7 (n=44)
8 (n=47)
9 (n=47)
10 (n=47)
minus strand (CTG) 1 (n=48) (3-14, 17-28)
2 (n=47) (23)
3 (n=47) (31)
4 (n=48) (5, 14, 20-30, 32-41)
5 (n=48)
6 (n=50) (17-23, 26-39, 42-45, 47-48)
7 (n=48) (15-20, 25-28, 35-46)
8 (n=47) (15-30, 33-46)
9 (n=47) (15-26, 28-33, 35-40)
10 (n=49)


Case 2
plus strand
1 (n=49)
2 (n=47)
3 (n=47)
4 (n=47)
5 (n=46)
6 (n=46)
7 (n=45)
8 (n=50)
9 (n=49)
10 (n=47)
11 (n=48)
12 (n=46)
13 (n=46)
14 (n=47)
minus strand
1 (n=50)
2 (n=42)
3 (n=45)
4 (n=47)
5 (n=46)
6 (n=48)
7 (n=48)
8 (n=47)
9 (n=44)
10 (n=48)
11 (n=51)


Case 3
plus strand
1 (n=47)
2 (n=48)
3 (n=45)
4 (n=47)
5 (n=49)
6 (n=44)
7 (n=46)
8 (n=49)
9 (n=44)
10 (n=48)
minus strand
1(n=47)
2(n=43)
3(n=50)
4(n=47)
5(n=48)
6(n=47)
7(n=48)
8(n=48)
9(n=50)
10(n=48)


Case 4
plus strand
1 (n=45)
2 (n=46)
3 (n=46)
4 (n=45)
5 (n=43)
6 (n=43)
7 (n=45)
8 (n=47)
9 (n=44)
10 (n=45)
minus strand
1 (n=45)
2 (n=47)
3 (n=47)
4 (n=47)
5 (n=45)
6 (n=45)
7 (n=43)
8 (n=48)


Case 5
plus strand
1 (n=44)
2 (n=47)
3 (n=48)
4 (n=47)
5 (n=45)
6 (n=46)
7 (n=47)
8 (n=48)
9 (n=47)
10 (n=49)
11 (n=46)
minus strand
1 (n=47)
2 (n=46)
3 (n=46)
4 (n=46)
5 (n=47)
6 (n=49)
7 (n=44)
8 (n=47)
9 (n=43)

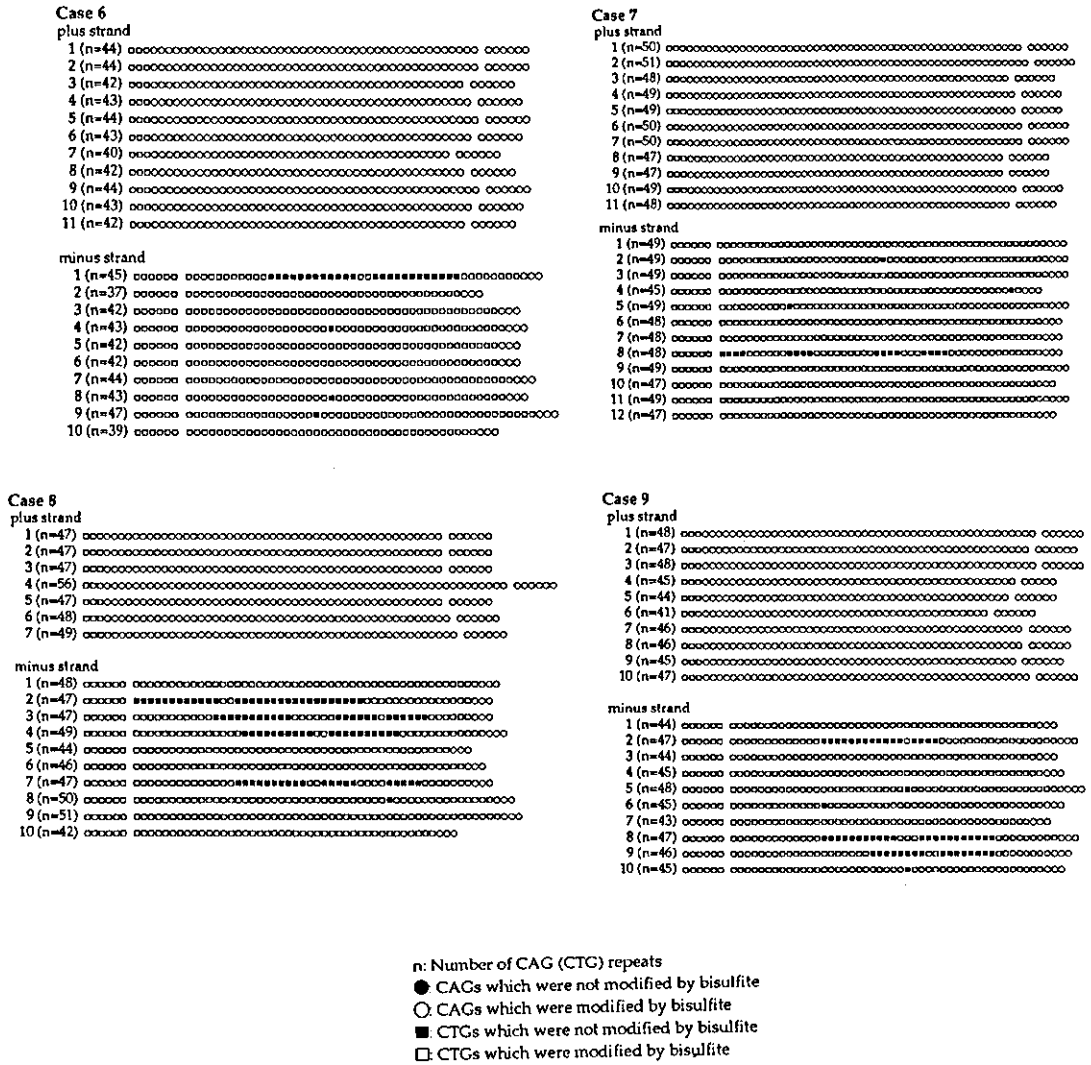



**Fig. 2.** Specific bisulfite modification in the hAR gene from lymphocytes and spinal cords of patients with SBMA. Circles and squares represent CAG and CTG triplets, respectively. Open circles and squares represent bisulfite-unmodified triplets while closed circles and squares represent bisulfite-modified triplets. Samples from cases 1 to 7 were obtained from lymphocytes of SBMA patients and samples from cases 8 and 9 (autopsy cases) are from spinal cord of patients with SBMA patients. In case 1, the numbers of unmodified triplets are shown in the parentheses on the right hand side. The first triplet of an amplified CTG repeat was numbered 1. In each case, approximately ten clones were sequenced and analyzed. In each clone, number of amplified triplets were shown in the parentheses after the clone number. In case 2, two independent bisulfite and PCR reactions were performed to eliminate the possibility of bias or incomplete bisulfite modification; for the plus strand, clones 1-10 (first time) and clones 11-14 (second time); for the minus strand, clones 1-8 (first time) and clones 9-11 (second time).