Allele-specific methylation and expression of an imprinted U2af1-rs1 (SP2) gene


National Cardiovascular Center Research Institute, 5-7-1 Fujishiro-dai, Suita, Osaka 565, Japan

Received October 19, 1994; Revised and Accepted December 1, 1994

ABSTRACT

The mouse U2af1-rs1(SP2) gene, which was cloned by a two-dimensional genome scanning method, is expressed exclusively from the paternally inherited chromosome. This gene has significant similarity to U2AF and located in chromosome 11, of which paternal duplication/paternal deficiency results in a small body. In this report, we cloned genomic U2af1-rs1(SP2) and found its promoter was methylated in a maternal-allele-specific manner. This allelic methylation was not established in parental gametes, but established between 1.5 d.p.c. and 12.5 d.p.c. on the contrary, the allele-specific expression occurred in the two-cell stage when transcription initiates. Absence of the methylation of the upstream region in this stage indicates that methylation is not necessary for inactivation of the expression.

INTRODUCTION

Genomic imprinting is the parental-allele-specific expression of genes. In mammals, parental imprinting ensures functional inequality of paternal and maternal genomes in the fertilized egg and causes developmental failure of embryos produced by parthenogenesis or by gynogenesis or androgenesis (1). Parental effect on particular chromosome regions involving embryo survival and gross phenotypic abnormalities were unequivocally documented by producing paternal or maternal disomes by means of Robertsonian and reciprocal translocations in the mouse (2,3). Such studies have established the fact that several autosomal chromosomes are concerned in imprinting.

In general, an inverse correlation between gene expression and DNA methylation at CpG dinucleotides has been established for a variety of vertebrate and viral genes (4). Therefore DNA methylation may be involved in regulation of genomic imprinting (5–8). We have recently cloned an imprinted gene SP2 by searching for parental-origin-specific CpG methylation using restriction landmark genome scanning (RLGS) (9,10). This gene is methylated at NdI site on the maternal allele and expressed from the paternal allele. SP2 encodes a putative 51 Kd protein with significant similarity to U2 small nuclear ribonucleoprotein auxiliary factor (U2AF), an essential mammalian splicing factor, and is located on mouse chromosome 11, of which maternal duplication/paternal deficiency results in a small body (2). SP2 expressed predominantly in the brain (9), specially in the pyramidal neurons in the hippocampus and dental gyrus (unpublished data). We renamed SP2 as U2af1-rs1 because several related genes have been isolated in mouse and human (unpublished data).

In this study, we cloned genomic U2af1-rs1 to identify the promoter and study whether methylation occurred in an allele-specific manner. Furthermore, we examined the methylation in gamete and expression and methylation in the two-cell embryo.

MATERIALS AND METHODS

Cloning

Genomic DNA was isolated from mouse liver and digested with XbaI and electrophoresed on Seakem GTG (FMC). Fragments around 8.3 kb were cut out from the gel and purified by PREP-A-GENE kit (Bio Rad) to perform ligation with λZAP XbaI digest. This library was screened with the 32P-labeled insert of pSP2 (9). Recombinant clone Xba5-1S was obtained from this library.

Sequencing

The insert of Xba5-1S was subcloned into pBluescript and the sequence around the mRNA coding region was determined by the primer walking method, employing the Taq dideoxyterminator kit (ABI).

RACE

RACE (11) was performed by using the 5’-AmpiiFINDER RACE kit (CLONTECH). The primer used for First-strand cDNA was 5’-CAGTCACCAGGTATCTGCA-3’ and that used for PCR was 5’-AGTTATCCGCAGTGTCGTG-3’.

* To whom correspondence should be addressed
CpG island analysis
The sequence of Xb5-1S was analyzed by using the ‘CpG Bio’ program (9).

Harrplot
Harrplot (12) was performed by using GENTYX-MAC HARR PLOT software.

Plasmid construction
pDra-Cp-CAT was constructed by inserting a 398 bp Dral–CspI fragment of pXb5-1S, which contains from 346 bp upstream to 51 bp downstream of the transcription initiation site, into the HindIII site of O/A CAT (13).

DNA transfection and CAT assay
L cells were cultured in minimum essential medium with 10% fetal calf serum. For each 3.5 cm dish, 1 mg of DNA was transfected by using LIPOFECTAMINE (Gibco-BRL). 48 h after transfection, 20 mg of total protein was assayed for CAT activity by the method described by Sleigh (14). Each assay was repeated at least three times.

Southern analysis
Hybridization of Southern analysis was performed for 20 h at 65°C in a solution of 6× SSC (0.9 M NaCl and 0.09 M sodium citrate), 0.5% (W/V) SDS, 5× Denhardt’s solution (1% each BSA, Ficoll and polyvinylpyrrolidone), 100 mg/ml of heat-denatured herring-sperm DNA and heat-denatured probes (1 × 10⁶ c.p.m./ml). Membranes were washed for 20 min at 65°C in 0.2× SSC, 0.1% SDS.

Isolation of sperm, oocytes and embryos
Standard techniques were used for obtaining sperm eggs and zygotes (15). Sperm were collected by squeezing resected epididymides.

Female mice were injected with 5 IU of pregnant mares’ serum (PMS), followed by an injection 44–48 h later of 5 IU of human chorionic gonadotropin (hCG). After injection, females were housed overnight with PWK males. Embryos were collected at the one-cell (0.5 d.p.c.) and two-cell (1.5 d.p.c.) stages. Unfertilized eggs were flushed from the oviducts of superovulated females 18 h after hCG injection and then treated with hyaluronidase to remove adherent cumulus cells. Embryos were flushed from the oviducts of mice into the medium and then treated with Acrid Tyrode to remove the zona pellucida. Eggs and embryos were collected by micropipette and washed several times with the medium.

HhaI sensitive PCR assay
DNA was obtained from a small number of eggs, and embryos by the guanidine–HCl extraction method (16). A group of 30 eggs or 30 embryos was incubated in lysis solution (6 M guanidine–HCl, 140 ml; 7.5 M ammonium acetate, 10 ml; 20% Sarkosyl, 10 ml; 20 mg/ml Protease K, 2 ml) at 60°C for 3 h. The lysate was passed through a 26-gauge hypodermic needle 10 times to shear the DNA to a minimal length. DNA was precipitated with ethanol after addition of 1 ml of ethachinamate (Nippongene) and 1 mg of pBluescript. The precipitate was collected by centrifugation, washed with 70% ethanol and dissolved in TE.

DNA was first cleaved at XbaI sites in order to reduce the DNA size and thus to facilitate the PCR reaction. One half of the digested DNA was further digested with HhaI and the other half was used as a non-digested control. To monitor completion of digestion, one tenth of the sample was electrophoresed in a 5% polyacrylamide gel. Polymerase chain reaction (PCR) was performed by using HhaI-digested and non-digested samples. Sequence of the primers used were 5’-CGTATTTACCCGGTTATCCCA-3’ and 5’-CCCCCTATCTTTTCCACAAG-AG-3’. The amplified DNA was detected by Southern hybridization.

Semiquantitative RT-PCR assay
Semiquantitative RT-PCR assay was based on the method described by Latham et al. (17). 20 mg of Escherichia coli (RNA as carrier and 0.25 pg of rabbit globin mRNA as internal control for RNA recovery and efficiency of the RT-PCR reaction was added to the 20 embryo, which were then lysed in 100 ml of guanidine isothiocyanate (GTC) solution [4 M GTC, 25 mM sodium citrate (pH 7), 0.5% sarcosyl, 0.1 M β-mercaptoethanol]. Sodium acetate [10 ml of 0.2 M (pH 4)] was added, and the sample was extracted with 100 ml of unbuffered phenol and 25 ml of chloroform: isomylalcohol (49:1). The resulting aqueous phase was collected, and the RNA was precipitated by the addition of 240 ml of ethanol. The precipitate was dissolved in water and treated with RQ1 DNase in the presence of RNasin. The RNA was extracted with phenol and ethanol-precipitated.

Reverse transcription was conducted on 8 embryos equivalents. The reaction was carried out in 50 mM Tris–HCl (pH 8.3), 75 mM KCl, 3 mM MgCl₂, 0.5 M dNTP, 10 mM dithiothreitol, 0.037 mg/ml pd(N)₆ and 400 U of SuperScript II (Gibco-BRL) at 42°C for 1 h. In case of negative control, reaction was carried out without SuperScript II. The samples were then boiled for 5 min and placed on ice.

PCR was conducted on 4 embryo equivalents. The sequence of primers used for U2af1-rs1 were 5’-AGTACATAGCCGTGCCCATG-3’ and 5’-AGATAAACCCGATACCTGG-3’, and for α-globin were 5’-CCAGCCACGGTGCGGATAT-3’ and 5’-GATGACCAGGAGCTTTGAAT-3’. For each set of primers, semilog plots of the amount of radiolabeled ampiclon product as function of cycle number was used to determine the range of cycle number over which ampiclon production was linear on such plots (data not shown). The number of cycles was 40 and 30 for U2af1-rs1 and α-globin, respectively. The amplified DNA was detected by Southern hybridization.

RESULTS
Genomic structure of U2af1-rs1 gene
Only a 8.3 kb fragment was detected in a XbaI digest of mouse genomic DNA by Southern analysis using a mouse cDNA probe (data not shown, 9). A mouse genomic library was then constructed from the XbaI digested genomic DNA fragment around 8.3 kb, and screened with the same cDNA probe yielding three putative genomic U2af1-rs1 clones. Three clones showed the same profile with restriction enzyme digestion. Therefore, we subcloned one of these clones, named Xb5-1S, into pBluescript.
The sequence around the mRNA coding region was determined by the primer walking method employing the Taq dideoxy-terminator kit (ABI) (Fig. 1, DDBJ accession number D26474). The transcription initiation site was determined by 5' RACE. For construction of 1st strand cDNA, an anti-sense primer derived from residues 26–45 of 2916-bp cDNA was used for reverse transcription of brain polyA+ RNA. Amplification was performed with the anti-sense primer derived from residues 7–26 of 2916-bp cDNA and linker primer. The amplified fragment was cloned into the EcoRV site of pBluescript and 6 independent clones were sequenced. All clones started at the location 906th from the beginning of the genomic sequence (Fig. 1). An TATA-like sequence (TAAAGAC) was located 33 bp upstream of the starting site. Comparison of the genomic sequence with cDNA sequence revealed that the U2af1-rs1 is an intronless gene with an unusually long 5' non-coding sequence (1168 bp).

To identify CpG island in this sequence, we plotted Obs/Exp CpG and %G(C) using a 100-bp window moving along the sequence at 1 bp intervals (Fig. 2) There is a big CpG island spanning from upstream of the transcription initiation site to upstream of the translation initiation site. The NotI site, which was previously shown to be methylated in a parental-specific manner, was included in this region.

Harrplot analysis (data not shown) revealed a highly repeated region spanning from upstream of the transcription initiation site to upstream of the translation initiation site. The unit of this repeated sequence is the FoK I family, which is one of those possessing a variable number of tandem repeats (VNTR).

Identification of the promoter region

Before examining the differential methylation and expression of this gene, the promoter activity was measured by using the CAT construct carrying the upstream region. pDNA-Cp-CAT plasmid was transfected into mouse L cell. High activity was observed as shown in Figure 3, suggesting that plasmid is sufficient for basal expression.

Parental specific methylation in promoter region

To study whether the upstream region of transcription initiation site is methylated in a parental specific manner, we first searched for a polymorphic restriction enzyme site in this region by Southern blot. A BanI site located 164 bp upstream of the

Figure 1. Nucleotide sequences of the region around the initiation site of the mouse genomic U2af1-rs1 gene. Transcription initiation site identified by 5' RACE is boxed. The putative TATA box is underlined.
transcription initiation site was found to exist in C57BL/10, but not to exist in Mus musculus molossinus (M.m.molossinus). Therefore, we were able to detect a 361-bp fragment in C57BL/10 and a 547-bp fragment in M.m.molossinus when genomic DNA was digested with BanI, SspI and BglII (Fig. 4). When digestion was performed with HhaI in addition to BanI, SspI and BglII, we found 547-bp M.m.molossinus specific fragment in (M.m.molossinus × C57BL/10)F1 and 361-bp C57BL/10 specific fragment in (C57BL/10 × M.m.molossinus)F1. This indicates methylation of the HhaI sites (H1 to H5, H1 and H2 are in the upstream region) was exclusively on the maternal allele, which is consistent with a functional role for DNA methylation, the repressed allele being methylated.

Analysis of methylation in gamete

Having shown that the upstream region of U2af1-rs1 is differentially methylated, we then set about to determine whether this pattern is established in gamete or not. If modification of this locus serves as an imprinting signal for identifying the parental alleles, these sites should be unmethylated in the sperm while methylated in the oocyte. In order to test this hypothesis, we first analyzed sperm by methylation-sensitive Southern blot using the same probe used for detection of parental specific methylation in the promoter region. We found that the upstream region is almost unmethylated in sperm (Fig. 5A). DNA from mature oocytes was assayed for methylation-sensitive PCR analysis. In contrast to our expectations, we could not amplify the upstream region from the HhaI-digested DNA (Fig. 5B). Thus, the differential methylation pattern found in somatic cells cannot be derived directly from the pre-fertilization parental gametes and must therefore be established at some stage during post-fertilization embryonic development.

Analysis of expression in pre-implantation embryo

By using semiquantitative RT-PCR assay (17) we determine whether U2af1-rs1 is expressed in pre-implantation. U2af1-rs1 RNA was not detected at the one-cell stage, but started to express at the two-cell stage when transcription initiates (18, Fig. 6A). We can distinguish the maternal and paternal allele in the embryos by cleaving the RT-PCR product with a polymorphic enzyme because embryos are F1 of C57BL/6 female and PWK male (BPF1). U2af1-rs1 was expressed from the paternal allele and not from the maternal allele (Fig. 6B). We expected that methylation in the upstream region begins at this stage. However, we could not amplify the upstream region from the HhaI-digested two-cell (1.5 d.p.c.) DNA (Fig. 6C). Thus, methylation in the upstream region seems not necessary for repression of the maternal allele in the two-cell embryo. We detected the parental specific methylation in the upstream region at 12.5 d.p.c. (Fig. 6D). Therefore, methylation is established between 1.5 d.p.c. and 12.5 d.p.c.
DISCUSSION

Analysis of genomic structure revealed that U2af1-rs1 is an intronless gene with an unusually long 5' untranslated sequence (from 906 to 2073, Fig. 1). The CpG island of U2af1-rs1 gene consists of 5' untranslated region and a part of the upstream region. The region of the CpG island corresponds to the highly repeated region revealed by Harplot analysis. The unit of this repeated sequence is FokI family, which is one of those possessing a variable number of tandem repeats (VNTR). B. Neumann and D. P. Barlow reported that imprinted genes and transgenes like the Igf2, the Xist and the TGA transgene contain iterated repeats and have a high GC density (18). Therefore, it is intriguing to think that the CpG island of U2af1-rs1 which consists of iterated repeats, may be involved in some part of the imprinting mechanism.

In order to understand the relationship between parental imprinting and DNA methylation, we have evaluated the methylation in the upstream region of U2af1-rs1 gene (Fig. 4). The upstream region of this gene is differentially methylated in an allele-specific manner in the newborn mouse. However, this modification pattern is not directly inherited from the gametes, but is rather added to the DNA at some stage between 1.5 d.p.c. and 12.5 d.p.c. (Figs. 5, 6). A similar result is observed in the Igf2r region 1 which overlaps the gene promoter, and which is allelically methylated in adult tissues like U2af1-rs1 (8, 20).

We expected that U2af1-rs1 gene would express in the early embryonic stages because this gene is a housekeeping gene, expressing in every tissue (9). This gene is actually expressed in the two-cell stage when transcription initiates (Fig. 6). This expression is from the paternal allele. This is different from Igf2r and Igf2 which express biallelically in preimplantation embryos (17). The component required for inactivating the imprinted allele of U2af1-rs1 seems to be present at the two-cell stage. Demethylation of the upstream region at the two-cell stage indicates methylation is not essential for inactivation of the maternal allele of the U2af1-rs1 gene at this stage. The methylation in upstream regions seems not to inactivate the expression directly, but rather stabilizes the inactivation because inactivation precedes the methylation.

ACKNOWLEDGEMENTS

We thank Dr. J. Forejt for providing PWK, Dr. K. Hidaka for her assistance in Cpg-islands analysis. This work was supported by Special Coordination Funds for promoting Science and Technology (Encouragement System of COE) from the Science and Technology Agency of Japan, a Research Grant for Aging from the Ministry of Health and Welfare and by a Grant-in-Aid for Creative Basic Research (Human Genome Program) from the Ministry of Education, Science and Culture in Japan.

REFERENCES