Proc. Natl. Acad. Sci. USA
Vol. 94, pp. 12509–12514, November 1997
Genetics

Multiple imprinted sense and antisense transcripts, differential methylation and tandem repeats in a putative imprinting control region upstream of mouse *Igf2*

T. MOORE*†, M. CONSTANCIA*, M. ZUBAIR‡, B. BAILLEUL*§, R. FEIL, H. SASAKI‡, AND W. REIK*

*Department of Development and Genetics, The Babraham Institute, Cambridge CB2 4AT, United Kingdom; and ‡Institute of Genetic Information, Kyushu University, Maidashi 3–1-1, Higashi-ku, Fukuoka 812, Japan

Edited by Shirley M. Tilghman, Princeton University, Princeton, NJ, and approved September 16, 1997 (received for review June 24, 1997)

ABSTRACT **The mouse insulin-like growth factor 2 (*Igf2*) locus is a complex genomic region that produces multiple transcripts from alternative promoters. Expression at this locus is regulated by parental imprinting. However, despite the existence of putative imprinting control elements in the *Igf2* upstream region, imprinted transcriptional repression is abolished by null mutations at the linked *H19* locus. To clarify the extent to which the *Igf2* upstream region contains autonomous imprinting control elements we have performed functional and comparative analyses of the region in the mouse and human. Here we report the existence of multiple, overlapping imprinted (maternally repressed) sense and antisense transcripts that are associated with a tandem repeat in the mouse *Igf2* upstream region. Regions flanking the repeat exhibit tissue-specific parental allelic methylation patterns, suggesting the existence of tissue-specific control elements in the upstream region. Studies in *H19* null mice indicate that both parental allelic methylation and monoallelic expression of the upstream transcripts depends on an intact *H19* gene acting in cis. The homologous region in human *IGF2* is structurally conserved, with the significant exception that it does not contain a tandem repeat. Our results support the proposal that tandem repeats act to target methylation to imprinted genetic loci.**

About twenty genes are currently known to exhibit parental allele-specific expression patterns during mammalian development. Mechanistically, the differential expression of the parental alleles is thought to result from parental germline-inherited differences in DNA methylation and chromatin structure: such differences are termed germline "imprints."

At present, the structure and regulation of three endogenous imprinted gene regions has been studied in detail in both mouse and man. These are the *IGF2-H19*, *IGF2R*, and Prader–Willi syndrome regions. Arising from these studies several common structural and mechanistic features of imprinted genes have been described that are relevant to this study. (*i*) Imprinted genes have associated parental allele-specific methylation patterns, in some cases of germline origin (1–5). (*ii*) Some imprinted genes contain tandem repeats associated with differentially methylated regions of potential regulatory significance (6, 7). (*iii*) There is a preponderance of untranslated RNAs associated with imprinted regions, including the untranslated *H19* gene (1, 2), and multiple *SNRPN* (8), and *Ipw* transcripts (9), in the Prader–Willi syndrome region.

The mouse *Igf2* maps to distal chromosome 7 and is flanked 5′ by *Ins2* and 3′ by *H19*, a linkage group that is conserved in the human. The maternal allele of *Igf2* is silenced during fetal development and early studies implicated parental allelic DNA methylation in the upstream region of the gene in the imprinting process. Specifically, a CpG-rich region, differentially methylated region 1, (DMR 1), situated 3 kb upstream of *Igf2* promoter 1 (P1) exhibits a mosaic pattern of methylation, with an excess of methylated sites in sperm-derived DNA (10, 11). DMR 1 is ≈3 kb 3′ of a tandem repeat, suggesting a functional motif with similarities to tandem repeat-associated DMRs in the *Igf2r* and *H19* genes. However, it is unclear whether the quantitative differences in allelic methylation in the *Igf2* DMR 1 contribute to differential parental allelic activity, particularly since subsequent studies of DNA hypersensitive sites (HS) failed to reveal parental allele specificity. Although the DMR 1 may function autonomously as a silencer (12), an autonomous role in the primary imprinting process is undermined by the fact that targeted disruption of the *H19* gene abolishes imprinting of *Igf2*, resulting in biallelic expression and fetal overgrowth (13). Therefore, the presence of differential methylation and tandem repeats in the *Igf2* upstream region notwithstanding, the dominant regulation of *Igf2* imprinting by *H19* argues against the existence of autonomous imprinting control elements in the *Igf2* upstream region. The function of DMR 1 therefore remains elusive. A second DMR (DMR 2) has been described near the 3′ end of *Igf2*, but is not implicated in the primary imprinting mechanism *per se*. Rather, DMR 2 is thought to act as a tissue-specific silencer because it is consistently methylated on the active (paternal) allele in tissues where *Igf2* is expressed (10).

It is unknown to what extent the structural features of the *IGF2* upstream region are conserved between the mouse and human. The presence of a potential functional equivalent of mouse DMR 2 in the human (14) coupled with the occurrence of parental allele-specific methylation in the promoter regions of the human, but not the mouse (11, 15), suggests a complicated picture of partial divergence of control mechanisms. Conservation of the DMR 1 in the human might suggest an important function for this region. To address these questions we have carried out a structural and functional comparison of the *IGF2* upstream region in the mouse and human. Our results suggest the existence of an *H19*-dependent placenta-specific *Igf2* imprinting control element in the mouse, and a comparison with the homologous region in the human sup-

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

This paper was submitted directly (Track II) to the *Proceedings* office.
Abbreviations: HS, hypersensitive sites; DMR, differentially methylated regions; RACE, rapid amplification of cDNA ends; UP, upper primer; LP, lower primer; RT, reverse transcription; AS, antisense; DSS, donor splice site; ASS, acceptor splice site; P1, promoter 1; RP, RACE primer; B6CBF$_1$, (C57BL/6 × CBA)F$_1$; BE, *Bam*HI–*Eco*RI; H1–5, *Hpa*II sites 1–5.
Data deposition: The sequence reported in this paper has been deposited in the GenBank database (accession no. Y13633).
†To whom reprint requests should be addressed. e-mail: tmoore@bbsrc.ac.uk.
§Present address: Institut National de la Santé et de la Recherche Médicale, Unité 124, Institut de recherches sur le cancer de Lille, Place de Verdun, 59045 Lille, France.

ports the proposal that tandem repeats act to target DNA methylation to imprinted genes.

MATERIALS AND METHODS

Human Tissues. DNA and RNA was extracted from normal human term placentae stored at −70°C. High molecular weight genomic DNA was obtained. Gel electrophoresis indicated that RNA was of good quality.

Mice. Expression studies were performed using embryos from (C57BL/6 × CBA)F_1, (B6CBF_1), females mated to B6CBF_1 males. SD7 mice were produced by repeated backcrossing of the *Mus spretus Igf2-H19* region to the C57BL/6 genetic background (four generations to date). Male or female SD7/SD7 mice bred to B6CBF_1 males or females to produce reciprocal (*M. domesticus* × SD7 and SD7 × *M. domesticus*) progeny, (B6CBF_1 × *M. spretus*)F_1 hybrids produced by *in vitro* fertilization and embryo transfer, and typed (B6CBF_1 × *M. spretus*)F_1 × B6CBF_1 backcross progeny, were used for imprinting studies. Embryos maternally disomic for distal chromosome 7 (matdi7 embryos) were produced using the T(7, 15)9H translocation (16). Male and female mice that were heterozygous for a *H19* null mutation (Δ*H19*, ref. 13) were mated to SD7/SD7 homozygotes.

Northern Analysis. Total RNA was extracted as per Chomczynski and Sacchi (17). Poly $(A)^+$ RNA was produced using the mRNA Purification kit (Pharmacia). Total RNA (5–20 μg), or poly$(A)^+$ (2 μg), was electrophoresed in low percentage formaldehyde gels, blotted onto Hybond N^+ (Amersham) membrane, and UV cross-linked. DNA probes were radioactively labeled using the Oligolabeling kit (Pharmacia) and purified using NAP-5 G-25 columns (Pharmacia). Riboprobes were made using the 2.2-kb *Bam*HI–*Eco*RI (BE) fragment (Fig. 1*A*) from cosmid clone CosIGF4, which was subcloned into pBluescript SK. Hybridization and washing of blots was performed with Church reagents (10).

Southern Analysis. Genomic DNA extraction and Southern blot analysis were performed as described (10). Mouse *Igf2* DNA probes were obtained from CosIGF4 and a bacteriophage λ DNA clone containing *Igf2* upstream sequences. Human *IGF2* DNA probes were obtained from CosINS-IGF2 (a gift from E. Holthuizen, Utrecht University, Utrecht, The Netherlands). Methylation levels were quantified by PhosphorImager (Molecular Probes) analysis.

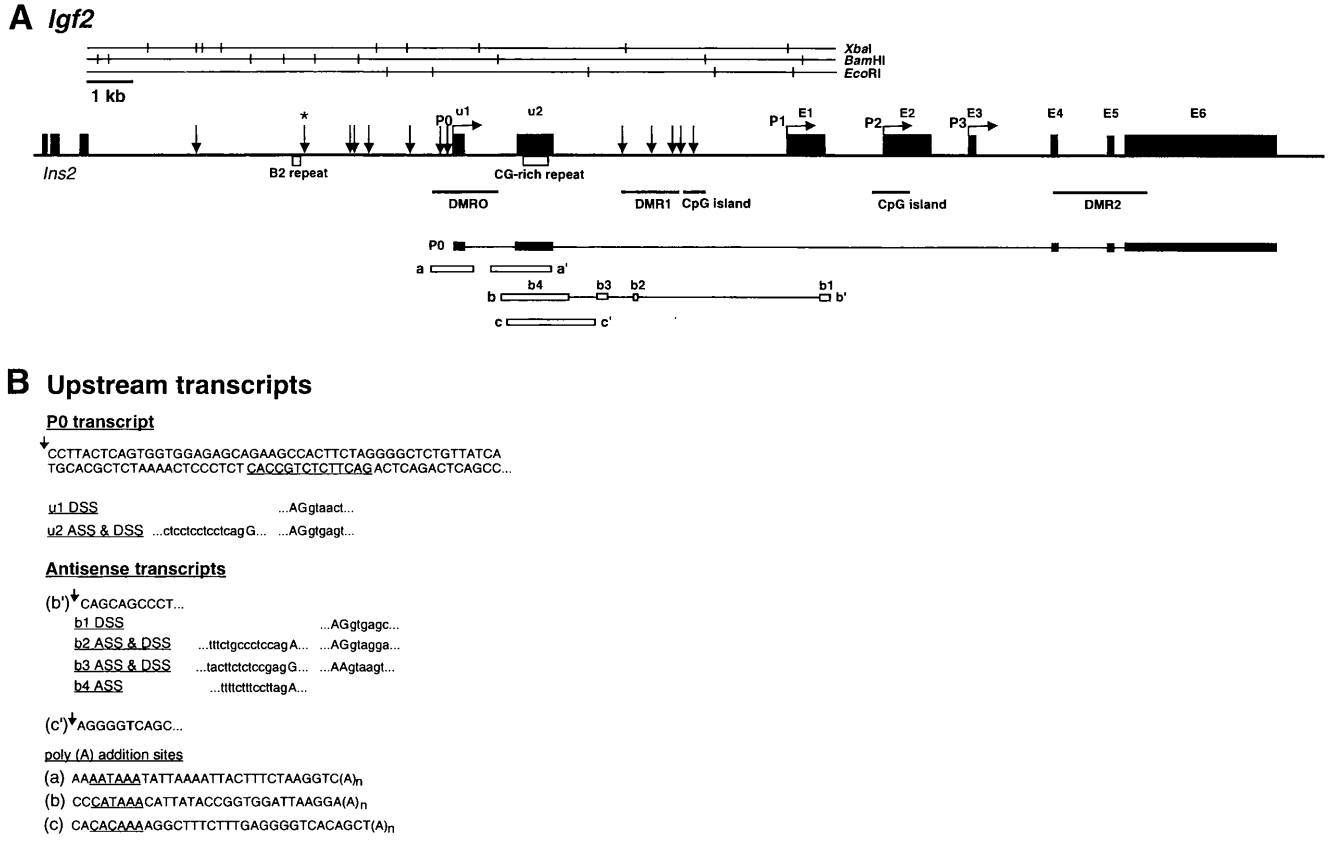

FIG. 1. (*A*) Genomic map of mouse *Igf2* and upstream transcripts. *Igf2* promoters (horizontal arrows): P0, P1, P2, and P3. *Ins2*; *Igf2* exons (black boxes): u1, u2, E1-E6. DNA repeats (shaded boxes): B2, CG-rich. *DNase* I HS (vertical arrows); ∗, embryo-specific HS. P0 transcript: exons (thick line), introns (thin line) includes exons u1, u2, E4-E6. A 1591-bp PCR product of P0 (primers UP2 and LP1) was cloned and sequenced. P0 poly$(A)^+$ addition sites were not mapped; however, on Northern blots P0 is identical in size to the 5-kb P2 transcript (data not shown). The identical size suggests similar inclusion of E6 sequences downstream of LP1. Transcribed AS regions were named according to the nearest associated poly$(A)^+$ addition sites—a, b, c: *Igf2as-a*, *Igf2as-b*, *Igf2as-c*. Putative transcription start sites, a′, b′, and c′. Open boxes represent sequenced AS RT-PCR products from primer pairs UP5 and LP7: b transcript, UP6 and LP6: c transcript, UP3 and LP9 or UP4 and LP5: a or a′ transcript. We were unable to clone or sequence a fragment from primer pair UP3 and LP5, and therefore cannot propose a final structure for a putative a-a′ AS transcript. Three hundred base pair regions upstream of putative transcription start sites were examined for regulatory sequences using the TFD software program at the Medical Research Council, Human Genome Mapping Project. Only the b′ site was associated with major promoter element motifs. (*B*) DNA sequence data for upstream transcripts. Nucleotide positions are numbered as per ref. 18. Putative transcription starts (vertical arrows): P0: 8,512; b′: 17,587; c′: 11,943. P0 acceptor splice sites (ASS) and donor splice sites (DSS): u1 putative ASS (underlined): 8,584–8,597; u1 DSS: 8,825–8,832; u2 ASS: 10,120–10,134; u2 DSS: 10,952–10,959. AS transcript splice sites and (underlined) poly (A^+) addition signal nucleotide positions numbered for the complementary, upper strand, as follows: b1 DSS: 17,406–17,399; b2 ASS, DSS: 12,932–12,918, 12,862–12,855; b3 ASS, DSS: 12,200–12,186, 11,976–11,969; b4 ASS: 11,273–11,259. Poly$(A)^+$ signals: a: 8,120–8,115; b: 9,651–9,646; c: 9,808–9,803.

Nuclease Sensitivity Assays. DNase I sensitivity studies were performed as described (11).

Blots were exposed to film or analyzed using a BAS2000 Image Analyzer (Fuji).

Reverse Transcription–PCR (RT-PCR). RT-PCR analysis was performed essentially as described (11). Standard reagents included Moloney murine leukemia virus reverse transcriptase, RNasin, random primers (Promega) and dNTPs, *Taq* polymerase (Pharmacia). Total RNA was treated with RQ1 *DNase* (Promega) to destroy genomic DNA. RT negative controls were run in parallel and were consistently negative. 3′ and 5′ rapid amplification of cDNA ends (RACE)-PCRs were performed using RACE kits from GIBCO/BRL. PCR products were subcloned using the Invitrogen TA cloning kit and sequenced by Applied Biosystems automatic sequencer. Sequences (5′ → 3′) of upper (UP) and lower (LP) primers (numbered according to ref. 18) are available at deposit@ndbserver.rutgers.edu. Primers included: UP1:8,682–8,702; UP2:8,545–8,565; UP3:8,167–8,187; UP4:9,474–9,494; UP5:9,742–9,762; UP6:9,932–9,952; UP7:22,677–22,697; UP8:19,896–19,915; UP9:11,698–11,718; UP10:8,605–8,625; LP1:24,469–24,449; LP2:8,756–8,737; LP3:8,675–8,655; LP4:8,633–8,613; LP5:10,936–10,913; LP6:11,938–11,918; LP7:17,582–17,562; LP8:24,650–24,630; LP9:9,027–9,007.

RESULTS

Paternal Allele-Specific Expression of Overlapping Sense and Antisense (AS) Transcripts in the *Igf2* Upstream Region is *H19*-Dependent. Transcription units were identified in the *Igf2* upstream region using the 2.2-kb BE probe on blots containing total RNA from mouse fetus and placenta (Figs. 1*A* and 2 *A* and *B*), and adult kidney (Fig. 2*C*). Expression of a predominant 5-kb transcript (Fig. 2*A*) was confined to e10 to e18 placenta as determined by Northern blot analysis (data not shown); however, smaller transcripts (3–4 kb), detected after longer exposures (1–3 weeks), were equally expressed in fetus and placenta. Hybridization of the 2.2-kb BE probe to blots containing RNA from normal and matdi7 embryos indicated that all of the transcripts are imprinted, with exclusive paternal allelic expression (Fig. 2*A*).

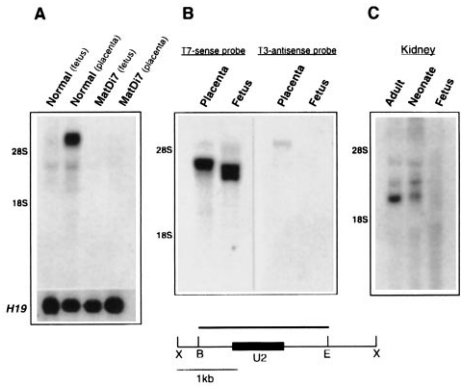

FIG. 2. Northern blot analysis of embryonic and adult RNA. (*A*) Age-matched, late gestation normal and matdi7 mouse fetal and placental RNA probed with the 2.2-kb BE DNA probe. Blot was autoradiographed for 1 week. Lack of expression of the 5-kb placenta-specific transcript and the 4-kb transcript in matdi7 samples is consistent with imprinted (paternal) expression. Hybridization with control *H19* probe confirms integrity of matdi7 RNA. (*B*) Identical aliquots of midgestation mouse fetal and placental poly(A)$^+$ enriched RNA were blotted and probed with sense and AS riboprobes spanning the 2.2-kb BE region. The sense probe detects multiple (3–4 kb) transcripts in both fetus and placenta; the AS probe detects a single placenta-specific 5-kb transcript. Following RNase treatment, blot was autoradiographed for 5 days. Relative band intensities between T3 and T7 probes do not reflect relative abundance of sense and AS transcripts. (*C*) Transcripts of widely different sizes were detected using the 2.2-kb BE DNA probe on Northern blots of adult mouse tissue total RNAs. Strongest expression was in kidney, which contains a predominant transcript of ≈3 kb. The blot was autoradiographed for 1 week.

Riboprobe analysis was carried out using sense and antisense 2.2 kb BE probes hybridized to poly(A)$^+$ RNA. The AS probe hybridized to a single, placenta-specific transcript, consistent with a 5-kb sense transcript. The sense probe detected multiple (two or three), "smeared" transcripts of 3–4 kb in both fetus and placenta, consistent with the occurrence of extensive AS transcription through the *Igf2* upstream region (Fig. 2*B*).

Overlapping Sense and AS Transcripts. We hypothesized that the 5-kb placental sense transcript contained transcribed pseudoexon 1 and 2 sequences spliced to *Igf2* exon 4. This was confirmed by RT-PCR between pseudoexon 1 and exon 6 (primers UP1 and LP1), using placental cDNA as template. Sequencing of the cloned PCR product indicated that, in addition to pseudoexon 2 sequence (19), the transcript contained an expanded exon that included ≈500 bases of adjacent pseudoexon 2 upstream sequence, consisting mainly of C+G-rich tandem repeats.

RACE-PCR was used to define the 5′ end of the transcript. In separate experiments, RT of total placental RNA using primers LP2 or LP5, and PCR using the RACE primer (RP) and LP3, followed by RP and nested LP4, gave a single product that identified a transcription start site <50 bases 5′ of published pseudoexon 1 sequences (Fig. 1*B*). The existence of a transcriptional start site upstream of pseudoexon 1 was confirmed by RT-PCR using primers UP2 and LP1 on placental cDNA. We designate this start site P0 and, to reflect the fact that they are transcribed, have renamed pseudoexons 1 and 2 as upstream exons 1 and 2 (exons u1 and u2), respectively. The abundance of the 5-kb placental transcript was determined by RNase protection. The results indicated that it accounts for <10% of total *Igf2* transcripts (data not shown).

We were unable to clone the nonabundant AS transcripts from a placental cDNA library, therefore we determined their structure by 3′ and 5′ RACE-PCRs, using a range of RT primers and nested PCR primers on several fetal and placental RNA samples. RACE-PCR defined three 3′ polyadenylation addition sites and three 5′ transcription start sites (Figs. 1 *A* and *B*).

Two transcription start sites were identified by sequencing 5′ RACE-PCR products (Figs. 1 *A* and *B*). Site b′, near the *Igf2* exon 1 DSS, is associated with upstream overlapping Sp1 and AP-2 motifs at −120 bp, a further Sp1 site at −160 bp, and a WT1/EGR-1 site at −215 bp, relative to the b′ start site. A second start site (c′) was identified 1 kb downstream of the exon u2 DSS. In addition, we infer the existence of a third start site (a′, Fig. 1*A*) close to the exon u2 DSS because we obtained two 5′ RACE-PCR products of ≈1.3 kb and ≈1 kb with RP and UP5, and RP and UP6, respectively. These PCR products are likely to be colinear with the genomic DNA because a 1-kb PCR product was obtained by PCR of RP-UP5 and RP-UP6 PCR products, and of placental cDNA using primers UP6 and LP5 (data not shown). We were unable to clone or direct sequence these PCR products, possibly due to secondary structures or rearrangements resulting from the presence of the tandem repeats near the LP5 primer.

3′ and 5′ cDNA ends that were identified by RACE-PCR were connected by RT-PCR of placental RNA. Fig. 1*A* indicates RT-PCR products that were successfully cloned and sequenced, and outlines the likely structure of the AS transcripts based on RT-PCR using all possible combinations of 3′ and 5′ end primers (UP3, UP4, UP5, UP6, LP5, LP6, and LP7).

Parental Allelic Expression of Upstream Transcripts in Wild-Type and *H19* Null Mutant Mice. To confirm that the upstream transcripts undergo parental imprinting, we used

RT-PCR to span restriction polymorphisms between *M. domesticus* and *M. spretus* alleles in reciprocal (*M. domesticus* × SD7)F_1 mice. Exclusive paternal expression of the P0 transcript was confirmed by RT-PCR using primers UP1 and LP1, followed by nested UP7 and LP8, which span a *Bsa*AI restriction fragment length polymorphism in exon 6 (Fig. 3). (The exon 2-specific primer, UP8, was used as a control.)

For the AS transcripts we used PCR primers spanning *M. domesticus/M. spretus* polymorphisms that we defined by Southern blot and PCR analysis. The first, *M. spretus*-specific *Bam*HI site, is just 3′ of exon u1. The second, *M. spretus*-specific *Xba*l site, is ≈60 bp 5′ of primer LP6. The results indicated exclusive paternal allelic expression of transcripts in these regions, including AS transcripts a and c (Fig. 3). We failed to identify a restriction fragment length polymorphism in the AS transcript b.

To determine the effect of *H19* deletion on imprinted expression of the transcripts, allele-specific RT-PCRs were repeated on RNA from mice doubly heterozygous for a (maternally inherited) *H19* null allele (18) and a (paternally inherited) SD7 (*M. spretus Igf2*) allele. This analysis indicated that, similar to *Igf2* P1-P3 transcripts, all analyzable upstream transcripts exhibit biallelic expression following deletion of the *H19* gene on the maternal chromosome. Paternal deletion of *H19* had no effect (Fig. 3).

Methylation Analysis Suggests That Mouse "DMR O" May Be a *H19*-Responsive Placenta-Specific Element. To investigate the epigenetic modifications of DNA and chromatin in the upstream region, we performed DNA methylation and nuclease sensitivity analyses. Nuclease sensitivity assays were performed on fetus and placenta, and adult kidney and spleen, using probes spanning the *Ins2* to DMR 1 region. We found a cluster of constitutive HS between the B2 repeat and exon u1, and confirmed the presence of a previously described HS cluster spanning DMR 1 (Figs. 1*A* and 4). Apart from a single, fetus-specific HS near the B2 repeat, no other tissue-specific HS were detected.

DNA methylation analysis was performed throughout the region, from *Ins2* to DMR 1 in fetus and placenta. Parental allele-specificity was analyzed using samples from matdi7 fetus

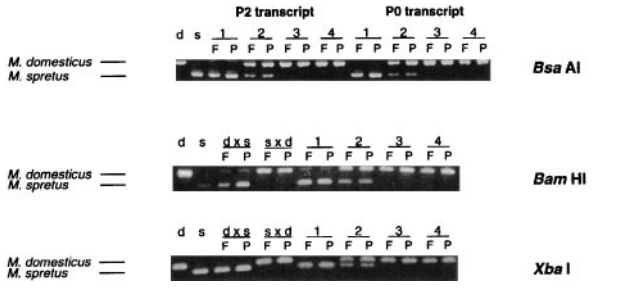

FIG. 3. RT-PCR analysis of parental allelic expression of *Igf2* upstream transcripts. Three *M. domesticus/M. spretus* polymorphisms were used to detect imprinted expression: *M. spretus* specific *Bsa*AI (primers UP7 and LP8) in exon 6 was used for P0 transcript. (This transcript is detectable in fetus using RT-PCR.) *M. spretus*-specific *Bam*HI (primers UP10 and LP9) and *Xba*I (primers UP9 and LP6) in the upstream region (see *Results* for approximate positions of polymorphic sites) were used for AS transcripts. d, s, d × s, s × d: cDNA samples from midgestation *M. domesticus*, *M. spretus,* (*M. domesticus* × SD7)F_1, and (SD7 × *M. domesticus*)F_1, respectively. cDNA samples 1 and 2: (+/+; +/SD7) and (Δ*H19*/+; +/SD7) littermates; 3 and 4: (+/SD7; +/+) and (+/SD7; Δ*H19*/+) littermates. F, fetus; P, placenta. Primer pairs UP7, LP8, and UP10, LP9 preferentially amplified *M. domesticus* alleles from (*M. domesticus* × *M.spretus*)F_1 genomic DNA (data not shown), which probably accounts for stronger *M. domesticus* bands in F_1 samples with biallelic expression. Results indicate paternal allelic expression of all transcripts and biallelic expression following maternal, but not paternal, inheritance of the Δ*H19* null allele.

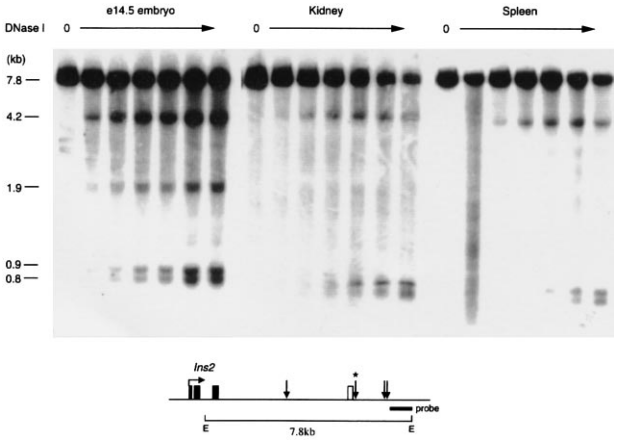

FIG. 4. *DNase* I sensitivity analysis of the region spanning the B2 repeat using a *Xho*I–*Eco*RI 0.78-kb fragment as a probe. Embryo-specific (1.9 kb) site was the only nonconstitutive HS detected in the *Igf2* upstream region. B2 repeat (open box); HS sites (vertical arrows); *, embryo-specific site.

and placenta, and species-specific restriction fragment length polymorphisms—*M. domesticus Bam*HI and *M. spretus Xba*I—in reciprocal F_1 matings (Fig. 5). The Hl (two *Hpa*II sites) and H2 (*Sma*I and five *Hpa*II sites) clusters were analyzed with probes A and B, respectively, by digestion of DNA with *Xba*I in combination with *Hpa*II or *Msp*I (H1 and H2), and with *Bam*HI and *Sma*I (H2). Single *Hpa*II sites H3, H4, and H5 were analyzed with probe C, using an *Xba*I in combination with *Hpa*II or *Msp*I digested DNA.

The results (Fig. 5) indicate that, exclusively in placenta, parental allele-specific methylation exists in the region extending from upstream of exon u1 (including the placenta-specific promoter region, P0) to the 5′ end of DMR 1. In contrast to the DMR 1 and 2 regions, which are paternally hypermethylated in fetus, the newly defined DMR 0 exhibits maternal hypermethylation. From an analysis of reciprocal F_1 crosses between *H19* null mutant and SD7 mice we conclude that an intact *H19* gene in cis is required to maintain hypermethylation of the maternal allele in placenta. Specifically, in (Δ*H19*/+; +/SD7) mice, sites H1, H2, H3, and H5 become demethylated to approximately the same extent as on the paternal allele (Fig. 5).

We also analyzed six *Hpa*II sites between *Ins2* and H1 (data not shown). These were relatively demethylated and did not exhibit parental allele-specific differences in fetus or placenta. In the fetus, the DMR 0 region is heavily methylated on both alleles, and remains so in reciprocal crosses involving *H19* null mice. The H4 site, which is close to the tandem repeat, was fully methylated on both parental alleles in all tissues, and was not responsive to *H19* deletion (Fig. 5*B*).

The Human *IGF2* Lacks a Tandem Repeat. To determine whether the structures and transcription associated with the mouse *Igf2* upstream region are conserved in the human *IGF2* we first sequenced a 5-kb region from *IGF2* exon 2 to 3′ of exon 3. Comparison with the mouse sequence indicated ≈50% identity between the region spanning exon u1 and the CpG island. Similar to the mouse, the human sequence contains a CpG island at the 3′ end of this fragment but lacks a major tandem repeat (Fig. 6).

We next sought to determine whether the human *IGF2* upstream region contains mouse equivalent sense and AS transcripts. RT-PCR analysis of cDNA from normal term placenta was performed using primer pairs flanking exon 2 and exon 3, respectively. We did not find evidence of transcription in this region (data not shown). To determine whether *IGF2* contains a promoter homologous to the mouse P0, RT-PCR was performed using an upper primer just upstream of the

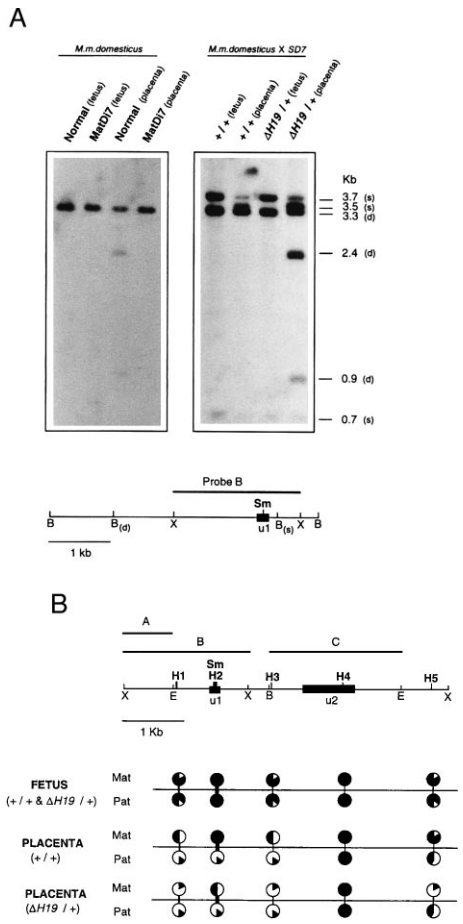

FIG. 5. DNA methylation analysis of region spanning exons u1 and u2. (*A*) Probe B on a Southern blot of *Bam*HI/*Sma*I digested DNA from samples as indicated. Polymorphic *Bam*HI sites and hybridization signals denoted by subscripts in parentheses: d, *M. domesticus*; s, *M. spretus*. Normal vs. disomy analysis: 2.4-kb and 0.9-kb fragments in normal, but not matDi7 placenta, indicates relative demethylation of paternal allele. Maternal deletion of *H19* in (*M. domesticus* × *M. spretus*)F_1 samples: relative demethylation of paternal (*M. spretus*) allele in placenta indicated by reduced intensity of 3.7-kb, and appearance of 3.5-kb, bands. Demethylation of maternal (*M. domesticus*) allele in placenta in response to *H19* deletion indicated by appearance of 2.4-kb and 0.9-kb bands. (*B*) Summary of analysis of fetus and placenta from normal and maternal *H19*-deficient midgestation embryos. Probes: A, *Xba*I–*Eco*RI; B, *Xba*I–*Xba*I; C, *Bam*HI–*Eco*RI fragments. H1–H5, *Hpa*II clusters; Sm, *Sma*I. Mat, maternal allele; Pat, paternal allele. Filled or partially filled circles indicate degree of methylation of sites. *H19* deletion has no effect on fetus. H4 is fully methylated in all samples.

exon 2 ASS, and lower primers in exons 3 and 9. Again, we did not find a human homologue of the mouse P0 transcript (data not shown). In addition, DNA methylation studies of the region homologous to the mouse DMR 1 were performed. Our preliminary results indicate that there is complete demethylation of this region in a range of tissues, suggesting the absence of a conserved functional equivalent of mouse DMR 1 (data not shown).

DISCUSSION

We have identified DMR 0, a placenta-specific, differentially methylated region upstream of mouse *Igf2*. In contrast to the previously described DMRs 1 and 2, which exhibit paternal hypermethylation (10, 11), both alleles of DMR 0 are extensively and equally methylated in this region in the fetus. In placenta, the maternal is the more methylated of the DMR 0

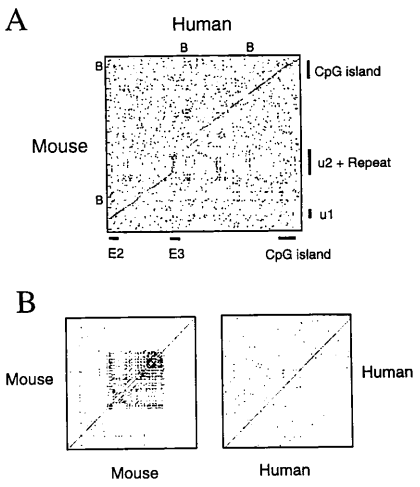

FIG. 6. Mouse-human DNA sequence comparisons. (*A*) DNA sequence comparison of 5 kb of mouse *Igf2* spanning exon u1 to CpG island, and homologous human region, using GCG sequence analysis program. B, *Bam*HI; u1, u2, mouse upstream exons 1 and 2; E2, E3, human exons 2 and 3; human sequence EMBL accession: Y13633. (*B*) Self-comparison of 1,200 bp of human and mouse *Igf2* sequences spanning human exon 3 and mouse exon u2 regions using STRIDER program. In contrast to human, the mouse sequence exhibits a large repeat block.

parental alleles, and undergoes demethylation when linked in cis to an *H19* null mutation. Recent methylation studies of DMR 1 and DMR 2 in placenta indicate that, here, the fetal pattern is reversed and, similar to DMR 0, both exhibit maternal hypermethylation (T. Forne, personal communication). However, in contrast to DMR 0, *H19* deletion has no effect on DMR 1 and DMR 2 methylation in placenta. We interpret these results as evidence for the existence of tissue-specific control regions in the *Igf2* gene, with separate fetal (DMR 1) and placental (DMR 0) elements in the upstream region. However, the *H19*-dependent nature of the parental allele-specific methylation patterns in these regions suggests that they do not act as autonomous imprinting elements. Instead, these elements may transduce the effects of *H19* expression on *Igf2* in a tissue-specific manner.

Recently, the effect of experimental deletion of the entire *Igf2* upstream region in mouse fibroblast cell lines has been reported (20). That study implicated the region in "imprint maintenance" and pointed to a possible mechanism of transduction of *Igf2* interhomolog and *Igf2-H19* interlocus interactions. Consistent with that study, our results suggest the presence of several potentially functionally distinct elements and, moreover, the additional complexity of tissue-specific regulation.

The region spanning mouse *Igf2* DMR 0 is transcribed in both directions. It is possible that the P0 transcript encodes functionally significant amounts of IGF2 peptide because there are precedents for *Igf2* transcripts of lesser abundance undergoing preferential translation (21, 22). Nor can we exclude a regulatory function for the P0 transcript in imprinted expression of *Igf2*, in view of the regulatory role ascribed to upstream exons of the *SNRPN* gene (8). Exclusive placental expression of the P0 transcript is consistent with relative demethylation of the paternal allele in placenta. However, the numerous *DNase* I HS between the B2 repeat and exon u1 are present in all tissues examined and are not correlated with P0 expression. Their significance in relation to the function of this complex genomic region therefore remains elusive.

At least three other transcripts are detectable, but in an AS orientation relative to P0 transcription. We designate these as *Igf2as-a*, *Igf2as-b* and *Igf2as-c*. Riboprobe analysis of Northern blots detected AS transcripts of 3–4 kb; however, RACE-PCR

mapping of transcripts in the exon u2 region did not detect transcription units appreciably larger than 2 kb. These differences may reflect selection for smaller or truncated amplicons in multiple rounds of nested RACE-PCR, or the occurrence of actual transcript heterogeneity in the region.

The AS transcripts have no ORFs. AS transcripts may encode functional RNAs or, given their low abundance, merely represent illegitimate transcription from regions of open chromatin. Of the AS transcripts described, the b transcript is most likely to be functionally significant because it is associated with recognizable promoter elements and contains four exons. In addition, all of the AS transcripts are transcribed through the tandem repeat and we cannot exclude a structural function for the repeat at the RNA level, particularly as certain transcripts of the mouse imprinted *Ipw*, *U2af*, and *Xist* genes also contain tandem repeats (9, 23, 24). At present, we do not know whether the P0 and P1 sense transcripts and the overlapping AS transcripts are coexpressed within individual cells, or whether there exists cellular mosaicism with respect to their expression. Presumably, such overlapping transcription could influence the activity of the P0 and P1 promoters: P0 and P1 promoter transcripts are much less abundant than those from P2 and P3. In addition, there are precedents for a role for the pairing of complementary sense and AS RNA sequences in regulating mRNA stability. In this context, we note that the b1 exon (Fig. 1*A*) overlaps the exon 1 DSS.

Our preliminary analysis of the human *IGF2* upstream region indicates conservation of several important structural features with the mouse region, including exonic sequences and the CpG island (Fig. 6). However, significant interspecies differences were also found. Significantly, the human lacks a tandem repeat equivalent to the expanded C+G-rich repeat, which forms part of exon u2 in the mouse. This is surprising because such repeats are strongly implicated in the imprinting mechanism (6, 7). Of course, we cannot rule out the presence of a repeat further upstream, between exons 1 and 2. Alternatively, imprinting of human *IGF2* may occur by a mechanism that does not require a repeat, or may use the variable number tandem repeat at the nearby insulin (*INS*) locus, different alleles of which influence the magnitude of *INS* expression (25, 26), and might influence *IGF2* expression. In any case, the *IGF2* gene departs from the situation in the *H19* and *IGF2R* genes, both of which have tandem repeats near DMRs that are associated with germline methylation imprints (5, 27).

The function of tandem repeats in the imprinting mechanism remains elusive. Consistent with a previous proposal (7), we find that the mouse *Igf2* upstream repeat is closely associated with DMRs. However, we note that the *Hpa*II site closest to the repeat (H4; Fig. 5*B*) does not exhibit parental allele-specific methylation and remains fully methylated in all contexts, including an *H19*-deficient genetic background. It may also be significant that the absence of the repeat in the human is associated with the absence of methylation of the region homologous to mouse DMR 1. In addition, bisulfite analysis of the *Igf2* and *H19* DMR-associated repeats in the mouse indicates that all CpG sites in the repeats are fully methylated on both parental alleles (J. Walter, personal communication). These observations support the proposal (28) that tandem repeats associated with the DMRs of imprinted genes may target either *de novo* or maintenance methylation to the region. Other epigenetic factors that interact with the repeat may then impose the observed parental allelic differences.

We thank E. Holthuizen for human probes and DNA sequence; S. Tilghman for Δ*H19* mice; S. Jones, J. Joyce, and P. Schofield for human RNA; M. S. H. Ko for advice on RACE-PCR; J. Coadwell and A. Murrell for DNA sequence analysis; W. Dean for mouse breeding; and T. Forne, J. Walter, and F. Ishino for unpublished results. We acknowledge Lloyd's Tercentenary Foundation and Babraham Research Fellowships to T. Moore; grants from the International Human Frontier Science Program Organization, Uehara Memorial Foundation and Ministry of Education, Science, Sports, and Culture of Japan to H. Sasaki; European Community Human Capital and Mobility Fellowship to B. Bailleul; JNICT/PRAXIS XXI Scholarship (Portugal) to M. Constancia; Action Research Grant to W.R.

1. Bartolomei M. S., Webber A. L., Brunkow M. E. & Tilghman S. M. (1993) *Genes Dev.* **7,** 1663–1673.
2. Brandeis M., Kafri T., Ariel M., Chaillet J. R., McCarrey J., Razin A. & Cedar H. (1993) *EMBO J.* **12,** 3669–77.
3. Razin A. & Cedar H. (1994) *Cell* **77,** 473–476.
4. Stoger, R., Kubicka, P., Liu, C.-G., Kafri, T., Razin, A., Cedar, H. & Barlow, D. P. (1993) *Cell* **73,** 61–71.
5. Tremblay K. D., Saam J. R., Ingram R. S., Tilghman S. M. & Bartolomei M. S. (1995) *Nat. Genet.* **9,** 407–413.
6. Chaillet, J. R., Bader, D. S. & Leder, P. (1995) *Genes Dev.* **9,** 1177–1187.
7. Neumann, B., Kubicka, P. & Barlow, D. P. (1995) *Nat. Genet.* **9,** 12–13.
8. Dittrich, B., Buiting, K., Korn, B., Rickard, S., Buxton, J., Saitoh, S., Nicholls, R. D., Poustka, A., Winterpacht, A., Zabel, B. & Horsthemke, B. (1996) *Nat. Genet.* **14,** 163–170.
9. Wevrick, R., Kerns, J. A. & Francke, U. (1997) *Hum. Mol. Genet.* **3,** 1877–1882.
10. Feil R., Walter J., Allen N. D. & Reik W. (1994) *Development (Cambridge, U.K.)* **120,** 2933–2943.
11. Sasaki H., Jones P. A., Chaillet J. R., Ferguson-Smith A. C., Barton S. C., Reik W. & Surani. M. A. (1992) *Genes Dev.* **6,** 1843–1856.
12. Ward, A., Fisher, R., Richardson, L., Pooler, J.-A., Squire, S., Bates, P., Shaposhnikov, R., Hayward, N., Thurston, M. & Graham, C. F. (1997) *Genes Funct.* **1,** 25–36.
13. Leighton P. A., Ingram R. S., Eggenschwiler J., Efstratiadis A. & Tilghman S. M. (1995) *Nature (London)* **375,** 34–39.
14. Schneid, H., Seurin, D., Vazquez, M.-P., Gourmelen, M., Cabrol, S. & Le Bouc, Y. (1993) *J. Med. Genet.* **30,** 353–362.
15. Issa, J.-P. J., Vertino, P. M., Boehm, C. D., Newsham, I. F. & Baylin, S. B. (1996) *Proc. Natl. Acad. Sci. USA* **93,** 11757–11762.
16. Searle, A. G. & Beechey, C. V. (1990) *Genet. Res.* **56,** 237–244.
17. Chomczynski, P. & Sacchi, N. (1987) *Anal. Biochem.* **162,** 156–159.
18. Sasaki, H., Shimozaki, K., Zubair, M., Aoki, N., Ohta, K., Hatano, N., Moore, T., Feil, R., Constancia, M., Reik, W. & Rotwein, P. (1996) *DNA Res.* **3,** 331–335.
19. Rotwein P. & Hall L. J. (1990) *DNA Cell Biol.* **9,** 725–735.
20. Hu, J.-F., Vu, T. H. & Hoffman, A. R. (1997) *J. Biol. Chem.* **272,** 20715–20720.
21. Graham, D. E., Rechler, M. M., Brown, A. L., Frunzio, R., Romanus, J. A., Bruni, C. B., Whitfield, H. J., Nissley, S. P., Seelig, S. & Berry, S. (1986) *Proc. Natl. Acad. Sci. USA* **83,** 4519–4523.
22. Newell, S., Ward, A. & Graham, C. (1994) *Mol. Reprod. Dev.* **39,** 249–258.
23. Hatada, I., Kitagawa, K., Yamaoka, T., Wang, X., Arai, Y., Hashido, K., Ohishi, S., Masuda, J., Ogata, J. & Mukai, T. (1995) *Nucleic Acids Res.* **23,** 36–41.
24. Brockdorff, N., Ashworth, A., Kay, G. F., McCabe, V. M., Norris, D. P., Cooper, P. J., Swift, S. & Rastan, S. (1992) *Cell* **71,** 515–526.
25. Bennett, S. T., Lucassen, A. M., Gough, S. C., Powell, E. E., Undlien, D. E., Pritchard, L. E., Merriman, M. E., Kawaguchi, Y., Dronsfield, M. J., Pociot, F., Nerup, J., Bouzekri, N., Cambon-Thomsen, A., Ronningen, K. S., Barnett, A. H. & Todd, J. A. (1995) *Nat. Genet.* **9,** 284–292.
26. Kennedy, C. G., Gwerman, M. S. & Rutter, W. J. (1995) *Nat. Genet.* **9,** 293–298.
27. Riesewijk, A. M., Schepens, M. T., Welch, T. R., van den Berg-Loonen, E. M., Mariman, E. M., Ropers, H.-H. & Kalscheuer, V. M. (1996) *Genomics* **31,** 158–166.
28. Turker, M. S. & Bestor, T. H. (1997) *Mutat. Res.* **386,** 119–130.