BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS **222,** 171–177 (1996)
ARTICLE NO. 0716

Absence of Imprinting in U2AFBPL, a Human Homologue of the Imprinted Mouse Gene *U2afbp-rs*

R. Scott Pearsall,* Hideo Shibata,† Anna Brozowska,‡ Kiyoshi Yoshino,† Kiyoji Okuda,§ Pieter J. deJong,‡ Christoph Plass,* Verne M. Chapman,* Yoshihide Hayashizaki,† and William A. Held*,[1]

**Department of Molecular and Cellular Biology, and ‡Department of Human Genetics, Roswell Park Cancer Institute, Buffalo, NY; †Genome Science Laboratory, Tsukuba Life Science Center, The Institute of Physical and Chemical Research (RIKEN), Tsukuba, Ibaraki, Japan; and §Department of Gynecology and Obstetrics, Osaka Medical College, Takatsuki, Osaka, Japan*

Received April 2, 1996

The mouse gene U2 auxiliary factor binding protein related sequence (*U2afbp-rs*) has previously been shown to be genomically imprinted with monoallelic expression from the paternal allele. To determine if the human homologue is imprinted and contains conserved structural features which regulate imprinting, we isolated genomic clones from a human P1-derived artificial chromosome (PAC) library that map to human chromosome 5q22-31, a region syntenic to the proximal portion of mouse chromosome 11 where *U2afbp-rs* resides. A genomic subclone was isolated which contained an open reading frame with high homology to the mouse gene. This subclone also maintained the intronless character of the mouse gene. A *Kpn*I polymorphism within the open reading frame of the gene was found to occur in 21% (8/38) of the alleles tested from human placental tissue samples. RT-PCR analysis of human placentas using the *Kpn*I polymorphism to determine the parental origin of the alleles indicates biallelic expression of the human chromosome 5 U2AFBPL gene.

Pronuclear transplantation studies in mice have shown that there is an absolute requirement for both parental genomes to ensure proper development (1). Genomic imprinting is an epigenetic phenomenon by which the two haploid parental genomes can be distinguished in the offspring. This modification can result in monoallelic expression that is both tissue and developmental stage specific (2,3). In mice, uniparental inheritance of some chromosomal regions (4) results in growth abnormalities or prenatal death indicating these regions contain important genes that are subject to imprinted regulation (5,6). Some human diseases are characterized by uniparental disomies and also exhibit growth and developmental abnormalities, indicating that there are regions of the human genome that also contain imprinted genes (7).

To date, seventeen genes have been identified in mouse and human that exhibit monoallelic expression due to genomic imprinting (8,9). Only four of these genes, Insulin-like growth factor 2 (*Igf2*), Igf2/Mannose-6-phosphate receptor (*Igf2r*), *H19* and a Small nuclear ribonucleoprotein-associated peptide (*Snrpn*), have been analyzed for imprinted expression in both mouse and human (10). *Igf2* and *Snrpn* show monoallelic expression from the paternal allele in both species, although there is biallelic expression in some tissues. *H19* shows expression exclusively from the maternal allele in mouse and human tissues. The *Igf2r* gene shows maternal specific expression in the mouse but shows biallelic expression in a majority of the human population (10). This may be a polymorphic trait since a small percent of the population seems to have monoallelic expression of *Igf2r* from the maternal allele (11). Evolutionary conservation of imprinted genes between mouse and human may provide a means of identifying important regulatory sequences required for the establishment of the imprint.

We have previously isolated a mouse gene, *U2afbp-rs*, which maps to the proximal region of

[1] To whom correspondence should be addressed. Fax: 716-845-8169.

0006-291X/96 $18.00

chromosome 11 and shows paternal specific expression in all tissues examined (12,13,14). A second mouse locus (*U2af1-rs2*) was identified that mapped to the X chromosome, but did not show parental specific monoallelic expression in adult mice (15). Fluorescent *in situ* hybridization (FISH) analysis of human chromosomes with cDNA clone from the imprinted mouse gene indicated homology to human 5q22-31, a region considered to be syntenic to mouse chromosome 11 where *U2afbp-rs* mapped (16). Recently, homologous sequences to the mouse *U2afbp-rs* gene were reported on human chromosomes 5, X and 19 (17). We have isolated genomic clones corresponding to the U2AFBPL locus on human chromosome 5 in the region that is syntenic to mouse chromosome 11. The chromosome 5 gene maintains the intronless feature of the mouse chromosome 11 gene, while the human X chromosome has been reported to contain introns (17). Analysis of the chromosome 5 gene as the homologous region to the imprinted mouse gene would allow us to determine if the monoallelic expression from the paternal allele was maintained and if there was sequence or structural similarity in the surrounding genomic region that may be controlling the imprinted phenotype. Our data show that there is biallelic expression of U2AFBPL in human placental tissue indicating an absence of imprinting regulation for this gene in humans.

MATERIALS AND METHODS

Library screening and analysis of genomic clones. A 600 bp *Eco*RI restriction fragment from the mouse *U2afbp-rs1* cDNA clone was [^{32}P] GTP labeled with a random priming kit (Boehringer Mannheim) and hybridized using standard conditions (18) against an arrayed human P1 Artificial Chromosome genomic library consisting of 200,000 clones (19). Six independent clones were isolated based on duplicate hybridization signals within the array. The clones were grown according to established conditions (19). DNA was isolated using a modified alkaline lysis procedure (18).

Two groups of clones were identified that mapped to either chromosome 5 or chromosome 19 on a mouse/human somatic cell hybrid panel (20). The panel was screened with the primer pair F32/R33 which was designed to be human specific and to distinguish the chromosome 5 and 19 PCR products by size (see below).

The PAC clones were sized and restriction mapped using a CHEF DRII (Biorad) pulse field gel apparatus. Agarose gels (1%) were electrophoresed at 170 volts for 18 hours with a switch time of 1–6 seconds in 1× TAE buffer. Single and double digests were performed to determine positions from known sites within the vector. Southern blot analysis using the [^{32}P] GTP labeled F32/R33 PCR product from PAC2/L1 (chromosome 5) was used to determine which fragments contain the *U2afbp-rs* homologue. T7 and Sp6 primers (located near the cloning site of the vector) were end-labeled and used as probes to determine the distance of restriction sites from the vector:insert junctions.

A 3.3 kb *Pst*I/*Hin*dIII fragment that cross hybridized with the PCR probe was isolated from PAC2/L1 (chromosome 5) and cloned into the multiple cloning site of the pBluescriptII KS+ vector (Stratagene) by standard methods (18). Recombinant plasmid DNA was isolated using spin columns (Qiagen).

The subclones were restriction mapped and sequenced using an ABI automated sequencing apparatus. Sequences were analyzed using the Wisconsin GCG software package (Version 8.0, 1994) run on a VAX computer and the DNA Strider program for the Macintosh. Self alignments were performed using the DottyPlot program (Version 1.0c, Gilbert 1989). Homology searches were performed using the BLAST algorithm available from Genbank. The sequence reported here is available via Genbank using accession number U51224.

DNA isolation and PCR assay conditions. Human placentas were obtained at normal delivery and parental bloods were prepared after obtaining written informed consent.

Genomic DNA from placenta or human blood was prepared with slight modifications of the previously described procedure (12). PCR assays were performed using standard conditions in a 50 μl final volume with combinations of the following oligonucleotide primers: F32; 5′-AGTGCAGGAGGGATGACT 3′, F38; 5′-ACGTTTCCCAAGAAGATGAC-3′, R33; 5′-GACTTCCTCCTCCCACGACTT-3′, R38; 5′ CAACACATCCTCATAGAAAT-3′. The primer positions within the pPAC2ΔPs/Hi clone are shown in Figure 1. The F32 primer was designed from the mouse sequence. The F32/R33 primer pair will amplify a 743 bp fragment corresponding to the chromosome 5 locus and a 1029 bp fragment corresponding to the chromosome 19 locus in human genomic DNA. The F38 primer has a 9 bp mismatch with the X chromosome cDNA sequence (17) at the 3′ end and should not amplify sequences from the X chromosome gene.

Genomic DNA was PCR amplified using a programmable thermal cycler (MJ Research) using cycling conditions of an initial denaturing step of 94°C for 5 min followed by 35 cycles of 58°C for 2 min, 72°C for 1 min and 30 sec, 94°C for 1 min. This was followed by a final annealing at 58°C for 3 min and final extension at 72°C for 5 min. For the F32/R31 and F32/R33 reactions 60°C annealing was used. Following PCR the products were digested at 37°C with *Kpn*I and visualized by ethidium bromide staining after separation using either 2% agarose gels or 4% polyacrylamide gels.

RNA isolation and RT-PCR assay conditions. Placental RNA was isolated and reverse transcribed (SuperScript II, BRL) with oligo dT primer as previously described (12). After cDNA synthesis PCR amplification was performed with the

addition of [^{32}P]dCTP using the F38/R38 primer pair with cycling conditions of an initial denaturing step at 95°C for 5 min, followed by 26 cycles of a 1 min denaturing at 94°, 55°C annealing for 2 min and 72°C extension for 3 min. This was followed by a final extension at 72°C for 5 min. RT-PCR products were digested with *Kpn*I and separated on a 4% polyacrylamide gel, dried onto filter paper and exposed to Kodak XAR film with an intensifying screen or BAS2000 phosphorimager.

RESULTS

Isolation of the Human Homologous Sequences

To isolate the human homologue of *U2afbp-rs*, a 600 bp *Eco*RI fragment from the mouse cDNA sequence was used to screen a human PAC library. Four clones were isolated and shown to map to chromosome 5q22-31 using mouse/human somatic cell hybrids and FISH analysis (see methods). This region of human chromosome 5 is thought to be syntenic to the region on proximal mouse chromosome 11 where the mouse gene is located (16). Clones from human chromosome 19 were also obtained from the library screen. Since chromosome 19 is not syntenic to mouse chromosome 11 and because sequence analysis of the PCR products from these clones showed an Alu insertion disrupting the putative ORF (see methods) our analysis was focused on the chromosome 5 clones as containing the true homologue to the imprinted *U2afbp-rs* gene.

Restriction mapping of the four chromosome 5 clones lead to the assembly of a contig covering 145 kb (Fig. 1A). A 3.3 kb *Pst*I/*Hin*dIII fragment which hybridized with the PCR product from the F32/R31 primer pair was subcloned from PAC clone 2/L1 and sequenced (Fig. 1B). A 1.4 kb region of the clone had both 80% nucleotide homology to the mouse *U2afbp-rs* sequence and a putative open reading frame (ORF) of 479 amino acids with 83% similarity (79% identity) to the mouse ORF. This region is 100% identical to the recently isolated cDNA sequence (17 and genbank acc. no. D49676) with two minor corrections of the reported genbank sequence. The reported genbank cDNA sequence is missing a T at position 1081 and contained an N at position 1426 that is an A when compared to the sequence reported here. Northern blot analysis of RNA from a human teratoma cell line (ATCC no. HTB-106) indicated a transcript size of approximately 1.5 kb (unpublished result) consistent with previous findings (17). It is interesting to note that the human gene maintains the intronless character of the mouse gene across the putative ORF although the potential 5′ and 3′ untranslated regions appear very short. The high homology is restricted only to the region representing the ORF. The homology to the mouse sequence drops below 40% in the regions surrounding the 1.4 kb coding region. The regions flanking the conserved region contain 2 Alu elements in the 5′ region (position 184–452 and 878–1147) and one in the 3′ end (position 2902–3161) (Fig. 1B).

Expression of U2AFBPL in Human Placenta

In order to determine if monoallelic expression of U2AFBPL occurs, we first needed to identify a polymorphism to distinguish the two alleles. We looked for polymorphism in the U2AFBPL gene by designing the primer pair F38/R38 that would specifically amplify a 726 bp product from the chromosome 5 gene. Amplification from the X chromosome gene is not expected to occur (see methods), but could be distinguished by size if the primers were able to bind. In addition, no amplification was detected in the chromosome 19 PAC clones with the F38/R38 primer pair. Chromosome 5 specific amplification of genomic DNA with this primer pair was further confirmed using the somatic cell hybrids (data not shown). PCR amplification of 19 placental DNAs was performed and the product was digested with various restriction enzymes. Eight of these samples were heterozygous for a *Kpn*I polymorphism. Thus, the allelic frequency of the *Kpn*I polymorphism in our 19 samples is 21% (8/38). Sequence analysis of the PCR product from heterozygous patients confirmed the site as being polymorphic (data not shown).

Three families were analyzed for allelic inheritance and for allelic expression patterns in pla-

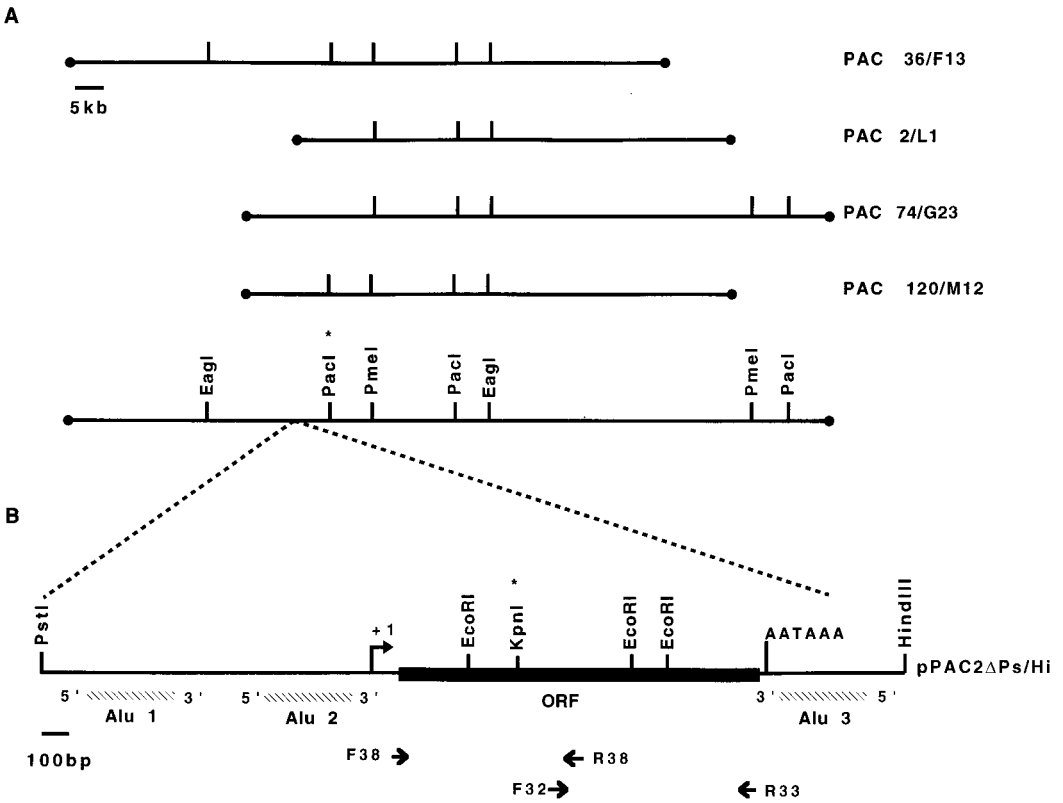

FIG. 1. Restriction map of PAC clones and pPAC2ΔPs/Hi subclone. (A) Restriction maps of the chromosome 5 PAC clone contig covering 145 kb surrounding U2AFBPL. The individual clones are followed by the consensus map for all the clones across the region. The closed circles are the vector:insert boundary sites created by partial *Mbo*I digest of human genomic DNA. The *Pac*I site (*) represents a polymorphic site in the human genomic source of the library that is present in two clones and absent in the others. (B) The position of the pPAC2ΔPs/Hi subclone within the contig is shown. The *Pst*I/*Hin*dIII genomic clone isolated from PAC2/L1 covers 3.3 kb and contains the U2AFBPL gene. The +1 and polyA signal represent the putative transcript of 1.4 kb based on the position of a TATA box in the 5′ untranslated region. The putative transcript shows homology to the mouse *U2afbp-rs* gene and contains an ORF of 479 amino acids (black bar), also with high homology to the mouse. The intronless character seen in the mouse gene is conserved, but there is only a short 5′ and 3′ untranslated region. The three Alu sequences flanking the gene and their orientation are shown (stippled bar). The position of the primers used for PCR analysis are indicated by arrows.

cental tissues. Genomic DNA from maternal and paternal blood samples and the placental sample were subjected to PCR amplification and digestion with *Kpn*I in order to follow the allelic inheritance. Figure 2 shows paternal inheritance of the *Kpn*I site in placental DNA sample H003p (lanes 3–5) and maternal inheritance of the site for samples H010p and H013p (lanes 8–10, 13–15). RNA from the three informative placentas was isolated and subjected to RT-PCR using oligo dT primer for the reverse transcription and the F38/R38 primer pair for PCR. Since the U2AFBPL gene is intronless, a negative control consisting of RNA without reverse transcriptase added was used to determine if there was genomic DNA contamination in the samples. No amplification was seen in the RNA control lane lacking reverse transcriptase (lanes 2,7,12). As expected (see above), there were no additional RT-PCR products corresponding to the X chromosome gene. Expression from the chromosome 5 gene was detected from both alleles in all of the placental RNA samples (lanes 1,6,11). The relative equal intensity of the bands indicates that both alleles were expressed at approximately the same level.

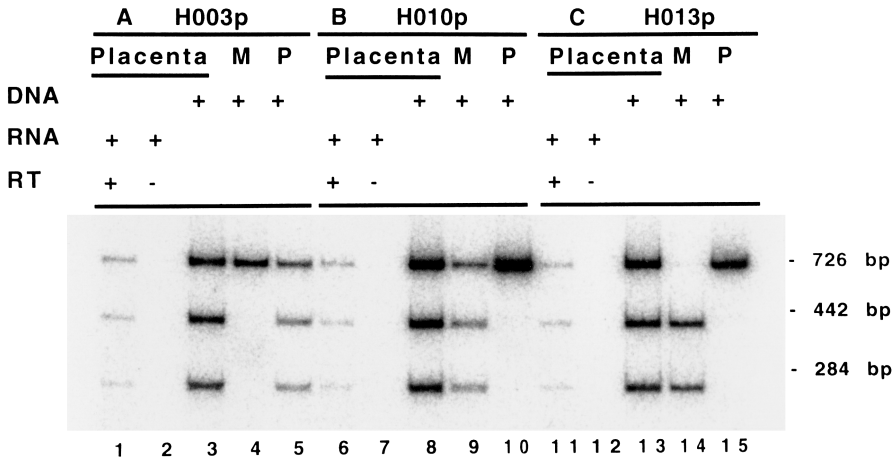

FIG. 2. Parental inheritance of a *Kpn*I polymorphism and allelic expression in placenta tissue for three informative families. Three families (A, B, C) were identified in which the parental allelic inheritance could be followed. Genomic DNA from placental tissues H003p, H010p and H013p was PCR amplified with the F38/R38 primer pair and shown to be heterozygous for a *Kpn*I site in U2AFBPL (lanes 3,8,13). The maternal (M) or paternal (P) origin of the *Kpn*I susceptible allele was determined by amplifying genomic DNA from blood of the parents of each patient. H003p shows paternal inheritance of the *Kpn*I site (lanes 3,4,5). H010p and H013p show maternal inheritance of the *Kpn*I site (lanes 8–10, 13–15). RNA from the placental tissues was subject to RT-PCR in the presence (lanes 1,6,11) or the absence (lanes 2,7,12) of reverse transcriptase (RT + or −). No expression is seen in the samples lacking RT. The size of the PCR products is shown on the right. Undigested products are 726 bp and products digested by *Kpn*I are 442 and 284 bps. The image was produced using a BAS2000 phosphorimager.

Self Alignment of U2afbp-rs and U2AFBPL

Since most imprinted genes identified contain a repeat element that may be an important regulatory feature (8), self alignments of the imprinted mouse gene and the non-imprinted human gene was performed. Figure 3A shows the self alignment of 2980 bp of the mouse *U2afbp-rs* cDNA sequence. A repeat structure can be seen within the 5′ untranslated region of the gene. Self alignment of the 3341 bp human U2AFBPL genomic sequence under the same, or lower stringency conditions does not show a similar repeat structure (Fig. 3B). There is repeat alignment of the Alu1 and Alu2 elements at positions 184–452 and 878–1147 of the sequence as would be expected since the elements are in the same 5′ to 3′ orientation.

DISCUSSION

In the mouse there are two members of the *U2afbp*-rs family. *U2afbp-rs1* is located on mouse chromosome 11 and shows monoallelic expression from the paternal allele (12,14). The second member, *U2afbp-rs2* is located on the X chromosome and does not show parental specific expression in adult mice (15). The two genes are 76% identical to each other at the nucleotide level and 72% at the amino acid level. In humans three homologous loci have been identified on chromosomes 5, 19 and the X (17). The cDNA sequences of the human chromosome 5 and chromosome X genes are over 90% identical to each other at the nucleotide and amino acid level. Both genes have nearly 80% homology to each of the mouse genes. No cDNA has been isolated from the chromosome 19 gene (17). We did not detect expression from this locus in our analysis and found the insertion of an Alu element in the ORF, suggesting this copy is a pseudogene.

We have isolated PAC genomic clones which cover 145 kb of human chromosome 5 and contain an expressed human homologue of the imprinted mouse gene *U2afbp-rs*. The clones map to human chromosome 5q21-31, a region syntenic to the mouse chromosome 11 *U2afbp-rs* region (16), making this a strong candidate for the true homologue of the mouse gene. A 3.3 kb *Pst*I/*Hin*dIII

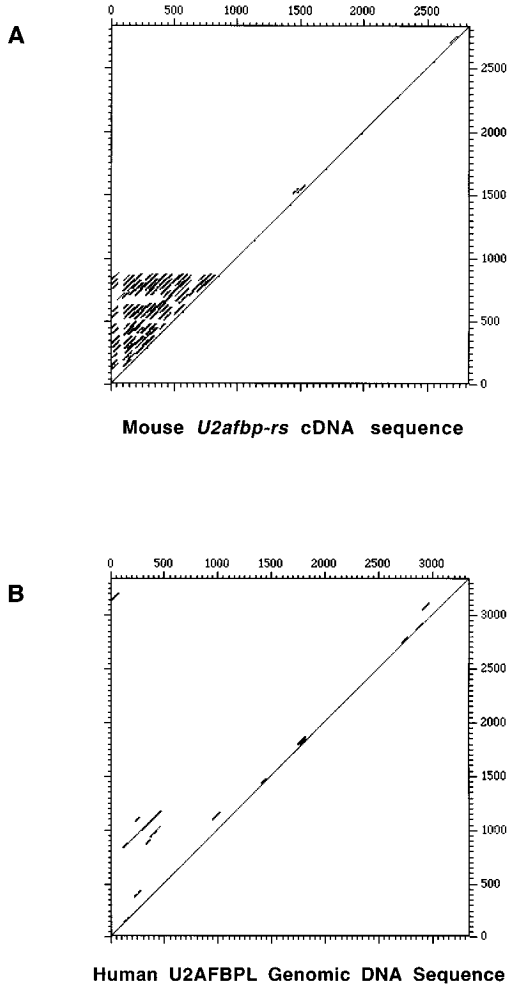

FIG. 3. Repeat structure in the mouse and human genomic clones. Self-alignment of sequences for (A) the mouse *U2afbp-rs* cDNA and (B) human U2AFBPL genomic DNA using the Dotty Plot program at 65% stringency. The mouse sequence contains a repeat element in the 5′ untranslated region while the human sequence lacks a similar repeat. The line representing a repeat in the human sequence represents alignment of the two Alu elements (1 and 2 in Fig. 1) that are in the same orientation (positions 184-452, and 878-1147).

subclone was isolated from one of these large fragment clones and found to contain nucleotide homology and a complete ORF with homology to the mouse *U2afbp-rs* gene. Human chromosome 5 specific PCR primers from this region identified a *KpnI* polymorphic site in genomic DNA. Expression and parental origin of each allele could be analyzed using RT-PCR analysis of three placental RNAs. Biallelic expression was observed in all three of the placental tissues tested, indicating that expression of U2AFBPL is not subject to regulation by genomic imprinting.

It has been observed (8,21) that imprinted genes contain a short direct repeat element and CpG islands that may contribute to a structural feature that is important as an “imprinting box”. Sequence analysis of the mouse gene indicates there is a direct repeat in the 5′ untranslated region, near the transcription start site, as well as a CpG island. The sequence conservation between the mouse *U2afbp-rs* and the human U2AFBPL genes is restricted exclusively to the ORF of the transcribed genes. The U2AFBPL gene does not contain a similar repeat structure or CpG islands in the comparable positions to the mouse gene and is not regulated by genomic imprinting.

The *Igf2r* gene is the only other known example where monoallelic expression is not conserved between mouse and human, except for a small portion of the population (10,11). However, Smrzka et al. (22) have shown that the human IGF2R gene contains a conserved repeat element in the second intron that is still subject to maternal specific methylation. Here, we demonstrate a case in which parental-specific monoallelic expression of a gene due to genomic imprinting has not been evolutionary conserved between mouse and human. Although we have only looked at human placental tissue, the fact that the mouse gene shows imprinted regulation in the early two cell stage (13) and in adult tissue (12) suggests that imprinting control may be absent in other human tissues. The loss of imprinted regulation of expression in U2AFBPL is coincident with lack of a repeat element found in *U2afbp-rs.* This suggests the repeat element may have an important role in the establishment of genomic imprinting (8,21).

ACKNOWLEDGMENTS

This work is dedicated to the memory of Verne Chapman and is supported by NIH Grant GM51881 to V.M.C. and a Research Grant for the Genome Exploration Research Project of the Science and Technology Agency of the Japanese Government, and by a Grant-in-Aid for a Creative Basic Research (Human Genome Program), a Grant-in-Aid for Developmental Scientific Research (B) on Priority Areas from the Ministry of Education and Culture, Japan, to Y.H. We thank T. Shows and R. Eddy for providing somatic cell hybrid DNA and A. Cairo for performing sequencing reactions. We also thank R. Haddad, M. Crowley and I. Kalcheva for critical reading of the manuscript.

REFERENCES

1. Solter, D. (1988) *Ann. Rev. Gen.* **22,** 127–146.
2. Szabo, P. E., and Mann, J. R. (1995) *Genes Dev.* **9,** 3097–3108.
3. Sasaki, H., Ferguson-Smith, A. C., Shum, S. W., Barton, S. C., and Surani, M. A. (1995) *Development* **121,** 4195–4202.
4. Beechy, C. V., and Cattanach, B. (1995) *Mouse Genome* **93,** 89–92.
5. Surani, M. A. H., Barton, S. C., Howlett, S. K., and Norris, M. L. (1988) *Development* **103,** 171–178.
6. Cattanach, B. M., and Kirk, M. (1985) *Nature* **315,** 496–498.
7. Ledbetter, D. H., and Enger, E. (1995) *Hum. Mol. Gen.* review **4,** 1757–1764.
8. Barlow, D. (1995) *Science* **270,** 1610–1613.
9. Kuroiwa, Y., Kaneko-Ishino, T., Kagitani, F., Kohda, T., Li, L., Tada, M., Suzuki, R., Yokoyama, M., Shiroishi, T., Wakana, S., Barton, S. C., Ishino, F., and Surani, M. A. (1996) *Nature Gen.* **12,** 186–190.
10. Kalscheuer, V. M., Mariman, E. C., Schepens, M. G., Rheder, H., and Ropers, H. H. (1993) *Nature Gen.* **5,** 74–78.
11. Xu, Y., Goodyear, C. G., Deal, C., and Polychronakos, C. (1993) *Biochem. Biophys. Res. Comm.* **197,** 747–754.
12. Hayashizaki, Y., Shibata, H., Hirotsune, S., Sugino, H., Okazaki, Y., Sasaki, N., Hirose, K., Imoto, H., Okuizumi, H., Muramatsu, M., Komatsubara, H., Shiroishi, T., Moriwaki, K., Katsuki, M., Hatano, N., Sasaki, H., Ueda, T., Mise, N., Takagi, N., Plass, C., and Chapman, V. M. (1994) *Nature Gen.* **6,** 33–40.
13. Latham, K. E., Rambhatla, L., Hayashizaki, Y., and Chapman, V. M. (1995) *Dev. Bio.* **168,** 670–676.
14. Hatada, I., Sugama, T., and Mukai, T. (1993) *Nucl. Acid Res.* **21,** 5577–5582.
15. Yamaoka, T., Hatada, I., Kitagawa, K., Wand, X., and Mukai, T. (1995) *Genomics* **27,** 337–340.
16. Kalcheva, I., Plass, C., Sait, S., Eddy, R., Shows, T., Watkins-Chow, D., Camper, S., Shibata, H., Ueda, T., Takagi, N., Hayashizaki, Y., and Chapman, V. (1995) *Cytogen. Cell Gen.* **68,** 19–24.
17. Kitagawa, K., Wang, X., Hatada, I., Yamaoka, T., Nojima, H., Inazawa, J., Abe, T., Mitsuya, K., Oshimura, M., Murata, A., Monden, M., and Mukai, T. (1995) *Genomics* **30,** 257–263.
18. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual (Ford, N., and Nolan, C., Eds.), 2nd ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
19. Ionnaou, P. A., Amenniya, C. T., Garnes, J., Korisel, P. M., Shizuya, H., Chen, C., Batzer, M. A., and deJong, P. J. (1994) *Nature Gen.* **6,** 84–89.
20. Shows, T. B., Sakaguchi, A. Y., and Naylor, S. L. (1982) Advances in Human Genetics (Harris, N., and Hirschhorn, K., Eds.), Vol. 12, pp. 341–452, Plenum Press, New York and London.
21. Neumann, B., Kubicka, P., and Barlow, D. P. (1995) *Nature Gen.* **9,** 12–14.
22. Smrzka, O. W., Fae, I., Stoger, R., Kurzbauer, R., Fischer, G. F., Henn, T., Weith, A., and Barlow, D. P. (1996) *Hum. Mol. Gen.* **4,** 1945–1952.