© 1997 Oxford University Press
Human Molecular Genetics, 1997, Vol. 6, No. 2 325–332

An imprinted mouse transcript homologous to the human imprinted in Prader–Willi syndrome (*IPW*) gene

Rachel Wevrick[1,+] and Uta Francke[1,2,*]

[1]Department of Genetics and [2]Howard Hughes Medical Institute, Stanford University School of Medicine, Stanford, CA 94305-5428, USA

Received September 19, 1996; Revised and Accepted November 15, 1996 GenBank accession numbers U69888, U77702–U77705

The Prader–Willi syndrome (PWS) is caused by genomic alterations that inactivate imprinted, paternally expressed genes in human chromosome region 15q11–q13. *IPW*, a paternally expressed gene cloned from this region, is not expressed in individuals with PWS, and is thus a candidate for involvement in this disorder. The *IPW* transcript does not appear to encode a polypeptide, suggesting that it functions at the level of an RNA. We have now cloned a mouse gene, named *Ipw*, that has sequence similarity to a part of *IPW* and is located in the conserved homologous region of mouse chromosome 7. The *Ipw* cDNA also contains no long open reading frame, is alternatively spliced and contains multiple copies of a 147 bp repeat, arranged in a head-to-tail orientation, that are interrupted by the insertion of an intracisternal A particle sequence. *Ipw* is expressed predominantly in brain. In an interspecies (*M.musculus*×*M.m.castaneus*) F1 hybrid animal, expression of *Ipw* is limited to the paternal allele. We propose that *Ipw* is the murine homolog of *IPW*.

INTRODUCTION

The *IPW* gene (1) is located within an imprinted region on human chromosome 15 that is deleted in 70% of cases of Prader–Willi syndrome (PWS) (2). Although the *IPW* transcript is widely expressed, spliced and polyadenylated, the 2.2 kb RNA sequence does not appear to have polypeptide-coding potential. The manifestations of PWS include mental retardation, neonatal hypotonia, hyperphagia leading to obesity, short stature and hypogonadism, suggesting a underlying hypothalamic defect. Since most non-deletion cases of PWS are due to uniparental maternal disomy for chromosome 15 (3,4) and are clinically indistinguishable from those with paternally derived deletions, the genes responsible for the phenotype must be imprinted, expressed only from the paternal allele and located in the deletion region.

IPW and two other paternally expressed genes have been identified within the deletion region, yet no causal connection has been made between the loss of expression of any particular gene and the manifestations of PWS. *SNRPN* codes for the small ribonucleoprotein-associated polypeptide N, a developmentally regulated protein component of the spliceosome expressed predominantly in neuronal tissues. It is paternally expressed in mice and humans (5–7). *ZNF127* (*D15S9*) encodes a zinc finger-containing protein, presumably a transcription factor, and is located in the proximal end of the deletion region (8). An allele-specific methylation difference has been documented for probes from this region (9), however, no expression studies have been published other than in abstract form (10).

A small number of PWS patients have deletions of <100 kb on the paternal chromosome that are located just centromeric of the *SNRPN* gene (11,12). These deletions result in loss of expression of imprinted genes in this region, including *SNRPN* (11) and *IPW* (13) and alter the differential methylation pattern of other loci, such as *PW71B* (14). For one family (PWS-U) lack of expression of *ZNF127* has also been reported (15).

There is growing evidence that untranslated RNAs play a role in establishment or maintenance of a chromatin configuration that is necessary for the expression of nearby genes. The alternative hypothesis is that the transcription of untranslated RNAs merely reflects the transcriptional state of their chromosomal region, while the RNAs themselves have no function. Deletion in mice of the imprinted *H19* gene, which encodes a non-translated RNA, disrupts imprinting in the region surrounding it (16). A disruption of *DGCR5*, a non-translated RNA from the DiGeorge syndrome region, is implicated in regional control of nearby genes (17). Finally, paternally expressed alternative *SNRPN* transcripts are proposed to be involved in imprint epigenotype switching. Non-translated exons upstream of *SNRPN* are disrupted in a subset of PWS and Angelman syndrome patients, whose disorder is apparently caused by a failure to reset the chromosome 15 imprint from one parental epigenotype to the other (18). In a further example, disruption of *Xist*, the inactive X-specific transcript, blocks X-inactivation in *cis*-configuration (19). In this last example, there is evidence that the *Xist* RNA may be directly involved in X chromosome inactivation by physical association between the *Xist* RNA and the chromosome (20).

IPW is located ~180 kb from the putative imprinting center required to maintain expression from this region on the paternal chromosome 15. The function of *IPW* is apparent neither from its expression pattern nor from its sequence. However, *IPW* may be involved in regional gene control in the PWS region, in a manner analogous to the proposed roles of other untranslated RNAs. In order to investigate further the function of *IPW* and its role in PWS and in gene regulation and chromatin configuration, we

*To whom correspondence should be addressed
+Present address: Department of Medical Genetics, University of Alberta, Edmonton, AB, Canada, T6G 2S2

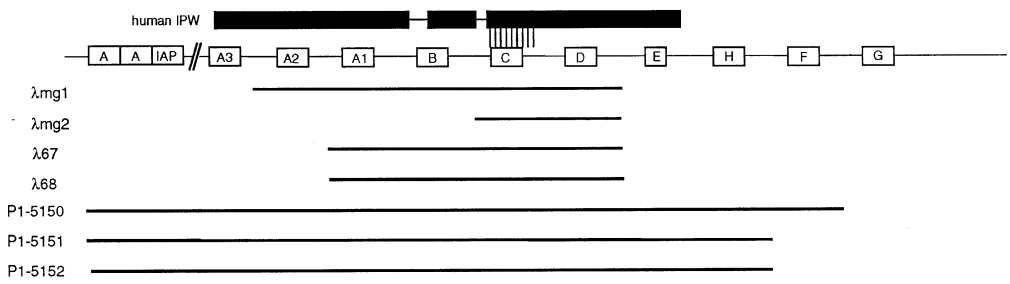

Figure 1. Genomic organization of the *Ipw* locus as determined from seven independent genomic clones. The exons contained within each genomic clone are illustrated, but the positions and sizes have not been determined for all introns. The location of exon G was inferred from a 3′ RACE clone. For comparison, the structure of the human *IPW* gene is shown at the top with vertical lines indicating the location of the conserved sequences.

have searched for a mouse homolog. We report here the identification of an imprinted gene (*Ipw*) with sequence similarity to *IPW* that contains tandem 147 bp repeats and appears not to code for protein.

RESULTS

Mouse genomic clones hybridizing to *IPW* map to a conserved syntenic region and contain a partially homologous gene

Seven overlapping clones (four λ phage clones and three P1 clones) were identified by hybridization to a probe from human *IPW* exon 3 and by PCR screening of mouse genomic DNA libraries (Fig. 1). Partial sequencing of clone λmg1 revealed a 319 bp region of similarity to *IPW* with 79% nucleotide identity and four gaps (Fig. 2). The 5′ end of this region (later named mouse exon C, 142 bp) is 70% similar to the 5′ end of *IPW* exon 3, and the 3′ end of mouse exon C corresponds to the middle of *IPW* exon 3. The 5′ splice sites are at identical positions when the human exon 3 and mouse exon C are aligned. A splice site is present in mouse DNA at position 143 that is not present in human DNA. It is most unusual, however, that the similarity between the human and mouse sequences extends through the 3′ splice site of mouse exon C into the adjacent intron for an additional 176 bp and diminishes thereafter.

A fragment from this region of sequence similarity was mapped by PCR to mouse chromosome 7 using a mouse × rodent somatic cell hybrid mapping panel (21). Furthermore, λ phage clones λmg1 and λmg2 were mapped to the middle region of mouse chromosome 7 by fluorescence *in situ* hybridization (data not shown). Thus, *Ipw* falls into a previously identified region of conserved synteny with proximal human chromosome 15 (22).

The sequence similarity of the mouse *Ipw* clones to human *IPW*, and the localization of these clones to a region of homology to human chromosome bands 15q11–q13, suggest that *Ipw* is indeed the mouse homolog of *IPW*. Furthermore, hybridization of either a human *IPW* exon 3 probe or a mouse exon C–intron C probe to restriction digests of mouse DNA on Southern blots resulted in identical band patterns (data not shown). Thus, *Ipw* appears to be the only locus with high sequence similarity to *IPW* in the mouse genome. No hybridizing fragments were seen, however, when mouse *Ipw* probes from upstream of exon C were hybridized to Southern blots of human DNA, indicating that the region of homology is limited to the sequence similarity between mouse *Ipw* exon C–intron C and *IPW* exon 3, including the common 5′ splice site. Together with the finding that this locus is similarly imprinted (see below), these results provide evidence that *Ipw* is the mouse gene equivalent to *IPW*, but that the sequence similarity between the two genes is limited.

Characterization of *Ipw* transcripts and genomic structure

Partial mouse brain cDNA clones were generated by three independent methods: rapid amplification of cDNA ends (RACE) (23) with primers derived from *Ipw* exon C, RT-PCR and cDNA library screening (Fig. 3). The RACE and RT-PCR cDNA clones contained overlapping sequences, termed exons A–G (Fig. 3A). The 5′ region of the *Ipw* cDNA contained multiple tandem repeats of a 147 bp sequence termed exon A. Two slightly different copies of exon A (A1 and A2) were found by sequencing subclones of λmg1. Exons A1 and A2 were present in phage λmg1 as single exons. The genomic structure of the other A repeats is unknown. Southern blot analysis with a probe from exon A revealed strongly hybridizing 0.8 and 2.5 kb *Eco*RI fragments in total mouse DNA and in DNA from the P1 clones. This finding suggests that the repeats are in separate exons, given the absence of restriction sites for *Eco*RI within each exon. It also suggests that the genomic sequence containing the A repeats is organized in tandem repeats of ~0.8 kb, defined by *Eco*RI restriction sites. It is noteworthy that the sequence of exon B bears some resemblance to that of exon A (Fig. 3C), indicating an ancient evolutionary relationship.

Hybridization screening of a mouse brain cDNA library yielded 12 cDNA clones (mbr0–2, 6–9, 11, 12, 15, 17 and 18) (Fig. 3A and B). The first clone isolated (mbr0) contained slightly more than two copies of exon A, followed by exon B, exon C and part of exon F (Fig. 3C). When this clone was used to rescreen the library, 11 additional clones were isolated. The homology of these cDNAs, however, was limited to exon A based on ~300 bp of sequences obtained from each end (Fig. 3B). The 147 bp tandemly repeated exon A sequences that were present in several RACE clones and in all mbr clones were not completely identical. They differed by 1–6 nucleotides corresponding to ~96–99% sequence conservation. In addition, ~1 in 20 of the exon A repeats had internal deletions of 18–22 contiguous nucleotides. Three of the mbr clones contained part of an intracisternal A particle repeat (IAP) adjacent to the exon A repeats. The full length of the IAP sequence (24) is 7096 bp (GenBank entry MUSFLIAP). Based on hybridization and partial sequence analysis, the nine other mbr clones appeared to be lacking IAP DNA, but they were not sequenced completely.

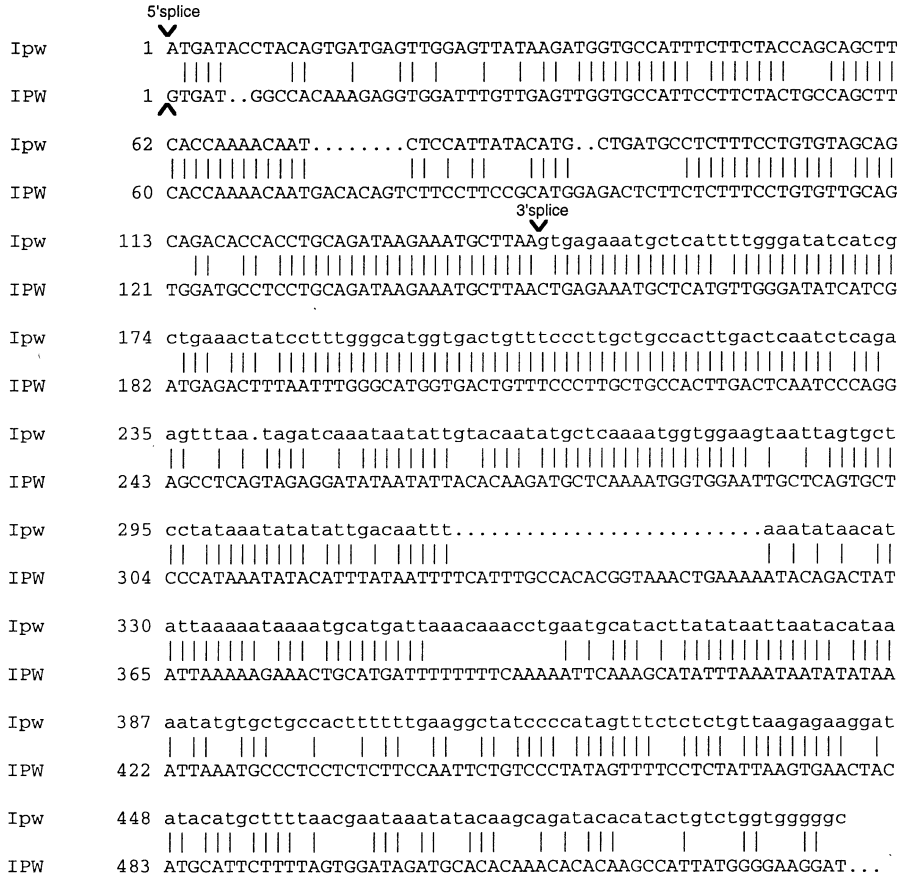

Figure 2. Alignment of mouse *Ipw* exon C–intron C and human *IPW* exon C sequences. Exon sequences are in upper case, intron sequences in lower case, and sequence identities are marked with vertical lines. The 5′ splice sites are at equivalent positions in the two cDNAs, whereas the 3′ splice sites differ.

A pattern of alternative splicing at the 3′ end of the cDNA clones (Fig. 3A) was confirmed by RT-PCR of mouse brain RNA. Reverse primers were designed from the sequence of exons D, E, F, G and H, and a forward primer was designed from the exon C sequence. RT-PCR between these reverse primers and the exon C forward primer yielded products from mouse brain RNA corresponding to the 3′ RACE clones. In all of these clones, the same exon C splice donor site was used. In addition, RT-PCR product clone R3-82-5 contained an alternatively spliced cDNA not represented among the 3′ RACE clones. Exons D, H and G appear to represent alternative 3′ ends to the message since 3′ RACE was performed with an oligo(dT) primer. Although the mbr clones were derived from a random and oligo(dT)-primed brain cDNA library, none of them had poly(A) tails. The order of exons D, E, F and H was established by hybridizing oligonucleotides complementary to each exon to Southern blots of restriction digested DNA from each of the genomic clones (Fig. 1). Exon G was present only in RACE clone R3-3 and not in any of the genomic clones. The sequence of exon F and its intron–exon boundaries was determined from clone mbr0 and from a subcloned fragment of the P1 clone P1-5150.

Sequence analysis reveals no coding potential

A similarity search of sequence databases revealed no significant similarity of the sequence shown in Figure 3C to known genes other than *IPW*. No sequence similarity between *Ipw* and human *IPW* was detected beyond the region noted in Figure 2. The longest open reading frame (ORF), starting with a potential initiator methionine, predicts 104 amino acids and covers part of exon A2, all of exon A1 and part of exon B. No similarities were found when the protein databases were searched with this translated sequence. The potential initiator methionine is not surrounded by the consensus sequence (25). Furthermore, the program CODONPREFERENCE, which analyses sequences for coding potential using a codon usage table, detected no evidence for coding potential within the *Ipw* cDNA sequence. Most remarkably, however, the putative ORFs for *IPW* (1) and *Ipw* are not within the conserved region of sequence similarity and, thus, the conceptual translation products are completely different. We conclude from this comparative analysis that both genes are non-coding and that their function must be at the RNA level.

The *Ipw* transcript is highly expressed in brain

The expression of *Ipw* in mouse tissues was examined by Northern blot hybridization of total RNA with a probe from exons B and C. There was no signal in heart, liver or skin fibroblasts, but a strong smear ranging in size from 0.5 to 12 kb was seen in total brain RNA (data not shown). By using RT-PCR with primers amplifying exons B and C, a specific product was obtained only from brain RNA. Expression in other tissues was detected

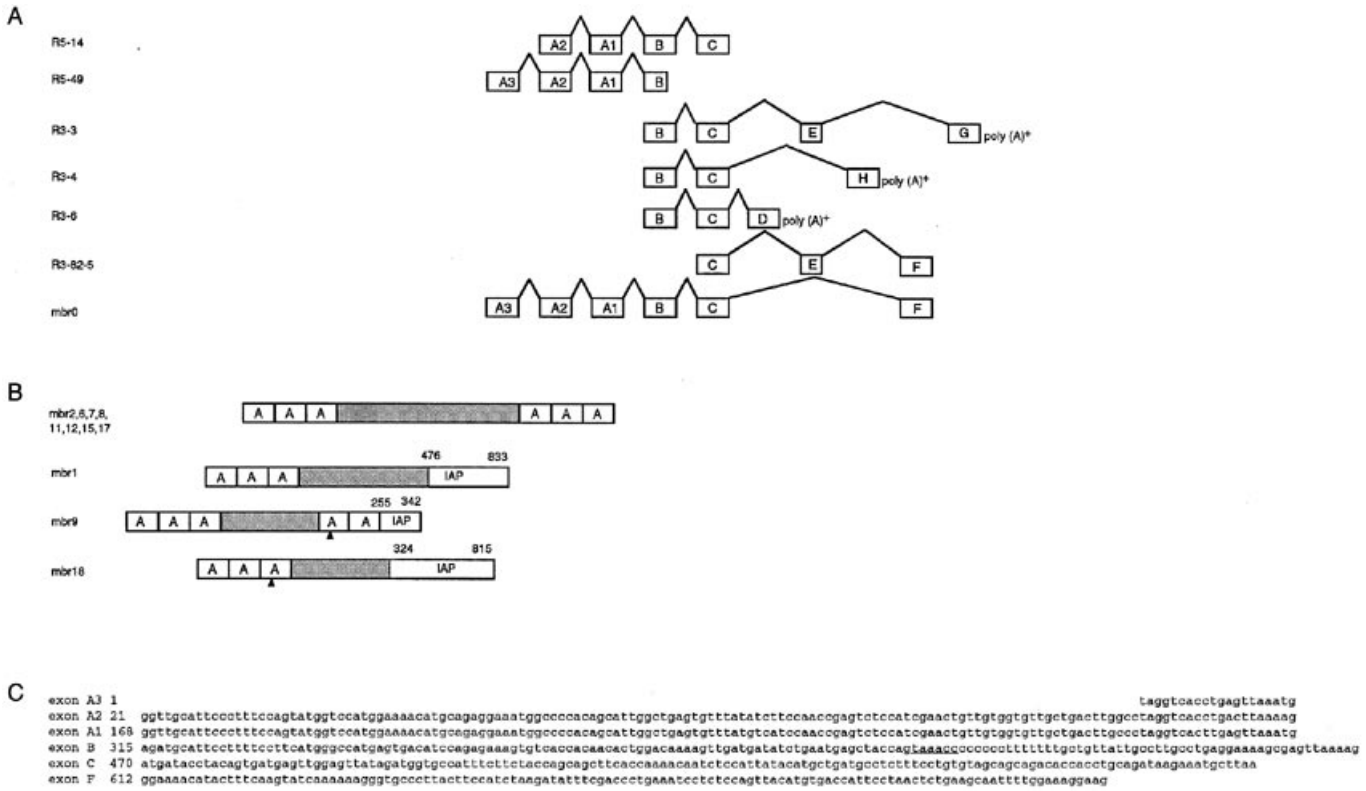

Figure 3. Transcripts and sequence of *Ipw* . The clones in (A) and (B) are aligned to show their location relative to each other. (**A**) 3′ cDNA clones. R5-14 and R5-49 are 5′ RACE clones, R3-3, R3-4 and R3-6 are 3′ RACE clones, R3-82-5 is an RT-PCR product, and mbr0 is a library cDNA clone that contained only part of exon F. Poly(A) tails on RACE clones are indicated. (**B**) 5′ cDNA clones. Eight independent clones had exon A repeats at both ends and ranged in size from 1.1 to 5 kb. Clones mbr1, mbr9 and mbr18 have IAP sequences at their 3′ ends. The position of the sequence present within the complete 7096 bp IAP sequence is noted for each clone. Gray areas represent regions not sequenced. Triangles represent small (18–22 bp) deletions within some copies of exon A. The exon structure of these eleven clones has not been determined. (**C**) Sequence of *Ipw* cDNA clone mbr0 that was extended by sequencing genomic clones to include exon F. Each exon sequence is written on a separate line and exons A1, A2 and A3 are aligned. The polymorphic site within exon B is underlined.

subsequently by using a hemi-nested RT-PCR, which detects very low abundance transcripts. This contrasts with human *IPW*, which is expressed at a low but relatively constant level in all tissues tested. Human *IPW* transcripts are easily detected by RT-PCR, without the use of the more sensitive hemi-nested procedure (1).

Ipw is expressed only from the paternally derived allele

To determine the imprinting status of *Ipw*, we searched for sequence polymorphisms among different strains of mice by direct sequencing of PCR products from exons B and C, amplified using primers from flanking introns. No sequence polymorphisms were found in three strains of *M.musculus* (ICR outbred, C57BL/6J, 129/SvJ) and NIH3T3 cells. However, between these strains of *M.musculus* and *M.m.castaneus* a polymorphism was identified in exon B that consisted of ‘taaacc’ in *M.musculus* and ‘taaccaccc’ in *M.m.castaneus.*

This polymorphism resides outside of the potential ORF. To look for parent-of-origin specific expression, we examined the transcripts in tissues from an F1 product of a *M.m.castaneus* male crossed to a *M.musculus* female mouse. When hemi-nested RT-PCR products containing the exon B polymorphism were sequenced directly, only the paternal *M.m.castaneus* allele was expressed in all tissues (stomach, heart, testes, muscle, spleen, small intestine, brain, kidney, liver and large intestine) (Fig. 4A). To ensure that the maternal allele was indeed present in the interspecies hybrid mouse, we performed PCR on liver DNA from the same animal with intron primers flanking exon B. Upon direct sequencing of this DNA PCR product, an overlapping signal from the two alleles was seen (Fig. 4B). Parallel control experiments on RNA from a *M.musculus* mouse demonstrated that the *M.musculus* allele was expressed and amplifiable in this reaction (Fig. 4C). We conclude that *Ipw* is an imprinted gene with expression limited to the paternal allele in all adult tissues examined, as is the case for *IPW*, its human counterpart.

DISCUSSION

Aside from the apparent lack of any other sequences homologous to *IPW* within the mouse genome, the specific evidence that *Ipw* is the murine homolog of *IPW* is 4-fold. The two genes share a region of 79% sequence similarity over 319 bp, although part of this region includes similarity between a human exon and a mouse intron. The *Ipw* gene is located in the central part of mouse chromosome 7, a region of known conserved synteny with human chromosome region 15q11–13, where *IPW* is located. Both genes lack protein-coding potential and both are imprinted, with expression limited to the paternal allele.

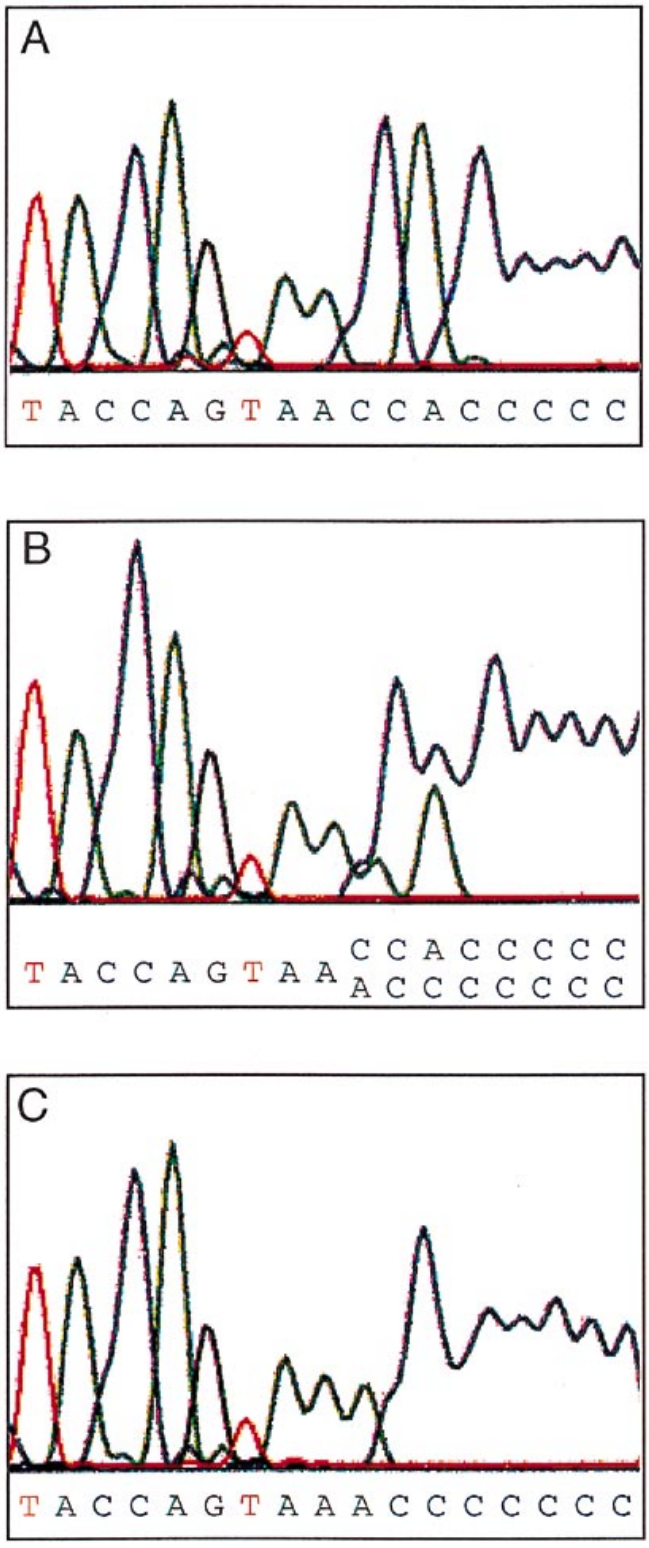

Figure 4. Imprinting analysis. Exon B sequence including the polymorphic nucleotides was obtained from RT-PCR products of tissue RNA and from genomic DNA PCR products. (**A**) Brain RNA from a [*M.musculus* (C57BL/10) × *M.m.castaneus*] F1 mouse contains only the *M.m.castaneus*-specific transcript. (**B**) DNA from the same mouse contains both alleles. (**C**) RNA from *M.musculus* brain reveals expression of the *M.musculus*-specific transcript, demonstrating its detectability by this assay.

In contrast to *IPW*, which is ubiquitously expressed at a moderate level in all human tissues, *Ipw* is expressed primarily in brain, and elsewhere at a very low level. The presence of tandem repeats within the cDNA, encoded by separate exons, is a distinctive feature of *Ipw*. Restriction digests of genomic clones containing the exon A repeats suggest a repetitive organization of a larger genomic unit. As recently suggested, sequence repeats may be a common feature of imprinted genes (26). The high degree of sequence conservation between the copies of exon A may indicate a recent expansion of these repeats or, less likely, a functional importance of their sequences. In contrast to *IPW*, the mouse gene is expressed in multiple alternatively spliced forms. Whether they have distinct functions, either in different regions of the brain or during mouse development, remains to be investigated.

Ipw joins a growing list of genes known to be imprinted in the mouse. There are now nine imprinted genes on mouse chromosome 7, including three in the central region [*Snrpn* (22), *Znf127* (10) and *Ipw*]. The other six genes are either in the proximal [*Peg3*, (27)] or distal [*Ins2*, *Igf2*, *H19*, *Mash2* and *p57kip2* (28,29)] region. One or more of these genes probably account for the imprinting effects observed in experimentally produced mice with uniparental disomies for regions of chromosome 7.

At least eight other mammalian genes produce spliced and polyadenylated transcripts with no long ORF and, therefore, appear to function at the level of RNA. Gene deletion experiments have been performed for two of these genes, *H19* and *Xist*. Deletion of the imprinted *H19* gene disrupts imprinting in the region surrounding it (16). *H19* has been proposed to regulate the imprinting status of the neighboring genes *Igf2* and *Ins2* (16,30,31). This model supposes that the transcription of *H19* blocks access of adjacent imprinted genes to shared enhancer elements, thus preventing their transcription. Lack of expression of *H19* allows access to these enhancers, and transcription of the other genes proceeds.

The murine inactive X-specific transcript, *Xist*, contains no conserved ORF and plays a major role in X-inactivation (32). In support of this hypothesis, disruption of *Xist* blocks X-inactivation in the *cis* configuration (19), and transposition of a genomic fragment containing *Xist* and flanking sequences to autosomes causes them to be counted and chosen for the X-inactivation process (20). In the early mouse embryo, *Xist* is expressed exclusively from the paternal X chromosome (33).

Both *H19* and *XIST* share >70% sequence conservation with their mouse homologs. Sequence comparison of additional mammalian homologs of *XIST* and *H19* suggests that the sequence of these two untranslated RNAs is more conserved than introns of coding genes, but less conserved than coding exons. Furthermore, although the human and mouse *XIST* genes are similar in sequence and gene structure, only the 5′ end of the *XIST* sequence appears to be conserved in other mammalian species (34).

Five other untranslated genes whose expression has not been assessed for imprinting have been described recently. The *gadd7* cDNA was isolated from Chinese hamster cells as one of a group of genes induced by DNA damage. Overexpression of *gadd7* leads to growth arrest in both hamster and human cells. The 900 nt *gadd7* cDNA does not cross-hybridize to RNA from human or mouse, and so is apparently poorly conserved (35). The murine *His1* locus is a common site of retroviral insertions leading to murine myeloid leukemia. *His1* is conserved across several species, yet its 3 kb transcript has no ORF (36). A gene proposed

to have a role in differentiation or sex determination (in acampomelic campomelic dysplasia) produces an untranslated 3.7 kb transcript (37). *DGCR5* is an alternatively spliced, untranslated transcript disrupted in an individual with DiGeorge syndrome (17). Finally, the introns but not the exons of the human U22 host gene (*UHG*) are highly conserved with the mouse homolog, consistent with the hypothesis that the function of the untranslated gene *UHG* is to provide a host gene for *UHG* intron-encoded small nucleolar RNAs (38).

Both *IPW* and *Ipw* are members of a family of untranslated RNA molecules. Although the possibility that the short ORF noted in the *Ipw* cDNA codes for a polypeptide has not formally been excluded, it is made less likely by the dissimilarity to the ORF in human *IPW*. Identification of *IPW* homologs in other organisms may help to identify regions of critical importance to gene function in this poorly conserved gene.

There are two possible roles for untranslated RNAs. First, each RNA product could have a separate cellular function, unrelated to its genomic location, perhaps as part of a ribonucleoprotein complex. The products of the *H19*, *gadd7* and *His1* genes all appear to be involved in regulation of cell growth (35,36,39). Expression of *H19* and *gadd7* inhibits cell growth, and *His1* is involved in murine leukemia when activated by retroviral insertion.

Second, genes encoding untranslated RNAs may be important regulators of regional gene expression. In favor of this idea, targeted disruptions of either *H19* or *Xist* in mice cause alterations in the expression of other genes in the *cis* configuration. Inherited microdeletions in PWS patients that include portions of alternative transcripts of the *SNRPN* gene causes loss of expression of paternally expressed genes when inherited in the *cis* configuration through the paternal germline (18). Transcription of these genes is thus implicated in control of regional gene expression. Similarly, *Ipw* could be a candidate for regulation of other genes in the imprinted region of central mouse chromosome 7. Gene knockout studies will be useful to address this possibility. Alternatively, transcripts from an imprinted region in which the chromatin conformation differs on the paternal and the maternal chromosome could simply reflect the opportunistic transcription from a region with open chromatin. Whether such transcription is the prerequisite or the consequence of the underlying mechanism that maintains parent-specific chromatin configuration becomes the crucial unanswered question.

MATERIALS AND METHODS

Identification of genomic clones

Two phage clones [λmg1 (13.3 kb), λmg2 (16.2 kb)] were isolated from a λFIX II mouse 129 library by hybridization to a probe from *IPW* exons 2 and 3 [PCR product C-D, (1)]. Two additional clones [λ67 (10.4 kb) and λ68 (9.5 kb)], were isolated from a 129SvJ embryonic stem cell library by hybridization to a subclone of λmg1 that included exon C and part of intron C. Three P1 clones [P1-5150 (60 kb), P1-5151 (55 kb), P1-5152 (60 kb)] were identified by PCR screening of a P1 library (Genome Systems) using primers derived from the exon C–intron C. P1 clone DNA was prepared by alkaline lysis. Chromosomal mapping of the *Ipw* gene by mouse × rodent somatic cell hybrid analysis and fluorescence *in situ* hybridization was performed as described (21).

RACE and identification of cDNA clones

Both 5′ and 3′ RACE were performed on mouse brain RNA essentially as described (23), with the exception that 5′ reverse transcription products were tailed with dGTP rather than dATP. RACE products were cloned into pBluescript (Stratagene) and sequenced. In 5′ RACE experiments, reverse primers in exon C were used to isolate clone R5-14 and identify exon B, then reverse primers from exon B were used generate clone R5-49. Forward primers from exon B were used in 3′ RACE to isolate clones R3-3, R3-4 and R3-6. Exon B and two copies of exon A were also identified during sequencing of λmg1. The RACE clones illustrated represent a subset of those identified. Poly(A) tails were found at the 3′ ends of clones R3-3, R3-4 and R3-6, indicating that multiple polyadenylation sites are used. Since 3′ RACE was performed using an oligo(dT) adapter, we assume that these correspond to 3′ ends of *Ipw*. Clone R3-82-5 was generated by RT-PCR with primers from exons C and F.

A cDNA probe containing parts of exons B and C was used to screen a mouse brain cDNA library (Stratagene #936309), yielding clone mbr0. This clone was used to rescreen the library, yielding 11 additional clones, ranging in size from 1.1 to 5 kb. Clone names and sizes were: mbr0, 652 bp; mbr1, 1.6 kb; mbr2, 1.1 kb; mbr6, 2.2 kb; mbr7, 5 kb; mbr8, 3 kb; mbr9, 1.6 kb; mbr11, 2.9 kb; mbr12, 1.3 kb; mbr15, 2 kb; mbr17, 1.6 kb; and mbr18, 2.3 kb. cDNA library screening was performed by filter lift hybridization, plaque purification and *in vivo* excision according to the instructions of the manufacturer.

Sequencing and sequence analysis

Sequence analysis was performed using the BLAST search algorithm (40) through the NCBI WWW site (http://www.ncbi.nlm.nih.gov/BLAST/) and the program CODON-PREFERENCE from the GCG package (Genetics Computer Group, Madison WI). Sequencing was performed either using a double stranded dideoxy chain termination method and Sequenase II (USB-Amersham) or by cycle sequencing (Perkin-Elmer Corp.) and an ABI 377 automated DNA sequencer.

RNA and DNA preparation

Total RNA was prepared from mouse tissues by first grinding under liquid nitrogen, then adding RNA-STAT (TelTestB Inc., Friendswood, TX) and proceeding according to the manufacturer's directions. DNA was extracted from tissues by grinding under liquid nitrogen, proteinase K–SDS treatment and phenol–chloroform extraction followed by ethanol precipitation.

PCR and RT-PCR

Reverse transcription-PCR was performed using Superscript reverse transcriptase (Gibco-BRL). For expression and imprinting analysis, a hemi-nested PCR strategy was used. RNA was reverse transcribed using random hexamers. The first round of PCR (95°C for 5′ then 20 cycles of 95°C for 30 s, 55°C for 30 s, 72°C for 30 s, then 72°C for 5 min) was performed with primers UF2092–UF2077. The second round of PCR, under the same conditions but for 30 cycles, was performed using primers UF2100–UF2077. Amplification of exons B and C from DNA was performed using primers UF2164–UF2165 and UF2135–UF2034 respectively, under the same PCR conditions, for 30 cycles. Primer sequences for imprinting analysis were: UF2092, 5′ACT-

GTTGTGGTGTTGCTGAC; UF2100, 5′GATGCATTCCTTTTC-CTTCA; UF2077 5′TGGTAGAAGAAATGGCACC; UF2164 5′CATCTCAATGTGCATGGTG; UF2165, 5′CTTTGGTCCA-CAAAGTCTTG; UF2135, 5′CTTTCCCTATGTAGTGTAACCA; and UF2034, 5′AGTTTGTTTAATCATGCATTTT. Oligonucleotides for hybridization and RT-PCR were: exon A, UF2092 (forward); exon B, UF2077 (forward); exon C, UF2100 (reverse); exon D, UF2183 (reverse), 5′ATCCTCCAGGTGGCTTCT; exon E, UF2181 (reverse), 5′GTGCACTGTGCAATTTCTTT; exon F, UF2182 (reverse), 5′AGGGCACCCTTTTTTGATA; exon G, UF2361 (reverse), 5′CCTTCCAAACTGGTTCAGAG; and exon H, UF2362 (reverse), 5′ATAGTTAGGTTTTATGGCCTTC.

ACKNOWLEDGEMENTS

We thank Xu Li for the FISH data, David Kingsley for the interspecific hybrid mouse and other mouse DNA samples, Anne Wang and Colin Stewart for genomic clones and helpful discussions and Julie Kerns and Davesh Sharma for technical assistance. R.W. was the recipient of a postdoctoral fellowship from the Medical Research Council of Canada and U.F. is an investigator of the Howard Hughes Medical Institute.

REFERENCES

1. Wevrick, R., Kerns, J.A. and Francke, U. (1994) Identification of a novel paternally expressed gene in the Prader–Willi syndrome region. *Hum. Mol. Genet.*, **3**, 1877–1882.
2. Holm, V.A., Cassidy, S.B., Butler, M.G., Hanchett, J.M., Greenswag, L.R., Whitman, B.Y. and Greenberg, F. (1993) Prader–Willi syndrome: consensus diagnostic criteria. *Pediatrics*, **91**, 398–402.
3. Mascari, M.J., Gottlieb, W., Rogan, P.K., Butler, M.G., Waller, D.A., Armour, J.A., Jeffreys, A.J., Ladda, R.L. and Nicholls, R.D. (1992) The frequency of uniparental disomy in Prader–Willi syndrome. Implications for molecular diagnosis. *N. Engl. J. Med.*, **326**, 1599–1607.
4. Robinson, W.P., Bottani, A., Xie, Y.G., Balakrishman, J., Binkert, F., Machler, M., Prader, A. and Schinzel, A. (1991) Molecular, cytogenetic, and clinical investigations of Prader–Willi syndrome patients. *Am. J. Hum. Genet.*, **49**, 1219–1234.
5. Özçelik, T., Leff, S.E., Robinson, W.P., Donlon, T., Lalande, M., Sanjines, E., Schinzel, A. and Francke, U. (1992) Small nuclear ribonucleoprotein polypeptide N (*SNRPN*), an expressed gene in the Prader–Willi syndrome critical region. *Nature Genet.*, **2**, 265–269.
6. Reed, M.L. and Leff, S.E. (1994) Maternal imprinting of human *SNRPN*, a gene deleted in Prader–Willi syndrome. *Nature Genet.*, **6**, 163–167.
7. Glenn, C.C., Porter, K.A., Jong, M.T., Nicholls, R.D. and Driscoll, D.J. (1993) Functional imprinting and epigenetic modification of the human *SNRPN* gene. *Hum. Mol. Genet.*, **2**, 2001–2005.
8. Mutirangura, A., Jayakumar, A., Sutcliffe, J.S., Nakao, M., McKinney, M.J., Buiting, K., Horsthemke, B., Beaudet, A.L., Chinault, A.C. and Ledbetter, D.H. (1993) A complete YAC contig of the Prader–Willi/Angelman chromosome region (15q11–q13) and refined localization of the *SNRPN* gene. *Genomics*, **18**, 546–552.
9. Driscoll, D.J., Waters, M.F., Williams, C.A., Zori, R.T., Glenn, C.C., Avidano, K.M. and Nicholls, R.D. (1992) A DNA methylation imprint, determined by the sex of the parent, distinguishes the Angelman and Prader–Willi syndromes. *Genomics*, **13**, 917–924.
10. Jong, M., Carey, A., Stewart, C., Rinchik, E., Glenn, C., Driscoll, D. and Nicholls, R. (1993) The ZNF127 gene encodes a novel C3H4 zinc-finger protein and its expression is regulated by genomic imprinting. *Am. J. Hum. Genet.*, **53**, A697.
11. Sutcliffe, J.S., Nakao, M., Christian, S., Orstavik, K.H., Tommerup, N., Ledbetter, D.H. and Beaudet, A.L. (1994) Deletions of a differentially methylated CpG island at the *SNRPN* gene define a putative imprinting control region. *Nature Genet.*, **8**, 52–58.
12. Buiting, K., Saitoh, S., Gross, S., Dittrich, B., Schwartz, S., Nicholls, R.D. and Horsthemke, B. (1995) Inherited microdeletions in the Angelman and Prader–Willi syndromes define an imprinting centre on human chromosome 15. *Nature Genet.*, **9**, 395–400.
13. Wevrick, R., Kerns, J.A. and Francke, U. (1996) The IPW gene is imprinted and is not expressed in the Prader–Willi syndrome. *Acta Genet. Med. Gemellol.*, 191–197.
14. Reis, A., Barbel, D., Greger, V., Buiting, K., Lalande, M., Gillessen-Kaesbach, G., Anvret, M. and Horsthemke, B. (1994) Imprinting mutations suggested by abnormal DNA methylation patterns in familial Angelman and Prader–Willi syndromes. *Am. J. Hum. Genet.*, **55**, 741–747.
15. Saitoh, S., Buiting, K., Rogan, P.K., Buxton, J.L., Driscoll, D.J., Arnemann, J., Konig, R., Malcolm, S., Horsthemke, B. and Nicholls, R.D. (1996) Minimal definition of the imprinting center and fixation of a chromosome 15q11–q13 epigenotype by imprinting mutation. *Proc. Natl Acad. Sci. USA*, **93**, 7811–7815.
16. Leighton, P.A., Ingram, R.S., Eggenschwiler, J., Efstratiadis, A. and Tilghman, S.M. (1995) Disruption of imprinting caused by deletion of the H19 gene region in mice. *Nature*, **375**, 34–39.
17. Sutherland, H.F., Wadey, R., McKie, J.M., Taylor, C., Atif, U., Johnstone, K.A., Halford, S., Kim, U.-J., Goodship, J., Baldini, A. and Scambler, P.J. (1996) Identification of a novel transcript disrupted by a balanced translocation associated with DiGeorge syndrome. *Am. J. Hum. Genet.*, **59**, 23–31.
18. Dittrich, B., Buiting, K., Korn, B., Rickard, S., Buxton, J., Saitoh, S., Nicholls, R.D., Poustka, A., Winterpacht, A., Zabel, B. and Horsthemke, B. (1996) Imprint switching on human chromosome 15 may involve alternative transcripts of the *SNRPN* gene. *Nature Genet.*, **14**, 163–170.
19. Penny, G.D., Kay, G.F., Sheardown, S.A., Rastan, S. and Brockdorff, N. (1996) Requirement for Xist in X chromosome inactivation. *Nature*, **379**, 131–137.
20. Lee, J., Strauss, W., Dausman, J. and Jaenisch, R. (1996) A 450 kb transgene displays properties of the mammalian X inactivation center. *Cell*, **86**, 83–94.
21. Li, X., Luna, J., Lombroso, P.J. and Francke, U. (1995) Molecular cloning of the human homolog of a striatum-enriched phosphatase (STEP) gene and chromosomal mapping of the human and murine loci. *Genomics*, **28**, 442–449.
22. Leff, S.E., Brannan, C.I., Reed, M.L., Ozcelik, T., Francke, U., Copeland, N.G. and Jenkins, N.A. (1992) Maternal imprinting of the mouse *Snrpn* gene and conserved linkage homology with the human Prader–Willi syndrome region. *Nature Genet.*, **2**, 259–264.
23. Frohman, M.A., Dush, M.K. and Martin, G.R. (1988) Rapid production of full-length cDNAs from rare transcripts: amplification using a single gene-specific oligonucleotide primer. *Proc. Natl Acad. Sci. USA*, **85**, 8998–9002.
24. Kuff, E.L., Feenstra, A., Lueders, K., Smith, L., Hawley, R., Hozumi, N. and Shulman, M. (1983) Intracisternal A-particle genes as movable elements in the mouse genome. *Proc. Natl Acad. Sci. USA*, **80**, 1992–1996.
25. Kozak, M. (1989) The scanning model for translation: an update. *J. Cell Biol.*, **108**, 229–241.
26. Neumann, B., Kubicka, P. and Barlow, D.P. (1995) Characteristics of imprinted genes [letter]. *Nature Genet.*, **9**, 12–13.
27. Kuroiwa, Y., Kaneko-Ishino, T., Kagitani, F., Kohda, T., Li, L.-L., Tada, M., Suzuki, R., Yokoyama, M., Shiroishi, T., Wakana, S., Barton, S.C., Ishino, F. and Surani, M.A. (1996) *Peg3* imprinted gene on proximal chromosome 7 encodes for a zinc finger protein. *Nature Genet.*, **12**, 186–189.
28. Hatada, I. and Mukai, T. (1995) Genomic imprinting of p57KIP2, a cyclin-dependent kinase inhibitor, in mouse. *Nature Genet.*, **11**, 204–205.
29. Guillemot, F., Caspary, T., Tilghman, S.M., Copeland, N.G., Gilbert, D.J., Jenkins, N.A., Anderson, D.J., Joyner, A.L., Rossant, J. and Nagy, A. (1995) Genomic imprinting of Mash2, a mouse gene required for trophoblast development. *Nature Genet.*, **9**, 235–242.
30. Bartolomei, M.S., Webber, A.L., Brunkow, M.E. and Tilghman, S.M. (1993) Epigenetic mechanisms underlying the imprinting of the mouse H19 gene. *Genes Dev.*, **7**, 1663–1673.
31. Brannan, C.I., Dees, E.C., Ingram, R.S. and Tilghman, S.M. (1990) The product of the H19 gene may function as an RNA. *Mol. Cell. Biol.*, **10**, 28–36.
32. Brockdorff, N., Ashworth, A., Kay, G.F., McCabe, V.M., Norris, D.P., Cooper, P.J., Swift, S. and Rastan, S. (1992) The product of the mouse Xist gene is a 15 kb inactive X-specific transcript containing no conserved ORF and located in the nucleus. *Cell*, **71**, 515–526.
33. Kay, G.F., Barton, S.C., Surani, M.A. and Rastan, S. (1994) Imprinting and X chromosome counting mechanisms determine Xist expression in early mouse development. *Cell*, **77**, 639–650.
34. Hendrich, B.D., Brown, C.J. and Willard, H.F. (1993) Evolutionary conservation of possible functional domains of the human and murine *XIST* genes. *Hum. Mol. Genet.*, **2**, 663–672.

35. Hollander, M.C., Alamo, I. and Fornace, A.J., Jr (1996) A novel DNA damage-inducible transcript, *gadd-7*, inhibits cell growth, but lacks a protein product. *Nucleic Acids Res.*, **24**, 1589–1593.
36. Askew, D., Li, J. and Ihle, J. (1994) Retroviral insertions in the murine *His-1* locus activate the expression of a novel RNA that lacks an extensive open reading frame. *Mol. Cell. Biol.*, **14**, 1743–1751.
37. Ninomiya, S., Isomura, M., Narahara, K., Seino, Y. and Nakamara, Y. (1996) Isolation of a testis-specific cDNA on chromosome 17q from a region adjacent to the breakpoint of t(12;17) observed in a patient with acampomelic campomelic dysplasia and sex reversal. *Hum. Mol. Genet.*, **5**, 69–72.
38. Tycowski, K.T., Shu, M.-D. and Steitz, J.A. (1996) A mammalian gene with introns instead of exons generating stable RNA products. *Nature*, **379**, 464–466.
39. Hao, Y., Crenshaw, T., Moulton, T., Newcomb, E. and Tycko, B. (1993) Tumour-suppressor activity of *H19* RNA. *Nature*, **365**, 764–767.
40. Altschul, S.F., Gish, W., Miller, W., Myers, E.W. and Lipman, D.J. (1990) Basic local alignment search tool. *J. Mol. Biol.*, **215**, 403.