Human Molecular Genetics, 2001, Vol. 10, No. 23 **2687–2700**

The IC-*SNURF–SNRPN* transcript serves as a host for multiple small nucleolar RNA species and as an antisense RNA for *UBE3A*

Maren Runte, Alexander Hüttenhofer[1], Stephanie Groß, Martin Kiefmann[1], Bernhard Horsthemke and Karin Buiting*

Institut für Humangenetik, Universitätsklinikum Essen, 45122 Essen, Germany and [1]Institut für Experimentelle Pathologie/Molekulare Neurobiologie, ZMBE, Münster, Germany

Received July 18, 2001; Revised and Accepted September 10, 2001

The imprinted domain on human chromosome 15 consists of two oppositely imprinted gene clusters, which are under the coordinated control of an imprinting center (IC) at the 5′ end of the *SNURF–SNRPN* gene. One gene cluster spans the centromeric part of this domain and contains several genes that are transcribed from the paternal chromosome only (*MKRN3*, *MAGEL2*, *NDN*, *SNURF–SNRPN*, *HBII-13*, *HBII-85* and *HBII-52*). Apart from the *HBII* small nucleolar RNA (snoRNA) genes, each of these genes is associated with a 5′ differentially methylated region (DMR). The second gene cluster maps to the telomeric part of the imprinted domain and contains two genes (*UBE3A* and *ATP10C*), which in some tissues are preferentially expressed from the maternal chromosome. So far, no DMR has been identified at these loci. Instead, maternal-only expression of *UBE3A* may be regulated indirectly through a paternally expressed antisense transcript. We report here that a processed antisense transcript of *UBE3A* starts at the IC. The *SNURF–SNRPN* sense/*UBE3A* antisense transcription unit spans more than 460 kb and contains at least 148 exons, including the previously identified *IPW* exons. It serves as the host for the previously identified HBII-13, HBII-85 and HBII-52 snoRNAs as well as for four additional snoRNAs (HBII-436, HBII-437, HBII-438A and HBII-438B), newly identified in this study. Almost all of those snoRNAs are encoded within introns of this large transcript. Northern blot analysis indicates that most if not all of these snoRNAs are indeed expressed by processing from these introns. As we have not obtained any evidence for other genes in this region, which, from the mouse data appears to be critical for the neonatal Prader–Willi syndrome phenotype, a lack of these snoRNAs may be causally involved in this disease.

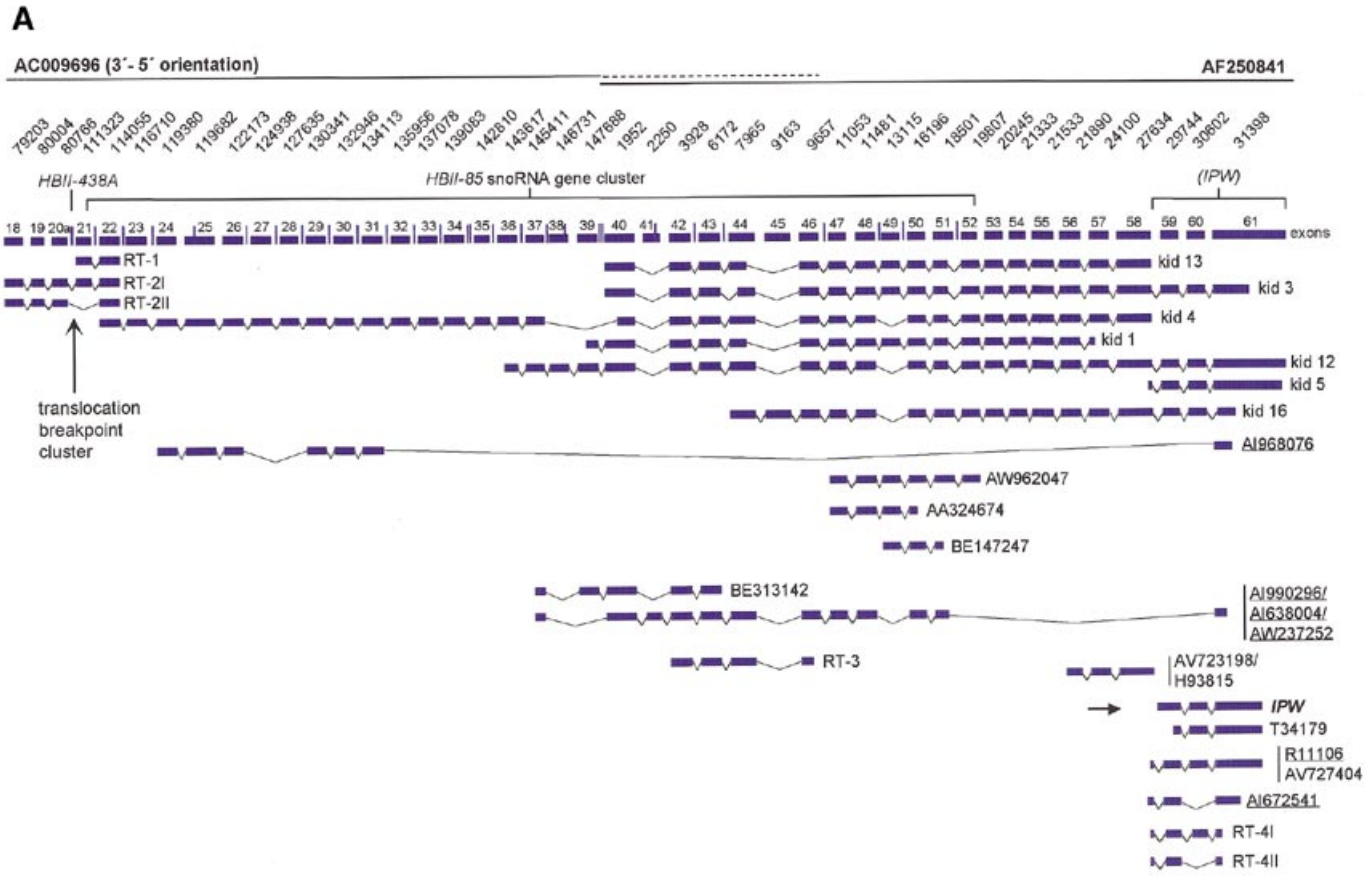
INTRODUCTION

In contrast to most other genes, imprinted genes are differentially expressed from the maternal and paternal allele. Imprinted gene expression can occur in all cells of an individual or in a temporally and spatially restricted manner. Whereas maternal silencing most often involves promoter methylation, paternal silencing does so less often. It has been argued that this is the evolutionary result of early epigenetic reprogramming in the zygote (1). As the paternal genome is actively demethylated in the oocyte, it may have developed other strategies to silence genes. One such strategy may be the expression of an antisense transcript, which indirectly silences the paternal allele. In fact, all antisense transcripts identified within imprinted regions so far are expressed from the paternal allele only, apart from *Tsix* (2). However, it is still unknown how an antisense transcript might silence a gene in *cis*. Whatever mechanism is involved in this process, indirect silencing through an antisense transcript may make gene expression amenable to temporal and spatial modulation.

A suitable system to study the regulation of imprinted gene expression is the Prader–Willi/Angelman syndrome (PWS/AS) region on human chromosome 15 (reviewed in 3). This region is under the coordinated control of an imprinting center (IC) at the 5′ end of the *SNURF–SNRPN* gene (4,5), which is expressed from the paternal allele in all tissues studied so far. *SNURF–SNRPN* and other paternally expressed genes in this region are associated with a differentially methylated region (DMR). In contrast, two genes located telomeric to *SNURF–SNRPN* (*UBE3A* and *ATP10C*) are imprinted in the opposite direction, and imprinted expression is restricted to certain tissues (6–9). Whereas *SNURF–SNRPN* is transcribed from centromere to telomere, *UBE3A* and *ATP10C* are transcribed from telomere to centromere.

Recently it has been reported that *UBE3A* was methylated in a monochromosomal hybrid cell line containing a paternal human chromosome 15, but unmethylated in a cell line containing a maternal chromosome 15, although the gene was expressed in both cell lines (10). However, systematic investigations by several labs have so far failed to detect a DMR at this locus in human tissues (11, A.C.Lossie and D.J.Driscoll, personal communication; unpublished data). On the other hand, a paternally expressed and intronless *UBE3A*

*To whom correspondence should be addressed. Tel: +49 201 723 4555; Fax: +49 201 723 5900; Email: karin.buiting@uni-essen.de
The authors wish it to be known that, in their opinion, the first two authors should be regarded as joint First Authors



antisense RNA fragment of ~20 kb has been detected by Rougeulle *et al.* (12). In the mouse, Chamberlain *et al.* (13) have demonstrated that this antisense transcript is under the control of the IC: in a mouse harboring an IC deletion on the paternal chromosome, the antisense transcript was absent and *Ube3a* was expressed biallelically in brain. However, it remained unclear, whether the antisense transcript was directly or indirectly controlled by the IC. Here we show that the antisense transcript starts at the IC.

Whereas *UBE3A* has been identified as the AS gene (14,15), the genes involved in PWS are less clear. Mouse data suggest that the region between *SNURF–SNRPN* and *IPW* may be critical. We have recently shown that this region encodes multiple copies of small nucleolar RNAs (snoRNAs) HBII-13, HBII-85 and HBII-52 (16). *HBII-85* gene copies have also been described by de los Santos *et al.* (17) and Meguro *et al.* (10). Whereas the *HBII-13* snoRNA is present as a single gene, *HBII-85* and *HBII-52* are present in 24 or 47 gene copies, respectively. Here we show that the *SNURF–SNRPN* sense/*UBE3A* antisense transcription unit serves as the host gene for these snoRNAs as well as for four newly identified candidates for snoRNAs, HBII-436, HBII-437, HBII-438 A and HBII-438B.

RESULTS

The *SNURF–SNRPN* transcription unit

Recently, we identified eight novel non-coding 3′ exons of the *SNURF–SNRPN* gene (exons 13–20) (18). Based on RT–PCR experiments, EST and UniGene cluster sequence data exon 20 appeared to be 15.8 kb in length, but it was unclear whether it was the true 3′ end of the *SNURF–SNRPN* transcription unit. We could not exclude at this point that this exon might contain an alternative splice donor site giving rise to a variant transcript extending much further. A similar situation had been encountered at exons 12 and 16 (18). To address this question, we searched for ESTs between exon 20 and the *UBE3A* locus. Using the sequence of the overlapping genomic BAC/PAC clones RP11–131I21, A17157+P0950, RP13–487P22 and pDJ373b1 (GenBank accession nos. AC009696.10, AF250841, AC084009 and AC004600) and the NIX software (http://www.hgmp.mrc.ac.uk/), we identified two EST clusters, one containing *IPW* sequences and another one covering most of the *UBE3A* gene in an antisense orientation to the latter gene. Both EST clusters contained spliced exons. By complete sequencing of six EST clones of the first cluster (AI968076, AI990296, AI638004, AW237252, R11106 and AI672541) and two RT–PCR products from human fetal brain RNA (RT-3 and RT-4; Fig. 1A), we identified 23 novel exons representing six different alternatively spliced variants of one transcript. This transcript includes *IPW* exons 1–3, which also were found to be subject to alternative splicing. The EST cluster in the *UBE3A* region comprises five overlapping ESTs with distinct spliced exons. We sequenced three of these cDNA clones (N52596, R19540 and W90408/W90381) and found 11 novel exons in three different alternatively spliced

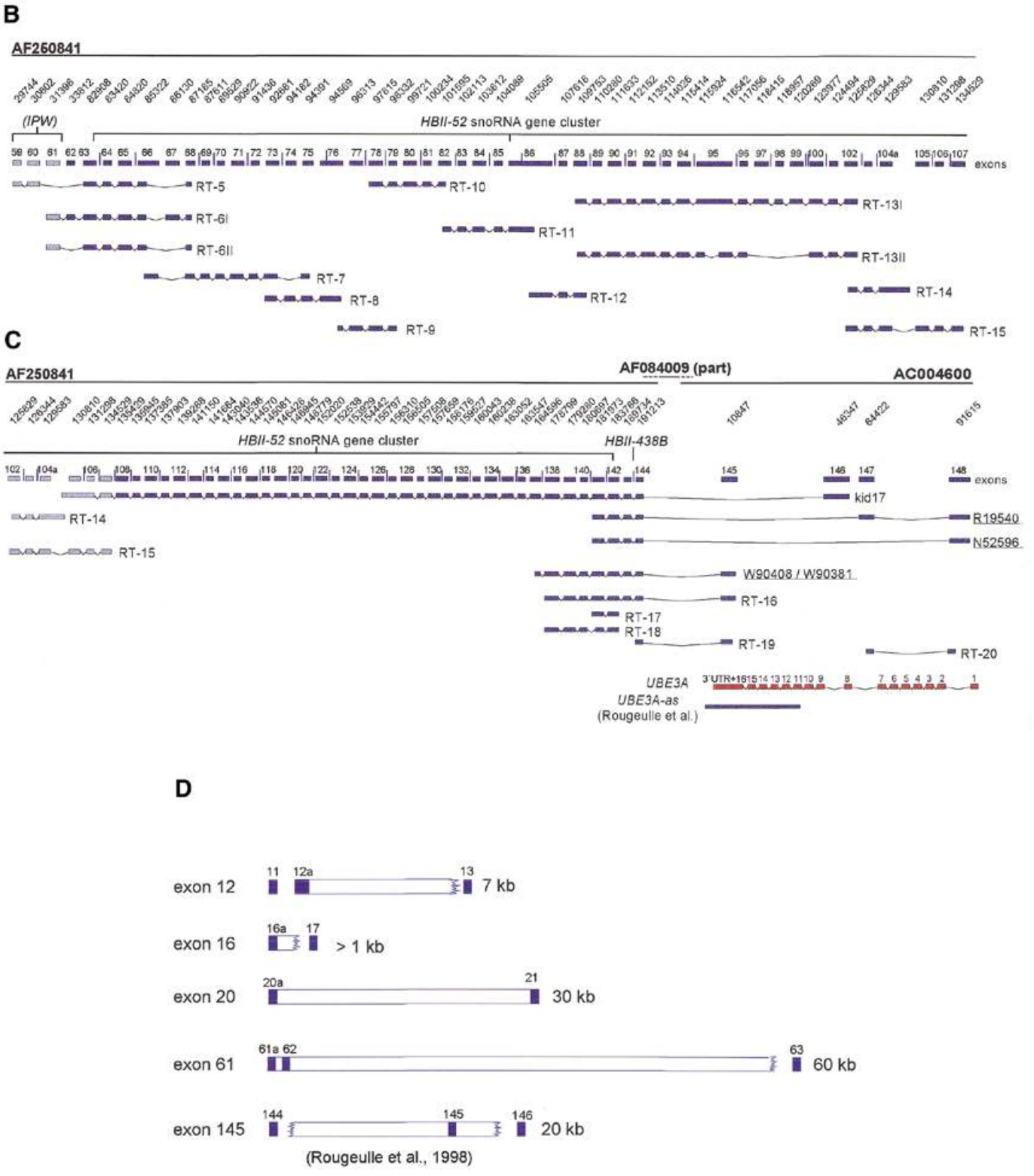


Figure 1. Schematic overview of the region spanning (**A**, opposite) *SNURF–SNRPN* exons 18–61, (**B**) exons 59–107 and (**C**) exons 102–148. Note that the maps in (A) and (B) overlap at exons 59–61 and that the maps in (B) and (C) overlap at exons 102–107 (indicated in light blue). *SNURF–SNRPN* exons are shown in blue, *UBE3A* exons are shown in red. snoRNA genes for HBII-438A, HBII-85, HBII-52 and HBII-438B are indicated as blue vertical lines. Top row, exon–intron organization; subsequent rows, cDNA clones from an adult human kidney cDNA library (kid), EST sequences and RT–PCR products. cDNA clones of ESTs which have been sequenced in completion are underlined. Nucleotide positions indicate the first base pair of the *SNURF–SNRPN* exons and correspond to the genomic sequences of either AC009696.10, AF250841 or AC004600. The translocation breakpoint cluster described by Wirth *et al.* (18) is indicated by an arrow. blue, paternally expressed; red, maternally expressed. (**D**) Schematic presentation of exons 12, 16, 20, 61 and 144, which contain alternative splice acceptor or donor sites. Unspliced RNA is shown in white. Blue boxes represent *SNURF–SNRPN* exons which use alternative splice sites inside the colinear expressed region.

isoforms. The four most downstream exons were found to cover most of *UBE3A* in an antisense manner with one exon between *UBE3A* exon 1 and 2, one between exon 7 and 8 and another one which overlaps with the *UBE3A* 3′ region (Fig. 1C).

By screening a human adult kidney cDNA library selected for large inserts with RT–PCR products for the *IPW* and *UBE3A* regions (RT-3 and RT-18), we obtained 16 cDNA clones from 1 to 4.9 kb (kid1–16) for the *IPW* region and one cDNA clone of 6.1 kb (kid17) for the *UBE3A* region. By complete sequencing of 8 cDNA clones, we identified 46 additional exons, 14 for the *IPW* and 32 for the *UBE3A* region (Fig. 1A and C). Interestingly, these exons map inside the previously reported snoRNA gene clusters *HBII-85* and *HBII-52*.

To find out whether all these exons belong to the *SNURF–SNRPN* transcription unit or are part of independent transcripts we searched for additional exons and tried to connect *SNURF–SNRPN* exon 20 with the *HBII-85/IPW* and the *HBII-52/UBE3A* exon clusters, respectively. For this purpose, we made use of the high sequence similarity of the exons inside the *HBII-52* and the *HBII-85* gene clusters and searched for more or less conserved splice donor and splice acceptor sites to predict putative exons for primer design. In fact, by extended exon-connection RT–PCR on human fetal brain RNA, we identified additional exons and could eventually link the *SNURF–SNRPN* 3′ exons with the exon cluster in the *HBII-85/IPW* region and in turn connect the *IPW* exons with the exons in the *HBII-52/UBE3A* region (Fig. 1A and B). Thereby we detected a splice donor site at nucleotide position 162 of exon 20. But, based on EST sequences of three ESTs (AI197860, BF672929 and AW294767) and RT–PCRs on DNaseI-treated RNA from fetal brain at seven different sites (data not shown), we found also that >30 kb of DNA contiguous with exon 20 is expressed as RNA (Fig. 1D).

Similarly to this situation, exon 61 (*IPW* exon 3) contains alternative splice donor and acceptor sites (Fig. 1D). Again, based on sequence data for some of the kidney cDNA clones, EST sequences of 18 ESTs (AI792942, AW299520, AW779767, N21972, AW893968, H63591, H85187, AW973432, H17549, AI537107, AV709519, AA001781, BF796272, AA719946, BF315994, AL719946, BF315994, AL137489) and RT–PCRs on DNaseI-treated RNA from fetal brain at five different sites, we found that >50 kb of DNA contiguous with exon 61 is expressed as RNA (data not shown; Fig. 1D).

In summary, we have identified 128 novel exons of the *SNURF–SNRPN* transcription unit. As for the previously reported 3′ exons 10–20 we could not find any significant open reading frame. Based on the cDNA clone kid17, there is at least one putative 3′ end in exon 146 with a polyadenylation site 21 nt upstream of a poly(A) tail.

Expression analysis

Northern blots containing poly(A)+ RNA from 16 different adult and four fetal tissues (Clontech) were hybridized with two different probes (RT–PCR products RT-3 and RT-18), representing exons 42, 43, 44, 46 and 138–142. Both probes failed to detect a distinct signal (data not shown). However, from EST sequences, cDNA clones and RT–PCR experiments, this transcript is expressed in various tissues. To investigate the imprinting status of the novel exons, we performed RT–PCR with primers AI990296 a and b (RT-3, exons 42, 43, 44, 46) and MRts 5–6F and R (RT-17, exons 141–142). As a template we used lymphoblastoid cell line RNA from a patient with AS and a maternal deletion of 15q11–q13 and a patient with PWS and a paternal deletion of this region. As shown in Figure 2, two RT–PCR products of the expected sizes of 501 and 120 bp, respectively, were obtained from the AS RNA, but not from the PWS RNA. These data indicate that these exons are expressed from the paternal chromosome only. This is in agreement with previously reported paternal only expression of the *IPW* exons, which we found to be part of the transcript unit.

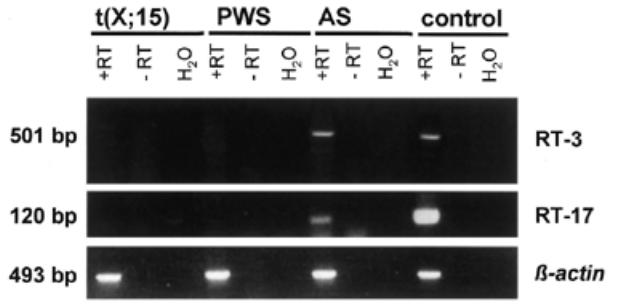


Figure 2. Expression analysis of the *SNURF–SNRPN* sense/*UBE3A* antisense transcript. RT–PCR primers for RT-3 and RT-17 were used to amplify a 501 and 120 bp RT–PCR product, respectively. Both RT–PCR products were absent in RNA from a patient with PWS and a paternal deletion of 15q11–q13 and in a patient with a balanced t(X;15) translocation (18), but present in a patient with AS and a maternal deletion of 15q11–q13 as well as in RNA from a normal control. The integrity of the RNA samples was shown by amplification of a 493 bp RT–PCR product from the β-*actin* locus. +RT, RT–PCR with reverse transcriptase; –RT, RT–PCR without reverse transcriptase; H_2O, RT–PCR without RNA.

To substantiate the notion that the newly identified exons are part of the *SNURF–SNRPN* transcription unit, we used the same primer pairs to investigate expression in a patient with a *de novo* translocation t(X;15)(q28;q12). In this patient, the 15q breakpoint is between exons 20a and 21. As previously shown for the *IPW* exons (18) which map between the two regions tested here, no expression was observed (Fig. 2).

Novel paternally expressed C/D box snoRNAs in the PWS/AS region

By computer aided analysis, using conserved sequence and structural motifs, we have identified four novel candidates for C/D box snoRNAs distal to *SNURF–SNRPN*, designated HBII-436, HBII-437, HBII-438A and HBII-438B (Fig. 3A). With one exception (HBII-437) the sequences contain all the sequence (C-, C′-, D′- and D-boxes) and structural motifs (short inverted repeats at their 5′- and 3′ ends) of bona fide C/D box snoRNAs. HBII-437 was found to contain a degenerate D-box: GTGA instead of CTGA, making it a less likely candidate for a bona fide snoRNA. HBII-436 maps ~3.5 kb proximal whereas HBII-437 maps ~0.9 kb distal to the HBII-13 snoRNA, inside the PAR-SN/PAR-5 region. With regard to the *SNURF–SNRPN* transcription unit, all three map inside intron 12. HBII-438A and HBII-438B are identical in sequence, but are located ~240 kb apart with one copy located within intron 20a of *SNURF–SNRPN*

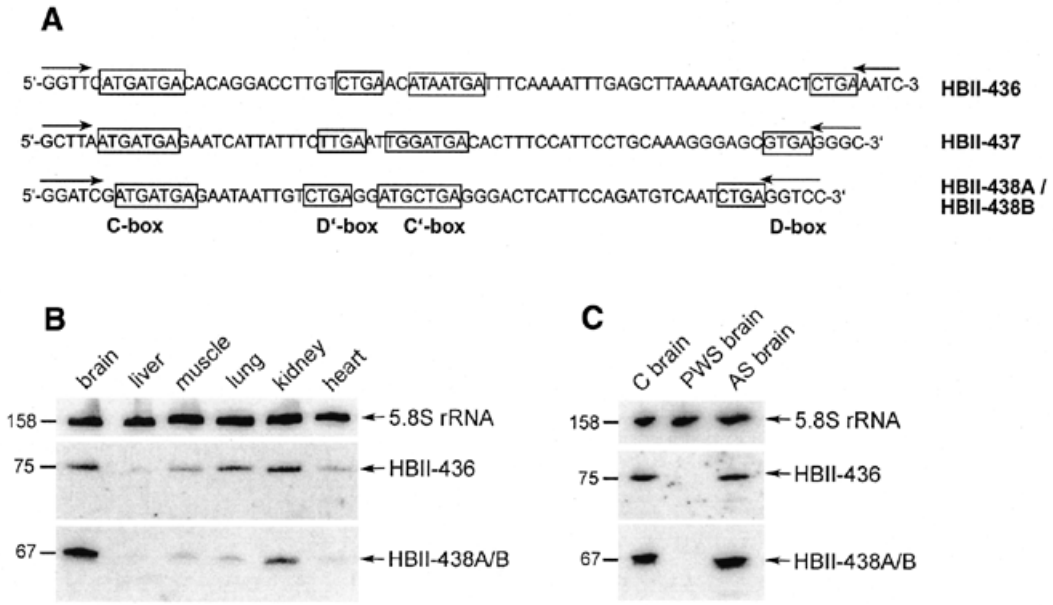


Figure 3. Novel snoRNA candidates and their expression. (**A**) Conserved sequence elements (C-, D′-, C′- and D-boxes) of canonical C/D box snoRNAs are indicated by rectangles. Short inverted repeats at the 5′ and 3′ ends of snoRNAs, able to form a short terminal stem structure, are indicated by arrows. Sequences of HBII-438A and HBII-438B snoRNAs are identical and therefore shown as one. (**B**) Northern blot analysis showing tissue-specific expression of *HBII-436* and *HBII-438A/B* snoRNAs in human. (**C**) Lack of expression of *HBII-436* and *HBII-438A/B* snoRNAs in the brain of a PWS patient. RNA samples were taken from the cortex of a human control sample (C brain), a PWS patient (PWS brain) and an AS patient (AS brain). Probing for ribosomal 5.8S rRNA provided an internal control. Sizes (in nt) of RNAs are indicated on the left.

just proximal to the *HBII-85* gene cluster, and the second copy within intron 143, just distal to the *HBII-52* gene cluster.

We have analyzed the tissue-specific expression of *HBII-436*, *HBII-437*, *HBII-438A* and *HBII-438B* by northern blot analysis containing total RNA of human brain, liver, muscle, lung, kidney and heart (Fig. 3B). *HBII-436* is expressed in brain, lung and kidney and, to a lower extent, in muscle and heart. After longer exposure of autoradiograms, expression in liver is also observed. The expression pattern closely resembles that of the recently described C/D box snoRNA *HBII-13*, which maps close to *HBII-436* (16). The expression pattern of *HBII-438A* and *HBII-438B* cannot be tested independently, since both sequences are identical. Their expression pattern resembles that of *HBII-85*, e.g. strongest expression in brain and kidney, weaker expression in muscle and lung and very low expression in liver and heart (Fig. 3B). *HBII-436* and *HBII-438A* and/or *B* are paternally expressed, imprinted snoRNAs, since they are not expressed in the brain of a PWS patient, but in the brain of an AS patient (Fig. 3C). The expression of *HBII-437*, which harbors a degenerate D-box (see above), could not be confirmed by northern blot analysis using three different oligonucleotides as probes which were derived from the *HBII-437* sequence (data not shown).

Expression of the *HBII-85* gene and *HBII-52* gene copies

We have previously identified 24 copies of the HBII-85 snoRNA species (16). Additionally, three copies (copies 24, 25 and 26; Fig. 3A) containing variant, but clearly related species to the HBII-85 snoRNA could be found in our screen. A sequence alignment of all 27 gene copies revealed that three main paralogous groups of *HBII-85* exist in the genome (Fig. 4A). Group I consists of gene copies 1–9, which is followed by Group II encoding copies 10–23 and Group III (copies 24–27) containing four degenerate copies, which map to the telomeric end of the *HBII-85* gene cluster.

Canonical snoRNAs targeting ribosomal RNAs for modification contain antisense boxes located immediately 5′ to the D- or D′-box, complementary to distinct regions within rRNAs. By this mechanism 2′-*O*-methylation of riboses within ribosomal RNA is achieved (reviewed in 19,20). Two unusual features can be observed within the antisense boxes of the HBII-85 snoRNA cluster (Fig. 4A). First, no complementarity to ribosomal RNA modification sites can be found indicative of other RNA targets, like for example mRNAs (16). Second, sequences of antisense boxes vary between all three groups. Using oligonucleotides specific for some members of each of the three groups (Fig. 4A, red dots), we performed northern blot analysis to determine whether representatives of each group were expressed in human brain. As shown in Figure 4B, this is indeed the case. Furthermore, the expression level of respective RNA species from the three groups, as assessed by northern blot analysis and subsequently quantitated by phosphoimaging, correlated well with their copy number. This is consistent with a model that all (or most) copies of *HBII-85* snoRNAs are indeed expressed in human brain.

No clear distinction of groups can be made for the 47 gene copies of the *HBII-52* snoRNA cluster, with the exception of three copies (copies 17–19), which deviate by three bases from the consensus motif of the antisense box (data not shown). Northern blot analysis using either a consensus oligonucleotide directed against 32 of the 47 copies (Fig. 4C, Group I)

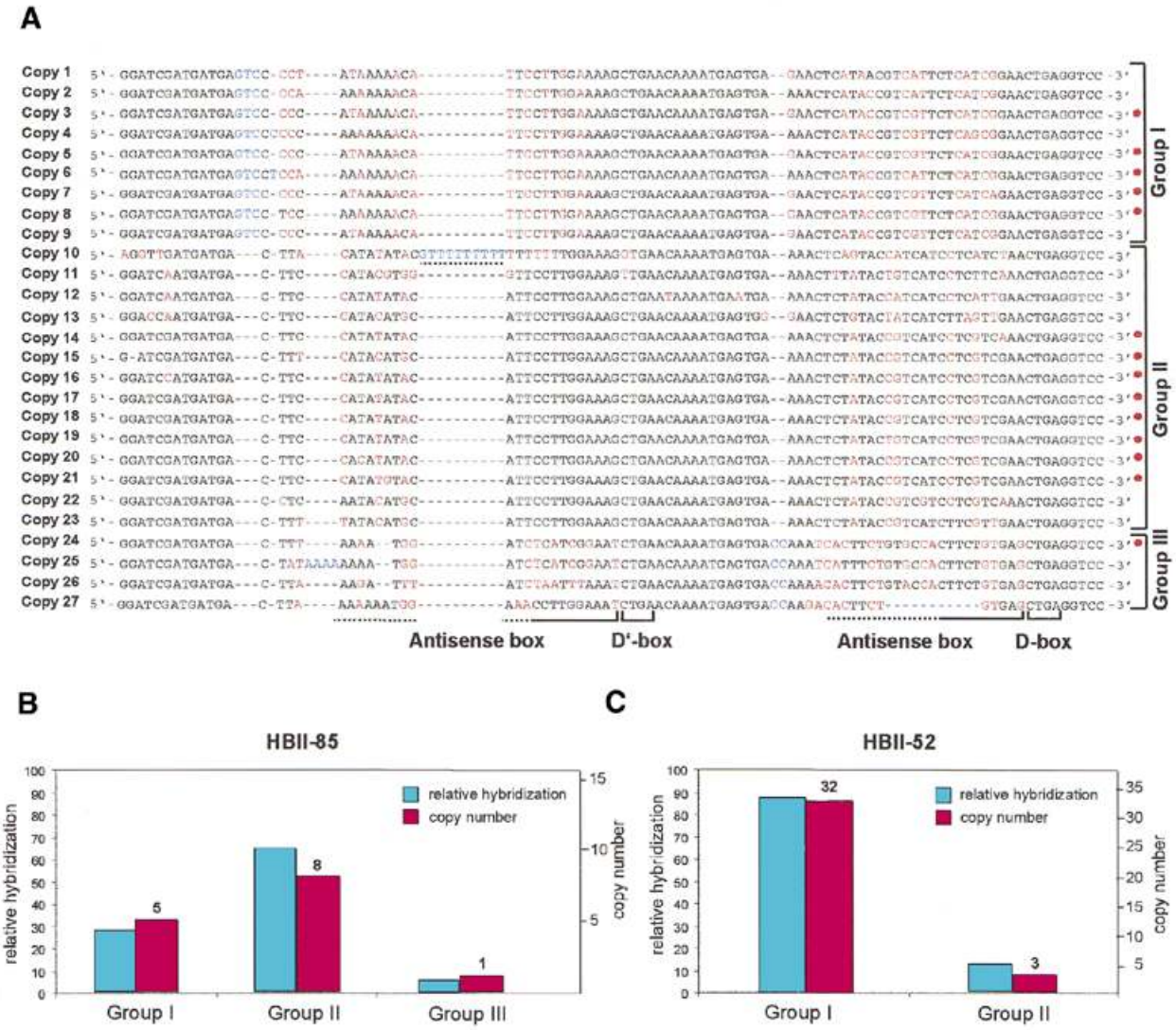


Figure 4. Expression analysis of the HBII-85 and HBII-52 snoRNAs. (**A**) Sequence alignment of the 27 *HBII-85* genomic copies. Base substitutions compared to the consensus sequence are indicated in red, base insertions or deletions are indicated in blue. Right: the 27 copies can be assigned to three main groups. For expression analysis (**B**), specific oligonucleotides were designed able to hybridize to distinct members of each group indicated by a red dot. Bottom: the locations of the D- or D′-box of the HBII-85 snoRNAs are indicated by black lines. In canonical C/D box snoRNAs, antisense boxes are located immediately 5′ to the D- or D′-boxes. Their complementarity to a RNA target (usually ribosomal or small nuclear RNAs) extends for a minimum of 10 nt to a maximum of 21 nt (20) which is indicated by a black (or dotted) line. (B) Expression analysis of the three groups of HBII-85 snoRNAs. (**C**) Expression analysis of the two groups of HBII-85 snoRNAs. Quantitation of hybridization signals was performed by phosphoimager analysis of northern blots (defined as relative hybridization). The number of copies within each group, which gives rise to the respective relative hybridization signal, is indicated by a red bar.

or a specifc oligonucleotide directed against the three deviating copies (Fig. 4C, Group II) demonstrates that the three copies were also expressed. As observed for the *HBII-85* snoRNA gene cluster, the expression level of the three versus 32 copies, correlated well with their respective copy number (Fig. 4C).

DISCUSSION

The *SNURF–SNRPN* locus on human chromosome 15 appears to be one of the most complex loci in the human genome. The core gene has 10 exons, which are transcribed into a 1.4 kb bicistronic mRNA. Whereas exons 1–3 encode the SNURF protein (21), exons 4–10 encode the SmN splicosomal protein (22). There are at least two alternative 5′ start sites and multiple untranslated upstream exons of unknown function (23,24). Furthermore, this locus harbors a bipartite IC, which controls the whole imprinted domain (5,23,25,26). Whereas the AS IC element, which spans the upstream exon u5, is necessary for maternal imprinting, the PWS IC element, which spans exon 1, is necessary for the postzygotic maintenance of the paternal imprint (27,28). We also detected additional 3′ exons of unknown function (18,29). Here we describe that the *SNURF–SNRPN* transcription unit extends much more 3′ and serves two additional functions: (i) it is the host for multiple snoRNA genes and (ii) it serves as the start site for the *UBE3A* antisense transcript (Fig. 5A). Although the role of the

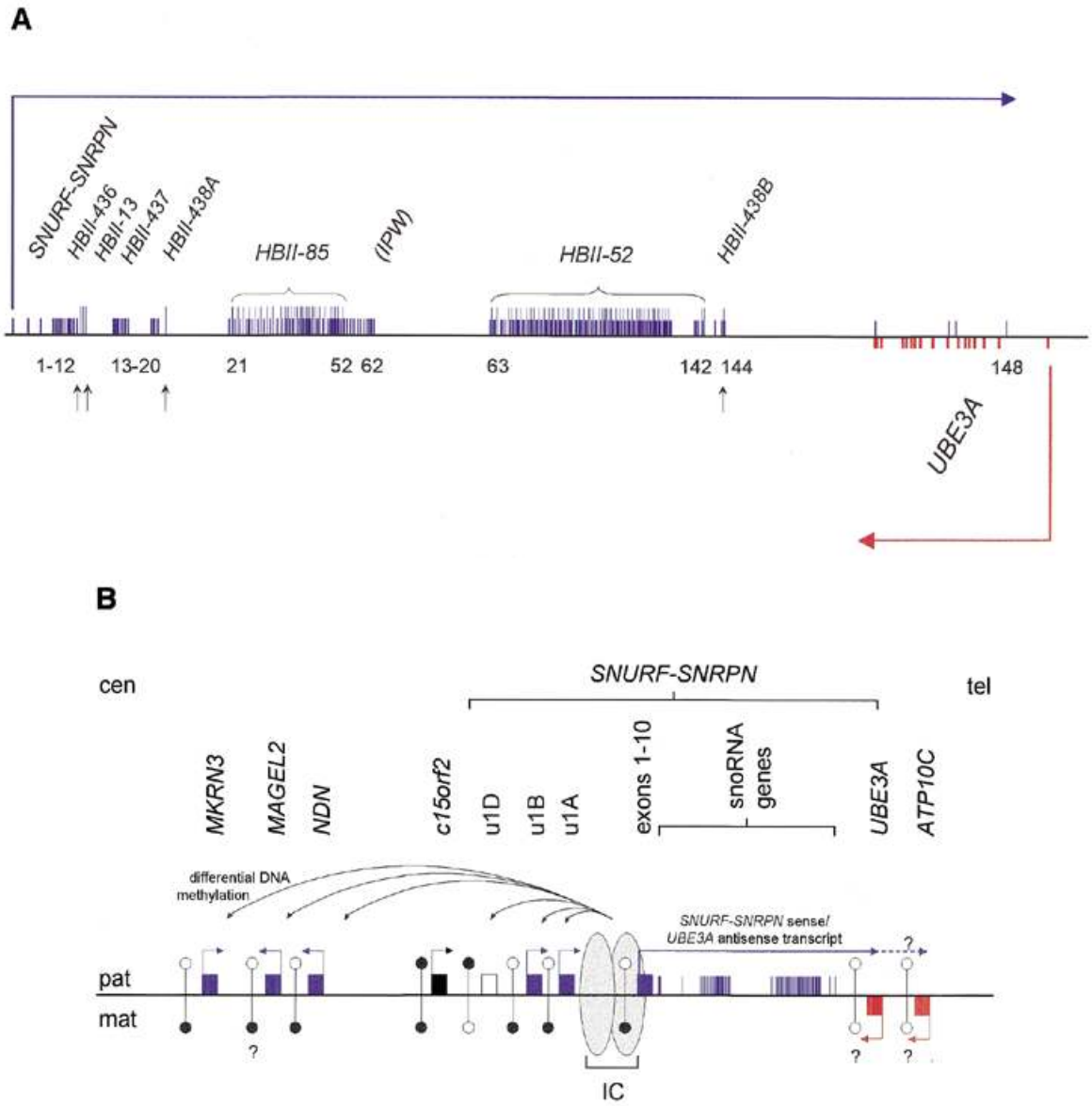


Figure 5. A model for regulation of imprinted gene expression in 15q11–q13. (**A**) Detailed map of the *SNURF–SNRPN/UBE3A* region. Exons of the *SNURF–SNRPN* sense/*UBE3A* antisense transcript are shown as short blue vertical lines and snoRNA genes are represented by long blue vertical lines. The numbers of exons are given below the horizontal line. Maternally expressed genes are represented by red vertical lines. Vertical arrows below show the positions of the novel snoRNAs HBII-436, HBII-437, HBII-438A and HBII-438B. Orientation of transcription is indicated by horizontal arrows. Not drawn to scale. (**B**) The IC regulates imprinted expression of upstream paternally expressed genes (blue boxes) by differential DNA methylation, whereas imprinted expression of maternally expressed genes (red boxes) is regulated through a paternally expressed antisense transcript initiated at the IC. The bipartite structure of the IC is indicated by two overlapping grey shadowed ellipses. Orientation of transcription for each gene is indicated by a horizontal arrow. The orientation of transcription of *MAGEL2* and *NDN* is based on physical maps and bioinformatics data (F.Muscatelli, personal communication). Cen, centromere; tel, telomere; pat, paternal chromosome; mat, maternal chromosome. Filled lollipop, methylated site(s); open lollipop, unmethylated site(s); question mark, methylation or expression status unknown. Not drawn to scale.

antisense transcript is unknown, there are some data compatible with the assumption that it may control imprinted expression of *UBE3A* (12,13). So far, we have failed to extend the transcript further downstream, but it is tempting to speculate that it may extend beyond exon 148 to the *ATP10C* locus and be associated with imprinted expression of this gene as well.

The evidence for a large *SNURF–SNRPN* sense/*UBE3A* antisense transcript initiated at the IC is based on the identification of overlapping cDNA clones and RT–PCR products. We have not been able to detect the transcript by northern blot analysis, but this may be due to its large size, low level of expression or unstable nature and the presence of multiple alternative splice forms. The evidence for this transcript is supported, although not proven, by three observations. (i) As shown by expression analysis in a patient with a balanced reciprocal translocation X;15 (18), exons distal to the translocation breakpoint are not expressed. (ii) In the mouse, a deletion of exon 1 on the paternal chromosome results in lack of expression of the *UBE3A* antisense transcript (13). (iii) There

Table 1. RT–PCR products and primer sequences

RT–PCR	RT–PCR primer pairs	Primer sequences (5′–3′ orientation)
RT-1	NG83	GCCCTGCAGAGTCCTGTAGT
	NG84	CACCTGCCAGCACACAGT
RT-2	NG95	GAGGTGGTACCAGTTTAAGAAGTGA
	NG96	AGAATCCAGGACTCCGTGTG
RT-3	AI990296a	TGTTGTGGCCATGGAAGTAA
	AI99026b	GCTGGTTCCCACAATGAACT
RT-4	IPWI	AATTTGGGCATGGTGACTGT
	IPWII	TTTACCGTGTGGCAAATGAA
RT-5	cPCR25F	GTGGCTCTCCATGCCTACCTGTGGT
	cPCR25R	CAAGGCTCAGTGGAAGAGACCAGTGTT
RT-6	cPCR24F	GTGATGGCCACAAAGAGGTGGATTTG
	cPCR25R	CAAGGCTCAGTGGAAGAGACCAGTGTT
RT-7	PE6F	GCCTTGAGCAGCATAGGTGA
	PE8R	CAGGGCAACAAAAGCTCTCT
RT-8[a]	cPCR19F	GACCCCAGAGGAAGACGTGCATT
	cPCR19R.2	AAGGGCTGGGCACCTGACTGATG
RT-9	PE4F	ATGGAAGACCCCTGTCATTG
	PE4R	TCACCTTTGCCAGTCAATCC
RT-10	cPCR20F	GGAAGAATTGCGTTAGGCCCTTTG
	cPCR20R	AGGAAGAGCCTGAGCTTCACCAC
RT-11	cPCR21F	GTGGTGAAGCTCAGGCTCTTCCT
	cPCR21R	CTTCCAGGTCTCCAGCCCAAAATAC
RT-12[a]	cPCR6F	CCTGAGTTGGGTCGATGATGAGA
	cPCR6R	CTCACCACAGCTCAGGGCAGGAG
RT-13	cPCR11F	AAAATGTCCCTCAGCCAGGT
	cPCR11R	CTATACCGGTCAATGCCAAGTG
RT-14	cPCR7F	CCAGTGTCTGTCAGCCAGTTTCC
	cPCR7R	CCCAACAGAAGTCTCACCATCTAGG
RT-15	Ex103F	CTGGTGCACTGAAGCTCAGGCCTT
	Ex106R	CTCAGTGCAAGAGACCAGGGAACCA
RT-16	MRts2-9F	TATGGAAGAAAAGCACTCTTTGG
	MRts2-9R	CAAAGTCTCCCCTTCGTGTT
RT-17	MRts5-6F	GGCACTGAAAATGTGGCATCCAGTC
	MRts5-6R	GGTGTGTCAGCTGTGCTGGTGTCAA
RT-18	MRts2-6F	CACTCTTTGGCCTGTTGTGA
	MRts2-6R	GTGTCAGCTGTGCTGGTGTC
RT-19	MRts8-9F	AAGGCCTGGAATCTGATCCT
	MRts8-9R	CCTAGATTTTAAATAGACAATCCAAAG
RT-20	MRts10–11F	AGAAAAGGCGCAATGAAAGA
	MRts10–11R	TTGGCAAGGAGAGCTTGTCT

[a]Due to the high sequence similarity of the *HBII-52* repeats, primer cPCR19R.2 and cPCR6R did not anneal to their complementary sequence, and RT-8 and RT-12 resulted from mispriming.

are no predicted CpG islands or other putative transcriptional start sites between *SNURF–SNRPN* and *UBE3A* (Buiting *et al.*, unpublished). This also suggests that there is no other gene in between, apart from the snoRNA genes (see below).

Although evidence for a *UBE3A* antisense RNA in humans and mice has been obtained before (12,13), both papers describe only short fragments that are colinear with genomic DNA. It is likely that the human antisense fragment detected

Table 2. Exon–intron boundaries of *SNURF–SNRPN* exons and intron size

SNURF–SNRPN	Size (bp)	Intron	Exon	Intron	Intron size (bp)
Exon 20a	172	tgtactctctcag	GTTCAAATCCAGAGA.....CCTGCTTTTTCGCAA	gtaatt	30385
Exon 21	135	ttccccggagaag	TTGTCATGGGAGGCC.....GGATGGCTTAGGACG	gtaagc	2597
Exon 22	140	ttacacttctcag	AGGCAGTTGCTGTGG.....AGGTGGCTCAGGACG	gtaagc	2515
Exon 23	140	ttacacttctcag	GGGCAGTTATCATGG.....GGATGGCTCAGGATG	gtaagc	2530
Exon 24	140	ttatacttctcag	GGGCAGTTGCCGTGG.....GGATTGCTCAGGATG	gtaagc	162
Exon 25	262	ctatcttccccag	GGTGCCGAAGGTCTT.....AGGCAGTTGCTGTGG	gtaaat	2229
Exon 26	139	ttacacttctcag	GGACAGTTGCCGTGG.....GGGTGGCCCAGGACG	gtaagc	2626
Exon 27	141	ttacacttttcag	GGGCAGTTGCCTTGA.....GGATGGCTCAGGACA	gtaagc	2556
Exon 28	120	cacttctcagcag	CAGTTGTCATGAGAG.....GTGCACCCAATGCTG	gtgagt	2586
Exon 29	140	ttacacttctaag	GGGCAGTTGCCGTGG.....GGATGGCTCAGGACG	gtaagc	2465
Exon 30	139	ttacacttctcag	GGGCAGTTGCCATGG.....GGGTGGCTCAGGACG	gtaagc	1028
Exon 31	159	tgtctttctgaag	CACACTCATTTCCTC.....GGATTTCTCCTGAAT	gtaagt	1684
Exon 32	132	ctgcacttcttag	GTGGCGTTGGTATGG.....GGATACCTCAGGATG	gtaagt	990
Exon 33	134	ctgtacctcccag	GTGGCGTTGGCATGG.....GGATCCGTCGTGAAT	gtaagc	1871
Exon 34	134	ctgcacttcccag	GTGGCATGGGCATGC.....GGATCCCTTCATACG	gtatgt	3593
Exon 35	117	ctgtacttcctag	GTGGTGTGGGCATGG.....CTTCTTCAGGTGCTG	gtaagt	690
Exon 36	132	ctgcacttcccag	GTGGTGTTGGCATGA.....GTATCCCTTCTGAAT	gtaagt	1662
Exon 37	133	ctgaactttccag	GTGGCGTGGGCATAG.....GGATTCCTTAGGATG	gtaagt	1187
Exon 38	107	tcctcctctgcag	GGACAAACACTGTGC.....CCTCGTCGAACTGAG	gtccag	850
Exon 39	133	ctgcacttcacag	GTGGTGTTGGCATGA.....GAAACCCTCCTGAAT	gtaagt	2014
Exon 40	253	tgtttatgaacag	GTGAGGCCAGAGACA.....GGATCCCTCCTGAAT	gtaagt	45
Exon 41	99	tcctcatttgcag	GGGCAAGGACTGGAT.....CCTCGTCGAACTGAG	gtccag	1579
Exon 42	149	ctgcacttcccag	GTGTTGTGGCCATGA.....TGGTGGATCCCACAG	gttggt	2114
Exon 43	141	ctctaattcctag	GTGGTCTGGCATGGA.....GGATCCCTCAGGATG	gtaaat	1652
Exon 44	202	tattttcctgtag	CTTTTCAAGGTTTTT.....CTTCTTTGTGTGCAG	gtattt	1490
Exon 44a	102	tattttcctgtag	CTTTTCAAGGTTTTT.....TCCAGAGAGGCGGAG	gtaaac	1590
Exon 44b	110	ttgtgttttccag	AGAGGCGGAGGTAAA.....CTTCTTTGTGTGCAG	gtattt	1490
Exon 45	203	tattttcctgtag	CTTTTCAAGGTTTTT.....CTTCTTTGTGTGCAG	gtattt	290
Exon 46	119	tctatttttgtag	TTCATTGTGGAACCA.....CAAAATGTGATATTG	gtaagc	1277
Exon 47	112	ctgtactctgtag	ATTGGGTGAGATAGA.....CAATACACTTCTAAG	gtaaca	316
Exon 48	134	tctcttcctacag	GACAGCATAGACCAA.....GGAACTCTAATATTG	gtaagc	1500
Exon 49	112	ctacactttgtag	ATTGGGTGAGATACA.....CAACACACTCCTGAG	gtagta	2969
Exon 50	112	atgtactttgtag	ATTGGTTTTCACAGA.....CAATACTCTCCTGAG	gtaaca	2193
Exon 51a	131	cctattttcaaag	GTTAATGTGGACCAA.....GGAATTCAAATATTG	gtaagc	1175
Exon 51b	59	cctattttcaaag	GTTAATGTGGACCAA.....CCTTCTACTTGAGGT	gtgaca	1247
Exon 52	117	ctgtgccttgcag	AACGGAAGTGATTTT.....GGACGCACTCCAAAG	gtaaca	321
Exon 53	129	actgttttcacag	ATCGGCATGAACCAA.....AGGAATTCTCCCATG	gtaagc	959
Exon 54	106	aattgatgtctag	GTGGACTTTACGGTT.....ACACCTTGTCAGTGA	gtatgt	94
Exon 55	206	ctgtggtttgcag	GTTCCATGTGATACT.....ATGAAACACGCTGAG	gtaaca	151
Exon 56	161	tgttcttttccag	AATTTTTGTGCTTTC.....CAGCAGTACCACCTG	gtgagc	2049
Exon 57	130	ttccccttttatag	ATTGGAAGTGATATT.....ACGATGCACTGCAAG	gttaac	3404
Exon 58	356	tctcgttcctcag	TGTGATTGGTCCAGA.....TACCCCTCAGGACCT	gtaagt	1754
Exon 59	106	attgacttgtcag	GAAGCAAAAGAATGA.....TGGACACCCTTGCAG	gtatgt	752
Exon 60	115	ttttaaacctcag	AAGATGACTTCCTGG.....GACCACCCACTAAAG	gtaaga	682
Exon 61[a]	139	tgtgtccttgcag	GTGATGGCCACAAAG.....TGCCTCCTGCAGATA	agaaat	2274
Exon 62	57	tattggtttacag	TTTTATCTTGCTGGG.....TCTTATTCCCAATAG	gtaagt	49039

Table 2. *Continued*

SNURF–SNRPN	Size (bp)	Intron	Exon	Intron	Intron size (bp)
Exon 63	131	cctgttcccccag	ATGGTGACCACAGAG.....CTGCACTGAGTTGTG	gtgagc	381
Exon 64	45	ctgtgtctttcag	TGAGCTCTTCTGCCC.....CATTGACCGGCATAG	gtgagt	1355
Exon 65	131	cttgtcccccag	ACAGTGAGCCTGGAG.....CTGCACTGAGCTGTG	gtgagc	371
Exon 66a	150	ttgtgtctttcag	TGAGCTCTTCCGCCC.....CCTGAATCCCCACTG	gtgagg	658
Exon 66b	45	ttgtgtctttcag	TGAGCTCTTCCGCCC.....CCTTGAGCAGCATAG	gtgagc	763
Exon 67	146	cctgtctctccag	ATGGTGACCCTGAAG.....CCCCAATAATTCAAC	gtaggt	909
Exon 68	58	atggcttcagcag	GTCCCTCCGTTTGGG.....TTCCACTGAGCCTTG	gtgagc	368
Exon 69	45	ttgtgtctttcag	TGAGCTCTTCCACCA.....CATTGACTGGCATAG	gtgagt	1873
Exon 70	43	ctgtgtctttcag	TGAGCTCTTCCCCCC.....GATTGACGAGCATAG	gtgagt	1350
Exon 71	132	cctgtcccccag	ATGGTGACTGTGGAA.....CTGCCCTAAGCTGTG	gtgagc	382
Exon 72	44	ctgtgtctttcag	CGAGCTCTTCTGCCC.....CCTTAAGTGGGATAG	gtaagt	1201
Exon 73	132	cccacccctccag	ATGGTGACCCCAGAG.....CTGCACTGAGCTGTG	gtgagc	1369
Exon 74	86	atgtgacttgcag	GTCTGCAGCGGCATC.....TGCCCAGTATATCTG	gtatga	123
Exon 75	87	cctgtccttccag	GGTTCGGTGGCTGAG.....ATGGATCCATGGAGG	gtgagc	91
Exon 76	558	cctgtttccccag	ATGTTGAGCCCAGAG.....CATTGACTGGCATAG	gtgagt	1186
Exon 77	135	cctgtcctttcag	GGTTCAGTGGCGGAG.....TGCCAGCATGGAAGG	gtgagg	1367
Exon 78	132	cccgtccctccag	ATGGTGAGCACAGAG.....CTGCACTGAGGTGTG	gtgagt	385
Exon 79	48	ctgtgtctttcag	TGAGCTCTTCCCCCC.....GACTGGCAAAGGTGA	gtggat	1341
Exon 80	131	cctgtcccccag	ATGGTGTCTGTGGAG.....CTGCACTGAGCTGTG	gtgagc	382
Exon 81	44	ctgtgtctttcag	TGACCTCTTTTGCCC.....CATTGACCAGCATAG	gtgagt	1317
Exon 82	132	tcctgtcccccag	ATAGTGAGACTGGAA.....CTGCACGGAGCTGCG	gtgagc	386
Exon 83	44	ctgtgtctttcag	TGAGCTCTTCTACCC.....CCTTGAGCGGCATAG	gtgagt	1455
Exon 84	111	cccacccttccag	ATGGTGACCCCAGAG.....CTGCACTGAGCTGCG	gtgagc	366
Exon 85	46	ttgtgtctttcag	TGAGCTCTTCCACCC.....CATTGACCAGCATAG	gtgagt	1370
Exon 86	559	cctgtttccccag	GTATTGAGCCTGGAG.....CATTGACTAGCATCG	gtgagt	1554
Exon 87	44	ctgtgtctttcag	TGAGCTCTTCTGCCC.....CAGTGACCAGCATAG	gtgagt	2091
Exon 88	132	cccacccctccag	ATGCTGACCCTGGAG.....CTGCACTGAGCTGTG	gtgagc	395
Exon 89	45	ttctgtctttcag	TGAGCTCTTCAGCCC.....CACTGATGCGCATAG	gtgagt	1308
Exon 90	133	cctgtcccccag	ATGGTGAGCCTGGAG.....CTACACTGAGCTGAG	gtgagc	386
Exon 91	46	ttgtgtctttcag	TGAGCTCTTCAACCC.....CATTCACCAGCATAG	gtgagt	1312
Exon 92	131	cctgttcccctag	ACTGTGAAATGGGAG.....CCTGCACTGAGCTTG	gtgagc	385
Exon 93	44	ctgtgtctttcag	TGAGCTCTTTTGCCC.....CATTGATCGGCATAG	gtgagt	1344
Exon 94	131	cctgtacccccag	ATGGCAGCGCTGGAG.....CCTGCACTGAGCTTG	gtgagc	381
Exon 95	749	ttgattctttcag	TGAGCTCTTCTGCTC.....CCTGCACTGAGCTTG	gtgagt	383
Exon 95a	45	ttgattctttcag	TGAGCTCTTCTGCTC.....CATTGACCAGCATAG	gtaagt	573
Exon 95b	131	cctgtcccccag	ATGGTGGCCCTGGAG.....CCTGCACTGAGCTTG	gtgagt	383
Exon 96	45	ctgtgtctttcag	TGAGCTCTTCTGCCC.....TCTTGATGGGCATAG	gtgagt	1344
Exon 97	158	ccccgcattgcag	GGTGAGGCCCTGTTT.....GTCTCCTGCACTGAG	gtgtgg	384
Exon 98	45	ctgtctctttcag	TGAGCTCTTCCGCCC.....CATTGACCGGCATAG	gtgagt	1266
Exon 99	132	cctcctgtcctag	ATGGTGAGCTCGGAG.....CTGCCCTGAGCTGTG	gtgagc	3577
Exon 100	132	cctgtccatccag	ATGGTGAGCCTGAAG.....CTGCACTGAGCTGTG	gtgagc	385
Exon 101	46	ttgtgtctttcag	TGAGCTCTTCTGCCA.....CATTGACCAACATAG	gtgagt	1289
Exon 102	136	cctgtcccccag	ATGGTGAGCCTGGAG.....CTGCACTGAGCTGTG	gtgagt	379
Exon 103	45	ttgtgtctttcag	TGAGTGCTTCTGCCC.....CATTGACCGGTATAG	gtgagt	3194
Exon 104a	130	tctgtccctctag	TGGTGAGCCTGGAGG.....CTGCACTGTGCTGTG	gtgagc	1097
Exon 105	125	cctgttcccctag	ATGGTGAGACTTCTG.....ACTGAGCTGTGATGA	gtacat	363

Table 2. *Continued*

SNURF–SNRPN	Size (bp)	Intron	Exon	Intron	Intron size (bp)
Exon 106	45	ctttgtctttcag	TGAGCTCTTTCACCT.....CATTGACTGGCATAG	gtgagt	3186
Exon 107	132	cctgtccctacag	ATGGTGACCCTGAAG.....TTGCACTGAGCTCTG	gtgagc	768
Exon 108	132	cctgtgcccccag	ATGGTGAGACTTGAG.....CTGAACTGAGCTGTG	gtgagc	384
Exon 109	44	ttgtgtctttcag	TGAGCTCTTCTGCCA.....CATTGACCAACATAG	gtgagt	1396
Exon 110	132	cctgtacccccag	ATGGTGAGCCTGGAG.....CTGCACTGAGCTGTG	gtgagt	382
Exon 111	45	ttgtgtctttcag	TGAGCTCTTCTGCCC.....AATTGATGGGTATAG	gtgagt	1344
Exon 112	132	cctgtctccccag	ATGGTGACCCCGGAG.....CTGCACTGAGCTCTG	gtgagt	1730
Exon 113	132	cctgtccctccag	ATGGTAAGCCTGGAG.....CTGCACTGAGCTGGG	gtaagc	382
Exon 114	41	ttgtgcctttcag	ATTTTCCACCTAGGA.....CATTGACTGACATAG	gtgagt	1335
Exon 115	112	tctgttcctccag	ATGGTGAGACTTCTG.....CTGGACTGAGCTGTG	gtgagc	384
Exon 116	45	ttgtgtctttcag	TGAGCTCTTCTGCCA.....CATTGACCAACATAG	gtgagt	989
Exon 117	132	cctgtacccccag	GTGGTGAGCCTGGAG.....CTGCACTGAGCTGTG	gtgagc	379
Exon 118	45	ctgtgtctttcag	TGAGCTCTTCGGCCC.....CCTTAAGCGGCATAG	gtgagt	1302
Exon 119	146	cctgtccctccag	ATGGTGAGCCTGGAG.....GATGACCATATCCAG	gtcctg	371
Exon 120	43	ttgtgtctttcag	TGAGCTCTTCCAAAC.....CATTGACCGACAGAG	gtgagt	1789
Exon 121	45	ttgtgtctttcag	TGAGCTCTGCCACCC.....CATTGACTGGCATAG	gtgagt	3196
Exon 122	132	cctgtccctgcag	ATGCTGACCCTGGAG.....CAGCACTGAGCTGTG	gtgagc	386
Exon 123	45	ttgtgtctttcag	TGAGCTTTTCCACCC.....CATTGACCGACATAG	gtgagt	1346
Exon 124	132	cctgtgcccccag	ATGATGAGACTGGAG.....CTGCACTAAGCTGTG	gtgagc	381
Exon 125	45	ttgtgtctttcag	TGAGCTCTTCCACCC.....CATTGACTGGCATAG	gtgagt	1310
Exon 126	131	acctctctcccag	ATGGTGAGACGGGAG.....CCTGCACTAGCTGTG	gtgagc	382
Exon 127	45	ccatgtctttcag	TGAGCTCTTCTGCCC.....CCTTACACGACATAG	gtgagt	150
Exon 128	137	aatctgatcctag	GAGCACGCATTTCTT.....CCCAGGCCACATGAG	gtgggc	866
Exon 129	72	cttgtctttccag	GGTTTGGTGGCCGGG.....CATGGATTCCATCAC	gtgggt	79
Exon 130	133	cctgtcaccccag	ATGGTGAGCCTTGAG.....CTGCACTGAGCTGTG	gtgagc	384
Exon 131	45	ttgtgtctttcag	TGAACTTTTCCTCCC.....CATTGACTGGCATAG	gtgagt	1306
Exon 132	132	acctctctcccag	ATGGTGAGACCGGAG.....CTGCACTGAGCTGTG	gtgagc	384
Exon 133	45	ccatgtctttcag	TGAGCTCTTCTGCCC.....CCTTAAGCGGCATAG	gtgagt	150
Exon 134	137	aatctgaccctag	GAGCATGCATTTCGT.....CCCAGGCCACATGAG	gtgggc	2677
Exon 135	131	cctgtcccttcag	ATGGTGAGTCTGGAG.....CTCCTGCACTGAGCT	gtggtg	364
Exon 136	44	ttgtgtctttcag	TGAGCTCTTCCACCC.....CATTGACTGGCATAG	gtgagt	1005
Exon 137	109	cttgtcccaccag	GTGGTGAGCCCGGAG.....CTGCGGACCCTCGCT	gtgagt	14095
Exon 138	184	atctgtcccttag	ATGATGATATGGAAG.....AAATCTTCTGATTTG	gtgaga	297
Exon 139	133	ttgtgtatttcag	TAAGACATGCTGCCA.....GTTTTACACCTTCAG	gtaatc	1284
Exon 140	125	cctatgcctgtag	ATAAAGACTGCTGAG.....AAGGATGCTATTCTG	gtaagg	1151
Exon 141	129	ttgtcttttacag	AAAAGACTGTGGAGG.....AGGAAACCATCTCTG	gtaagc	1684
Exon 142	117	ctcctcccttag	ATAAGGATGACTGAG.....AAGGATGCCACTCTG	gtaggt	5831
Exon 143	43	tattttgctgcag	GTTAAAAGCTGAAAC.....CTTCAGGGAAAAGAG	gtgagc	1437
Exon 144	38	tttgctgagatag	AAGGCCTGGAATCTG.....CTTCAGAGAACAGGG	gtgagt	59162
Exon 145[a]	189	cattcatttccag	GTCAGCTTACTGTAT.....AAAACTCTATCTTAA	aaaaaa	35312
Exon 146[a]	1955	actctgttgccag	GTTGGAATGCAGTGG.....TACCTTTAAAATCAA	aaaaaa	16136
Exon 147	196	tactgctccccag	AGAAAAAAGTACATG.....AGACGGCAACCTGAG	gtaagg	26997
Exon 148[a]	675	atttttttcatcag	GTGGTGGAGTCTATG.....ACATTTCCATACAAA	aaaaaa	
Consensus		Py(10)cag		gtaagt	

[a]Exons 145, 146 and 148 may represent alternative 3′ ends of the transcript or may be due to false priming of an oligo d(T) primer within a larger exon.

by Rougeulle *et al.* (12) is contiguous with exon 144. As described previously (18,29) and in this report, exons 12, 16, 20 and 61 and possibly other exons contain an alternative splice acceptor and donor site. As a consequence, several kilobases of DNA contiguous with these exons are expressed as RNA. The same situation appears to be true for the orthologous locus on mouse chromosome 7. Whereas Chamberlain *et al.* (13) have described an unprocessed *Ube3a* antisense RNA fragment, database searches with genomic DNA sequences from this locus suggest the presence of distinct exons and a processed transcript (unpublished data). Coexistence of a spliced and unspliced form of an antisense transcript has been reported for the mouse *Xist* antisense transcript, *Tsix* (2,30).

It is unclear how the antisense transcript might regulate *UBE3A* expression in *cis*. However, well established mechanisms such as tissue-specific RNA splicing may generate different isoforms of the antisense transcript, which may or may not silence *UBE3A*. This may explain the observation that *UBE3A* is expressed from both alleles in certain tissues and preferentially from the maternal allele in others.

Our findings have important implications for understanding the function of the 15q IC. The data suggest that there is a functional correlation between the regulation and the spatial organization of the paternally expressed genes and the maternally expressed genes. Imprinted expression of the paternally expressed genes, which are located upstream of the IC, is regulated by domain-wide differential DNA methylation, which is set by the IC. Imprinted expression of the maternally expressed genes, which are located downstream of the IC, may be regulated through a paternally expressed antisense transcript, which is initiated at the IC. In this model (Fig. 5B), the *SNURF–SNRPN* gene is the master gene, which has acquired two different mechanisms to control genes located 5′ and 3′ to it. The model predicts that most of the imprinted domain is transcriptionally open on the paternal chromosome and closed on the maternal chromosome. In fact, most of the imprinted domain replicates early at S phase (31,32). We are aware that the replication timing data do not necessarily support the model shown in Figure 5B and that other models are conceivable.

Three C/D box snoRNA genes, *HBII-13*, *HBII-85* and *HBII-52* have been previously mapped to the region between *SNURF–SNRPN* and *UBE3A* (16). Here we report on four additional candidates for C/D box snoRNA genes. Three of the four novel snoRNA genes, *HBII-436*, *HBII-438A* and *HBII-438B*, were found to be expressed in different tissues with predominant expression in brain and are subject to imprinting. *HBII-438A* and *HBII-438B* are identical in sequence, but map ~240 kb apart from each other within the PWS/AS region. Their conserved sequences give additional support for a proposed functional role of these snoRNAs. The fourth novel candidate snoRNA, HBII-437, which was not detected by northern blot analysis, contains a degenerate D-box. It has been shown previously, that the C- and D-box of C/D-box snoRNAs are essential for processing, stability and function of these molecules (33–35). Accordingly, a mutation within the D-box would render the snoRNA instable and subject to rapid degradation by cellular RNases.

In mammals, most snoRNAs known to date are processed from introns of host genes by splicing and subsequent exonucleolytic cleavage (34,35, reviewed in 36,37). The fact, that all but a few of the snoRNA genes lie within introns of the *SNURF–SNRPN* sense/*UBE3A* antisense transcription unit indicates that it serves as the host for these snoRNA genes and that these RNA molecules are indeed processed from this transcription unit. For snoRNA genes *HBII-85* and *HBII-52*, which are present in multiple copies (27 and 47, respectively) in the genome, it is not possible to confirm the expression of single copies. However, since in the case of *HBII-85* the different copies from this gene cluster can be assigned to three distinct groups differing in sequence, we were able to confirm the expression of at least one member of each group by northern blot analysis. Since the expression level of different groups correlated well with their respective copy number (Fig. 4B), it is likely that most, if not all, copies from the *HBII-85* cluster are expressed. Similar data were obtained for the *HBII-52* snoRNA gene cluster (Fig. 4C).

Although one copy of the snoRNA *HBII-85* genes and two of the *HBII-52* genes were found to lie in exons of the transcription unit, we cannot exclude that these copies map inside intronic sequences of yet undetected spliced isoforms of the transcripts as found for other copies of the two snoRNA gene clusters. On the other hand, intron 12, 34, 39 and 99 contain more than one snoRNA gene each. So far, only the presence of a single snoRNA gene per intron has been reported in the literature. Therefore, this suggests that there may be additional, not yet detected exons between the snoRNA copies or alternatively that multiple snoRNAs can be processed from a single intron.

In the case of the *HBII-85* and, less so, the *HBII-52* snoRNA gene clusters, sequences of antisense elements complementary to a potential RNA target differ between the respective copies of one cluster (Fig. 3A and data not shown). Therefore, two possibilities can be envisioned. First, only some of the snoRNA copies may have an actual RNA target, whereas the others are not functional. Alternatively, all copies are functional by being directed against different RNA targets. Currently we are employing bio-computational as well as biological approaches to identify potential targets for these RNAs.

Data from mouse models suggest that a paternal deletion from *Snrpn* to *Ube3a* causes hypotonia, growth retardation and partial lethality (38). This suggests the presence in this interval of a gene or genes involved in PWS. Our data indicate that there is no other gene in this interval except for the snoRNA genes, which are encoded within introns of the *SNURF–SNRPN* sense/*UBE3A* antisense transcription unit. Specifically, *IPW* is not an independent gene, but part of this unit. Furthermore, in this region the snoRNA genes appear to be the only conserved entities between mouse and human (unpublished data; 39). This strengthens the hypothesis that loss of the snoRNAs identified by our groups (16, and this work) contributes to PWS. To elucidate the underlying mechanism, it will be of utmost interest to detect their target molecules.

MATERIALS AND METHODS

RNA preparation

Total RNA from human brain, liver, muscle, lung, kidney, heart and lymphoblastoid cell lines was prepared by the trizol method (Gibco/BRL). For RT–PCR experiments, RNA was treated with DNase I to remove residual traces of genomic DNA.

RT–PCR

RT–PCRs were performed with the GeneAmp RNA PCR Kit (Perkin Elmer). Total RNA from brain or lymphoblastoid cell lines (1 μg) was reverse transcribed using random hexamers. The cDNA products were amplified by 35 cycles of PCR.

To determine the expression status of selected gene fragments, we performed RT–PCR with primers which anneal to exons 42 and 46 (RT-3), and primers which anneal to exons 141 and 142 (RT-17; Table 1). To check the integrity of the RNA, we used primers for β-*actin* (40). In control experiments, reverse transcriptase or cDNA was omitted from the reactions.

Exon-connection PCR was performed on Marathon ready cDNA of human fetal brain or on reverse transcribed human fetal brain RNA (Clontech). PCR products were verified by sequencing with PCR primers. In the case of multiple PCR products, these product were subcloned and sequenced with vector-specific primers. The subset of RT–PCR primers used to amplify a minimal contig of the transcription unit are shown in Table 1.

Northern blot analysis

For northern blot analysis of the novel exons, human multiple tissue northern blots (Clontech) were hybridized according to the manufacturer's protocol. The final wash was in 2× SSC, 0.1% SDS at 50°C for 10 min. As probes, we used the 501 bp RT–PCR product RT-3 and the 608 bp RT–PCR product RT-18 (Table 1). To determine the expression and imprinting status of the novel snoRNAs, total RNA was separated on an 8% denaturing polyacrylamide gel (7 M urea, 1× TBE buffer) and transferred onto a nylon membrane (Qiabrane Nylon Plus; Qiagen) using the Bio-Rad semi-dry blotting apparatus (Trans-blot SD; Bio-Rad). After immobilizing of RNAs by the Stratagene crosslinker, the nylon membrane was pre-hybridized for 15 min in 1 M sodium phosphate buffer pH 6.2, 7% SDS. Oligonucleotides complementary to the respective RNA species were end-labeled with ^{32}P-ATP and T4 polynucleotide kinase; hybridization was carried out at 58°C in 1 M sodium phosphate buffer pH 6.2, 7% SDS for 12 h. A final wash was carried out twice at room temperature for 15 min in 2× SSPE buffer (20 mM sodium phosphate pH 7.4, 0.3 M NaCl, 2 mM EDTA), 0.1% SDS and subsequently for 1 min at 58°C in 0.1× SSPE, 0.5% SDS. Membranes were exposed to Kodak MS-1 film and autoradiographed for 12 h. Oligonucleotides used as probes for the novel snoRNAs are as follows: 5′-AAATCATTATGTTCAGACAAGGTCCT-3′ (*HB-436*), 5′-ACTCCAGCAAATTACTTTGATCATGA-3′, 5′-TCACGCTCCCTTTGCAGGAATGG-3′ and 5′-TTTGCAGGAATGGAAAGTGTCATCC-3′ (*HB-437*), CAGATTGACATCTGGAATGAGTCCCTC-5′ (*HB-438*). For expression studies of the different *HBII-85* and *HBII-52* gene copies the following oligonucleotides were used: 5′-AAAACTCTATACCGTCATCCTCGTC-3′ (*HBII-85* gene copies 3, 5, 6, 7 and 8), 5′-GAACTCATACCGTCGTTCTCATCG-3′ (*HBII-85* gene copies 14–21), 5′-CCAAATCACTTCTGTGCCACTTCTG-3′ (*HBII-85* gene copy 24), 5′-CATTCTCAAAAGGATTATGC-3′ (*HBII-52* gene copies 17–19) and 5′-CATGCTCAATAGGATTACGC-3′ (32 gene copies of HBII-52). Quantification of hybridization signals for HBII-85 and HBII-52 snoRNAs was performed by analysis of northern blots on a Fujix BAS 1000 phosphoimager (Fuji) using the Mac BAS V1.0 program.

cDNA clones

EST clones were provided to us from the Resource Centre of the Human Genome Project (RZPD), Berlin. A total of 10^6 phages from a size-selected poly(A)-tailed human adult kidney library (courtesy of L.Schomburg, Hannover, Germany) were screened using probes RT-3 and RT-18. Positive plaques were purified in two rounds. Insert sequences of cDNA clones were determined with vector and sequence-specific primers.

Sequence analysis

RT–PCR products were purified with Microcon-100 microconcentrators (Amicon) or exctracted from agarose gels (Minielute; QIAgen). Sequencing reactions on cDNA clones and RT–PCR products were performed with fluorescence-tagged dideoxynucleotides (BIGDye kit) and the *Taq cycle* sequencing procedure (ABI). Sequences were analyzed on an ABI 3100 DNA Sequencer.

Data deposition

The sequences of a minimal contig of the transcription unit represented by RT–PCR products RT-2I, RT-5, RT-6I, RT-7, RT-8, RT-9, RT-10, RT-11, RT-12, RT-13I, RT-13II, RT-15, RT-16, complete sequence of EST clone R19540 and cDNA clones kid4, kid12, kid16 and kid17 have been deposited in the GenBank database (accession nos AF400485, AF400489, AF400490, AF400491, AF400492, AF400493, AF400494, AF400495, AF400496, AF400497, AF400498, AF400499, AF400500, AF400502, AF400486, AF400487, AF400488 and AF400501, respectively). The GenBank accession nos for the newly identified snoRNAs are AY055806 (HBII-436), AY055807 (HBII-437) and AY055808 (HBII-438A/B).

ACKNOWLEDGEMENTS

We would like to thank Robert Nicholls for helpful discussions and critical reading of the manuscript, Jürgen Brosius for his steady encouragement and interest in this work, Christina Lich for expert technical assistance, Amy Lossie, Daniel Driscoll and Francoise Muscatelli for unpublished data, Jörn Walter and Sabine Engemann for genomic sequencing and Marc Lalande for brain RNA from PWS and AS patients. This work was supported by the Deutsche Forschungsgemeinschaft (Ho949/12-3) to B.H. and by the German Human Genome Project through the BMBF (no. 01KW9966) and an IZKF grant (Teilprojekt IZKF3 G6, Münster) to A.H.

NOTE ADDED IN PROOF

After submission of the manuscript we have identified additional splice variants of the IC-*SNURF–SNRPN* transcript. In one variant, 383 bp of exon 76 including the snoRNA gene *HBII-52-7* are spliced out. In another variant, a novel exon between the genes for *HBII-52-21* and *HBII-52-22* was detected. Thus, there are more exons than previously found.

REFERENCES

1. Reik, W. and Walter, J. (2001) Evolution of imprinting mechanisms: the battle of the sexes begins in the zygote. *Nat. Genet.*, **27**, 255–256.
2. Sado, T., Wang, Z., Sasaki, H. and Li, E. (2001) Regulation of imprinted X-chromosome inactivation in mice by *Tsix*. *Development*, **128**, 1275–1286.
3. Nicholls, R.D. and Knepper, J.L. (2001) Genome organization: function and imprinting in Prader–Willi and Angelman syndromes. *Annu. Rev. Genomics Hum. Genet.*, **2**: 153–175
4. 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. *Nat. Genet.*, **8**, 52–58.
5. 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. *Nat. Genet.*, **9**, 395–400.
6. Albrecht, U., Sutcliffe, J.S., Cattanach, B.M., Beechey, C.V., Armstrong, D., Eichele, G. and Beaudet, A.L. (1997) Imprinted expression of the murine Angelman syndrome gene, *Ube3a*, in hippocampal and Purkinje neurons. *Nat. Genet.*, **17**, 75–78.
7. Rougeulle, C., Glatt, H. and Lalande, M. (1997) The Angelman syndrome candidate gene, *UBE3A/E6-AP*, is imprinted in brain. *Nat. Genet.*, **17**, 14–15.
8. Meguro, M., Kashiwagi, A., Mitsuya, K., Nakao, M., Kondo, I., Saitoh, S. and Oshimura, M. (2001) A novel maternally expressed gene, *ATP10C*, encodes a putative aminophospholipid translocase associated with Angelman syndrome. *Nat. Genet.*, **28**, 19–20.
9. Herzing, L.B.K., Kim, S.-J., Cook, E.H. and Ledbetter, D.H. (2001) The human aminophospholipid-transporting ATPase gene *ATP10C* maps adjacent to *UBE3A* and exhibits similar imprinted expression. *Am. J. Hum. Genet.*, **68**, 1501–1505.
10. Meguro, M., Mitsuya, K., Nomura, N., Masakazu, K., Kashiwagi, A., Nishigaki, R., Yoshioka, H., Mitsuyoshi, N., Oishi, M. and Oshimura, M. (2001) Large-scale evaluation of imprinting status in the Prader–Willi syndrome region: an imprinted direct repeat cluster resembling small nucleolar RNA genes. *Hum. Mol. Genet.*, **10**, 383–394.
11. Lossie, A.C., Whitney, M.M., Amidon, D., Dong, H.J., Chen, P., Theriaque, D., Hutson, A., Nicholls, R.D., Zori, R.T., Williams, C.A. and Driscoll, D.J. (2001) Distinct phenotypes distinguish the molecular classes of Angelman syndrome. *J. Med. Genet.*, in press.
12. Rougeulle, C., Cardoso, C., Fontés, M., Colleaux, L. and Lalande, M. (1998) An imprinted antisense RNA overlaps *UBE3A* and a second maternally expressed transcript. *Nat. Genet.*, **19**, 15 –16.
13. Chamberlain, S.J. and Brannan, C.I. (2001) The Prader-Willi syndrome imprinting center activates the paternally expressed murine *Ube3a* antisense transcript but represses paternal *Ube3a*. *Genomics*, **73**, 316–322.
14. Kishino, T., Lalande, M. and Wagstaff, J. (1997) *UBE3A/E6-AP* mutations cause Angelman syndrome. *Nat. Genet.*, **15**, 70–73.
15. Matsuura, T., Sutcliffe, J.S., Fang, P., Galjaard, R.J., Jiang, Y.H., Benton, C.S., Rommens, J.M. and Beaudet, A.L. (1997) *De novo* truncating mutations in E6-AP ubiquitin-protein ligase gene (*UBE3A*) in Angelman syndrome. *Nat. Genet.*, **15**, 74–77.
16. Cavaillé, J., Buiting, K., Kiefmann, M., Lalande, M., Brannan, C.I., Horsthemke, B., Bachellerie, J.-P., Brosius, J. and Hüttenhofer, A. (2000) Identification of brain-specific and imprinted small nucleolar RNA genes exhibiting an unusual genomic organization. *Proc. Natl Acad. Sci. USA*, **97**, 14311–14316.
17. de los Santos, T., Schweizer, J., Rees, C.A. and Francke, U. (2000) Small evolutionarily conserved RNA, resembling C/D box small nucleolar RNA, is transcribed from PWCR1, a novel imprinted gene in the Prader–Willi deletion region, which is highly expressed in brain. *Am. J. Hum. Genet.*, **67**, 1067–1082.
18. Wirth, J., Back, E., Hüttenhofer, A., Nothwang, H.-G., Lich, C., Groß, S., Menzel, C., Schinzel, A., Kioschis, P., Tommerup, N. *et al.* (2001) A translocation breakpoint cluster disrupts the newly defined 3′ end of the *SNURF–SNRPN* transcription unit on chromosome 15. *Hum. Mol. Genet.*, **10**, 201–210.
19. Smith, C.M. and Steitz, J.A. (1997) Sno storm in the nucleolus: new roles for myriad small RNPs. *Cell*, **89**, 669–672.
20. Bachellerie, J.P., and Cavaillé, J. (1998) Small nucleolar RNAs guide the ribose methylations of eukaryotic rRNAs. In Grosjean, H. and Benne, R. (eds), *Modification and Editing of RNA: The Alteration of RNA Structure and Function*. ASM Press, Washington, pp. 255–272.
21. Gray, T.A., Saitoh, S. and Nicholls, R.D. (1999) An imprinted, mammalian bicistronic transcript encodes two independent proteins. *Proc. Natl Acad. Sci. USA*, **96**, 5616–5621.
22. Özcelik, T., Leff, S., Robinson, W., 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. *Nat. Genet.*, **2**, 265–269.
23. 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. *Nat. Genet.*, **14**, 163–170.
24. Färber, C., Dittrich, B., Buiting, K. and Horsthemke, B. (1999) The chromosome 15 imprinting centre (IC) region has undergone multiple duplication events and contains an upstream exon of *SNRPN* that is deleted in all Angelman syndrome patients with an IC microdeletion. *Hum. Mol. Genet.*, **8**, 337–343.
25. Ohta, T., Buiting, K., Kokkonen, H., McCandless, S., Heeger, S., Leisti, H., Driscoll, D.J., Cassidy, S.B., Horsthemke, B. and Nicholls, R.D. (1999) Molecular mechanism of Angelman syndrome in two large families involves an imprinting mutation. *Am. J. Hum. Genet.*, **64**, 385–396.
26. Ohta, T., Gray, T.A., Rogan, P.K., Buiting, K., Gabriel, J.M., Saitoh, S., Muralidhar, B., Bilienska, B., Krajewska-Walasek, M., Driscoll, D.J. *et al*. (1999) Imprinting-mutation mechanisms in Prader–Willi syndrome. *Am. J. Hum. Genet.*, **64**, 397–413.
27. Bielinska, B., Blaydes, S.M., Buiting, K., Yang, T., Krajewska-Walasek, M., Horsthemke, B. and Brannan, C.I. (2000) *De novo* deletions of *SNRPN* exon 1 in early human and mouse embryos result in a paternal to maternal imprint switch. *Nat. Genet.*, **25**, 74–78.
28. El-Maarri, O., Buiting, K., Peery, E.G., Kroisel, P.M., Balaban, B., Wagner, K., Urman, B., Heyd, J., Lich, C., Brannan, C.I. *et al*. (2001) Maternal methylation imprints on human chromosome 15 are established during or after fertilization. *Nat. Genet.*, **27**, 341–344.
29. Buiting, K., Dittrich, B., Endele, S. and Horsthemke, B. (1997) Identification of novel exons 3′ to the human *SNRPN* gene. *Genomics*, **40**, 132–137.
30. Lee, J.T., Davidow, L.S. and Warshawsky, D. (1999) Tsix, a gene antisense to Xist at the X-inactivation centre. *Nat. Genet.*, **21**, 400–404.
31. Kitsberg, D., Selig, S., Brandeis, M., Simon, I., Keshet, I., Driscoll, D.J., Nicholls, R.D. and Cedar, H. (1993) Allele-specific replication timing of imprinted gene regions. *Nature*, **364**, 459–463.
32. Knoll, J.H., Cheng, S.D. and Lalande, M. (1994) Allele specificity of DNA replication timing in the Angelman/Prader–Willi syndrome imprinted chromosomal region. *Nat. Genet.*, **6**, 41–46.
33. Huang, G.M., Jarmolowski, J., Struck, C.R. and Fournier, M.J. (1992) Accumulation of U14 small nuclear RNA in *Saccharomyces cerevisiae* requires box C, box D and a 5′, 3′ terminal stem. *Mol. Cell. Biol.*, **12**, 4456–4463.
34. Cavaillé, J. and Bachellerie, J.-P. (1996) Processing of fibrillarin associated snoRNAs from pre-mRNA introns: an exonucleolytic process exclusively directed by the common stem-box terminal structure. *Biochimie*, **78**, 443–456.
35. Cafarelli, E., Fatica, A., Prislei, S., De Gregorio, E., Fragapane, P. and Bozzoni, I. (1996) Processing of the intron encoded U16 and U18 snoRNAs: the conserved C and D boxes control both the processing and the stability of the mature snoRNA. *EMBO J.*, **15**, 1121–1131.
36. Tollervey, D. and Kiss, T. (1997) Function and synthesis of small nucleolar RNAs. *Curr. Opin. Cell. Biol.*, **9**, 337–342.
37. Smith, C.M. and Steitz, J.A. (1998) Classification of gas5 as a multi-small-nucleolar-RNA (snoRNA) host gene and a member of the 5′-terminal oligopyrimidine gene family reveals common features of snoRNA host genes. *Mol. Cell. Biol.*, **18**, 6897–6909.
38. Tsai, T.F., Jiang, Y.H., Bressler, J., Armstrong, D. and Beaudet, A.L. (1999) Paternal deletion from *Snrpn* to *Ube3a* in the mouse causes hypotonia, growth retardation and partial lethality and provides evidence for a gene contributing to Prader–Willi syndrome. *Hum. Mol. Genet.*, **8**, 1357–1364.
39. Hüttenhofer, A., Kiefmann, M., Meier-Ewert. S., Lehrach, H., Bachellerie, J.P. and Brosius, J. (2001) RNomics: an experimental approach that identifies 201 candidates for novel, small, non-messenger RNAs in mouse. *EMBO J.*, **20**, 2943–2953.
40. Buiting, K., Barnicoat, A., Lich, C., Pembrey, M., Malcolm, S. and Horsthemke, B. (2001) Disruption of the bipartite imprinting center in a family with Angelman syndrome. *Am. J. Hum. Genet.*, **68**, 1290–1294.