 Human Molecular Genetics, 2002, Vol. 11, No. 13 1527–1538

Identification of tandemly-repeated C/D snoRNA genes at the imprinted human 14q32 domain reminiscent of those at the Prader–Willi/Angelman syndrome region

Jérôme Cavaillé[1,*], Hervé Seitz[1], Martina Paulsen[2], Anne C. Ferguson-Smith[3] and Jean-Pierre Bachellerie[1]

[1]LBME-CNRS (UMR5099), Université P. Sabatier, 118 Route de Narbonne, 31062 Toulouse Cedex, France, [2]Universität des Saarlandes, FR Genetik, Postfach 151150, D-66041 Saarbrücken, Germany and [3]Department of Anatomy, University of Cambridge, Downing Street, Cambridge CB2 3DY, UK

Received February 18, 2002; Revised and Accepted April 17, 2002

A human imprinted domain at 14q32 contains two co-expressed and reciprocally imprinted genes, *DLK1* and *GTL2*, which are expressed from the paternally and maternally inherited alleles, respectively. We have previously shown that another imprinted locus, on human 15q11–q13, contains a large number of tandemly repeated C/D small nucleolar RNA genes (or C/D snoRNAs) only expressed from the paternal allele. Here we show that the region downstream from the *GTL2* gene is also characterized by a transcription unit spanning many repeated intron-encoded C/D snoRNA genes, most of them arranged into two tandem arrays of 31 and 9 copies. Intriguingly, these snoRNAs depart from previously reported rRNA or snRNA methylation guides by their tissue-specific expression and by their lack of complementarity to rRNA or snRNA within their sequences. Analysis of the orthologous region in the mouse shows that the previously reported maternally expressed *Rian* gene, located downstream of *Gtl2* on the distal 12 chromosome, encodes at least nine C/D snoRNAs. Through a systematic search in rodents, we could identify other C/D snoRNA genes in this domain. All snoRNAs identified on mouse distal 12 are brain-specific and only expressed from the maternally inherited allele. The human imprinted 14q32 domain therefore shares common genomic features with the imprinted 15q11–q13 loci. This link between tandemly repeated C/D snoRNA genes and genomic imprinting suggests a role for these snoRNAs and/or their host non-coding RNA genes in the evolution and/or mechanism of the epigenetic imprinting process.

INTRODUCTION

A subset of mammalian genes undergo genomic imprinting, an epigenetic phenomenon that determines gene expression (or repression) according to parental origin (only the allele inherited from the father or that from the mother is expressed). Imprinted genes are often clustered, and multiple regulatory gene expression networks are predicted to operate within the same imprinted locus, involving allele-specific methylation, chromatin modification, antisense RNAs, chromatin boundaries and silencers (1–4). However, if common genomic features conferring local epigenetic control exist within imprinted domains, they have not yet been identified.

Among the most extensively studied imprinted domains are the insulin-like growth factor-2 (IGF2)/H19 domain and the PWS/AS region involved in the Beckwith–Wiedemann and Prader–Willi/Angelman syndromes, respectively. The IGF2/H19 domain is located at the human 11p15.5 locus (mouse chromosome 7F) and possesses two co-expressed and reciprocally imprinted genes located 70 kb apart at one end of a larger cluster of imprinted genes. In IGF2/H19, the parental-specific gene expression is mediated by a germline-specific differentially methylated region (DMR) located between *IGF2* and *H19* that functions as an insulator element upon the binding of CTCF, a methylation-sensitive DNA-binding protein (5–8). The Prader–Willi/Angelman syndrome chromosomal region at human 15q11q13 (mouse chromosome 7C) is characterized by another gene cluster containing several paternally expressed genes (the bicistronic *SNURF–SNRPN* locus and the *MKRN3*, *NDN* and *MAGEL2* genes), while two genes, *UBE3A* and

*To whom correspondence should be addressed. Tel: + 33 5 61335934; Fax: + 33 5 61335886; Email: cavaille@ibcg.biotoul.fr
Correspondence may also be addressed to J.-P. Bachellerie.

ATP10C, are expressed from the maternally inherited alleles (9). Imprinted expression within this large (approximately 2–3 Mb) domain is coordinated by a bipartite *cis*-acting element encompassing a DMR and located upstream from the paternally expressed *SNURF–SNRPN* gene (10–13).

We and others have recently shown that the Prader–Willi/Angelman syndrome chromosomal region and its counterpart in mouse also encode multiple, tandemly repeated C/D boxes containing small nucleolar RNA genes (or C/D snoRNAs) expressed only from the paternal chromosome (14–17). C/D snoRNAs guide the formation of 2′-*O* methylations of both rRNA and U snRNAs through a specific RNA duplex at each modification site (18–20). However, the C/D snoRNAs at PWS lack a canonical antisense segment against rRNA and/or U snRNAs, suggesting that they can have other functions or modify other cellular RNAs such as mRNA (14). Remarkably, Runte and colleagues (21) have recently shown that all the human PWS-encoded snoRNAs are processed from a single paternally expressed transcription unit. It starts at the IC and overlaps the maternally expressed gene *UBE3A* in an antisense direction. Deletion of the PWS-IC on the paternal allele in mouse results in the loss of *Ube3a* antisense expression and

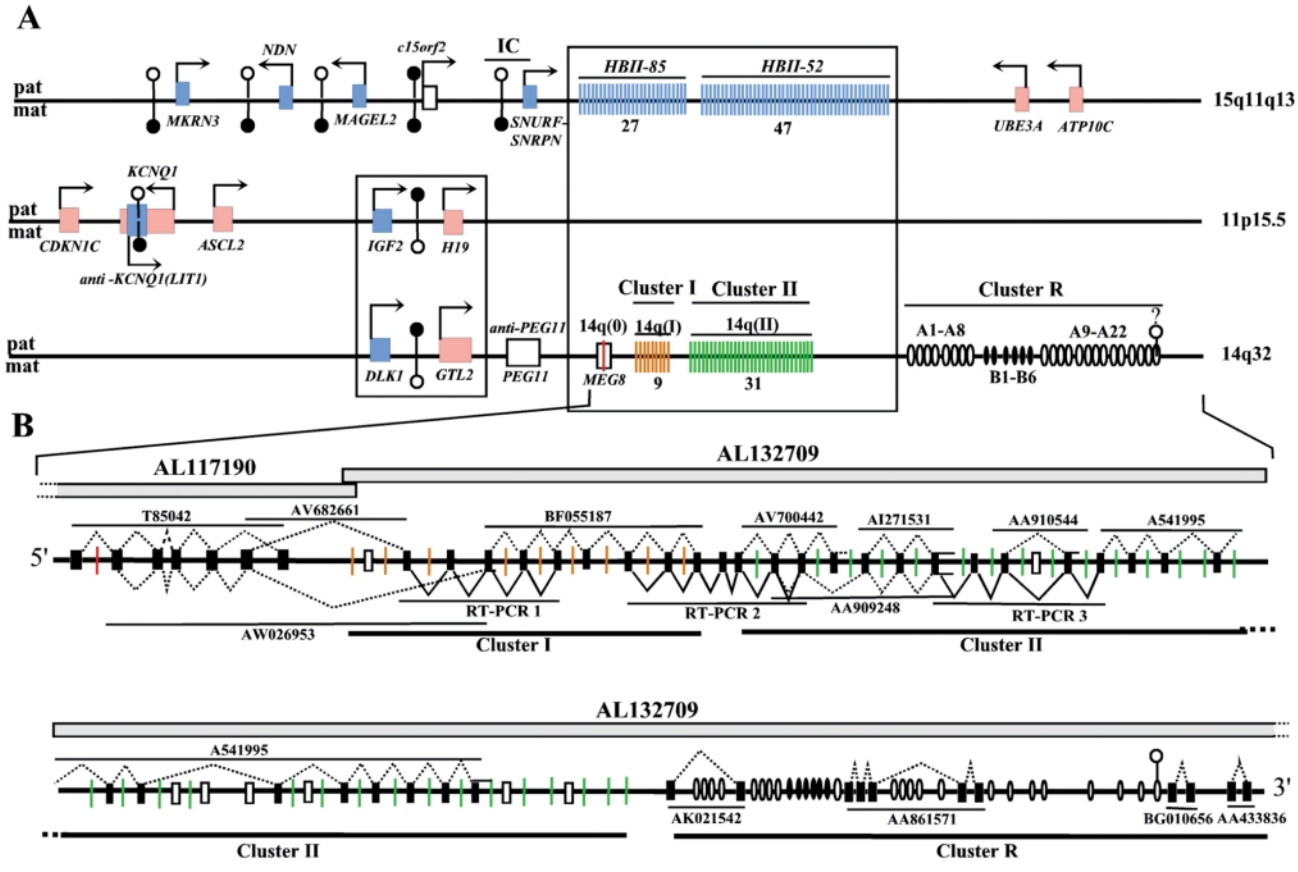

Figure 1. Identification of tandemly repeated C/D snoRNA genes at the human 14q32 imprinted locus. **(A)** Common imprinting features identified at the three human imprinted loci at 11p15.5, 14q32 and 15q11q13. The position and direction of transcription of the several imprinted genes within each human imprinted locus are shown by squares and arrows, respectively. Blue genes are paternally expressed while pink ones are expressed from the maternal allele only. The imprinting status of white squares remains to be checked in human. The C/D snoRNA genes at 14q32 and at 15q11q13 are depicted by vertical bars [note that HBII-13 (14) and HBII-436, HBII-437 and HBII-438A/B snoRNAs (21) are not indicated], with the number of snoRNA genes being indicated below. Several differentially methylated regions characterized within each domain loci are represented by circles (filled, methylated regions; open, unmethylated regions). IC, imprinting center; mat and pat, maternally and paternally inherited chromosomes. Repeated sequences of A-type (22 copies) or B-type (6 copies) within the repeat cluster (cluster R) are represented by open and filled ovals, respectively. Common structural gene organization features between these imprinted loci are boxed. **(B)** C/D snoRNAs at human 14q32 are intron-encoded and processed from a single transcript. Schematic representation of the genomic region spanning the three clusters of repeated sequences (derived from AL117190 and AL132709 clones). C/D snoRNA genes are indicated by vertical bars [red, 14q(0) snoRNA; brown, 14q(I) snoRNAs; green, 14q(II) snoRNAs], while exons detected by genomic/ESTs comparison or predicted by *in silico* analysis are represented by black or white boxes, respectively. Dotted and filled lines denote splicing events identified by ESTs analysis or RT–PCR experiments, respectively. Nucleotide positions of exons within AL117190: (?–134517); (138432–138562); (145574–145640); (146416–146610); (149709–149772); (152908–152998); (154149–?); within AL132709 (reverse complementary strand): *(5510–5582);* (7041–7109); (12940–12993); (14714–14827); (16345–16430); (17822–17898); (19307–19392); (21704–21780); (23751–23814); (25976–26062); (27920–27979); (28466–28538); (29344–29414); (30374–30436); (31363–?); (32906–32986); (33864–33948); (34827–34891); (37204–37285) (43295–43381); *(44257–44334)*; (45587–45677); (46581–46641); (47590–47666); (48401–48487); (51206–51288); (53068–53145); (54353–54438); *(55008–55094); (55897–55983);(56547–56636)*; (59519–59605); *(60516–60602)*; (61446–61527); (62403–62484); (63290–63376); (64597–64682); (65561–65648); *(66486–66572)*; (*68636–68714)*; (?–89220); (106288–106357); (?–123635); (123746–123858); (123948–124126); (130009–130211); (130544–?); (?–148997); (149384–?); (?–150933); (151301–?). Nucleotide positions of predicted exons are indicated in italics. Other symbols are as in (A). The cartoons are not drawn to scale.

activation of the paternal *Ube3a* allele (22). Although an RNA-independent chromatin mechanism might be involved in this regulation, this observation raises the possibility of a role for this snoRNA host gene in the silencing of paternal *UBE3A* gene expression (21–23).

Indeed, there are several examples of imprinted, protein-coding genes being physically linked to non-coding RNAs and exhibiting a reciprocally imprinted expression pattern. Transcribed from the other parental chromosome, these mono-allelically expressed, non-coding RNAs are often transcribed as antisense RNAs relative to the sense, protein-coding genes: *Igf2r/Air* (24), *KVLQT/LIT1* (25), *UBE3A/UBE3A-AS* (21,23) and *Nesp/Nesp-AS* (26). Reciprocally imprinted non-coding RNAs can also be transcribed in the same orientation as the protein-encoding gene, as described for both the Igf2/H19 and Dlk1/Gtl2 imprinted domains (3). In most cases, the function of these RNAs remains to be established, but the available evidence suggests that this particular gene organization relative to regional controlling elements is crucial in the regulation of the two pairs of reciprocally imprinted genes (3,27–29).

Like the two better-characterized clusters of imprinted genes at chromosomes 11p15.5 and 15q11q13, recent data indicate that the *DLK1–GTL2* region on chromosome 14q32/mouse distal 12 may also be part of a larger cluster of imprinted genes. Although still under investigation, the gene organization at this locus has been studied in the sheep and human, where two genes downstream of *GTL2*, namely *PEG11* and *MEG8*, have been described and shown in the sheep to exhibit paternal-specific and maternal-specific allelic expression, respectively (30,31). In the mouse, *Rian*, a brain-specific non-coding RNA expressed from the maternal allele, has also been mapped to this region, although its precise location relative to *Gtl2* has not been determined (32). Comparative genomic analysis of imprinted domains between or within species can identify conserved genomic features that may be functionally important in imprinting control (31,33–36). Using this approach, we describe here the identification of tandemly repeated C/D snoRNA genes located at the human 14q32 locus, adjacent to the imprinted *DLK1/GTL2* domain. Most of these are intron-encoded and processed from a complex transcription unit mapping to the *MEG8* gene. Furthermore, we have identified other putative C/D snoRNA genes at orthologous loci in rodents. All of these C/D snoRNAs are brain-specific and expressed from the maternal allele only. Thus, the human imprinted 14q32 domain shares common genomic features with the Prader–Willi/Angelman syndrome chromosomal region (Fig. 1A).

RESULTS

Searching for tandemly repeated C/D snoRNA genes at human 14q32

We have recently identified a rat C/D snoRNA, RBII-36, encoded in the *Bsr* gene (37). The gene organization and tissue-specific expression pattern of RBII-36 are highly reminiscent of those of imprinted C/D snoRNAs at human 15q11–13/mouse 7C loci (14,38). The RBII-36 snoRNA gene maps to 6q31q32, a locus that is orthologous to mouse distal chromosome 12 and human 14q32, which harbour imprinted genes. Among the most well-characterized imprinted genes at this locus are *Dlk1*, a paternally expressed gene encoding a cell-surface membrane protein of the Notch–Delta family, and *Gtl2/Meg3*, a maternally expressed gene located 80 kb downstream and lacking conserved open reading frames (39,40). To identify putative human homologues of RBII-36 snoRNA genes, we searched the human 14q32 imprinted locus for the presence of direct tandemly repeated C/D snoRNAs. Three direct-repeat clusters (two snoRNA clusters I and II and one cluster of direct repeats, cluster R) have been detected in BAC sequences AL132709 (Fig. 1B). Cluster I (nucleotide position ~3700–24700) and cluster II (nucleotide position ~28750–72260), are located approximately 40 kb downstream from *PEG11* and contain 9 and 31 copies, respectively, of 70–75 nt-long related snoRNA-like sequences (Fig. 2A). Copies of these snoRNA-like sequences contain the characteristic C/D′/C′/D motifs (in that order) and also a terminal 5′–3′ stem structure exhibiting the conserved sequence 5′-GGACC . . . GGTCC-3′. Despite a relatively high overall similarity of the snoRNA gene copies within each cluster, their sequence tracts upstream from either box D or box D′ exhibit a substantial number of nucleotide differences among copies, as exemplified for snoRNA copies within cluster I, and do not represent potential antisense elements against rRNA or U snRNAs (Fig. 2A). Moreover, in agreement with our failure to detect specific RBII-36 homologues in human (37), none of these sequences shows significant similarities with rat RBII-36 snoRNA. These sequences therefore represent novel human C/D snoRNA genes, and we propose to name them 14q(I) and 14q(II) for snoRNAs encoded in clusters I and II, respectively. By focusing on sequences conserved between human and ovine *MEG8*, we were able to identify a third intron-encoded C/D snoRNA gene present as a single copy and termed 14q(0) since it does not belong to a cluster (Fig. 1). Again, this C/D snoRNA, conserved between human and sheep (Fig. 2A), does not contain any obvious antisense element against rRNAs or U snRNAs. Further analysis downstream of the snoRNAs (nucleotide position ~102260–144600) resulted in identification of two classes of ~80–100 nt-long direct repeats (22 and 6 copies of A- and B-type, respectively) that do not exhibit snoRNA hallmarks (Figs 1 and 2B). All are in head-to-tail orientation, with a relatively irregular spacing between A repeats, while five of six B repeats are clustered and separated from each other by a 350–380 nt-long spacer sequence. Interestingly, one of these repeats, A22, is embedded within a conserved CpG island at the end of cluster R (Fig. 1B).

The C/D snoRNAs are intron-encoded and processed from the tissue-specific non-coding human *MEG8* RNA

To learn more about the snoRNA gene organization and mode of biosynthesis, GenBank was searched for expressed sequence tags (ESTs) covering the snoRNA gene clusters. We identified several human spliced ESTs and putative exons overlapping this region, all of which appear to be devoid of significant protein coding potential (Fig. 1B). Comparisons of genomic and cDNA sequences and RT–PCR analysis revealed that most of the tandemly repeated C/D snoRNA genes and also several A and B repeats are positioned within introns (Fig. 1B).

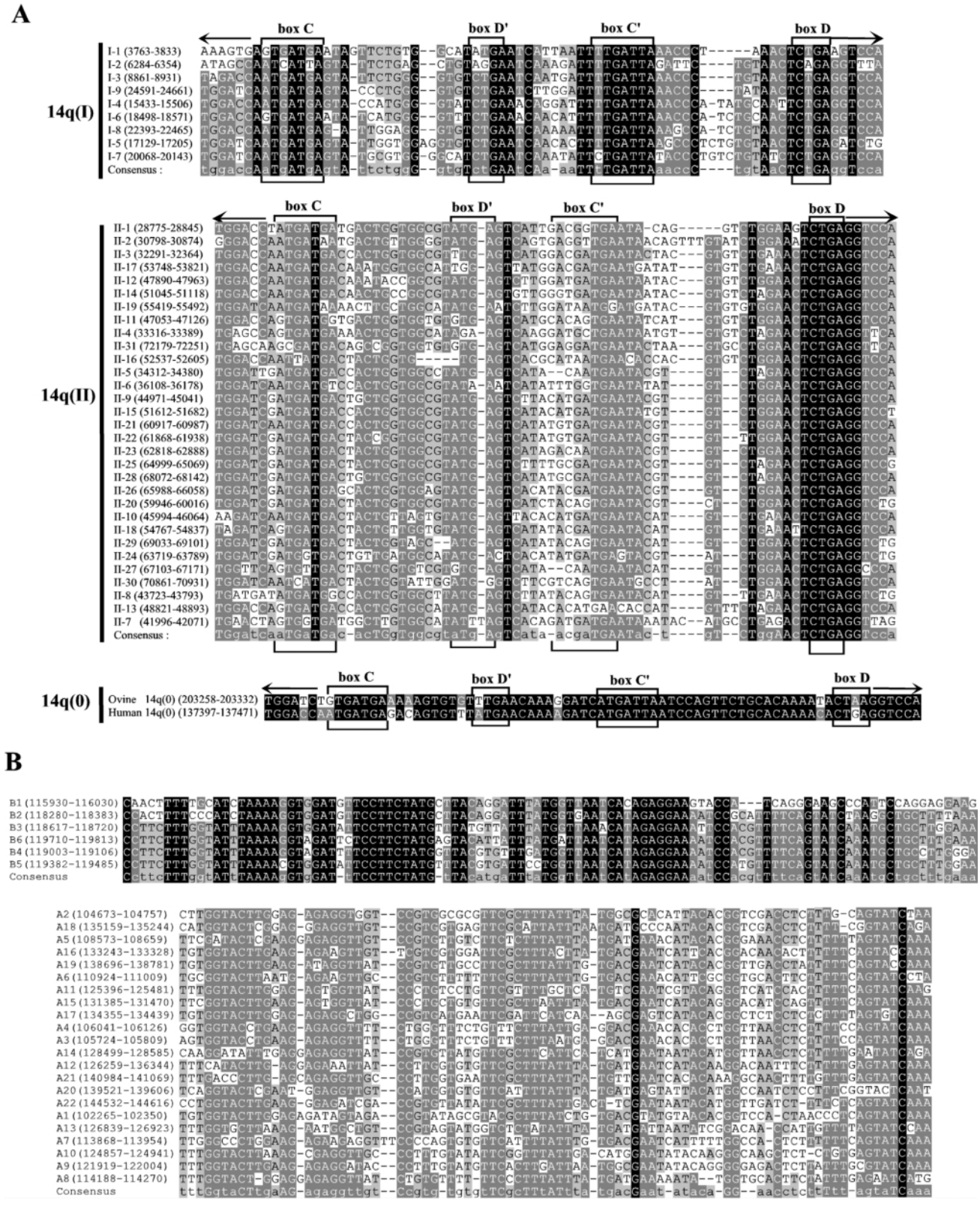

Figure 2. Sequence alignment of the repeated human C/D snoRNA genes and A and B repeats. **(A)** Sequence alignment of human 14q(I) and 14q(II) snoRNA gene copies within clusters I and II, respectively, and of 14q(0) human/ovine snoRNA genes. Nucleotide positions of snoRNA genes within BAC AL117190 [14q(0)]; AL132709 [14q(I) and 14q(II)] and AF354168 [ovine 14q(0)] are given in brackets. Box C/C′ and box D/D′ motifs are boxed and the terminal 5′–3′ stem is indicated by arrows in the opposite orientation. **(B)** Sequence alignment of A and B repeats. Nucleotide positions of these repeats within AL132709 are shown in brackets. Conservation of each nucleotide position is denoted by shading according to GeneDoc software: black (100%), dark grey (80–99%) and clear grey (60–79%).

Interestingly, the AW026953 EST clearly connects the 3′ part of *MEG8* RNA with the 5′ end of the snoRNA gene cluster I. Clusters I and II also appear to be transcribed from the same unit, as suggested by sequencing of a RT–PCR product 2 (Fig. 1B) overlapping the two types of snoRNA genes. Taken together, this analysis suggests that the three novel types of human snoRNAs are processed from a single transcription unit that includes the *MEG8* non-coding RNA gene. Northern blot analysis performed on total RNA extracted from a panel of human tissues showed that the three C/D snoRNAs, 14q(0), 14q(I) and 14q(II), exhibit the same tissue-specific expression pattern. The strongest expression was observed for the brain and the uterine mucous membrane and to a lesser extent within the heart, while only a very faint signal could be detected for all other human tissues analysed (Fig. 3). This expression pattern was compared with that of the snoRNAs encoded at 15q11–q13. While HBII-52 snoRNA was only detected in the brain, HBII-85 was detected at relatively high levels in the brain, uterine mucous membrane and kidney and at low levels in several other tissues (Fig. 3) (14).

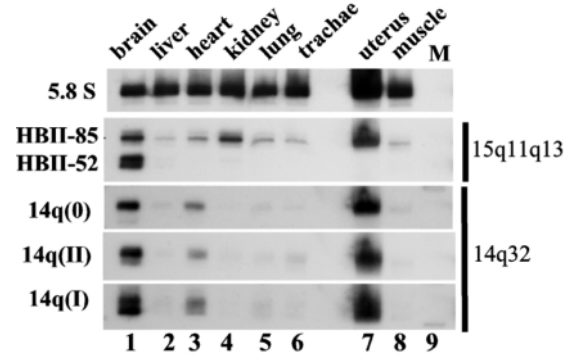

Figure 3. Tissue-specific expression of the novel snoRNAs analysed by northern hybridization. Total RNAs (5 μg/lane) purchased from Clontech (lanes 1–6) or extracted from normal adult human tissues (lanes 7 and 8) were fractionated on a 6% acrylamide gel/7 M urea and analysed by northern blot with ^{32}P labelled oligonucleotide probes against 14q32- or 15q11q13-encoded snoRNAs (see Materials and Methods). A 5.8S rRNA probe has also been used for gel loading control. M, DNA marker.

The maternally expressed *Rian* transcript at mouse distal 12 chromosome encodes at least 9 C/D snoRNAs

Hatada *et al.* (32) recently described a novel 5.4 kb maternally expressed brain-specific non-coding RNA gene, named *Rian*, that is closely linked to the *Gtl2* gene on distal mouse chromosome 12. BLAST N database searches identified several mouse ESTs positive for *Rian* (data not shown), among them AK017440 cDNA harbouring four segments exhibiting 100% identity to the *Rian* transcript. Sequence analysis revealed that these segments of homology correspond to exons (exons a–d), since they are interrupted by sequences displaying splice donor and acceptor consensus intronic hallmarks at both ends (data not shown; Fig. 4A). Moreover, AK017440 also displays homology with a 43 kb-long mouse contig 474407. By further sequence analysis, we could identify within AK017440

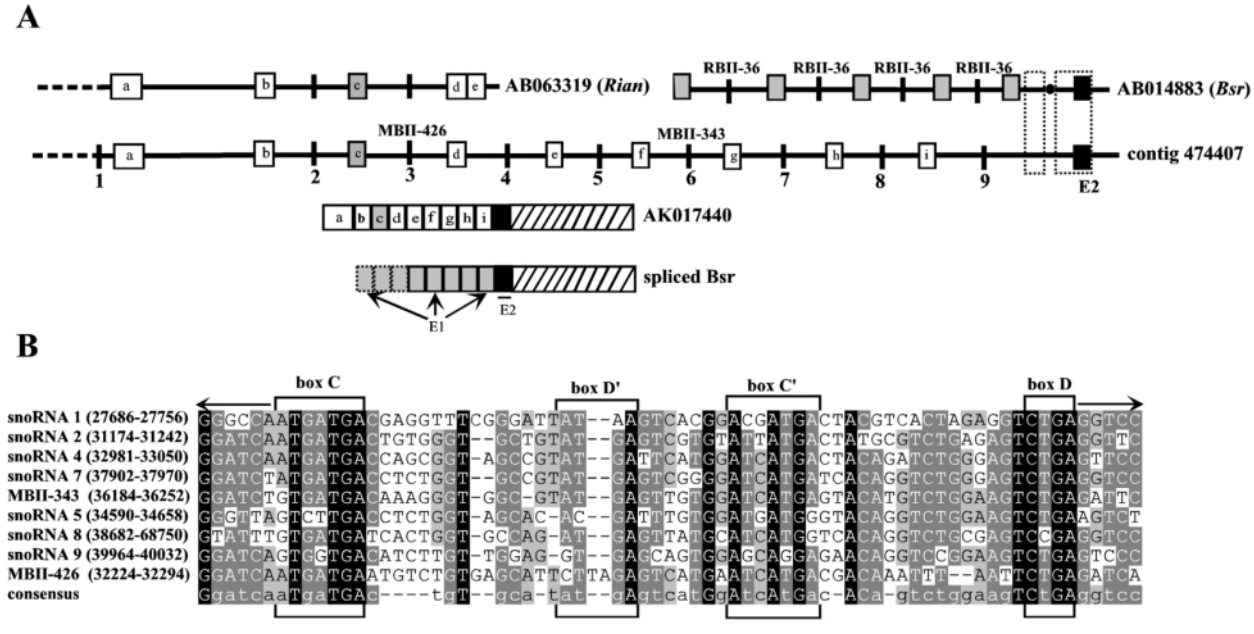

Figure 4. The maternally expressed mouse *Rian* gene encodes nine related C/D snoRNAs. **(A)** Schematic representation of the *Rian* gene (contig 474407), hemi-processed *Rian* RNA (AB063319) as described in (32), fully spliced *Rian* RNA (AK017440), the *Bsr* gene (AB014883) and spliced *Bsr* RNAs as described in (37). Exons are depicted by boxes, introns by a thick horizontal line and snoRNAs by a vertical bar. *Rian* exons closely similar to repeated E1 and E2 Bsr exons (37) are filled in grey and black, respectively. The repetitive element Lx7LINE/L1 at the 3′ end of both spliced Bsr and Rian RNAs is denoted by a hatched box. Intronic sequences conserved between mouse *Rian* and rat *Bsr* genes are denoted by dotted boxes (the black oval in between depicts a repetitive ID element in the *Bsr* intron not found within the *Rian* gene). **(B)** Sequence alignment of C/D snoRNA genes encoded in *Rian* introns. Nucleotide positions of C/D snoRNAs within contig 474407 are shown in brackets. Symbols are as in Fig. 2. Average percentages of identity to the consensus sequence of the *Rian*-encoded snoRNAs and RBII-36 snoRNA are 70.6% and 88.7%, respectively (according to GeneDoc software).

additional spliced exons (e–i), suggesting strongly that the *Rian* gene, as originally described, might represent an RNA-processing intermediate. Remarkably, AK017440 cDNA also shares some sequence and structural features in common with the rat *Bsr* cDNA. First, *Rian* exon c displays 80% identity to the repeated E1 exons from *Bsr* RNA (37) (Fig. 4A). Second, *Rian* and *Bsr* cDNAs both exhibit a conserved exon E2 sequence (82% identity) immediately followed by an Lx7/LINE1 repetitive element. Third, sequences of the last *Bsr* and *Rian* introns are very similar to each other (with two ~264 and 155 nt-long segments exhibiting 72% and 76% identity; Fig. 4A). These common features prompted us to look for the presence of RBII-36-related snoRNAs within the *Rian* gene. We could identify nine C/D snoRNAs genes, all of which are intron-encoded (Fig. 4A). Surprisingly, they do not display significant homology with RBII-36 and are relatively divergent from each other (Fig. 4B). C/D snoRNAs located within introns 3 and 6 correspond to mouse MBII-426 and MBII-343

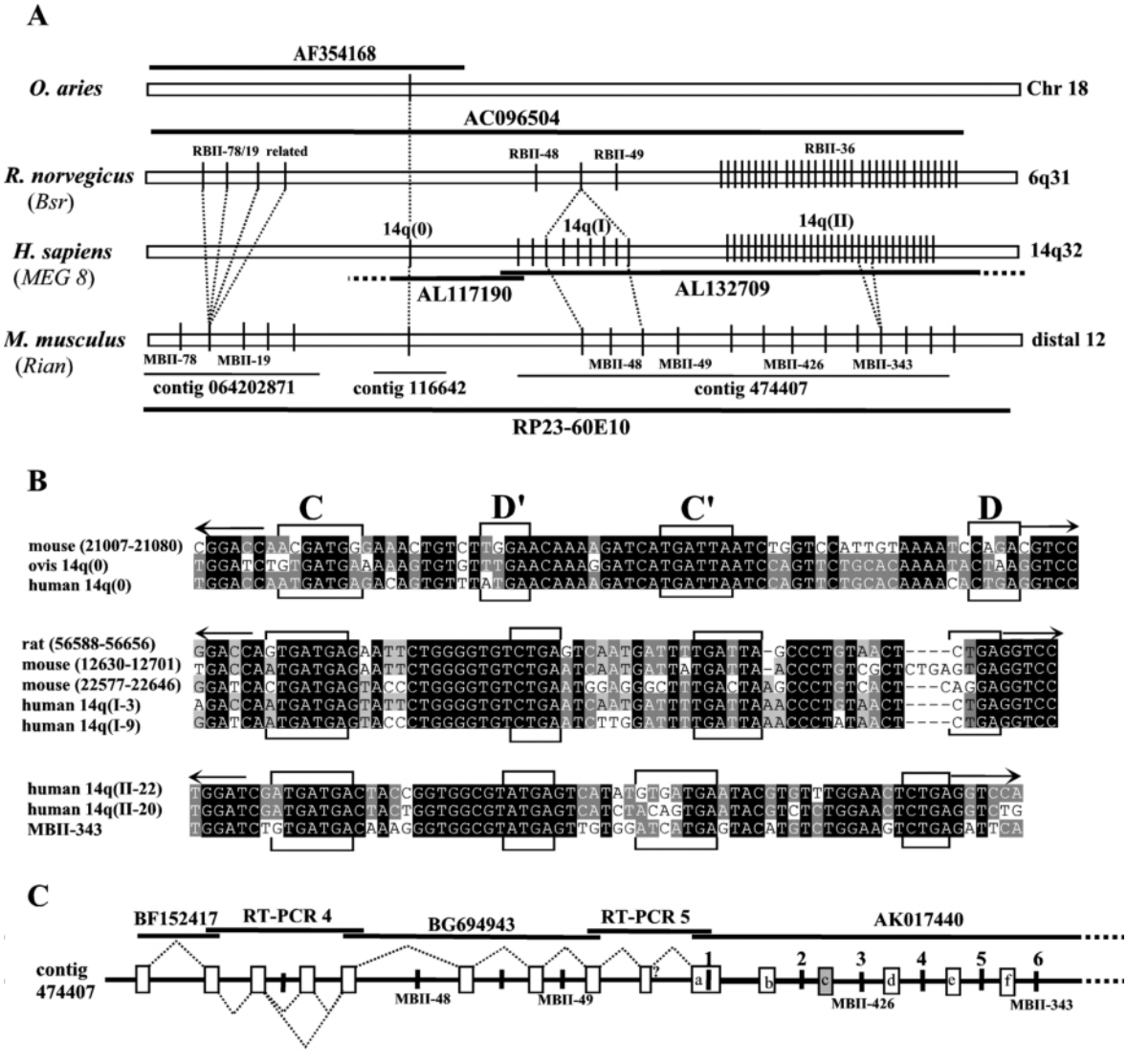

Figure 5. Identification of C/D snoRNA genes at mammalian chromosomal loci orthologous to human 14q32. **(A)** Genomic location of C/D snoRNAs at several mammalian loci. C/D snoRNA genes are represented by vertical black bars. The relative locations of snoRNAs within BAC AC096504 clone (48 unordered pieces) and the positions of mouse contigs 064202871, 116642 and 474407 have been chosen arbitrarily. Sequence homology between rodents and human C/D snoRNAs are denoted by dotted lines (only several homologies are shown). Within AC096504, RBII-78/19-related snoRNAs are located at 45772–45843, 46637–46708, 77115–77186 and 77980–78049, RBII-49 at 58503–58566 and RBII-48 at 37220–37289. Within contig 064202871, MBII-78/19-related snoRNAs are located at 11698–11764, 12553–12625, 13272–13344, 14137–14209 and 14747–14819. Within contig 474407, MBII-48 snoRNA is located at 17568–17637 and MBII-49 at 24393–24459. The cartoons are not drawn to scale. **(B)** Sequence alignment of putative mammalian C/D snoRNA homologues of 14q(0) snoRNA (top), 14q(1) snoRNA (middle) and 14q(II) snoRNA (bottom). Symbols are as in Fig. 2A. **(C)** MBII-48, MBII-49 and other flanking C/D snoRNAs are also intron-encoded within the *Rian* gene. Nucleotide positions of exons within contig 474407: (1867–1954); (6656–6730); (8603–8708); (15831–15923); (16949–17012); (19394–19481); (24120–24207); (24969–25016); (26817–26872 ?); (27404–28312); (30711–30799); (31629–31714); (32570–32656); (33416–33495); (35843–35980); (37406–37491); (38345–38388); (39610–39678); (41280–41364). Underlined exon might represent an alternatively spliced, snoRNA1-containing product. Other symbols are as in Figs 1B and 4A.

snoRNAs, respectively, previously characterized by analysis of a cDNA library (41). As for the other imprinted C/D snoRNAs, we could not detect potential cellular RNA targets based on appropriate complementarity. Interestingly, several copies of the MBII-343/426-related snoRNA cluster exhibit significant homology with several snoRNA copies within human cluster II, suggesting that the *Rian*-encoded snoRNAs might be the rodent counterparts of human 14q (II) snoRNAs (Fig. 5A, B).

Identification of additional C/D snoRNAs at rat 6q31 and mouse distal chromosome 12

To gain more information about the imprinting status of snoRNA genes encoded at human 14q32, orthologous mammalian loci were searched systematically for the presence of putative C/D snoRNAs. During the course of this work, a portion of the rat *Bsr* genomic locus was deposited in GenBank (accession no. AC096504); 33 copies of RBII-36 (Fig. 5A) as well as 7 other C/D snoRNA genes were detected in this portion. Four of these are related to each other and similar to the previously identified mouse MBII-19 and MBII-78 snoRNAs (41), which seem likely to represent variants of the same snoRNA species (data not shown). The three other rat C/D snoRNA genes display strong homology to the mouse C/D snoRNA MBII-48 and MBII-49 (Fig. 5) (41) and also interestingly to the several human 14q(I) snoRNA gene copies (Fig. 5B). Through BLAST searches, we then focused our attention on the genomic organization of the mouse homologues. Interestingly, MBII-48 and MBII-49 snoRNAs were both detected within contig 474407 containing the *Rian* gene (Fig. 5A), in addition to several other C/D snoRNAs that are more particularly related to some of the human 14q(I) snoRNA copies (Fig. 5B). By using a mouse BLAT search at http://genome.ucsc.edu/, a putative mouse homologue of the human 14q(0) snoRNA (Fig. 5B) was also detected within contig 116642 embedded within a region of strong homology to human *MEG8*. Finally, five related C/D snoRNAs, including MBII-19 and MBII-78 snoRNAs were also found within mouse contig 064202871. PCR analysis of a partially characterized BAC containing *Gtl2* (RP23-60E10) confirmed that these snoRNA genes are actually encoded by the mouse distal 12 imprinted region (data not shown). RT–PCR experiments and analysis of mouse ESTs overlapping the mouse snoRNA genes showed that MBII-48, MBII-49 and other flanking C/D snoRNAs are also intron-encoded and processed from a transcript physically linked to *Rian* RNA (Fig. 5C). Several of these rodent snoRNA genes contain one or two base substitutions within the C/D motifs (Fig. 5B), and could possibly be unstable or non-functional. Consistent with this, expression of several of them was undetectable on a northern blot (data not shown).

The brain-specific C/D snoRNAs are imprinted in mouse

Northern blot analysis of total RNA extracted from mouse adult tissues shows that all the mouse C/D snoRNAs located at mouse distal chromosome 12 are only expressed within the brain. These include MBII-19, MBII-343 and MBII-426, which were previously reported as ubiquitously expressed, possibly as a result of probe cross-hybridization (Fig. 6A) (41). Expression levels of C/D snoRNAs at mouse distal chromosome 12 have also been studied during brain development. Northern blot analysis performed at various developmental stages (Fig. 6B) showed that all C/D snoRNAs located at distal chromosome 12 are expressed within the embryo (E11–E15), the newborn (postnatal days P0–P7) and the adult. During brain ontogeny, the level of snoRNA expression is roughly constant or slightly increased. This developmental expression pattern is very different from that observed for MBII-52 and MBII-85 snoRNAs (encoded at chromosome 7C), which are very poorly

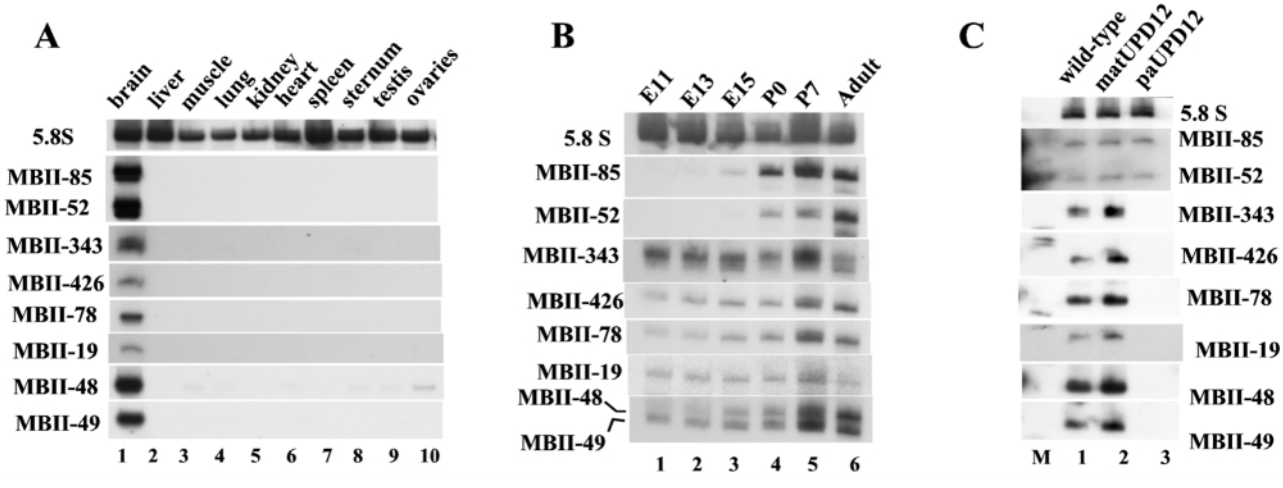

Figure 6. C/D snoRNA genes encoded at mouse distal chromosome 12 are brain-specific and expressed only from the maternal chromosomes. Each lane has been loaded with 10 μg of total RNAs extracted from different mouse adult tissues as specified at the top of each lane in **(A)**, or from developing mouse embryos (E11–E15), newborn (P0–P7) or adult brains in **(B)**, or from embryos (E15.5) with either maternal or paternal uniparental disomy for chromosome 12 (matUPD12 and patUPD12, respectively (55)) in **(C)**. Samples were fractionated on a 6% acrylamide/7 M urea gel and analyzed by northern blot with ^{32}P-labelled oligonucleotide probes (see Materials and Methods). A 5.8S rRNA probe has been used as gel loading control. Signal intensities (ratio snoRNA/5.8S rRNA signal) were quantified with a Fuji-Bas-1000 imager. M, DNA marker.

expressed between E11 and E15 but are highly upregulated after birth.

We then examined whether the C/D snoRNAs were subjected to genomic imprinting as predicted for genes encoded within this mouse distal chromosome 12 region. As shown in Figure 6C, E15.5 embryos with two maternal chromosomes 12 (matUPD12) express all the C/D snoRNAs analysed. Moreover, for most of them, the expression level is approximately twice that seen for normal embryos as determined by quantification of the hybridization signal (data not shown). Conversely, none of those snoRNAs was detected in total RNA extracted from embryos having two paternal chromosomes 12 (patUPD12), showing that all of these snoRNA genes are imprinted and expressed only from the maternal allele. As a control, we also checked the expression of the two C/D snoRNAs located at 7C. As expected, they were unaffected in UPD12 embryos.

DISCUSSION

Comparative genomic organization and expression of C/D snoRNAs at human 15q11–q13 (mouse 7C) and 14q32 (mouse distal 12)

Imprinted *DLK1/GTL2* and *IGF2/H19* domains at human 14q32 and 11p15.5, respectively, appear to share several genomic and epigenetic features that may be involved in the regulation of these two pairs of reciprocally imprinted genes (3,35,36,39,40,42). In this work, multiple tandemly repeated C/D snoRNA genes at human imprinted 14q32 and at its orthologous region on mouse distal 12 chromosome were identified. These C/D snoRNAs have a gene organization strikingly reminiscent of that previously reported for the PWS/AS region (14–16) (Fig. 1 and Table 1). Indeed, both imprinted loci are characterized by a complex transcription unit spanning arrays of tandemly encoded C/D snoRNA genes, giving rise to a spectrum of alternatively spliced non-coding RNAs, mostly consisting of short repeated exons that are not conserved between human and rodents. All 14q32-encoded C/D snoRNAs appear to be derived from a unique transcription unit encompassing *MEG8*, similar to the PWS-encoded C/D snoRNA (21) (Fig. 1B). However, despite their related gene organizations, the two loci differ in both snoRNA expression pattern and genomic features around the snoRNA host genes. While the three different snoRNA species encoded in human 14q32 have the same tissue-specific expression pattern, the expression patterns of HBII-52 and HBII-85 snoRNAs at the PWS/AS region differ considerably from each other (Fig. 3). Since the PWS-encoded snoRNAs are produced from the same primary transcript (21), this might reflect brain-specific differential processing of the more distal part of the *SNRPN–SNURF*–snoRNA host transcript. Unlike human snoRNAs, all mouse distal 12-encoded C/D snoRNAs are mainly detected within the brain, and their level of expression is variable according to the snoRNA species – but always lower than that of 7C-encoded snoRNAs, which are among the most abundant C/D snoRNA within the brain (37,41). Mouse 7C-encoded C/D snoRNAs are predominantly expressed in the adult from the paternal allele (Fig. 6B) (14), while the mouse distal 12 snoRNAs are transcribed from the maternal allele and are expressed during brain development (Fig. 6B, C). The snoRNA host genes at 14q32 and at the PWS/AS domain are dramatically different in their content of common interspersed

Table 1. Imprinted snoRNA genes at human 15q11q13 and 14q32 loci

Locus and snoRNA features		15q11q13			14q32	
snoRNA gene organization:						
C/D snoRNA gene cluster		HBII-85 (27 copies) and HBII-52 (47 copies)			14q(I) (9 copies) and 14q(II) (31 copies)	
Cluster size (kb)		48	99		21.3	44.3
% identity snoRNA copies/consensus[a]		90.2	93.9		82.3	80.8
% GC content		50.1	58.36	(41)	34.1	33.98
Repetitive elements (core of snoRNA clusters)	SINEs	0.62	0	(13.14)	4.72	6.23
	LINEs	1.43	0	(20.42)	13.76	1.62
	LTR elements	0	1.2	(8.29)	2.53	0.68
	DNA elements	0	0	(2.84)	2.76	1.21
snoRNA mode of biosynthesis:						
Located within repeated introns		Yes			Yes	
A single transcription unit		Yes			Yes	
% identity repeated spliced exons/consensus[a]		59.3	79.9 and 67.5[b]		47.1	31.2
Imprinted snoRNA genes in human/mouse		Yes/Yes (paternal)			?/yes (maternal)	
snoRNA expression						
Brain-specific in human		HBII-85 (no) and HBII-52 (yes)			Mainly expressed in brain/heart/mucous uterin	
Brain-specific in rodents		Yes			Yes	
Expression upon embryonic mouse development		No or very low			Yes	
snoRNA features						
Common terminal stem 5′-GG(A/G)TC . . .GG(C/T)CC-3′		Yes			Yes	
Antisense element conservation		Strong			Moderated	
Identified putative RNA targets		No (5-HT2c mRNA and HBII-52?)			No	

[a]As defined by GeneDoc software.
[b]The repeat unit spanning the HBII-52 snoRNA gene contains two different repeated exons. The complete human genome values (57) are given in parentheses.

repeats and overall base composition (Table 1). In both cases, a sharp transition in G + C content is observed at the borders of the snoRNA clusters, corresponding to a striking increase (14q32) or decrease (15q11q13) in the flanking sequences (Fig. 7). The reciprocal parental expression of the snoRNA gene clusters with regard to their G + C content might reflect the evolutionary maintenance or function of the snoRNA regions in ways that are not yet understood.

C/D snoRNA function and evolutionary divergence

Methylation guides are evolutionary ancient molecules (detected both in Archaea and eukaryotic cells) required to modify a large panel of cellular RNAs, including ribosomal RNAs (rRNAs) and small nuclear RNAs (snRNAs) (20), tRNAs (43) and possibly mRNAs (14). The C/D snoRNAs identified at the two imprinted loci depart from other known mammalian methylation guides by their tissue-specific expression (Figs 3 and 6A). Moreover, in contrast to most of the other C/D snoRNAs, no statistically significant complementarity to rRNA or U snRNA within snoRNAs of the imprinted loci were identified. HBII-52 was the only exception to this (14). Characterization of the cellular RNAs potentially targeted by these imprinted C/D snoRNAs represents a worthwhile challenge, since lack of HBII-85 snoRNA expression has been proposed to contribute to PWS phenotypes (44). Although the human (*MEG8*), mouse (*Rian*) and rat (*Bsr*) snoRNA host genes share a common physical localization, their DNA sequences are not highly conserved, similar to the human and mouse host genes at the PWS/AS region (14–16). However, in contrast to the PWS-encoded snoRNA gene copies which are relatively well conserved among each other, the *Rian*-encoded mouse snoRNA gene copies, and also to some extent the human 14q(I) and 14q(II) snoRNAs, are much more divergent from each other (Figs 2A and 4B and Table 1). The lack of strong selection pressure on the snoRNA sequence segments upstream from D/D′ between mammals or even between snoRNA copies within the same gene cluster might suggest that these imprinted C/D snoRNAs are rapidly evolving molecules with multiple potential RNA targets. Alternatively, this may strengthen the hypothesis that they confer a genomic regulatory function rather than the alternative previously described snoRNA functions (see below).

Tandemly repeated C/D snoRNAs and imprinting

Interestingly, the 5′–3′ terminal stem of most of the tandemly repeated snoRNAs at 15q11q13 and 14q32 loci exhibit an identical sequence (Fig. 2A and Table 1) not shared by any other snoRNA among the scores of ubiquitously expressed, non-repeated C/D snoRNAs reported for mammals (18,41,45). This could indicate that the novel tandemly repeated snoRNAs at the two imprinted loci have all evolved from a common snoRNA ancestor gene during mammalian evolution. Several observations have suggested that C/D snoRNA genes may represent mobile genetic elements (46,47). The tandemly repeated snoRNA arrays may well result from retrotransposition of a snoRNA gene into an intron followed by tandem duplications of the intron and flanking exon. Generation of the repeated snoRNA arrays could have occurred after establishment of the imprinted status of these loci, merely reflecting the fact that they represent ‘hot spots’ for DNA integration events in the germline, in relation to a particular chromatin structure. Alternatively, their occurrence could have predated imprinting and even played a role in its establishment and maintenance.

Many imprinted loci harbour non-coding RNAs having an imprinted expression opposite to that of linked protein-coding genes on the same chromosomes, raising the possibility of a role for these RNAs in silencing (3,27–29). C/D snoRNA host-gene transcripts could therefore be involved in gene regulation, perhaps in a manner similar to that of other chromosomal RNAs involved in dosage compensation or, like Air RNA, in the regulation of imprinted genes at mouse chromosome 17 (48,49). However, the snoRNA genes at 15q11q13 do not seem to play a central role in the epigenetic process, since a paternal deletion from the *Snrpn* to the *Ube3A* gene in the mouse does not affect the imprinted status of the upstream *Ndn* gene (50). Rather, it has been proposed that transcription of the distal part of the snoRNA host gene might silence in *cis* the *Ube3A* paternal expression, either by RNA antisense mechanisms or by

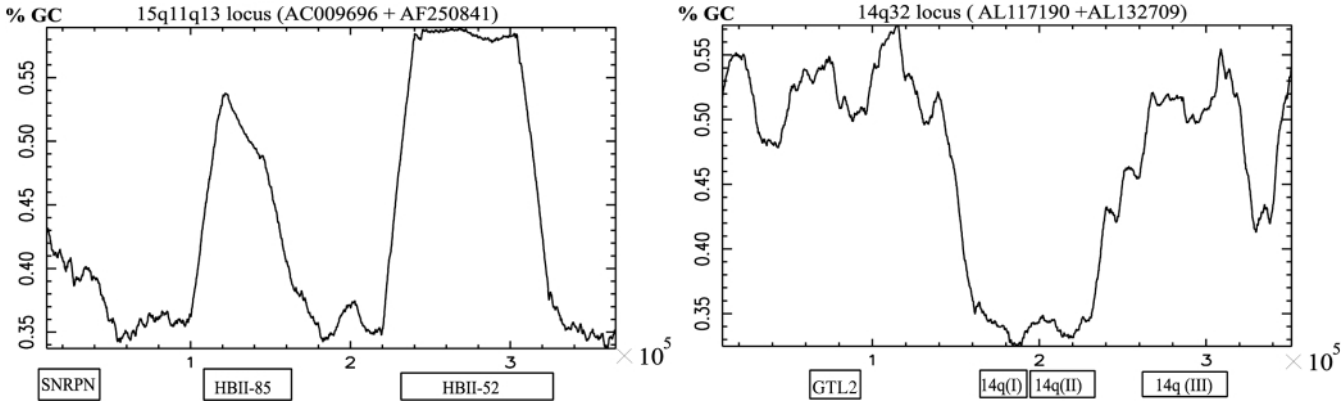

Figure 7. DNA isochores showing differences in G + C content within and around snoRNA gene clusters at the two human imprinted 15q11q13 and 14q32 loci. At the PWS/AS region, the paternally expressed snoRNA genes lie within a (G + C)-rich environment, while at 14q32 the maternally expressed snoRNA genes are located within an (A + T)-rich environment. Accession numbers of the sequences are given in bracket and the relative position of imprinted genes (boxed) is schematized.

altering local chromatin structure (21–23). Whether C/D snoRNA host genes at 14q32 also control the expression of neighbouring imprinted genes is an attractive hypothesis that can be tested using mouse models.

Finally, it is possible that the presence of these snoRNAs reflects a brain-specific function for these imprinted genes. The PWS and AS loci, which cause distinct neurological disorders, are linked to a locus containing snoRNAs. BWS (linked to the imprinted cluster on 11p15.5, where no snoRNAs have been noted to date) is a somatic overgrowth syndrome with no evidence of neurological problems. Aberrant imprinting on chromosome 14 is associated with both growth and neurological anomalies in mUPD14 and pUPD14 individuals. The former exhibit precocious puberty and are growth-retarded, while pUPD14 patients have defects of the axial skeleton and mental retardation (51–53). It is tempting to make functional associations between developmental growth regulators on chromosome 11p and on 14q, and between the neural-specific snoRNA genes on chromosome 15q and on 14q.

In conclusion, we have shown that a second imprinted locus in humans, 14q32, in addition to the previously reported 15q11q13, encodes arrays of tandemly repeated C/D snoRNA genes, and that this outstanding feature of the locus is conserved in human and rodents. Imprinted snoRNAs have only been detected within eutherian mammals so far (15,37). Given that genomic imprinting might be restricted to marsupials and placental mammals (54), a detailed sequence analysis of the orthologous 14q32 and 15q11q13 loci in non-eutherian mammals together with systematic sequence searches of a large set of imprinted loci in mammalian genomes should provide further insight into the potential significance of this intriguing association between genomic imprinting and the presence of these unusual tandemly repeated snoRNA genes.

MATERIALS AND METHODS

Oligonucleotides

These were all synthesised by Y. de Préval (LBME, Toulouse) on a PerSeptive Biosystems Expedite apparatus. 14q(0): 5′-GACCTCAGTGTTTTGTGCAGAAC-3′; 14q(I): 5′-TGGACCTCAGAGTTCCAGACACGTATTCA-3′; 14(II): TGATTCAGAC/TACCCCAG/CG/A G/ATACTCATC-3′; MBII-48: 5′-ATAAGGGTTTAATCACTGTCCT CGGTCA-3′; MBII-49: 5′-AATCCAGTATGTTGTCATCGTCTATG-3′; MBII-19: 5′-CAGACATCTG TTCTCATGGCT-3′; MBII-78: 5′-ACCTCAGATATCTGTTCATGTCA-3′; MBII-52: 5′-CTGACGTAATCCTATTGAGCAT; MBII-85: 5′-ACA-GAGTTTTCACTCATTTTGTTC-3′; MBII-343: 5′-TCTCAG ACTTCCAGACATGTACT-3′; MBII-426: 5′-TGATCTCA-GAATTAAATTTGTCG-3′, 5′ RT–PCR-2: 5′-CGCGGATCC-GACGAGATTGGATTTGGTCATTTCC-3′; 3′ RT–PCR-2: 5′-CCGGAATTCCAGGCTCCTACCCAGAGGCAACTG-3′; 3′ RT–PCR1: 5′-CGCGGATCCTATTTGTCTCTATGCTCCT-TACTC-3′; 5′ RT–PCR1: 5′-GCGGAATTCTACATGGATCC-CACTTTGGACAAAG-3′; 5′ RT–PCR3: 5′-GCGGAAT-TCTGGATGCAATGAGCTGATCA-3′; 3′ RT–PCR3: 5′-CGCGGATCCTGAGGCTCACAGAGGACGGCAG-3′; 5′ RT–PCR4: 5′-CCGGAATTCGGATGGTTACTGTCTGA-GACTGAG-3′; 3′ RT–PCR4: 5′-CCGGAATTCTCATC-TATCCTCCTGACTCAGGAC-3′; 5′ RT–PCR5: 5′-CCGGAATTCTTGCATGCAAGCTCTACAGTTATGC-3′; 3′ RT–PCR5: 5′-CCGGAATTCAGCTGCCGAGCTCCATCCA-CATGGT-3′; 5′ MBII-19/78-cluster: 5′-GGCTTTGATCCTT-CGGTTGGA-3′; 3′ MBII-19/78-cluster: 5′-CCCTTCT GCTCCAAGTTTGCCTA-3′.

Search for tandemly repeated C/D snoRNAs and sequence analysis

BAC sequences around DLK1/GTL2 positions have been retrieved from the Human Genome Project Working Draft (http://genome.ucsc.edu/), purged for common interspersed repeats by using Repeat Masker (http://repeatmasker.genome.washington.edu/cgi-bin/RepeatMasker) and systematically compared with themselves by dotplot analysis conducted by PipMaker (http://bio.cse.psu.edu/). C/D snoRNA-like sequences and other repeats have been subsequently detected by BLAST2 sequences (http://www.ncbi.nlm.nih.gov/blast/bl2seq/bl2.html). CpG islands have been defined according to the Human Genome Project Working Draft annotations. Sequence alignment have been obtained by either Clustal W (http://clustalw.genome.ad.jp/) or MultiAlin (http://prodes.toulouse.inra.fr/multalin/multalin.html), and conserved nucleotides have been shaded by GeneDoc (http://www. Cris.com/~ketchup/genedoc.shtml). Analysis of DNA isochores has been performed using Isochore (EMBOSS) at http://bioweb.pasteur.fr/seqanal/interfaces/isochore.html.

Mice with matUPD12 and patUPD12

E15.5 embryos with matUPD12 and patUPD12 and normal control littermates were generated by intercrossing parent animals of two different genetic backgrounds that were double heterozygotes for Robertsonian translocations with monobrachial homology for chromosome 12, as described previously (55).

RNA isolation and northern blot analysis

Total RNA was isolated by the method of Chomczynski and Sacchi (56), adapted to our conditions as follows. Rodent tissues freshly prepared according to French institutional and UK HO guidelines were quickly frozen in liquid nitrogen and stored at −80°C. Frozen samples were then homogenized and resuspended in a 1 : 1 (vol : vol) mixture of (i) 4 M guanidinium thiocyanate, 25 mM sodium citrate (pH 7.0), 0.5% Sarkosyl 0.1 M 2-mercaptoethanol and (ii) water-saturated phenol. This mixture was then supplemented with sodium acetate (pH 4.0) to 0.1 M. The lysate was mixed with 0.1 volume of chloroform, vigorously vortexed and incubated for 10–15 min at room temperature. RNA was precipitated with 2 volume of ethanol and stored at −20°C. RNAs were fractionated by electrophoresis in 6% acrylamide–7 M urea gels, transferred electrophoretically (120 min in 0.5 × TBE at 1.0 A) to a nylon membrane (Hybond N +, Amersham) and crosslinked by ultraviolet irradiation (1200 J/cm^2, Stratalinker, Stratagene, La Jolla, Calif.). Northern blot hybridizations were carried out in the presence of a 5′-end ^{32}P-labelled oligodeoxynucleotide probe (500 000 c.p.m./ml) by overnight incubation at 50°C in

5 × SSPE, 1% SDS, 5 × Denhardt's, 150 µg/ml yeast tRNA. Membranes were washed twice for 15 min at room temperature in 0.1 × SSPE/0.1%SDS.

RT–PCR and cDNA clones

Ten micrograms of total human or mouse brain RNA was reverse-transcribed at 42°C for 120 min, using Superscript II (Gibco BRL) with random hexamer primers and 1/40 of cDNA products amplified by 40 cycles of PCR with *Taq* polymerase (Promega). PCR product cloned into in pGEM-T easy system I vector (Promega) and IMAGE ESTs clones (T85042, AW026953, BF055187, AA451995, AA861571, AA910544, AA680166, AA910544 and AA433836) provided by UK-HGMP RC (http://www.hgmp.mrc.ac.uk/) have been sequenced by the CEQ 2000 DNA analysis system (Beckman).

ACKNOWLEDGEMENTS

We thank A. Hüttenhofer for helpful discussions throughout this work and C. Gaspin and P. Thebault for help with the use of sequence analysis soft-wares. We also thank N. Joseph for technical assistance in DNA sequencing and Y. de Préval for oligonucleotide synthesis (IEFG 109). We are grateful to Neil Youngson and Shau-Ping Lin for comments on the manuscript. This work was supported by laboratory funds from the Centre National de la Recherche Scientifique and Université Paul-Sabatier, Toulouse, and by grants from the Toulouse Genopole/Pole Santé (to J.P.B.), from La Ligue Contre le Cancer/Comité de Haute Garonne (to J.C.) and from the UK Medical Research Council (to A.C.F.S.).

REFERENCES

1. Tilghman, S.M. (1999) The sins of the fathers and mothers: genomic imprinting in mammalian development. *Cell*, **96**, 185–193.
2. Reik, W. and Walter, J. (2001) Genomic imprinting: parental influence on the genome. *Nat. Rev. Genet.*, **2**, 21–32.
3. Ferguson-Smith, A.C. and Surani, M.A. (2001) Imprinting and the epigenetic asymmetry between parental genomes. *Science*, **293**, 1086–1089.
4. Paulsen, M. and Ferguson-Smith, A.C. (2001) DNA methylation in genomic imprinting, development, and disease. *J. Pathol.*, **195**, 97–110.
5. Szabo, P., Tang, S.H., Rentsendorj, A., Pfeifer, G.P. and Mann, J.R. (2000) Maternal-specific footprints at putative CTCF sites in the H19 imprinting control region give evidence for insulator function. *Curr. Biol.*, **10**, 607–610.
6. Bell, A.C. and Felsenfeld, G. (2000) Methylation of a CTCF-dependent boundary controls imprinted expression of the Igf2 gene. *Nature*, **405**, 482–485.
7. Hark, A.T., Schoenherr, C.J., Katz, D.J., Ingram, R.S., Levorse, J.M. and Tilghman, S.M. (2000) CTCF mediates methylation-sensitive enhancer-blocking activity at the H19/Igf2 locus. *Nature*, **405**, 486–489.
8. Kanduri, C., Pant, V., Loukinov, D., Pugacheva, E., Qi, C.F., Wolffe, A., Ohlsson, R. and Lobanenkov, V.V. (2000) Functional association of CTCF with the insulator upstream of the H19 gene is parent of origin-specific and methylation-sensitive. *Curr. Biol.*, **10**, 853–856.
9. 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.
10. 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.
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. *Nat. Genet.*, **8**, 52–58.
12. Shemer, R., Hershko, A.Y., Perk, J., Mostoslavsky, R., Tsuberi, B., Cedar, H., Buiting, K. and Razin, A. (2000) The imprinting box of the Prader–Willi/Angelman syndrome domain. *Nat. Genet.*, **26**, 440–443.
13. Ben-Porath, I. and Cedar, H. (2000) Imprinting: focusing on the center. *Curr. Opin. Genet. Dev.*, **10**, 550–554.
14. Cavaille, J., Buiting, K., Kiefmann, M., Lalande, M., Brannan, C.I., Horsthemke, B., Bachellerie, J.P., Brosius, J. and Huttenhofer, 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.
15. 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.
16. Meguro, M., Mitsuya, K., Nomura, N., Kohda, M., Kashiwagi, A., Nishigaki, R., Yoshioka, H., Nakao, M., 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.
17. Filipowicz, W. (2000) Imprinted expression of small nucleolar RNAs in brain: time for RNomics. *Proc. Natl Acad. Sci. USA*, **97**, 14035–14037.
18. Kiss-Laszlo, Z., Henry, Y., Bachellerie, J.P., Caizergues-Ferrer, M. and Kiss, T. (1996) Site-specific ribose methylation of preribosomal RNA: a novel function for small nucleolar RNAs. *Cell*, **85**, 1077–1088.
19. Cavaille, J., Nicoloso, M. and Bachellerie, J.P. (1996) Targeted ribose methylation of RNA *in vivo* directed by tailored antisense RNA guides. *Nature*, **383**, 732–735.
20. Kiss, T. (2001) Small nucleolar RNA-guided post-transcriptional modification of cellular RNAs. *EMBO J.*, **20**, 3617–3622.
21. Runte, M., Huttenhofer, A., Gross, S., Kiefmann, M., Horsthemke, B. and Buiting, K. (2001) The IC–SNURF–SNRPN transcript serves as a host for multiple small nucleolar RNA species and as an antisense RNA for UBE3A. *Hum. Mol. Genet.*, **10**, 2687–2700.
22. 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.
23. Rougeulle, C., Cardoso, C., Fontes, M., Colleaux, L. and Lalande, M. (1998) An imprinted antisense RNA overlaps UBE3A and a second maternally expressed transcript. *Nat. Genet.*, **19**, 15–16.
24. Wutz, A., Smrzka, O.W., Schweifer, N., Schellander, K., Wagner, E.F. and Barlow, D.P. (1997) Imprinted expression of the Igf2r gene depends on an intronic CpG island. *Nature*, **389**, 745–749.
25. Smilinich, N.J., Day, C.D., Fitzpatrick, G.V., Caldwell, G.M., Lossie, A.C., Cooper, P.R., Smallwood, A.C., Joyce, J.A., Schofield, P.N., Reik, W. *et al.* (1999) A maternally methylated CpG island in KvLQT1 is associated with an antisense paternal transcript and loss of imprinting in Beckwith–Wiedemann syndrome. *Proc. Natl Acad. Sci. USA*, **96**, 8064–8069.
26. Wroe, S.F., Kelsey, G., Skinner, J.A., Bodle, D., Ball, S.T., Beechey, C.V., Peters, J. and Williamson, C.M. (2000) An imprinted transcript, antisense to Nesp, adds complexity to the cluster of imprinted genes at the mouse Gnas locus. *Proc. Natl Acad. Sci. USA*, **97**, 3342–3346.
27. Reik, W. and Walter, J. (2001) Evolution of imprinting mechanisms: the battle of the sexes begins in the zygote. *Nat. Genet.*, **27**, 255–256.
28. Kelley, R.L. and Kuroda, M.I. (2000) Noncoding RNA genes in dosage compensation and imprinting. *Cell*, **103**, 9–12.
29. Sleutels, F., Barlow, D.P. and Lyle, R. (2000) The uniqueness of the imprinting mechanism. *Curr. Opin. Genet. Dev.*, **10**, 229–233.
30. Charlier, C., Segers, K., Karim, L., Shay, T., Gyapay, G., Cockett, N. and Georges, M. (2001) The callipyge mutation enhances the expression of coregulated imprinted genes in cis without affecting their imprinting status. *Nat. Genet.*, **27**, 367–369.
31. Charlier, C., Segers, K., Wagenaar, D., Karim, L., Berghmans, S., Jaillon, O., Shay, T., Weissenbach, J., Cockett, N., Gyapay, G. *et al.* (2001) Human–ovine comparative sequencing of a 250-kb imprinted domain encompassing the callipyge (clpg) locus and identification of six imprinted transcripts: DLK1, DAT, GTL2, PEG11, antiPEG11, and MEG8. *Genome Res.*, **11**, 850–862.
32. Hatada, I., Morita, S., Obata, Y., Sotomaru, Y., Shimoda, M. and Kono, T. (2001) Identification of a new imprinted gene, Rian, on mouse chromosome

12 by fluorescent differential display screening. *J. Biochem. (Tokyo)*, **130**, 187–190.
33. Ishihara, K., Hatano, N., Furuumi, H., Kato, R., Iwaki, T., Miura, K., Jinno, Y. and Sasaki, H. (2000) Comparative genomic sequencing identifies novel tissue-specific enhancers and sequence elements for methylation-sensitive factors implicated in Igf2/H19 imprinting. *Genome Res.*, **10**, 664–671.
34. Onyango, P., Miller, W., Lehoczky, J., Leung, C.T., Birren, B., Wheelan, S., Dewar, K. and Feinberg, A.P. (2000) Sequence and comparative analysis of the mouse 1-megabase region orthologous to the human 11p15 imprinted domain. *Genome Res.*, **10**, 1697–1710.
35. Paulsen, M., Takada, S., Youngson, N.A., Benchaib, M., Charlier, C., Segers, K., Georges, M. and Ferguson-Smith, A.C. (2001) Comparative sequence analysis of the imprinted Dlk1–Gtl2 locus in three mammalian species reveals highly conserved genomic elements and refines comparison with the Igf2–H19 region. *Genome Res.*, **11**, 2085–2094.
36. Wylie, A.A., Murphy, S.K., Orton, T.C. and Jirtle, R.L. (2000) Novel imprinted DLK1/GTL2 domain on human chromosome 14 contains motifs that mimic those implicated in IGF2/H19 regulation. *Genome Res.*, **10**, 1711–1718.
37. Cavaille, J., Vitali, P., Basyuk, E., Huttenhofer, A. and Bachellerie, J.P. (2001) A novel brain-specific box C/D small nucleolar RNA processed from tandemly repeated introns of a noncoding RNA gene in rats. *J. Biol. Chem.*, **276**, 26374–26383.
38. Komine, Y., Tanaka, N.K., Yano, R., Takai, S., Yuasa, S., Shiroishi, T., Tsuchiya, K. and Yamamori, T. (1999) A novel type of non-coding RNA expressed in the rat brain. *Brain Res. Mol. Brain Res.*, **66**, 1–13.
39. Schmidt, J.V., Matteson, P.G., Jones, B.K., Guan, X.J. and Tilghman, S.M. (2000) The Dlk1 and Gtl2 genes are linked and reciprocally imprinted. *Genes Dev.*, **14**, 1997–2002.
40. Takada, S., Tevendale, M., Baker, J., Georgiades, P., Campbell, E., Freeman, T., Johnson, M.H., Paulsen, M. and Ferguson-Smith, A.C. (2000) Delta-like and gtl2 are reciprocally expressed, differentially methylated linked imprinted genes on mouse chromosome 12. *Curr. Biol.*, **10**, 1135–1138.
41. Huttenhofer, A., Kiefmann, M., Meier-Ewert, S., O'Brien, J., 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.
42. Takada, S., Paulsen, M., Tevendale, M., Tsai, C.E., Kelsey, G., Cattanach, B.M. and Ferguson-Smith, A.C. (2002) Epigenetic analysis of the Dlk1–Gtl2 imprinted domain on mouse chromosome 12: implications for imprinting control from comparison with Igf2–H19. *Hum. Mol. Genet.*, **11**, 77–86.
43. d'Orval, B.C., Bortolin, M.L., Gaspin, C. and Bachellerie, J.P. (2001) Box C/D RNA guides for the ribose methylation of archaeal tRNAs. The tRNATrp intron guides the formation of two ribose-methylated nucleosides in the mature tRNATrp. *Nucleic Acids Res.*, **29**, 4518–4529.
44. Wirth, J., Back, E., Huttenhofer, 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.
45. Bachellerie, J.P. and Cavaille, J. (1997) Guiding ribose methylation of rRNA. *Trends Biochem. Sci.*, **22**, 257–261.
46. Maxwell, E.S. and Fournier, M.J. (1995) The small nucleolar RNAs. *Annu. Rev. Biochem.*, **64**, 897–934.
47. Bachellerie, J.P., Nicoloso, M., Qu, L.H., Michot, B., Caizergues-Ferrer, M., Cavaille, J. and Renalier, M.H. (1995) Novel intron-encoded small nucleolar RNAs with long sequence complementarities to mature rRNAs involved in ribosome biogenesis. *Biochem. Cell Biol.*, **73**, 835–843.
48. Park, Y. and Kuroda, M.I. (2001) Epigenetic aspects of X-chromosome dosage compensation. *Science*, **293**, 1083–1085.
49. Sleutels, F., Zwart, R. and Barlow, D.P. (2002) The non-coding Air RNA is required for silencing autosomal imprinted genes. *Nature*, **415**, 810–813.
50. 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.
51. Healey, S., Powell, F., Battersby, M., Chenevix-Trench, G. and McGill, J. (1994) Distinct phenotype in maternal uniparental disomy of chromosome 14. *Am. J. Med. Genet.*, **51**, 147–149.
52. Georgiades, P., Chierakul, C. and Ferguson-Smith, A.C. (1998) Parental origin effects in human trisomy for chromosome 14q: implications for genomic imprinting. *J. Med. Genet.*, **35**, 821–824.
53. Cotter, P.D., Kaffe, S., McCurdy, L.D., Jhaveri, M., Willner, J.P. and Hirschhorn, K. (1997) Paternal uniparental disomy for chromosome 14: a case report and review. *Am. J. Med. Genet.*, **70**, 74–79.
54. Killian, J.K., Byrd, J.C., Jirtle, J.V., Munday, B.L., Stoskopf, M.K., MacDonald, R.G. and Jirtle, R.L. (2000) M6P/IGF2R imprinting evolution in mammals. *Mol. Cell*, **5**, 707–716.
55. Georgiades, P., Watkins, M., Surani, M.A. and Ferguson-Smith, A.C. (2000) Parental origin-specific developmental defects in mice with uniparental disomy for chromosome 12. *Development*, **127**, 4719–4728.
56. Chomczynski, P. and Sacchi, N. (1987) Single-step method of RNA isolation by acid guanidinium thiocyanate–phenol–chloroform extraction. *Anal. Biochem.*, **162**, 156–159.
57. Lander, E.S., Linton, L.M., Birren, B., Nusbaum, C., Zody, M.C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W. *et al.* (2001) Initial sequencing and analysis of the human genome. *Nature*, **409**, 860–921.