The EMBO Journal Vol.16 No.18 pp.5742–5751, 1997

Clusters of multiple different small nucleolar RNA genes in plants are expressed as and processed from polycistronic pre-snoRNAs

David J.Leader, Gillian P.Clark, Jenny Watters, Alison F.Beven[1], Peter J.Shaw[1] and John W.S.Brown[2]

Cell and Molecular Genetics, Scottish Crop Research Institute, Invergowrie, Dundee DD2 5DA and [1]John Innes Centre, Norwich Research Park, Colney, Norwich NR4 7UH, UK

[2]Corresponding author
e-mail: jbrown@scri.sari.ac.uk

Small nucleolar RNAs (snoRNAs) are involved in many aspects of rRNA processing and maturation. In animals and yeast, a large number of snoRNAs are encoded within introns of protein-coding genes. These introns contain only single snoRNA genes and their processing involves exonucleolytic release of the snoRNA from debranched intron lariats. In contrast, some *U14* genes in plants are found in small clusters and are expressed polycistronically. An examination of *U14* flanking sequences in maize has identified four additional snoRNA genes which are closely linked to the *U14* genes. The presence of seven and five snoRNA genes respectively on 2.05 and 0.97 kb maize genomic fragments further emphasizes the novel organization of plant snoRNA genes as clusters of multiple different genes encoding both box C/D and box H/ACA snoRNAs. The plant snoRNA gene clusters are transcribed as a polycistronic pre-snoRNA transcript from an upstream promoter. The lack of exon sequences between the genes suggests that processing of polycistronic pre-snoRNAs involves endonucleolytic activity. Consistent with this, U14 snoRNAs can be processed from both non-intronic and intronic transcripts in tobacco protoplasts such that processing is splicing independent.

Keywords: box C/D snoRNAs/box H/ACA snoRNAs/polycistronic expression/2-*O*′-ribose methylation/rRNA processing

Introduction

Pre-ribosomal RNA (pre-rRNA) processing and ribosome assembly occur in the nucleolus of eukaryotic cells. The pre-rRNA transcript undergoes processing to generate mature 18S, 5.8S and 28S rRNAs (Eichler and Craig, 1994; Maxwell and Fournier, 1995; Sollner-Webb *et al.*, 1995; Venema and Tollervey, 1995). Several small nucleolar RNAs (snoRNAs) have been shown to be essential for processing steps which lead to production of 18S rRNA (U3, U14 and U22 in vertebrates; U3, U14, snR10 and snR30 in yeast) and for generation of 5.8S and 28S rRNAs [U8 in mammals and MRP in both vertebrates and yeast] (Eichler and Craig, 1994; Lafontaine and Tollervey, 1995; Maxwell and Fournier, 1995; Sollner-Webb *et al.*, 1995; Venema and Tollervey, 1995; Tollervey and Kiss, 1997). In addition to these essential snoRNAs, genes encoding >80 snoRNAs have been isolated recently and it is expected that a large number of snoRNAs are present in eukaryotes (Balakin *et al.*, 1996; Kiss-László *et al.*, 1996; Nicoloso *et al.*, 1996; Ganot *et al.*, 1997a).

All of the snoRNAs characterized to date fall into three groups: box C/D, box H/ACA and MRP snoRNAs. The latter class contains MRP RNA as its sole member which, as the ribonucleoprotein particle RNase MRP, cleaves the pre-rRNA upstream of 5.8S (Lygerou *et al.*, 1996, and references therein). The box C/D snoRNAs are characterized by the presence of conserved nucleotide sequences, box C (consensus UGAUGA) and box D (consensus CUGA) usually positioned near the 5′ and 3′ ends of the snoRNA, adjacent to terminal inverted repeats (Maxwell and Fournier, 1995; Kiss-László *et al.*, 1996; Nicoloso *et al.*, 1996). The formation of a stem between these inverted repeats and binding of protein factors to the adjacent box C and D sequences are thought to block cleavage by exonucleases in the formation of mature snoRNA (Tycowski *et al.*, 1993; Caffarelli *et al.*, 1996; Cavaillé and Bachellerie, 1996; Watkins *et al.*, 1996; Xia *et al.*, 1997). Boxes C and D, therefore, are required for stability and accumulation of snoRNAs and are likely to be involved in binding snoRNP proteins, such as fibrillarin, a nucleolar protein common to box C/D snoRNPs (Huang *et al.*, 1992; Peculis and Steitz, 1994; Caffarelli *et al.*, 1996; Cavaillé and Bachellerie, 1996; Watkins *et al.*, 1996; Xia *et al.*, 1997). They are also required for hypermethylation of the 5′ cap and nuclear retention of certain box C/D snoRNAs such as U3 which are transcribed from their own promoters (Terns *et al.*, 1995). A second feature of intron-encoded box C/D snoRNAs is the presence of regions of at least 10 nucleotides of complementarity to 18S or 28S rRNAs lying adjacent to box D or an internal box D-like sequence, D′ (Bachellerie *et al.*, 1995a,b; Cavaillé *et al.*, 1996; Kiss-László *et al.*, 1996; Nicoloso *et al.*, 1996; Tykowski *et al.*, 1996b). SnoRNAs containing such complementary regions act as guide RNAs to determine the position of 2′-*O*-ribose methylation of rRNAs, with methylation occurring on the rRNA transcript, within the complementary region and five nucleotides from the box D or D′ sequence (Cavaillé *et al.*, 1996; Kiss-László *et al.*, 1996; Nicoloso *et al.*, 1996; Tykowski *et al.*, 1996b).

The second major group of snoRNAs (box H/ACA) are characterized by secondary structures consisting of a hairpin–hinge–hairpin–tail. The hinge region contains the conserved box H sequence (AnAnnA) and the 3′ tail has the sequence ACA three nucleotides from the 3′ end of the molecule (Balakin *et al.*, 1996; Ganot *et al.*, 1997a). Both sequences are required for snoRNA accumulation and are postulated to form at least part of a binding site

for protein factors such as Gar1p (Balakin *et al.*, 1996; Ganot *et al.*, 1997a). Box H/ACA snoRNAs are involved in the site-specific pseudouridylation of rRNAs (Bousquet-Antonelli *et al.*, 1997; Ganot *et al.*, 1997b; Tollervey and Kiss, 1997). In the absence of box C and D sequences and terminal repeats, the 5′ and 3′ stem–loops and protein association may protect and stabilize these snoRNAs during processing, such that both box C/D and box H/ACA snoRNAs are processed by a similar mechanism but utilize different intrinsic structures and protein factors.

Although some snoRNAs are expressed from their own promoters, the majority of snoRNAs are intron encoded. In both vertebrates and yeast, only a single snoRNA is found in any particular intron sequence (Maxwell and Fournier, 1995; Kiss and Filipowicz, 1995). Extreme examples of such organization are vertebrate U22 host genes (*UHG*) which contain a different snoRNA gene (*U22–U31*) in each intron, but whose spliced mRNAs lack an open reading frame (ORF) (Tycowski *et al.*, 1996a). The presence of only a single snoRNA per intron is important because processing of both mammalian and yeast intronic snoRNAs is largely splicing dependent and involves exonucleolytic trimming of linearized snoRNA-containing intron lariats (Kiss and Filipowicz, 1995; Cavaillé and Bachellerie, 1996; Kiss *et al.*, 1996; D.Tollervey, personal communication). Nevertheless, in *Xenopus* oocytes, splicing-independent release of intronic U16 and U18 snoRNAs occurs (Caffarelli *et al.*, 1996).

We have shown previously that some plant U14snoRNA genes are tightly clustered and transcribed polycistronically (Leader *et al.*, 1994b). This novel genomic organization of snoRNA genes implies that processing of the pre-U14snoRNA transcript requires endonucleolytic cleavage between individual U14s (Leader *et al.*, 1994b). Here we have characterized sequences flanking the *U14* genes on isolated maize genomic clones and, instead of exonic sequences, we have identified a further four novel plant snoRNA genes (both box C/D and the first plant box H/ACA snoRNA) tightly linked to the *U14* genes. The multiple snoRNA genes are non-intronic and are expressed from an upstream promoter as a polycistronic transcript, from which they are processed. We also demonstrate that processing of U14 is splicing independent. Therefore, the evolution of a novel snoRNA gene organization in plants is complemented by a processing mechanism which allows release of individual snoRNAs from polycistronic transcripts. The organization of these multiple snoRNA genes is so far unique to plants.

Results

5′ Regions flanking maize U14snoRNAs contain other small RNA genes

Isolation of two maize genomic clones (λMzU14.1 and λMzU14.4) containing *U14* genes has been described previously (Leader *et al.*, 1994b). A total of 1.96 and 0.65 kb of 5′ and 3′ sequence flanking the maize *U14.1* gene cluster and 1.57 and 0.32 kb flanking the maize *U14.4* gene have been generated. These sequence data have been submitted to the EMBL database under accession Nos Y11048 and Y11049. A number of short ORFs were observed in the flanking sequences but none matched sequences in the database, or contained suitable combinations of splice site sequences to suggest that the *U14* genes were contained within introns. Four regions of nucleotide homology (90–340 bp) (Figure 1) were observed between the 5′-flanking regions of the two genomic clones, but they did not correspond to potential ORFs nor were putative amino acid sequences conserved.

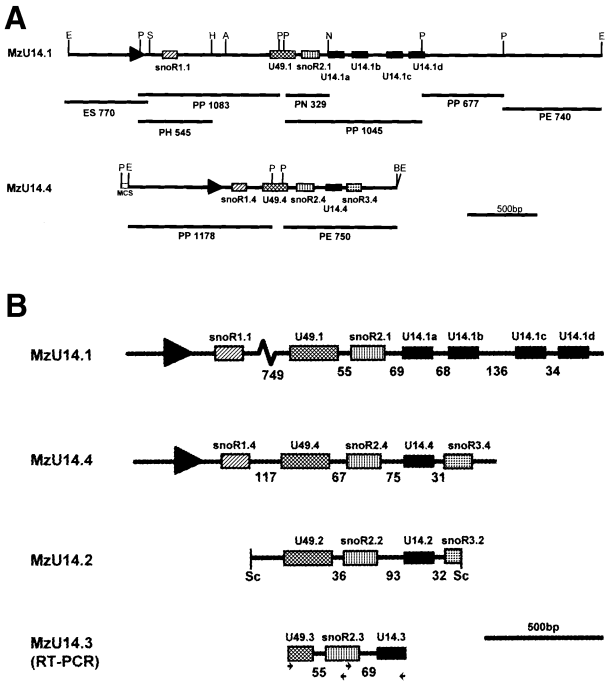

Fig. 1. Organization of maize snoRNA genes. (**A**) The position of snoRNA genes on the 4.3 kbp *Eco*RI fragment of λMzU14.1 and the 2.2 kbp *Eco*RI–*Bam*HI fragment of λMzU14.4 are shown with fragments used in Northern analyses (Figure 2). (**B**) Detailed gene organization of the genomic clones λMzU14.1, λMzU14.2 and λMzU14.4 and the cloned RT–PCR fragment (MzU14.3). Intergenic distances are given, and the gene sizes are presented in Figure 4. The approximate positions of gene-specific primers used in intergenic RT–PCR (Figure 5) are indicated by arrowheads on MzU14.3. Genes are represented by boxes, and flanking and intergenic regions by solid lines. Conserved upstream regions representing putative promoter elements are indicated as large arrowheads. B, *Bam*HI; E, *Eco*RI; H, *Hin*dIII; N, *Nsi*I; P, *Pst*I; S, *Sph*I; Sc, *Sac*I; MCS, multiple cloning site.

To investigate whether sequences flanking the maize U14snoRNA genes corresponded to protein-coding exons, subfragments of the genomic clones (Figure 1A) were hybridized to total maize seedling RNA in Northern analyses. Initial formaldehyde–agarose Northerns failed to detect higher molecular weight RNAs expected for host gene mRNAs but instead clearly showed hybridization to small RNA species. Further analysis was therefore carried out using denaturing 6% polyacrylamide gels to obtain better resolution of the small RNAs (Figure 2). The 1045 bp *Pst*I–*Pst*I fragment containing the majority of the *U14.1* gene cluster hybridized to a major band of ~120 nt (U14), to three bands in the 135–150 nt range, a band of 195 nt and a faint band of 250 nt in maize leaf total RNA (Figure 2, lane 5). The 1083 bp *Pst*I–*Pst*I fragment of λU14.1 lying upstream of the *U14* gene cluster hybridized to two bands of ~90 and 195 nt (Figure 2, lane 8). The 329 bp *Pst*I–*Nsi*I fragment (λU14.1) hybridized to the 195 nt RNA and to the group of bands of 135–150 nt (not shown). The 545 bp *Pst*I–*Hin*dIII fragment (λU14.1)

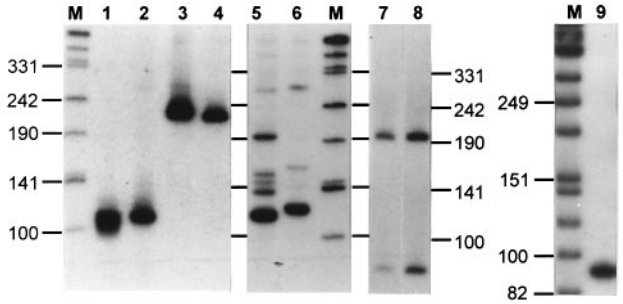

Fig. 2. Northern analysis of maize and potato total RNA. Total plant RNA was probed with subfragments of the λU14.1 and λU14.4 genomic clones indicated in Figure 1A: PP 1045 (lanes 5 and 6), PP 1083 (lanes 8); PP 1178 (lane 7) and PH 545 bp (lane 9). Control hybridizations were performed with maize U5snRNA (lanes 1 and 2) and U3snoRNA probes (lanes 3 and 4) (Leader *et al.*, 1993, 1994a). Hybridization was to maize RNA (lanes 1, 3, 5 and 7–9) and potato RNA (lanes 2, 4 and 6). M = ^{32}P-end-labelled *Hpa*II-digested pUC13 or *Hin*fI-digested ϕX174 DNA markers.

hybridized to the 90 nt RNA only (Figure 2, lane 9). The 770 bp E*co*RI–S*ph*I fragment at the 5′ end of the λU14.1 clone and the *Pst*I–*Pst*I 677 bp and *Pst*I–E*co*RI 740 bp fragments at the 3′ end (Figure 1A) did not hybridize to maize RNA (not shown). Finally, the 1178 bp *Eco*RI–*Pst*I fragment of the λU14.4 genomic clone hybridized to only the 90 and 195 nt RNAs (Figure 2, lane 7). Taken together, these results indicated that other small RNA genes were located upstream of the *U14* genes in both genomic clones. Furthermore, the patterns of hybridization were consistent with three of the four regions of homology between the two genomic clones representing genes encoding the various small RNAs. In addition to the conserved regions 5′ to the *U14* genes, the 3′ regions were searched for potential snoRNA sequences on the basis of identifying box C/D sequences and adjacent inverted repeats, and a putative small RNA gene 3′ to the *U14.4* gene was identified. The identification of these genes was confirmed by Northern analysis with gene-specific probes (results not shown). The genes 5′ to *U14* were called *snoR1*, *U49* and *snoR2*, and that downstream of *U14.4* was called *snoR3*.

Nucleolar localization of the small RNAs

To demonstrate that the above genes encoded snoRNAs, gene-specific antisense probes were used in *in situ* hybridizations on maize root tip tissue sections. The nuclei were counterstained with the DNA-specific dye 4′,6′-diamidino-2-phenylindole (DAPI), showing the nucleoli clearly as dark, unstained holes (Figure 3B). An antisense U14 probe clearly labelled the nucleoli (Figure 3A). The four novel snoRNAs were all localized to the nucleolus, although the detailed patterns of hybridization within the nucleolar region varied among the different genes (Figure 3C–F). SnoR1 was detected at relatively low levels throughout the nucleolus and particularly in a region at the centre of the nucleolus (Figure 3C). The labelling patterns seen with U49 and snoR2 probes (Figure 3D and E) were similar to that seen with U14 (Figure 3A), which we previously showed to label the dense fibrillar component (Beven *et al.*, 1996). SnoR3 was present throughout the nucleolus (Figure 3F). As previously observed with U3, U14 and 7-2/MRP in plant nuclei (Beven *et al.*, 1996),

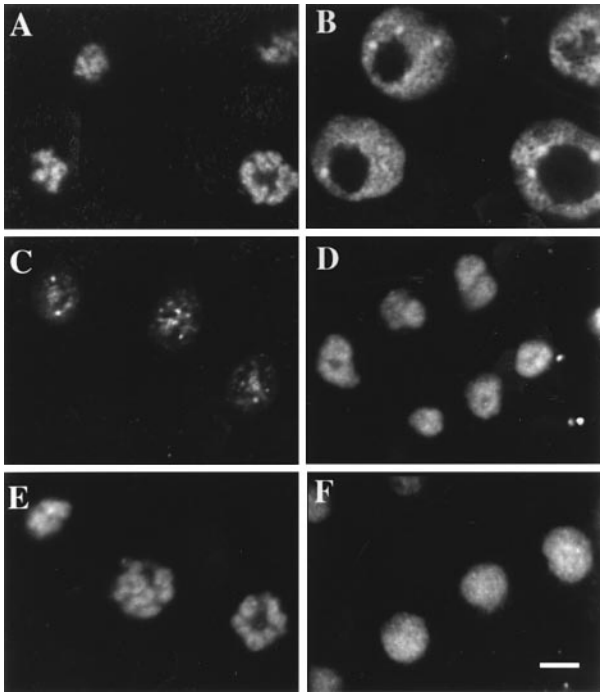

Fig. 3. Localization of individual snoRNAs by *in situ* hybridization and confocal fluorescence microscopy of maize root vibratome sections. (A and B) The same field of cells was labelled with an antisense probe to U14 (**A**) and stained with DAPI (**B**). The DAPI image shows nuclear chromatin, and nucleoli are visible as dark regions in the nuclei (B) which are the sites of *in situ* labelling with U14 (A). (C–F) *In situ* labelling of nucleoli with snoR1 (**C**), U49 (**D**), snoR2 (**E**) and snoR3 (**F**). Bar = 5 μm.

each of the snoRNAs often showed a concentration of labelling in the central nucleolar cavity.

The novel snoRNA genes encode box C/D- and box H/ACA-type snoRNAs

The order of snoRNA genes was conserved between the λMzU14.1 and λMzU14.4 genomic clones (Figure 1B). Southern analysis suggested that there are 8–10 genomic locations of *U14* genes in maize (results not shown). Isolation of a third genomic clone (λMzU14.2) and an RT–PCR product (MzU14.3) (see below) confirmed conservation of gene order at these different loci (Figure 1B). The 5′ and 3′ ends of *snoR1*, *U49* and *snoR2* were determined by a combination of primer extension and RNase A/T1 mapping using probes spanning the whole gene and 3′-specific probes (results not shown). The sequences of the gene variants from the three genomic clones and the RT–PCR product are aligned in Figure 4. *snoR1*, *U49* and *snoR3* genes, like *U14*, encode box C/D-type snoRNAs, containing conserved box C and D sequences, inverted repeats at their 5′ and 3′ ends and regions of complementarity to rRNAs which specify sites of ribose methylation (Figure 4). The regions of complementarity of U14, snoR1 and U49 RNAs coincided with methylation sites conserved in yeast and human rRNAs (Figure 4). The U14 sequence is highly conserved in all eukaryotes and the base-pairing interaction is identical to that of human (Maden, 1990a; Kiss-László *et al.*, 1996; Nicoloso *et al.*, 1996). This sequence would specify ribose methylation at position Cm418 in maize (Messing *et al.*, 1984) and Cm416 in *Arabidopsis* 18S

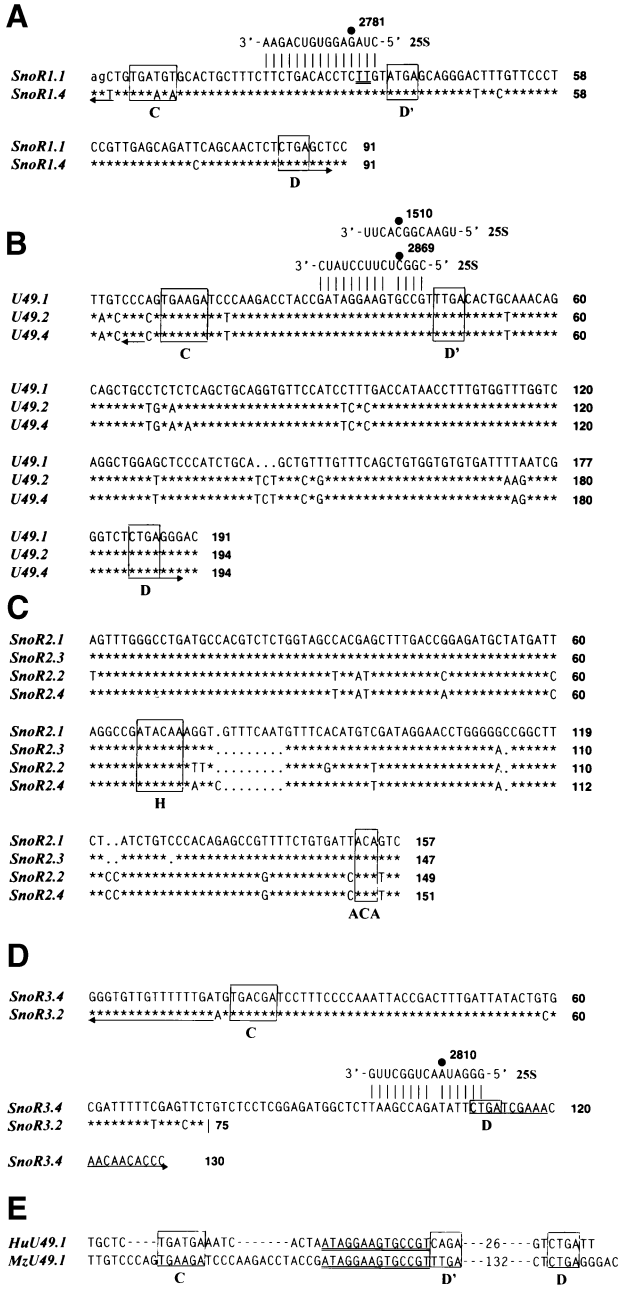

Fig. 4. Sequence comparison and rRNA complementarities of different variants of the maize snoRNAs. (**A**) *snoR1*, (**B**) *U49*, (**C**) *snoR2*, (**D**) *snoR3* and (**E**) alignment of human and maize *U49*. The 5′ and 3′ ends of the *snoR1*, *U49* and *snoR2* gene products were mapped using a combination of primer extension analysis and RNase A/T1 protection with probes specific to the 5′ and 3′ ends of the molecules (not shown). The ends of the *snoR3* products have not been determined experimentally and the sequences presented are defined by the box C/D sequences and extensive inverted repeats. Nucleotide identities are shown by asterisks, the sizes of coding regions are given and inverted repeat sequences are arrowed. Nucleotides which are not present in the mature snoRNA but form potential inverted repeats are shown in lower case. Conserved snoRNA motifs box C, box D, box D′, box H and ACA are boxed. Potential base-pairing interactions between complementary regions of maize snoRNAs and plant 25S rRNA (*Arabidopsis thaliana*) are presented in (A), (B) and (D), and these regions are double underlined in (E). The putative sites of plant rRNA ribose methylation are indicated by black dots. In (A), additional bases in the *snoR1* complementary sequence, compared with that of human *U60*, are double underlined. In (B), two potential regions of complementarity to 25S are shown.

rRNA (Unfried *et al.*, 1989). The sequence of complementarity in snoR1 is similar to that of human U60 but contains two extra nucleotides which would specify ribose methylation of maize 25S rRNA two nucleotides upstream of the equivalent position in human 28S (Figure 4A). This position is not methylated in human 28S rRNA, but is methylated (Gm2788) in yeast 26S (Maden, 1990a). Thus, snoR1 would specify methylation at position Gm2781 in *Arabidopsis* 25S rRNA (Unfried and Gruendler, 1990) (there is currently no published maize 25S sequence). A sequence in the maize U49 is complementary to two regions of 25S rRNA. The first, specifying methylation at position Cm2869 in *Arabidopsis* 25S rRNA, is virtually identical to that of human U49 (Kiss-László *et al.*, 1996), even containing the same U–U mismatch in the base-paired region (Figure 4B). The equivalent position is methylated in both yeast and human. Interestingly, although the maize and human U49s contained the same complementary sequence, besides this sequence and boxes C and D, the RNAs were otherwise unrelated (Figure 4E). The maize U49 (195 nt) is more than twice the length of its human counterpart (71 nt; Kiss-László *et al.*, 1996). The second region of potential complementarity in U49, although only 10 bp long, would specify methylation at position Cm1510 in *Arabidopsis* 25S rRNA (Figure 4B), and the equivalent position in human 28S (Cm2399) is methylated but is unmethylated in yeast (Maden, 1990a). The region of complementarity found in snoR3 would specify methylation at position Am2810 in *Arabidopsis* 25S rRNA (Figure 4D). This position is not known to be methylated in other eukaryotes, although it lies close to a known site of methylation in yeast 25S rRNA (position Gm2811) (Maden, 1990a) such that the complementary sequence overlaps that of yeast snR38 (Kiss-László *et al.*, 1996; Nicoloso *et al.*, 1996). Thus, for the most part, the plant box C/D snoRNAs contain complementary regions to rRNAs at known sites of ribose methylation in other eukaryotes. This strongly suggests that these snoRNAs are also involved in this post-transcriptional modification and that the mechanism of determination of the position of methylation is conserved in plants.

SnoR2 contained the sequence ACA, three nucleotides from the predicted 3′ end of the molecule (Figure 4C). Secondary structure models predict stem–loop structures in the 5′ and 3′ halves of the molecule and a box H sequence (consensus AnAnnA) is present in the putative single-stranded hinge region, as found in other box H/ACA-snoRNAs (Balakin *et al.*, 1996; Ganot *et al.*, 1997a). Thus, snoR2 appears to belong to the box H/ACA class and represents the first plant box H/ACA snoRNA gene to be isolated.

Polycistronic expression of multiple snoRNAs

The original cluster of *U14* genes isolated on genomic clone λMzU14.1 were shown to be expressed on the same polycistron by RT–PCR (Leader *et al.*, 1994b). The discovery of multiple, closely linked snoRNAs suggests that these genes may also be polycistronically expressed. This was supported by the lack of known plant promoter elements, particularly those of UsnRNA genes (TATA box; upstream sequence element and monocot-specific element; Vankan and Filipowicz, 1992; Connelly *et al.*, 1994), and RNA polymerase III elements, between adja-

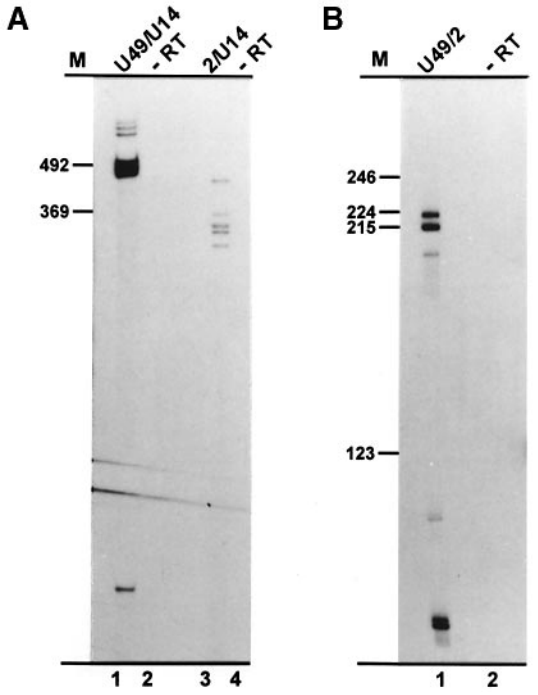

Fig. 5. Detection of polycistronic snoRNA transcripts by RT–PCR. RT–PCR amplification (**A**) between *U14* and *U49* (lane 1), *U14* and *snoR2* (lane 3), and (**B**) between *U49* and *snoR2* (lane 1). Lanes 2 and 4 in (A) and lane 2 in (B) (–RT) are control RT–PCR reactions carried out without reverse transcriptase.

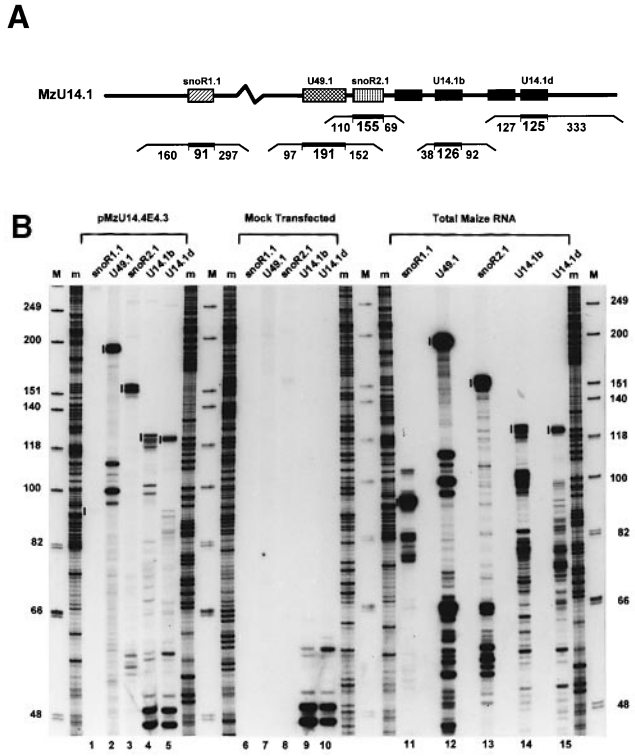

Fig. 6. Expression of multiple snoRNAs from the maize *U14.1* genomic clone in transfected tobacco protoplasts. (**A**) Schematic representation of RNase A/T1 probes specific to each of the snoRNA genes encoded by MzU14.1. The sizes of expected full-length protected products (thicker bars) and flanking sequences complementary to cloned sequences (thin lines) are given. Angled lines represent additional transcribed vector-derived sequences. (**B**) RNase A/T1 analysis of RNA isolated from tobacco protoplasts transfected with pMU14.1E4.3 (lanes 1–5), mock-transfected tobacco protoplasts (lanes 6–10) or maize leaf total RNA (lanes 11–15) with probes specific to *snoR1.1* (lanes 1, 6 and 11), *U49.1* (lanes 2, 7 and 12), *snoR2.1* (lanes 3, 8 and 13), *U14.1b* (lane 4, 9 and 14) or *U14.1d* (lanes 5, 10 and 15). Vertical bars show full-length protected products. M = ^{32}P-end-labelled *Hin*fI-digested ϕX174 DNA markers; m = DNA sequence markers.

cent genes. In addition, the small intergenic distances which ranged in size from 31 to 136 bp (with the exception of *snoR1.1* and *U49.1* which lie 749 bp apart) (Figure 1) further suggested that each individual gene did not contain its own promoter elements. RT–PCR reactions were therefore carried out using primers designed to amplify between different combinations of snoRNA genes: *U49/snoR2*, *U49/U14* and *snoR2/U14* (Figures 1B and 5A and B). RT–PCR using labelled primers produced a number of bands which corresponded largely to the sizes of products expected from the sequences of *U14.1* and *U14.4*. For the *U49/U14* and *snoR2/U14* primer combinations, bands of 489/502 bp and 344/348 bp were expected and bands of similar size were observed (Figure 5A, lanes 1 and 3 respectively). RT–PCR products of ~200–225 bp were obtained with the *U49/snoR2* primer combination (Figure 5B, lane 1), corresponding to the sizes expected from the genomic clones of 213/224 bp. In all cases, a number of bands of varying sizes were obtained, reflecting the amplification of transcripts from different genetic loci. No products were observed in any of the RT–PCR controls lacking reverse transcriptase (Figure 5A, lanes 2 and 4; B, lane 2). An RT–PCR product from the *U49/U14* reaction was cloned and sequenced, and contained the 3′ half of a *U49* allele (*U49.3*), a complete *snoR2* allele (*snoR2.3*) and most of a *U14* allele (*U14.3*) (Figures 1B and 4). Thus, the novel snoRNAs are transcribed on the same polycistronic transcript as the *U14* genes.

Expression of clustered plant snoRNAs from upstream promoters

The tight linkage of the multiple snoRNAs and their polycistronic transcription suggests that the snoRNAs are transcribed from an upstream promoter. While it is feasible that the snoRNA clusters lie within an intron of a protein-coding gene, no ORF (exon) sequence could be identified in either of the U14.1 or U14.4 clones. In addition, a highly conserved region of 340 bp (76% identical) which did not hybridize to small RNAs on Northern analyses (data not shown) was present upstream of the *snoR1* genes and could represent promoter regions. In order to investigate whether the cloned plant snoRNA gene loci contain all of the sequences necessary for expression of the snoRNA polycistrons, tobacco protoplasts were transfected with plasmids containing each of the snoRNA gene clusters. The plasmid pMU14.1E4.3, which encodes *snoR1.1*, *U49.1*, *snoR2.1* and the maize *U14.1* gene cluster (Figure 6A), contained 830 bp of sequence upstream of the *snoR1.1* gene, including the 340 bp region of sequence homology, and 1.52 kbp downstream of the *U14* gene cluster. Protoplasts from three transfections were pooled and RNA isolated. In addition, RNA was isolated from three sets of mock-transfected protoplasts from the same batches used for the transfections, and total maize leaf seedling RNA was included as positive control. RNA from ~2.5×10^5 transfected (Figure 6B, lanes 1–5) or

untransfected (Figure 6B, lanes 6–10) protoplasts or 5 μg of total maize RNA (Figure 6B, lanes 11–15) was used in RNase A/T1 analysis with gene-specific probes to *snoR1.1* (lanes 1, 6 and 11), *U49.1* (lanes 2, 7 and 12), *snoR2.1* (lanes 3, 8 and 13), *U14.1b* (lanes 4, 9 and 14) and *U14.1d* (lanes 5, 10, and 15). Full-length protected products were detected for four of the snoRNAs encoded by pMU14.1E4.3. The *U49.1* probe protected products of ~195 nt (lane 2), the *snoR2.1* probe detected bands of ~150–155 nt (lane 3) and the two *U14* probes each detected bands of ~117–124 nt (lanes 4 and 5). The full-length products protected for each probe corresponded to products protected by total maize RNA (lanes 12–15). The similarity in banding patterns, where some probes produce multiple bands in both tobacco and maize (e.g. *snoR2.1* and *U14.1b*; lanes 3 and 13, and 4 and 14 respectively), suggests that the bands represent transcripts which differ slightly in length due to variation in the extent of processing, as observed previously for other snoRNAs (Balakin *et al.*, 1994; Kiss and Filipowicz, 1995). In addition to products corresponding to transcripts from the genes under study, maize total RNA also protected a number of shorter fragments due to the presence of other sequence variants. No full-length protected products were detected with RNA from mock-transfected protoplasts (lanes 6–10), but shorter fragments representing partial protection of endogenous tobacco snoRNAs were observed with the *U14* probes (lanes 9 and 10). The probe specific to *snoR1.1* protected a faint product of ~103 nt and a much stronger band of 85–90 nt with total maize RNA (lane 11). The latter is the expected size of full-length protected product for *snoR1.1*. A similar product is visible with RNA from transfected protoplasts (lane 1) but is present at a level ~100-fold less than the other snoRNAs. Similar results were obtained with RNase A/T1 protection mapping of RNA from tobacco protoplasts transfected with the plasmid pMzU14.4E2.2 (results not shown). This plasmid contains the *snoR1.4*, *U49.4*, *snoR2.4*, *U14.4* and *snoR3.4* genes with 880 bp upstream of *snoR1.4* and 150 bp downstream of *snoR3.4*. Full-length protected products were obtained for all of the genes but, as with pMzU14.1E4.3 (Figure 6), protected products for *snoR1.4* were extremely faint.

These results indicate that both plasmids contain sufficient sequences for expression and processing of at least four of the snoRNAs of the gene clusters. In addition, the results demonstrate that the maize (monocot) promoter functions in tobacco (dicot) protoplasts. Given the fairly limited sequences upstream of the snoRNAs on the two plasmids, it seems likely that the conserved sequences may form at least part of the promoter region directing polycistronic transcription of the snoRNA gene clusters. The absence of strong protected products corresponding to *snoR1.1* may indicate that this snoRNA is unstable in tobacco protoplasts (although this appears unlikely from the presence of strong products in total maize RNA). Alternatively, the transcription start site of the maize promoters may be inaccurate in tobacco protoplasts, interfering with expression of *snoR1.1* by, for example, truncating the 5′ end of the snoRNA.

U14.4 is processed from both non-intronic and intronic transcripts

Vertebrate snoRNAs are each contained within a single intron and, with the exception of some *Xenopus* snoRNAs,

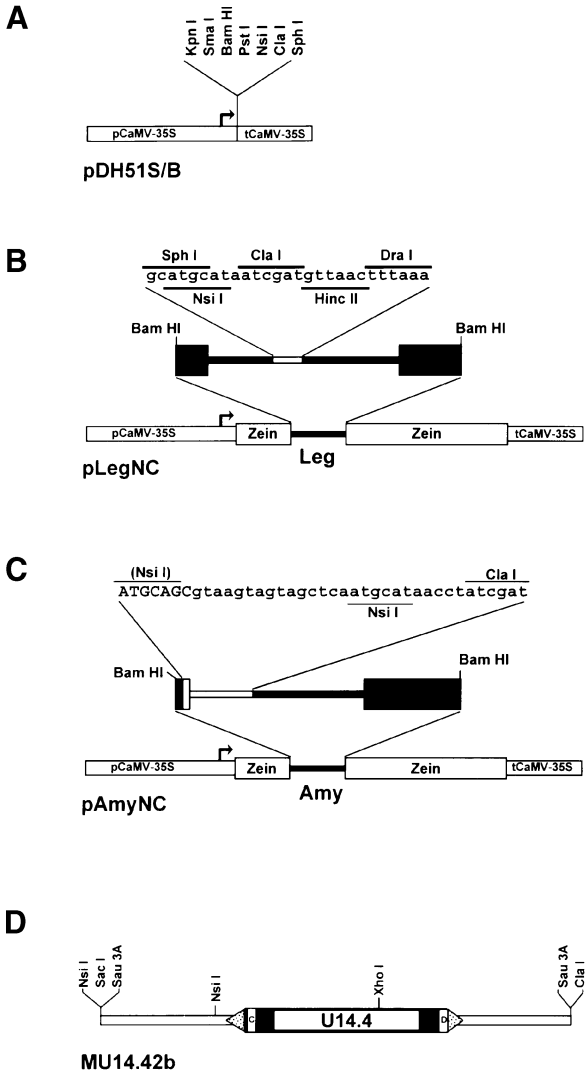

Fig. 7. Constructs for the expression of the maize *U14.4* gene in different transcriptional contexts. The *U14.4* gene fragment MU14.42b (**D**) *was introduced as an Nsi*I–*Cla*I fragment into (**A**) pDH515S/B for non-intronic transcription, (**B**) pLegNC and (**C**) pAmyNC for intronic transcription.

are processed in a largely splicing-dependent manner involving exonuclease activity. The small intergenic distances and the lack of suitable intron splice sites, from either classical or AU–AC introns (Wu *et al.*, 1996), between the clustered maize snoRNA genes suggest that the individual snoRNAs are not separated by mini-exons and must, therefore, be processed from the polycistronic transcript without the requirement for splicing. To demonstrate that processing was splicing independent, a 308 bp *Nsi*I–*Cla*I fragment containing the maize *U14.4* gene with 83 bp upstream of box C and 64 bp downstream of box D (Figure 7D; MU14.42b) was introduced into three different plant expression vectors: pDH51S/B (a modified version of pDH51; Pietrzak *et al.*, 1986), and pLegNC and pAmyNC [modified versions of pL and pA (Simpson *et al.*, 1996); see Materials and methods; Figure 7A–C]. The U14.4 construct would be transcribed as part of an mRNA transcript from the cauliflower mosaic virus (CaMV) 35S promoter either as a non-intronic transcript in pDH51S/B (pDMU14.42b) or within the efficiently

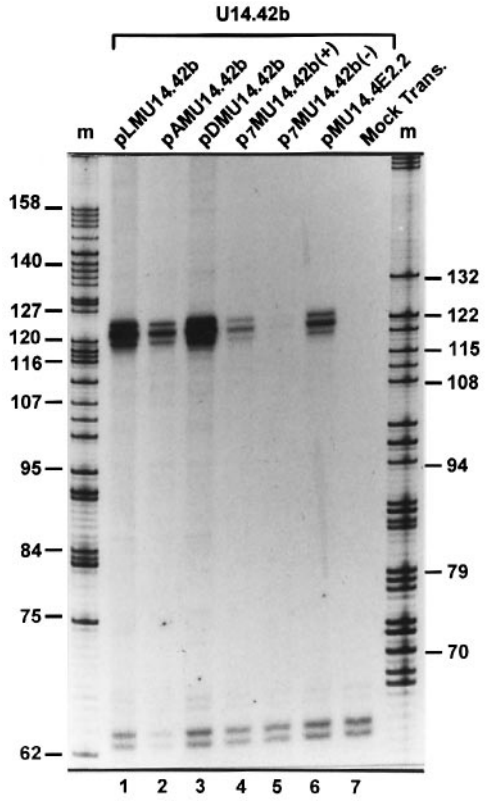

Fig. 8. Processing of U14.4 snoRNA from intronic and non-intronic transcripts. RNase A/T1 analyses of RNA from tobacco protoplasts transfected with different plasmids containing the MU14.42b fragment. Lanes 1–3, the U14.4 fragment in different transcriptional contexts: in the legumin intron in pLegNC (lane 1), in the amylase intron in pAmyNC (lane 2) or directly transcribed from the CaMV 35S promoter in pDH51S/B (lane 3). Lanes 4 and 5, U14.4 in pGEM3Zf(+) in both orientations; lane 6, U14.4 genomic clone; and lane 7, RNA from mock-transfected protoplasts. M = DNA sequence markers.

spliced legumin intron in pLegNC (pLMU14.42b) or the poorly spliced amylase intron in pAmyNC (pAMU14.42b) (Figure 7). The legumin and amylase introns differ in their efficiencies of splicing in tobacco protoplasts: >90 and 2–5% efficiency respectively (Simpson *et al.*, 1996). RNA was extracted and analysed by RNase A/T1 protection mapping (Figure 8) with a labelled antisense probe complementary to the *U14.4* insert. All three expression constructs, pLMU14.42b (lane 1), pAMU14.42b (lane 2) and pDMU14.42b) (lane 3), produced three major protected products of 119, 121 and 123 nt, where the 121 nt product was the most abundant. Protoplasts from the same batch were transfected with the plasmid pMU14.4E2.2 and full-length U14.4 was again detected (Figure 8A, lane 6). In initial experiments, the plasmid pMU14.42b(+), containing the U14.4 insert in pGEM7Zf(–), was included as a negative expression control and gave weak fully processed U14 bands, as seen in Figure 8 (lane 4). However, reversing the insert orientation to give pMU14.42b(–) gave virtually no product (lane 5). Thus, the *U14.4* gene fragment does not contain active promoter sequences and, presumably, the weak products seen with one orientation are due to processing of transcripts initiated by sequences within pGEM7Zf(–).

In conclusion, fully processed U14.4 accumulated from both intronic and non-intronic transcripts. Clearly, although U14 accumulated from intronic transcripts, it was not possible to distinguish between direct processing from the mRNA transcript or from spliced, debranched intron lariats, and presumably both can occur. The high levels of accumulation of U14 from the non-intronic transcripts (Figure 8, lane 3) contrasts the reduced accumulation of vertebrate snoRNAs *in vivo* using analogous constructs (Kiss *et al.*, 1995; Cavaillé and Bachellerie, 1996) and thereby demonstrates that processing of U14 is splicing independent in plant cells.

Discussion

We have shown here that U14snoRNA genes in maize are found in gene clusters alongside other different snoRNA genes. In addition, we have isolated four other clusters of at least four snoRNA genes from *Arabidopsis* (D.J.Leader and J.W.S.Brown, unpublished results). Therefore, the occurrence of multiple, different snoRNA genes in tightly linked groups is a general feature of plant snoRNA gene organization which is as yet unique to plants. The detection of polycistronic transcripts suggests that the snoRNA gene clusters are transcribed together as pre-snoRNA transcripts, which further implies that release of individual snoRNAs will require endonuclease activity. We have demonstrated that sequences upstream of the gene clusters are sufficient for production of a series of different fully processed snoRNAs. Thus, some plant snoRNAs have a mode of snoRNA expression distinct from the expression strategies employed by animals and yeast: transcription from classical promoter elements or intron encoded (Maxwell and Fournier, 1995).

The snoRNA genes found linked to *U14* encode either box C/D snoRNAs or box H/ACA snoRNAs. The box C/D snoRNAs contain regions of complementarity to rRNAs at sites of ribose methylation in human and yeast. Very little information exists on plant rRNA ribose methylation (Maden, 1990b) and to date no methylation sites have been mapped. However, detailed studies on yeast, human and *Xenopus* have shown many methylation sites to be conserved although each organism can contain unique sites (Maden, 1990a). These snoRNAs are likely, therefore, to function in rRNA ribose methylation, as recently demonstrated for a number of animal snoRNAs (Cavaillé *et al.*, 1996; Kiss-László *et al.*, 1996; Nicoloso *et al.*, 1996; Tykowski *et al.*, 1996b) and strongly suggests that the mechanism of determination of sites of ribose methylation is conserved in plants. The *snoR2* alleles are the first box H/ACA-type snoRNA genes to have been isolated from plants. A number of examples have been identified in vertebrates and yeast (Balakin *et al.*, 1996; Ganot *et al.*, 1997a) and they appear to act as guide RNAs determining sites of pseudouridine modifications in rRNAs (Bousquet-Antonelli *et al.*, 1997; Ganot *et al.*, 1997b; Tollervey and Kiss, 1997). The likely conservation of these snoRNAs in plants again suggests conservation of function in rRNA modification.

The small distances between adjacent snoRNA genes and the lack of known promoter elements suggest that they are not transcribed individually from classical snRNA promoters. The absence of classical and AU–AC intron splice site consensus sequences in the intergenic regions

further suggests that unlike the majority of vertebrate and many yeast snoRNAs, the plant snoRNAs are not intron encoded. The detection of polycistronic RNAs by RT–PCR and of fully processed snoRNAs from along the gene clusters following transfection of the maize genomic fragments into tobacco protoplasts showed that the gene clusters were transcribed from upstream promoters. The activity of the maize promoter regions in tobacco is of interest because of the clear distinction which exists between the genes encoding spliceosomal snRNAs and nucleolar U3 and MRP snoRNAs in the two angiosperm classes (monocots and dicots) (Connelly *et al.*, 1994). Expression of these genes in monocots requires a unique element, the monocot-specific promoter (MSP) element which is absent in dicot genes and is not needed for expression in dicots (Connelly *et al.*, 1994). Transcription from both genomic fragments, each containing ~800 bp of sequence upstream of the *snoR1* genes, suggests that the 340 bp conserved regions in the two promoters may represent or contain elements essential to expression. These regions did not contain classical snRNA promoter elements, but contained a putative TATA box at around –200 in both genomic clones and were GC rich (62%), reminiscent of housekeeping genes (Dynan, 1986).

The expression of polycistronic pre-snoRNAs suggests that processing is splicing independent, and this was shown by the successful processing and accumulation of U14.4 from a transcript lacking exons and transcribed from the CaMV 35S promoter. In contrast, accumulation of vertebrate U17 and U20 was greatly reduced when analogous constructs, consisting of the snRNA gene and flanking intron sequences, but either lacking exons or flanked by exons with mutated splice sites, were introduced into simian COS or mouse cells respectively (Kiss and Filipowicz, 1995; Cavaillé and Bachellerie, 1996). Notably, the maize pre-snoRNAs contain both box C/D and box H/ACA snoRNAs, which are processed from the same transcript. While there are clear structural differences between the classes of snoRNAs, their processing in vertebrates and yeast follows essentially the same pathway: exonucleolytic degradation of flanking sequences following intron debranching, with the snoRNA protected from degradation by secondary structural features and binding of proteins (Kiss and Filipowicz, 1995; Balakin *et al.*, 1996; Caffarelli *et al.*, 1996; Cavaillé and Bachellerie, 1996; Kiss *et al.*, 1996; Watkins *et al.*, 1996; Ganot *et al.*, 1997a; Xia *et al.*, 1997). Furthermore, the same processing machinery may be utilized for both snoRNA classes, and in yeast involves exonucleases which are also required for mRNA degradation and exonucleolytic trimming of precursor RNAs in pre-rRNA processing (Lafontaine and Tollervey, 1995; Tollervey and Kiss, 1997). Processing of different classes of snoRNA from the same transcript in plants suggests that the snoRNP assembly factors and processing machinery required by each type of snoRNA are in the same cellular and, presumably, nuclear/nucleolar region.

The non-intronic clustering of the plant snoRNA genes demonstrates that plants have evolved a gene organization which differs fundamentally from the intron-encoded organization of the majority of vertebrate and some yeast snoRNAs (Maxwell and Fournier, 1995). The transcription of a polycistronic pre-snoRNA and splicing-independent processing suggest that endonucleolytic activity releases individual snoRNAs, probably still containing flanking sequences which are removed subsequently by exonucleolytic activity. Although some examples of splicing-independent release of intronic snoRNAs (U16 and U18 in *Xenopus*) also require an endonucleolytic activity to cleave in the flanking sequences prior to exonuclease trimming (Caffarelli *et al.*, 1996), the organization of plant snoRNAs is more reminiscent of that of yeast *snoR190* and *U14* (Zagorski *et al.*, 1988). These genes are closely linked and are likely to be expressed as a dicistron requiring endonucleolytic cleavage for maturation of these snoRNAs (Lafontaine and Tollervey, 1995; D.Tollervey and M.Fournier, personal communications). Thus, yeast may reflect an ancestral arrangement of snoRNA gene organization in containing snoRNA genes transcribed from their own promoters, intron-encoded snoRNAs produced in a splicing/exonuclease-dependent manner and co-transcribed snoRNAs requiring endonuclease processing. To date, polycistronic snoRNA genes have not been found in vertebrates, and they, therefore, appear to have lost the ability to process such transcripts.

Finally, the intronic location of these snoRNA genes permits co-ordinated regulation of the host protein and the various snoRNAs via splicing-dependent release of snoRNAs by exonucleolytic digestion of debranched lariats (Kiss and Filipowicz, 1995; Kiss *et al.*, 1996). The plant snoRNA gene organization and mode of expression permits the co-ordinated expression of multiple, different snoRNAs (including members of both of the major classes of snoRNAs) through transcription of polycistronic pre-snoRNAs. This mode of snoRNA production represents the second major difference in snoRNA transcription between plants and animals: U3snoRNA is transcribed by RNA polymerase III in plants and RNA polymerase II in animals (Kiss *et al.*, 1991). How transcription of snoRNAs is coordinated with expression of nucleolar or ribosomal proteins and the nature of the processing machinery remain to be addressed.

Materials and methods

Oligonucleotides

The following oligonucleotides were used: snS1F (5′-AAGGATCCGATAGCCTCCTAGTAGCTGTGATGTGC-3′); snS1R (5′-AAGAATTCGAGGGAGGAGCTCAGAGAGTTGC-3′); snL1F (5′-AAGAATTCTTGTTTTAGTTTCCTGAATCTGCC-3′); snL1R (5′-AAGAATTCTACAGTCCCTCAGAGACC-3′); snM1F (5′-AACTGCAGTAGTTTGGGCCTGATGC-3′); snM1R (5′-AAGAATTCATGTAATCACAGAAAACGGC-3′); snoR3Hd (5′-GGGAAGCTTGTCCAGCTCTTCATCGTTC-3′); L1Ns1 (5′-GGATGCATTTGCCTAGCTTCCTGATG-3′); snoR3Spe (5′-CCCACTAGTTTCGATCAGAATATCTGGC-3′); DH51Sph (5′-CATCGATGATGCATCTGCAG-3′); DH51Bam (5′-GATCCTGCAGATGCATCATCGATGCATG-3′); LegDra (5′-CATAATCGATGTTAACTTT-3′); LegSph (5′-AAAGTTAACATCGATTATGCATG-3′); AMYNC (5′-ACCCTGCAGCGTAAGTAGTAGCTCAATGCATAACCTTATCGATAACATTTC-3′); MU14.1NsSp (5′-CCGCATGCATTAGCTTGAACTAGAATGCCTGCG-3′). Incorporated restriction sites are underlined.

Construction of snoRNA gene plasmids

For initial Northern analyses, various fragments of the genomic clones were isolated and subcloned (Figure 1A) by standard techniques. Plasmids for production of antisense probes specific to each of the snoRNA genes for use in *in situ* hybridization were prepared by PCR from appropriate subclones and contained the coding sequence of the gene with minimal flanking sequences. The *snoR1.1*-specific plasmid pgSnoR1.1 contained

a 125 bp PCR product generated with primers snS1F and snS1R and subcloned in pGEM3Zf+. Plasmid p7U49.1 contained a 191 bp *Eco*RI–*Bam*HI fragment generated by PCR with primers snL1F and snL1R subcloned in pGEM7Zf–. pgSnoR2.1 contained a 163 bp *Eco*RI fragment generated by PCR with primers snM1F and snM1R. An antisense probe to the *snoR3.4* gene was transcribed from plasmid pgSnoR3.4 which contained a *Hin*dIII site immediately upstream of the *snoR3.4* 5′ inverted repeat sequence as a result of PCR with primer snoR3Hd and universal sequencing primer. In addition to the coding sequence of *snoR3.4*, pgSnoR3.4 also contains ~170 bp downstream of the gene to an *Eco*RI site derived from the lambda vector. The plasmid pgMU14.1d has been described previously (Leader *et al.*, 1994) and encodes the *U14.1d* gene from the maize *U14.1* gene cluster.

Additional transcription plasmids for the preparation of RNase A/T1 probes contained more extensive 3′- and 5′-flanking sequences. Plasmid pgS1PH contained a 548 bp *Hin*dIII–*Pst*I fragment containing the complete *snoR1.1* coding sequence and 160 and 297 bp of 5′- and 3′-flanking sequence. Plasmid pgU49.1 was made using PCR with the primer L1Ns1 which introduced an *Nsi*I site 104 bp upstream of the *U49.1* gene. Digestion with *Nsi*I and *Taq*I produced a 443 bp fragment which was subcloned in pGEM3Zf(+). Probes specific to *snoR2.1* were transcribed from pgM1PN and were complementary to *snoR2.1*, a 55 bp fragment of *U49.1*, 55 bp of intergenic sequence and 69 bp of sequence downstream of *snoR2.1*. Plasmid psnoR1.4 contained a 357 bp *Sac*I–*Bgl*II fragment cloned in pGEM3Zf(+).

Construction of expression vectors

The plasmid pDH51S/B was a derivative of pDH51 (Pietrzak *et al.*, 1986) in which part of the polylinker between the *Sph*I and *Bam*HI sites was replaced with a novel polylinker produced by annealing the oligonucleotides DH51Sph and DH51Bam. The new polylinker contained *Bam*HI, *Pst*I, *Nsi*I, *Cla*I and *Sph*I sites downstream of the CaMV 35S promoter (Figure 7A). Plasmids pAmyNC and pLegNC were derivatives of the previously described plasmids pA and pL respectively (Simpson *et al.*, 1996). Both of these plasmids were based on pDH51, into which an intron-less zein gene was introduced between the CaMV 35S promoter and terminator regions. This gene contained a unique *Bam*HI site into which either an amylase intron (pA) or a legumin intron (pL) were introduced. Plasmid pLegNC was generated by replacing a 27 nt *Dra*I–*Sph*I fragment in the legumin intron with an artificial linker made by annealing oligonucleotides LegDra and LegSph. This generated *Nsi*I and *Cla*I sites 52 and 59 nt downstream of the legumin 5′ splice site respectively (Figure 7B) and resulted in the substitution of five nucleotides within the intron with no change in AU content. Suitable restriction sites were inserted into the amylase intron by PCR with pA, using primers AMYNC and O9 (Simpson *et al.*, 1996) which is complementary to part of the downstream zein sequence. Following digestion with *Pst*I and *Bam*HI, the 189 bp PCR fragment was used to replace the equivalent *Nsi*I–*Bam*HI fragment of pA. This destroyed the *Nsi*I site located seven nucleotides upstream of the amylase intron 5′ splice site and generated novel *Nsi*I and *Cla*I sites 16 and 27 nt downstream of the 5′ splice site respectively (Figure 7C). These changes resulted in the substitution of seven nucleotides within the intron, although the overall AU content was only increased by 1% and the length of the intron was maintained.

The maize *U14.4* gene was subcloned into the above vectors as the MU14.42b fragment (Figure 7C). This initially was isolated as a 263 bp *Sau*3AI fragment, containing 93 bp of sequence upstream of box C and 64 bp of sequence downstream of box D, and was subcloned into the *Bam*HI site of pGEM7Zf(–) to give plasmids pMU14.42b(+) and pMU14.42b(–) differing in the orientation of the cloned fragment. For cloning into the expression vectors, a 302 bp *Nsi*I–*Cla*I fragment was isolated from pMU14.42b(+). This fragment contained an additional 39 bp of sequence derived from the pGEM polylinker. Sense and antisense probes for RNase A/T1 were produced from pMU14.42b(+).

RNA isolation and Northern analysis

RNA was isolated from leaf material of maize and potato as described previously (Chomczynski and Sacchi, 1987). Ten μg of total maize seedling RNA was separated on a denaturing 6% polyacrylamide gel and electroblotted onto Hybond N+. The blot was cut and individual strips were hybridized to ^{32}P-labelled probes. Probes were labelled with [^{32}P]dCTP using a random priming kit (Boehringer-Mannheim).

In situ hybridization

Root tips were excised from 3-day-old maize seedlings, fixed in 4% formaldehyde in a buffer containing 50 mM PIPES–KOH, pH 6.9, 5 mM EGTA, 5 mM $MgSO_4$. After thorough washing with TBS (25 mM Tris–HCl, pH 7.4, 140 mM NaCl and 30 mM KCl), 50 μm sections were cut with a vibratome and dried onto microscope slides. *In situ* hybridization was carried out as described previously (Beven *et al.*, 1995, 1996) with digoxigenin-labelled antisense RNA probes. The hybridized probe was detected by a primary anti-digoxigenin FAB fragment and secondary fluorescein- or Cy3-coupled antibody. Confocal data sets were collected using a BioRad MRC600 or MRC 1000 UV confocal microscope.

Protoplast transfections

Protoplasts were isolated from fully expanded leaves of 6- to 8-week-old *Nicotiana tabacum* cv. Xanthii as described by Chupeau *et al.* (1994). Approximately 2.5×10^5 protoplasts were transfected with 15 μg of doubly caesium chloride-purified plasmid DNA and incubated for 24 h prior to RNA extraction. Protoplasts were also co-transfected with the plasmid pMzU5.3 which contains the maize *U5.3* gene (Connelly *et al.*,1994). RNA was isolated from protoplasts or from plant tissues by the guanidinium isothiocyanate method according to Goodall *et al.* (1990). RNA was treated twice with RNase-free DNase, prior to use in RNase A/T1 and RT–PCR analysis.

RT–PCR and RNase A/T1 protection mapping

RT–PCR was carried out with 5′-end-labelled primers as described previously (Simpson *et al.*, 1996). RNase A/T1 protection analysis was performed as described (Goodall *et al.*, 1990) using RNA isolated from ~2.5×10^6 tobacco protoplasts or 5 μg of total maize seedlings (cultivar Kelvedon Glory). ^{32}P-labelled antisense or sense probes for RNase A/T1 protection analysis were transcribed from plasmids with either T7 or SP6 RNA polymerase following appropriate linearization of the plasmids. Probes were treated with RQ1-RNase-free DNase (Promega), gel purified and protected products were separated on an 8% polyacrylamide denaturing gel with sequencing reactions or a ^{32}P-end-labelled *Hin*fI-digested ϕX174 DNA as size markers.

Acknowledgements

We thank Drs Maurille Fournier, Tamas Kiss and David Tollervey for making unpublished results available. This research was supported by Gene Shears PTY Ltd, by grant-in aid from the Scottish Office Agriculture, Environment and Fisheries Department and by grant-in-aid to the John Innes Centre from the Biotechnology and Biological Sciences Research Council.

References

Bachellerie,J.-P., Michot,B., Nicoloso,M., Balakin,A., Ni,J. and Fournier M.J. (1995a) Antisense snoRNAs: a family of nucleolar RNAs with long complementarities to rRNA. *Trends Biochem. Sci.*, **20**, 261–264.

Bachellerie,J.-P., Nicoloso,M., Qu,L.-H., Michot,B., Caizergues-Ferrer,M., Cavaillé,J. and Renalier,M.-H. (1995b) Novel intron-encoded small nucleolar RNAs with long sequence complementarities to mature rRNAs involved in ribosome biogenesis. *Biochem. Cell. Biol.*, **73**, 835–843.

Balakin,A.G., Lempicki,R.A., Huang,G.M. and Fournier,M.J. (1994) *Saccharomyces cerevisiae* U14 small nuclear RNA has little secondary structure and appears to be produced by post-transcriptional processing. *J. Cell Biol.*, **269**, 739–746.

Balakin,A.G., Smith,L. and Fournier,M.J. (1996) The RNA world of the nucleolus: two major families of small RNAs defined by different box elements with related functions. *Cell*, **86**, 823–834.

Beven,A.F., Simpson,G.G., Brown,J.W.S. and Shaw,P.J. (1995) The organization of spliceosomal components in nuclei of higher plants. *J. Cell Sci.*, **108**, 509–518

Beven,A.F., Lee,R., Razaz,M., Leader,D.J., Brown,J.W.S. and Shaw,P.J. (1996) The organization of ribosomal RNA processing correlates with the distribution of nucleolar snRNAs. *J. Cell Sci.*, **109**, 1241–1251.

Bousquet-Antonelli,C., Henry,Y., Gélunge,J.-P., Caizergues-Ferrer,M. and Kiss,T. (1997) A small nucleolar RNP protein is required for pseudouridylation of eukaryotic ribosomal RNAs. *EMBO J.*, **16**, 4769–4775.

Caffarelli,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 reaction and the stability of the mature snoRNA. *EMBO J.*, **15**, 1121–1131.

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.

Cavaillé,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.

Chomczynski,P. and Sacchi,N. (1987) Single-step method of RNA isolation by acid guanidinium thiocyanate–phenol–chloroform extraction. *Anal. Biochem.*, **162**, 156–159.

Chupeau,Y., Bourgin,J.P., Missonier,C., Dorion,N. and Morel,G. (1974) Préparation et culture de protoplastes de divers *Nicotiana*. *C.R. Acad. Sci. Paris, Ser. D*, **278**, 1565–1568.

Connelly,S. Marshallsay,C., Leader,D.J., Brown,J.W.S. and Filipowicz,W. (1994) Small nuclear RNAs genes transcribed by either RNA polymerase II or RNA polymerase III in monocot plants share three promoter elements and use a strategy to regulate gene expression different from that used by their dicot plant counterparts. *Mol. Cell. Biol.*, **14**, 5910–5919.

Dynan,W.S. (1986) Promoters for house-keeping genes. *Trends Biochem. Sci.*, **11**, 196–197.

Eichler,D.C. and Craig,N. (1995) Processing of eukaryotic ribosomal RNA. *Prog. Nucleic Acid Res. Mol. Biol.*, **49**, 197–239.

Ganot,P., Caizergues-Ferrer,M. and Kiss,T. (1997a) The family of box ACA small nucleolar RNAs is defined by an evolutionarily defined secondary structure and ubiquitous sequence elements essential for RNA accumulation. *Genes Dev.*, **11**, 941–956.

Ganot,P., Bortolin,M.-L. and Kiss,T. (1997b) Site-specific pseudouridine formation in pre-ribosomal RNA is guided by small nucleolar RNAs. *Cell*, **89**, 799–809.

Goodall,G.J., Wiebauer,K. and Filipowicz,W. (1990) Analysis of pre-mRNA processing in transfected plant protoplasts. *Methods Enzymol.*, **181**, 148–161.

Huang,G.M., Jarmolowski,A., Struck,J.C. 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.

Kiss,T. and Filipowicz,W. (1995) Exonucleolytic processing of small nucleolar RNAs from pre-mRNA introns. *Genes Dev.*, **9**, 1411–1424.

Kiss,T., Marshallsay,C. and Filipowicz,W. (1991) Alteration of the RNA polymerase specificity of U3 snRNAs during evolution and *in vitro*. *Cell*, **65**, 517–526.

Kiss,T., Bortolin,M.-L. and Filipowicz,W. (1996) Characterization of the intron-encoded U19 RNA, a new mammalian small nucleolar RNA that is not associated with fibrillarin. *Mol. Cell. Biol.*, **16**, 1391–1400.

Kiss-László,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.

Lafontaine,D. and Tollervey,D. (1995) Trans-acting factors in yeast pre-rRNA and pre-snoRNA processing. *Biochem. Cell. Biol.*, **73**, 803–812.

Leader,D.J., Connelly,S., Filipowicz,W., Waugh,R. and Brown,J.W.S. (1993) Differential expression of U5snRNA gene variants in maize (*Zea mays*) protoplasts. *Plant Mol. Biol.*, **21**, 133–143.

Leader,D.J., Connelly,S., Filipowicz,W. and Brown,J.W.S. (1994a) Characterization and expression of a maize U3snRNA gene. *Biochim. Biophys. Acta*, **1219**, 145–147.

Leader,D.J., Sanders,J.F., Waugh,R., Shaw,P.J. and Brown,J.W.S. (1994b) Molecular characterization of plant U14 small nucleolar RNA genes: closely linked genes are transcribed as a polycistronic U14 transcript. *Nucleic Acids Res.*, **22**, 5196–5200.

Lygerou,Z., Allmang,C., Tollervey,D. and Seraphin,B. (1996) Accurate processing of a eukaryotic precursor ribosomal RNA by ribonuclease MRP *in vitro*. *Science*, **272**, 268–270.

Maden,B.E.H. (1990a) The numerous modified nucleotides in eukaryotic ribosomal RNA. *Prog. Nucleic Acid Res. Mol. Biol.*, **39**, 241–303.

Maden,B.E.H. (1990b) The modified nucleotides in ribosomal RNA of man and other eukaryotes. In Gehrke,C.W. and Kuo,K.C.T. (eds), *Chromatography and Modification of Nucleosides*. Elsevier, Amsterdam, pp. B265–B301.

Maxwell,E.S. and Fournier,M.J. (1995) The small nucleolar RNAs. *Annu. Rev. Biochem.*, **35**, 897–934.

Messing,J., Carlson,J., Hagen,G., Rubenstein,I. and Oleson A. (1984) Cloning and sequencing of the ribosomal RNA genes in maize: the 17S region. *DNA*, **3**, 31–40.

Nicoloso,M., Qu,L.-H., Michot,B. and Bachellerie,J.-P. (1996) Intron-encoded, antisense small nucleolar RNAs: the characterization of nine novel species points to their direct role as guides for the 2′-*O*-ribose methylation of rRNAs. *J. Mol. Biol.*, **260**, 178–195.

Peculis,B.A. and Steitz,J.A. (1994) Sequence and structural elements critical for U8 snRNP function in *Xenopus* oocytes are evolutionarily conserved. *Genes Dev.*, **8**, 2241–2255.

Pietrzak,M., Shillito,R., Hohn,T. and Potrykus,I. (1986) Expression in plants of two bacterial antibiotic resistance genes after transformation with a new plant expression vector. *Nucleic Acids Res.*, **14**, 5857–5868.

Simpson,C.G., Clark,G., Davidson,D., Smith,P. and Brown,J.W.S. (1996) Mutation of putative branchpoint consensus sequences in plant introns reduces splicing efficiency. *Plant J.*, **9**, 369–380.

Sollner-Webb,B., Tycowski,K.T. and Steitz,J.A. (1995) Ribosomal RNA processing in eukaryotes. In Zimmerman,R.A. and Dahlberg,A.E. (eds), *Ribosomal RNA*. CRC Press, Boca Raton, FL, pp. 469–490.

Terns,M.P., Grimm, C., Lund,E. and Dahlberg,J.E. (1995) A common maturation pathway for small nucleolar RNAs. *EMBO J.*, **14**, 4860–4871.

Tollervey,D. and Kiss,T. (1997) Function and synthesis of small nucleolar RNAs. *Curr. Opin. Cell Biol.*, **3**, 337–342.

Tycowski,K.T., Shu,M.-D. and Steitz,J.A. (1993) A small nucleolar RNA is processed from an intron of the human gene encoding ribosomal protein S3. *Genes Dev.*, **6**, 1120–1130.

Tycowski,K.T., Shu,M.-D. and Steitz,J.A. (1996a) A mammalian gene with introns instead of exons generating stable RNA products. *Nature*, **379**, 464–466.

Tycowski,K.T., Smith,C.M., Shu,M.-D. and Steitz,J.A. (1996b) A small nucleolar RNA requirement for site-specific ribose methylation of rRNA in *Xenopus*. *Proc. Natl Acad. Sci. USA*, **93**, 14480–14485.

Unfried,I. and Gruendler,P. (1990) Nucleotide sequence of the 5.8S and 25S rRNA genes and of the internal transcribed spacers from *Arabidopsis thaliana*. *Nucleic Acids Res.*, **18**, 4011.

Unfried,I., Stocker,U. and Greundler,P. (1989) Nucleotide sequence of the 18S rRNA gene from *Arabidopsis thaliana* Col0. *Nucleic Acids Res.*, **17**, 7513.

Vankan,P. and Filipowicz,W. (1992) A UsnRNA gene-specific upstream element and a –30 TATA box are required for transcription of the U2snRNA gene of *Arabidopsis thaliana*. *EMBO J.*, **8**, 3875–3882.

Venema,J. and Tollervey,D. (1995) Processing of pre-ribosomal RNA in *Saccharomyces cerevisiae*. *Yeast*, **11**, 1629–1650.

Watkins,N.J., Leverette,R.D., Xia,L., Andrews,M.T. and Maxwell,E.S. (1996) Elements essential for processing intronic U14 snoRNA are located at the termini of the mature snoRNA sequence and include conserved nucleotide boxes C and D. *RNA*, **2**, 118–133.

Wu,H.-J., Gaubier-Comella,P., Delseny,M., Grellet,F., Van Montagu,M. and Rouzé,P. (1996) Non-canonical introns are at least 109 years old. *Nature Genet.*, **14**, 383–384.

Xia,L., Watkins,N.J. and Maxwell,E.S. (1997) Identification of specific nucleotide sequences and structural elements required for intronic U14snoRNA processing. *RNA*, **3**, 17–26.

Zagorski,J., Tollervey,D. and Fournier M.J. (1988) Characterization of an SNR gene locus in *Saccharomyces cerevisiae* that specifies both dispensable and essential small nuclear RNAs. *Mol. Cell. Biol.*, **8**, 3282–3290.

Received on April 29, 1997; revised on June 20, 1997

Note added in proof

While this paper was in press Ni *et al*. [(1997) *Cell*, **89**, 563–573] published similar results to Ganet *et al*. (1997b) on the direction of pseudouridylation of rRNA by box H/ACA snoRNAs.