



Splicing-independent processing of plant box C/D and box H/ACA small nucleolar RNAs

David J. Leader^{1,3}, Gillian P. Clark¹, Jennifer Watters¹, Alison F. Beven², Peter J. Shaw² and John W.S. Brown^{1,*}

¹Cell and Molecular Genetics, Scottish Crop Research Institute, Invergowrie, Dundee DD2 5DA, UK (*author for correspondence); ²John Innes Centre, Norwich Research Park, Norwich NR4 7UH, UK; ³Current address: Zeneca Plant Science, Plant Biotechnology Section, Jealott's Hill Research Station, Bracknell RG42 6ET, UK

Received 14 May 1998; accepted in revised form 13 October 1998

Key words: box C/D, box H/ACA, intron, processing, small nucleolar RNA

Abstract

Small nucleolar RNAs (snoRNAs) are involved in various aspects of ribosome biogenesis and rRNA maturation. Plants have a unique organisation of snoRNA genes where multiple, different genes are tightly clustered at a number of different loci. The maize gene clusters studied here include genes from both of the two major classes of snoRNAs (box C/D and box H/ACA) and are transcribed as a polycistronic pre-snoRNA transcript from an upstream promoter. In contrast to vertebrate and yeast intron-encoded snoRNAs, which are processed from debranched introns by exonuclease activity, the particular organisation of plant snoRNA genes suggests a different mode of expression and processing. Here we show that single and multiple plant snoRNAs can be processed from both non-intronic and intronic transcripts such that processing is splicing-independent and requires endonucleolytic activity. Processing of these different snoRNAs from the same polycistronic transcript suggests that the processing machineries needed by each class are not spatially separated in the nucleolus/nucleus.

Introduction

Transcription and processing of ribosomal RNA (rRNA) and ribosome assembly occur in the nucleoli of eukaryotic cells and are dependent on numerous small nucleolar RNAs (snoRNAs) and nucleolar and ribosomal proteins (reviewed in [12, 26, 30, 36]). The nucleolar localisation of snoRNAs, their association with nucleolar proteins such as fibrillarin and Gar1p, and the presence, in many snoRNAs, of short sequences with complementarity to rRNAs suggest that they are involved in aspects of pre-mRNA processing and ribosome biogenesis. Several snoRNAs are essential for processing of the pre-rRNA to produce 18S, 5.8S and 28S mature rRNAs (reviewed in [12, 13, 26, 27, 30, 31, 36]).

The more than 80 eukaryotic snoRNA genes characterised to date can be classified into three groups: box C/D, box H/ACA and MRP snoRNAs. The latter group contains only the RNA moiety of RNase MRP

[25 and references therein]. Box C/D snoRNAs contain conserved box C and D sequences, usually at their 5' and 3' ends respectively and often adjacent to terminal inverted repeats, and regions of complementarity to rRNAs [1, 26]. Boxes C and D are required for accumulation of the snoRNAs [5, 6, 16, 29, 37, 38]. The regions of complementarity to 18S and 28S rRNA determine the positions of 2'-O-ribose methylation of rRNA [7, 20, 28, 35]. The box H/ACA class of snoRNAs is characterised by 5' and 3' stem-loop structures, a conserved box H sequence between the stems and the presence of the sequence 'ACA' 3 nt from the 3' end of the snoRNA [3, 14]. The box H and ACA sequences are required for accumulation of this class of snoRNAs and are likely to form part of a site of protein binding [3, 14]. Box H/ACA snoRNAs appear to be involved in pseudouridylation of rRNA [4] although U17 and E3 may also be involved in the early cleavage event in the 5' external transcribed spacer of pre-rRNA [27].

Besides their potential functions in pre-rRNA processing and maturation, snoRNAs have been of interest due to their distinct modes of expression. Some of the essential and more abundant snoRNAs, such as U3, U8 and U13, are expressed from classical snRNA promoter sequences [11, 32]. However, the majority of vertebrate and some yeast snoRNAs (box C/D and box H/ACA) are encoded within introns of genes encoding proteins which are often involved in nucleolar or ribosomal function [26]. In both vertebrates and yeast only a single snoRNA is found in any particular intron sequence [19, 26]. This organisation is exemplified by vertebrate *UHG* genes which contain a different snoRNA gene (*U22–U31*) in each intron although the spliced mRNAs of the host genes lacked an open reading frame [34]. The presence of only a single snoRNA in any particular intron is consistent with processing of intronic snoRNAs being largely splicing-dependent via exonucleolytic trimming of linearised snoRNA-containing intron lariats, as demonstrated for the vertebrate box H/ACA snoRNAs, U17 and U19 [8, 17, 19] and the box C/D U20 snoRNA [6]. Splicing-dependent processing of intronic snoRNAs in yeast has been demonstrated by the lack of U24 and reduced U18 production in a yeast strain carrying a mutation in the intron debranching enzyme ([21]; D. Tollervey, personal communication). That some intronic snoRNAs can be produced when in non-intronic contexts or, like yeast U18 (above), when debranching is blocked by mutation, suggests that non-specific degradation can also give rise to linear snoRNA-containing RNAs which can be subsequently processed [6]. In *Xenopus* oocytes, there is evidence of endonucleolytic cleavage in the release of intron-encoded U16 and U18snoRNAs [5 and references therein]. In this case, snoRNA processing and production of the host gene mRNA are mutually exclusive events. This mode of snoRNA production may, however, reflect the particular nature of the *Xenopus* oocyte system in containing low intron debranching activity or represent a regulatory step in the production of these snoRNAs and host protein [5]. Finally, processing of both major snoRNA classes is related in that the secondary structure of snoRNAs and association of proteins with conserved sequences are thought to block exonucleolytic trimming of the linearised intron to generate the mature snoRNAs in the form of snoRNPs [3, 5, 6, 8, 14, 33, 37, 38]. For box C/D snoRNAs, these signals (inverted repeats and adjacent boxes C and D) are positioned at the 5' and 3' termini. Extensive internal sequences are, therefore, not

required for processing and stability, such that large internal deletions of U14 and U20 are tolerated [6, 37, 38].

We have shown previously that some plant snoRNA genes are non-intronic, tightly clustered and transcribed polycistronically [22, 24]. This novel genomic organisation of snoRNA genes implies that processing of the pre-snoRNA transcript would require endonucleolytic cleavage between individual snoRNAs [22, 24]. In this paper, we have demonstrated that both single and multiple U14 (box C/D) and a box H/ACA snoRNA are processed from both spliceable and non-spliceable transcripts, consistent with the involvement of endonucleases. No evidence for specific sites of endonucleolytic cleavage was obtained. Therefore, plants have evolved a novel mode of snoRNA expression where the gene organisation is complemented by processing machinery which allows release of individual snoRNAs from polycistronic transcripts.

Materials and methods

Oligonucleotides

MU144Cl_a (5'-GCCGATCGATATGGCTCAGACATC CAAGG-3'); MU14L14 (5'-CATAACCTTTGTGGTT TGGTC-3'); MU141Cs (5'-CCTTTCGAATCACACA GGCAATAAACTGGAGGTGC-3').

Construction of plasmids expressing different snoRNA gene fragments

Construction of vectors for expression of pre-snoRNAs in either a non-intronic or intronic context have been described previously [22] and are shown in Figure 2. The maize U14.4 gene was subcloned into the above vectors with either extended flanking sequences (MU14.42b) or minimal flanking sequences (MU14.41a). The MU14.42b fragment was initially isolated as a 263 bp *Sau3* AI fragment, containing 93 bp of sequence upstream of box C and 64 bp of sequence downstream of box D, and subcloned into the *Bam*HI site of pGEM7Zf(–) to give plasmids p7MU14.42b(+) and p7MU14.42b(–) 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 p7MU14.42b(+). This fragment contained an additional 39 bp of sequence derived from the pGEM polylinker.

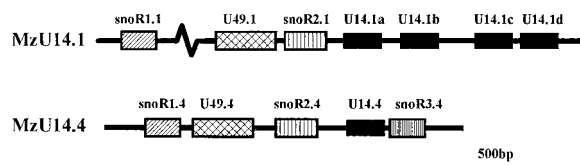


Figure 1. Organisation of maize snoRNA gene clusters. Schematic representation of snoRNA genes contained on maize genomic clones MzU14.1 and MzU14.4 respectively [32]. Genes are represented by differently shaded boxes, and flanking and intergenic regions by solid lines. The relative positions of the genes are drawn to scale with the exception of the intergenic region between *snoR1.1* and *U49.1* (749 bp) which is shown by a wavy line.

The MU14.41a fragment contained the U14.4 coding sequence, adjacent 6 bp inverted repeats and 8 bp and 2 bp of wild-type sequence upstream of the 5'- and 3'-terminal inverted repeats respectively. A *Cla*I site was introduced downstream of box D by PCR with primer MU144Cla and the natural *Nsi*I site located upstream of *U14.4* were used for cloning into the expression vectors. The entire maize U14.1 gene cluster was isolated as a single fragment, MU141ad, by PCR with primers MU14L14 and MU141Cs on plasmid pMU14-1E4.3 [22]. After digestion with *Nsi*I and *Csp*45I the 854 bp product contained each of the four maize U14.1-coding sequences with 13 bp of sequence upstream of the box C sequence of *U14.1a*, including the *Nsi*I site, and 115 bp of sequence downstream of the box D sequence of *U14.1d* including the artificial *Csp*45I site introduced by the primer.

The *snoR2.1* gene was subcloned into pGEM3Zf(+) as a 334 bp *Pst*I/*Nsi*I to give pgsnoR2.1. This fragment contained the 155 bp coding sequence of the *snoR2.1* gene with 110 bp and 69 bp of 5'- and 3'-flanking sequence respectively. The snoR2.1 fragment was subsequently re-isolated following digestion with *Pst*I and *Acc*I and subcloned into the expression vectors. This fragment contained 6 nt of additional sequence derived from the pGEM polylinker.

Protoplast transfections

Protoplasts were isolated from fully expanded leaves of 6–8 week old *N. tabacum* cv. Xanthi as described previously [10]. About 2.5×10^5 protoplasts were transfected with 15 μ g of doubly caesium chloride purified plasmid DNA and incubated for 24 h prior to RNA extraction. Protoplasts were co-transfected with 15 μ g of the plasmid pMzU5.3 which contains the maize U5.3 gene which is expressed at high levels in tobacco protoplasts [23]. RNA was isolated from protoplasts or from plant tissues by the guanidinium

iso-thiocyanate method [15]. RNA was treated twice with RNase-free DNase, prior to use in RNase A/T₁ and RT-PCR analysis.

RNase A/T₁ protection analysis

Probes for the detection of expression from vectors containing the MU14.42b fragment were transcribed with from the plasmid p7MU14.42b(-). This probe was complementary to the entire MU14.42b fragment with additional nucleotides derived from the pGEM vector. Specific transcription constructs were made for producing probes specific to each of the MU14.41a constructs. *Bam*HI fragments of 315 bp and 320 bp were isolated from pAMU14.41a and pLMU14.41a respectively and subcloned into pGEM7Zf(-) to give p7AMU14.41a and p7LMU14.41a respectively. Probes transcribed from p7AMU14.41a were complementary to the MU14.41a sequence as well as 37 nt and 152 nt of 5'- and 3'-flanking amylose sequences respectively. Probes transcribed from p7LMU14.41a were complementary to the MU14.41a sequence as well as 76 nt and 118 nt of 5'- and 3'-flanking legumin sequence respectively. The plasmid pgDMU14.41a contained a 360 bp fragment from pDMU14.41a cloned into pGEM3Zf(+) and was complementary to 20 nt and 207 nt of 5'- and 3'-flanking sequence derived from pDH51S/B. Radiolabelled antisense probes to detect U14.1b and U14.1d transcripts were transcribed from plasmids pgMU14.1b and pg MU14.1d [24] and each probe included antisense sequences to intergenic regions flanking the U14 genes.

RNase A/T₁ protection analysis was performed as described [15] with 5 mg of total RNA isolated from maize seedlings (cv. Kelvedon Glory). RNA was isolated from leaf material of maize as described previously [9]. [³²P]-labelled antisense probes for RNaseA/T₁ protection analysis were transcribed from plasmids with either T7 or SP6 RNA polymerase after appropriate linearisation 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 [³²P]-end-labelled *Hinf*I-digested ϕ X174 DNA as size markers.

Results

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

To examine the processing of the maize snoRNAs (Figure 1), a number of different snoRNA constructs were made in each of three plant expression vectors [22]. Briefly, the three different plant expression vectors allowed expression of single or multiple snoRNAs directly from the CaMV 35S promoter as part of a non-intronic transcript (pDH51S/B, Figure 2a), or within one of two different introns located within the coding sequence of a maize zein gene (pLegNC and pAmyNC; Figure 2b and c respectively; [22]). All constructs were transfected into tobacco protoplasts, RNA isolated and examined by RNase A/T₁ protection mapping. Three transfection experiments were carried out for each construct and protoplasts were co-transfected with the maize *U5-3* plasmid to allow the efficiency of individual transfections to be monitored. Although some variation in the intensities of processed products was evident within and between experiments, this was largely reflected in small variations in transfection efficiencies (results not shown). RT-PCR across the introns showed that splicing of intronic transcripts occurred (results not shown).

The initial construct was the *U14.4* gene containing approximately 100 bp of 5' and 3' authentic flanking sequences (Figure 3a, MU14.42b). Fully processed *U14.4* snoRNAs (products of 121-125 nt) accumulated to similar levels from all three expression constructs (Figure 4, lanes 1, 3 and 5; [22]). In addition, larger protected fragments representing unprocessed transcripts, and various smaller products due to protection of fragments of endogenous tobacco *U14* (compare lane 7 with lanes 1, 3 and 5) were also visible. Full-length *U14.4* was not observed with a sense probe (Figure 4, lanes 2, 4, 6 and 8) nor in untransfected protoplasts (lane 7). Thus, fully processed *U14.4* accumulates to similar levels from both intronic (Figure 4, lanes 1 and 3) and non-intronic (lane 5) transcripts. The same results were also obtained with a *U14.4* fragment with minimal flanking sequences (8 nt upstream and 2 nt downstream of the terminal inverted repeats) (MU14.41a, Figure 3b) (results not shown). This suggests that authentic flanking sequences are not required for *U14* processing, in agreement with results from vertebrates, where box C/D and inverted repeats are the only sequences required for processing [5, 6, 37, 38].

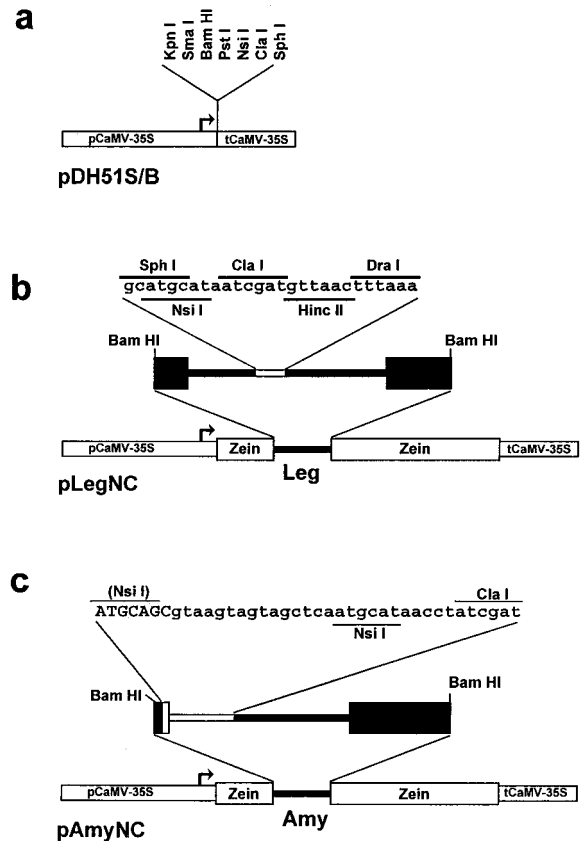


Figure 2. Vectors for expression of snoRNAs in different transcriptional contexts. a. pDH51S/B for non-intronic transcription. b,c. pLegNC (b) and pAmyNC (c) for intronic transcription (see [32]). The CaMV 35S promoter and terminator regions are shown as narrow open boxes, and the transcription start as an arrow. The zein protein coding sequence is shown as a larger open box and the legumin and amylase intron insertions as large filled boxes (authentic exon sequence) and thick black lines (introns). The sequences of polylinkers introduced into the introns are given. SnoRNA constructs (Figure 3) are introduced as *Nsi*/*Cla*I fragments.

To examine whether endonucleolytic cleavage of the authentic flanking sequences occurred at a preferred (conserved site) site processing of *U14.4* (MU14.42b, Figure 3a) was investigated using 5'- and 3'-specific probes (Figure 5a). The 5'-specific probe protected multiple products of 80–83 nt corresponding to the expected size of processed *U14.4*, with all three constructs (Figure 5b, lanes 1, 3 and 5). Larger protected fragments of ca. 170 nt corresponded to protection of unprocessed *U14.4*. Faint protected products of > 180 nt in lanes probed with the sense probe were indicative of residual plasmid DNA in the RNA preparations and the product of around 178 nt in these lanes appeared to be a tobacco-specific

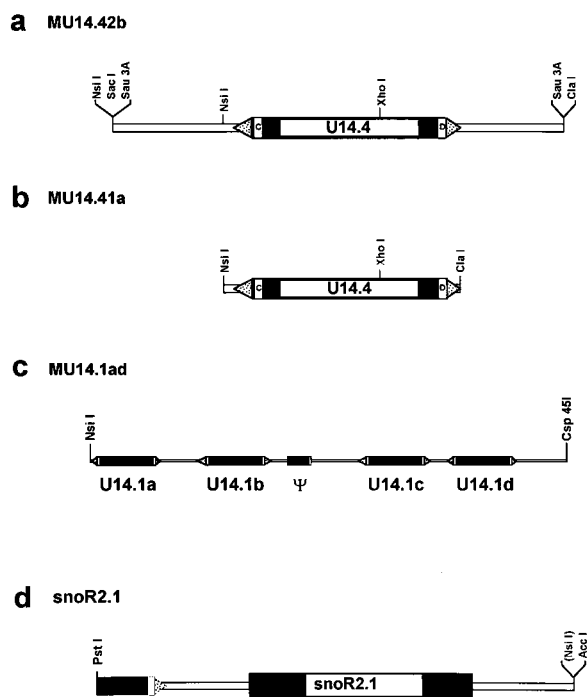


Figure 3. SnoRNA gene fragments for analysis of processing *in vivo* in protoplasts. a. MU14.42b: the *U14.4*-coding region plus flanking sequences. b. MU14.41a: the *U14.4*-coding region with minimal flanking sequences. c. MU14.1a–d: the MzU14.1 gene cluster containing the four complete U14 genes (U14.1a–d) and a gene fragment (Y). d. snoR2.1: the box H/ACA *snoR2.1* gene plus flanking sequences. Coding regions are shown as filled boxes; boxes C and D as small open boxes and shaded triangles represent inverted repeats.

product as it was found in untransfected protoplasts probed with the sense probe to pDMU14.42b (Figure 5b, lane 8). The bands of 39–63 nt in lanes 1, 3, 5 and 7 were due to endogenous tobacco U14s as they are found in untransfected protoplast RNA (Figure 5b, lane 7). The 3'-specific probe generated similar results in that the only strongly protected products represented fully processed U14.4 (products of 41–45 nt) and unprocessed U14.4 (products of ca. 120 nt). These products are not observed with sense probes (Figure 5b, lanes 10, 12, 14 and 16) and they overlap products due to endogenous tobacco U14 (Figure 5b, lane 15). The sizes of the fully processed products are consistent with processing being halted at or within the terminal inverted repeats. No strongly protected bands representing specific cleavage intermediates were observed with either the 5'- or 3'-specific probes. Instead a faint smear with a number of bands is observed. Thus, processing of U14.4 does not appear to occur via well-defined intermediates.

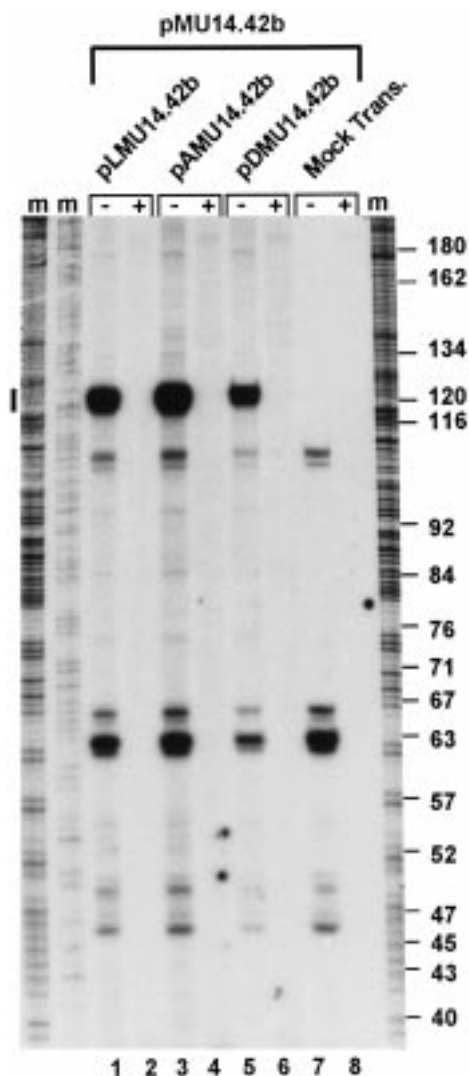


Figure 4. Processing of U14.4 with authentic flanking sequences from intronic and non-intronic transcripts. RNase A/T₁ protection mapping of RNA from protoplasts transfected with pLMU14.42b (lanes 1 and 2), pAMU14.42b (lanes 3 and 4), pDMU14.42b (lanes 5 and 6) and mock-transfected protoplasts (lanes 7 and 8). RNA was probed with antisense (–) or sense (+) MU14.42b probes. Full-length protected fragments are indicated by the bar on the left hand side. m, DNA sequence markers.

Polycistronic plant U14s can be efficiently processed from introns and non-intronic transcripts

To demonstrate that individual U14s can be processed from a single transcript containing multiple U14s, a fragment containing the entire U14.1 gene cluster (Figure 3c) was cloned into the three expression vectors described above. The fragment was generated by PCR and contained 13 bp upstream of the inverted

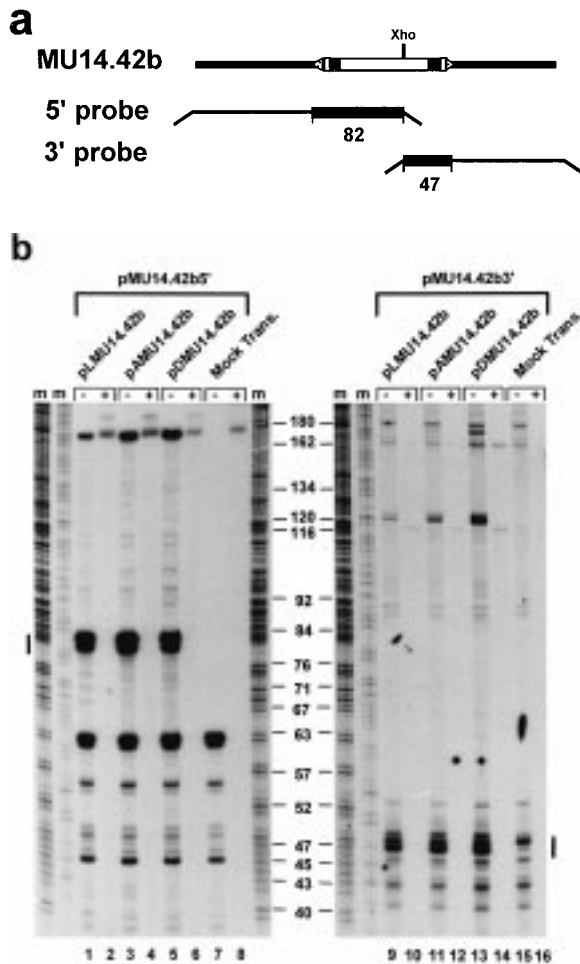


Figure 5. Processing of 5' and 3' ends of U14.4 from intronic and non-intronic transcripts. **a.** Schematic representation of the MU14.42b gene fragment and the positions of the 5'- and 3'-specific probes. The regions complementary to the coding region are shown as thick black lines and the expected product size is given. **b.** RNase A/T₁ protection mapping of RNA from protoplasts transfected with pLMU14.42b (lanes 1, 2, 9 and 10), pAMU14.42b (lanes 3, 4, 11 and 12), pDMU14.42b (lanes 5, 6, 13 and 14) and mock-transfected protoplasts (lanes 7, 8, 15 and 16). RNA was probed with antisense (-) or sense (+) 5'-specific (lanes 1–8) or 3'-specific (lanes 9–16) probes. Full-length protected fragments are indicated by the bars on the left and right hand sides. m, DNA sequence markers.

repeat sequence of the *U14-1a* gene and 110 bp downstream of the box D sequence of the *U14-1d* gene located at the 3' end of the cluster (Figure 3c). Aliquots of the same batches of isolated RNA were probed with antisense RNA probes to U14.1b (Figure 6, lanes 1–3) and U14.1d (Figure 6, lanes 6–8). Full-length protection products of ca. 120–125 nt were detected with both probes for all three constructs. Interestingly, the patterns of protected products from transfected to-

bacco were identical to those with total maize RNA (Figure 6, lanes 5 and 10) showing again that a single gene can give rise to multiple products, and that these presumably reflect the particular sequence attributes (box C and inverted repeats) of individual U14 sequence variants. No protected products were observed with RNA from mock-transfected protoplasts (Figure 6, lanes 4 and 9), while maize seedling total RNA protected full-length products (Figure 6, lanes 5 and 10). Thus, individual U14snoRNA variants were fully processed from both intronic and non-intronic transcripts containing multiple U14 genes. While the possibility exists that only a single snoRNA would be produced from each polycistronic transcript as a result of exonuclease digestion through adjacent U14s, it seems unlikely in the light of the general features of the multiple snoRNA gene clusters in plants and the likely association of proteins with each snoRNA sequence.

Processing of snoR2, a box H/ACA snoRNA, is also splicing-independent

The maize snoRNA gene clusters contain *snoR2* alleles are the only box H/ACA snoRNA genes to have been isolated from plants. To demonstrate that processing of this class of snoRNAs in plants is also splicing-independent, a fragment containing the *snoR2.1* gene (Figure 3d) was subcloned into each of the expression vectors to produce pLsnoR2.1, pAsnoR2.1 and pDsnoR2.1. Several full-length protected products of between 148 and 156 nt (products of 152 nt and 153 nt were the most prominent) were observed in all cases with the antisense probe (Figure 7, lanes 2, 4 and 6). These products were not detected by the sense probe (Figure 7, lanes 1, 3 and 5) and were equivalent to full-length protected bands produced by protection with total RNA isolated from maize leaf tissue (Figure 7, lane 10). Mock transfected protoplasts did not contain full-length protected products (Figure 7, lane 8).

In addition to the full-length *snoR2.1* products produced from pDsnoR2.1, a group of products was detected between 184 nt and 191 nt. These products are about 30–40 nt longer than the expected full-length products. The identity of these bands and whether they represent 5' or 3' extensions or both is unknown at present. If they were the products of specific endonucleolytic cleavage, their size would place the cleavage sites within the authentic *snoR2* flanking sequences. However, these sequences are also present in the ac-

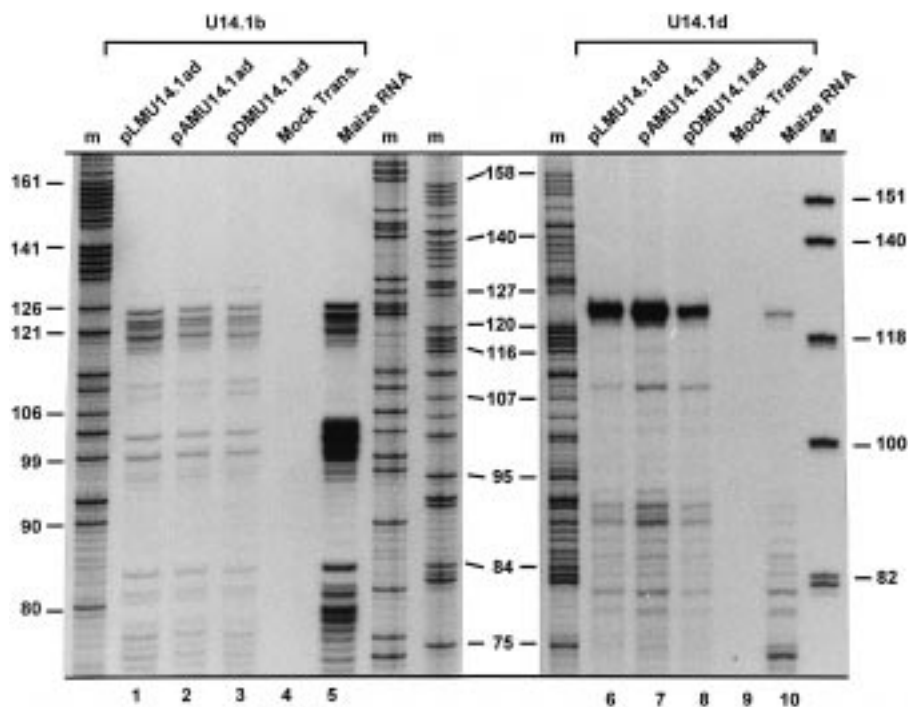


Figure 6. Processing of multiple U14 genes from intronic and non-intronic transcription constructs. RNase A/T₁ protection mapping of RNA from tobacco protoplasts transfected with pLMU14.1ad (lanes 1 and 6), pAMU14.1ad (lanes 3 and 7), pDMU14.1ad (lanes 5 and 8), mock-transfected protoplasts (lanes 7 and 9) and RNA from total maize leaf (lanes 5 and 10). RNA was probed with probes specific to U14.1b (lanes 1–5) or U14.1d (lanes 6–10). M, [³²P]-end-labelled *Hinf*I-digested fX174 DNA markers; m, DNA sequence markers.

comparing intronic constructs and in maize, but the larger protected fragments were not observed (Figure 7, lanes 2 and 4). Thus, the additional bands with the pDsnoR2 construct appear to be an artefact of transcription/processing from this particular construct, resulting, for example, from an alternative secondary structure of the transcript providing protection from the snoRNA processing machinery for a longer transcript. Thus, the box H/ACA snoR2.1 can be correctly processed from either the legumin intron, the amylase intron or from a non-intronic transcript.

Discussion

The occurrence of multiple, different snoRNA genes in tightly linked groups is as yet unique to plants. Besides the maize U14-containing gene clusters (this paper and see [22, 24]), we have isolated four other clusters of snoRNA genes in plants, suggesting that this organisation is a general feature of plant snoRNA genes (unpublished). In this paper we have demonstrated that mature snoRNAs (both box C/D and box H/ACA) can be processed from either single or mul-

tle snoRNA-containing transcripts irrespective of whether they are contained within splicable or unsplicable transcripts. All of the snoRNAs examined here accumulated to similar levels when expressed from either non-intronic or intronic constructs, transcribed from the CaMV 35S promoter. An analogous non-intronic construct containing vertebrate U17 without exons and transcribed from the cytomegalovirus promoter, produced U17 at greatly reduced levels compared to when the gene was present in a splicable intron [19]. Similarly, expression of non-intronic versions of U20 led to a 100-fold lower accumulation than when U20 was expressed in an intron [6]. The plant snoRNA intronic constructs also accumulated snoRNAs in plant cells, although it is not possible to distinguish between direct processing from the pre-mRNA transcript or processing from the linearised intron, and presumably both can occur.

In vertebrates, 5'-3' exonucleolytic processing of pre-snoRNAs has been demonstrated since a 5' cap structure prevents degradation and leads to accumulation of stable intermediates with 5' extensions to the cap [6, 8, 19]. Although all of the plant snoRNA-containing transcripts from the CaMV 35S promoter

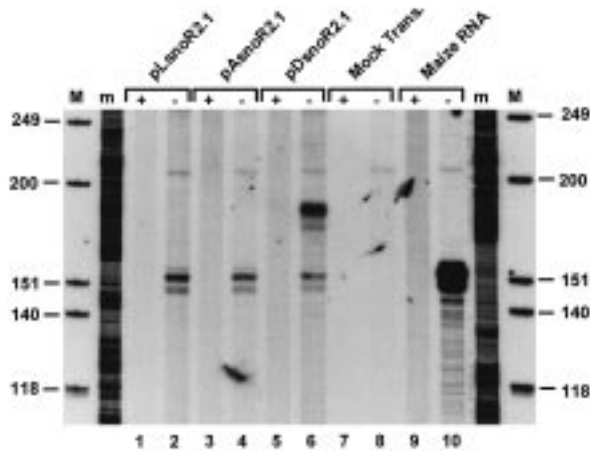


Figure 7. Processing of snoR2.1 from intronic and non-intronic transcripts. RNase A/T₁ protection mapping of RNA from protoplasts transfected with pLsnoR2.1 (lanes 1 and 2), pAsnoR2.1 (lanes 3 and 4), pDsnoR2.1 (lanes 5 and 6), mock-transfected protoplasts (lanes 7 and 8) and total maize leaf RNA (lanes 9 and 10). RNA was probed with antisense (-) or sense (+) snoR2.1 probes. M, [³²P]-end-labelled *Hinf*I-digested fX174 DNA markers; m, DNA sequence markers.

would be capped, no evidence of longer-lived products extending to the 5' cap site was obtained. In addition, 5'-3' and 3'-5' exonucleolytic processing has also been shown by analysing constructs containing tandem vertebrate U17 or U20 snoRNAs. In these cases, there is accumulation of high levels of transcripts containing both snoRNAs, where the 5' end of the upstream snoRNA and the 3' end of the downstream snoRNA are processed correctly [6, 19]. In addition, very low levels of the single mature snoRNAs were produced. In contrast, the maize multiple U14 construct produced only fully processed single U14s. Thus, processing of plant snoRNAs *in vivo* is not affected by the presence of a cap, whether the host transcript is splicable or unsplicable nor whether the snoRNAs are expressed polycistronically.

These results are consistent with a model whereby the plant snoRNA gene clusters are transcribed as a polycistronic pre-snoRNA from which individual pre-snoRNAs are released by endonucleolytic cleavage in the intergenic sequences (Figure 8). The resulting molecules are further processed by exonucleases and the snoRNAs protected by virtue of sequence and secondary structure elements, and protein association, as already described for many vertebrate and yeast intron-encoded snoRNAs [5, 6, 14, 38]. This is in contrast to the largely splicing-dependent processing of intronic snoRNAs of vertebrates and yeast which are

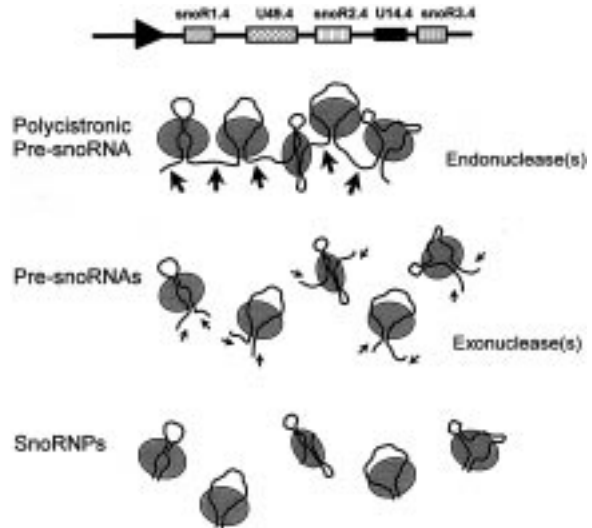


Figure 8. Model of processing of plant polycistronic pre-snoRNAs as illustrated by the MzU14.4 gene cluster. Box C/D (snoR1.4, U49.4, U14.4 and snoR3.4) and box H/ACA (snoR2.4) snoRNAs are associated with proteins (shaded ovals) and processed by endonucleolytic and exonucleolytic activities.

released via exonucleolytic digestion of debranched intron lariats ([6, 17, 19]; D. Tollervey, personal communication).

Clear parallels exist between the proposed endonucleolytic processing of plant snoRNAs and the processing of U16 and U18 in *Xenopus* oocytes, where introns carrying these snoRNAs are cleaved at U-rich regions prior to exonucleolytic trimming [5]. However, unlike the situation in oocytes, no evidence of distinct intermediate products representing specific endonuclease cleavage sites was obtained in the various RNase A/T₁ analyses of the plant snoRNA transcripts. Instead, the smear observed in the RNase A/T₁ analyses argues for endonucleolytic cleavage at random sites in the intergenic regions, with the flanking sequences being rapidly degraded by exonucleases. Thus, the nature of the endonucleolytic processing of plant snoRNAs differs from that of U16 and U18 in *Xenopus* oocytes. A further difference between vertebrate and plant is that *in vitro* and *in vivo* studies on other vertebrate box C/D snoRNAs have detected longer-lived transcripts with short 3' extended sequences, which are removed more slowly than the rest of the flanking sequence [8, 18, 19, 33]. In the processing of the various plant snoRNAs, corresponding short 3' extended products were again not observed. On the other hand, all of the snoRNAs do produce multiple bands differing in size by a small number of

nucleotides at the 5' and/or 3' ends. The length heterogeneity of snoRNAs may reflect differential degrees of exonucleolytic trimming at the termini of the RNA in the snoRNP complex [2].

Currently, the organisation of the yeast *snR190* and *U14* genes is the most reminiscent of the non-intronic clustering of plant snoRNA genes [39] although a cluster of different yeast snoRNAs has recently been identified in the database [M. Fournier, personal communication]. *SnR190* and *U14* lie 67 bp apart and transcription as a dicistron would again imply that endonucleolytic cleavage is needed to generate the two snoRNAs ([21]; M. Fournier and D. Tollervey, personal communications). Thus, all eukaryotes contain snoRNAs which are expressed from their own promoters; vertebrates and yeast contain intron-encoded snoRNAs largely processed by exonucleases; and plants and yeast contain polycistronic snoRNAs which require endonucleolytic cleavage. The difference between plants and vertebrates may, therefore, reflect differences either in the RNA degradatory enzymes themselves, in compartmentalisation of these enzymes and thereby accessibility to pre-snoRNAs, or that animals have taken advantage of the intron degradation pathway to further achieve co-regulation of snoRNA and host gene expression. It has previously been suggested that coiled bodies, which are often associated with the nucleolus, may be sites of turnover of excised introns and could deliver processed intronic snoRNAs to the nucleolus. We are currently investigating the nucleolar distribution of the polycistronic pre-snoRNAs to attempt to determine where pre-snoRNA processing occurs in the plant cell.

Acknowledgements

We are very grateful to Drs Maurille Fournier, Tamas Kiss, and David Tollervey for making unpublished work available. This research is supported by Gene Shears PTY Ltd, the Scottish Office Agriculture, Environment and Fisheries Department (SCRI) and the Biotechnology and Biological Sciences Research Council (JIC).

References

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

2. Balakin AG, Lempicki RA, Huang GM, Fournier MJ: *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 (1994).
3. Balakin AG, Smith L, Fournier MJ: The RNA world of the nucleolus: two major families of small RNAs defined by different box elements with related functions. *Cell* 86: 823–834 (1996).
4. Bousquet-Antonelli C, Henry Y, Gélunge J-P, Caizergues-Ferrer M, Kiss T: A small nucleolar RNP protein is required for pseudouridylation of eukaryotic ribosomal RNAs. *EMBO J* 16: 4770–4776 (1997).
5. Caffarelli E, Fatica A, Prislei S, De Gregorio E, Frapapan P, Bozzoni I: 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 (1996).
6. Cavaillé J, Bachelierie J-P: Processing of fibrillar-in-associated snoRNAs from pre-mRNA introns: and exonucleolytic process exclusively directed by the common stem-box terminal structure. *Biochimie* 78: 443–456 (1996).
7. Cavaillé J, Nicoloso M, Bachelierie J-P: Targeted ribose methylation of RNA *in vivo* directed by tailored antisense RNA guides. *Nature* 383: 732–735 (1996).
8. Cecconi F, Mariottini P, Amaldi F: The *Xenopus* intron-encoded U17 snoRNA is produced by exonucleolytic processing of its precursor in oocytes. *Nucl Acids Res* 23: 4670–4676 (1995).
9. Chomczynski P, Sacchi N: Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Anal Biochem* 162: 156–159 (1987).
10. Chupeau Y, Bourgin JP, Missonier C, Dorion N, Morel G: Préparation et culture de protoplastes de divers *Nicotiana*. *CR Acad Sci Paris* 278D: 1565–1568 (1974).
11. Dahlberg JE, Lund E: The genes and transcription of major small nuclear RNAs. In: Birnstiel ML (ed) *Structure and Function of Major and Minor Small Nuclear Ribonucleoprotein Particles*, pp. 38–70. Springer-Verlag, Berlin/Heidelberg/New York (1988).
12. Eichler DC, Craig N: Processing of eukaryotic ribosomal RNA. *Prog Nucl Acid Res Mol Biol* 49: 197–239 (1995).
13. Enright CA, Maxwell ES, Sollner-Webb B: 5= ETS rRNA processing facilitated by four small RNAs: U14, E3, U17 and U3. *RNA* 2: 1094–1099 (1996).
14. Ganot P, Caizergues-Ferrer M, Kiss T: 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 (1997).
15. Goodall GJ, Wiebauer K, Filipowicz W: Analysis of pre-mRNA processing in transfected plant protoplasts. *Meth Enzymol* 181: 148–161 (1990).
16. Huang GM, Jarmolowski A, Struck JC, Fournier MJ: 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 (1992).
17. Kiss T, Bortolin M-L, Filipowicz W: Characterisation of the intron-encoded U19 RNA, a new mammalian small nucleolar RNA that is not associated with fibrillar. *Mol Cell Biol* 16: 1391–1400 (1996).
18. Kiss T, Filipowicz W: Small nucleolar RNAs encoded by introns of the human cell cycle regulatory gene *RCC1*. *EMBO J* 12: 2913–2920 (1993).

19. Kiss T, Filipowicz W: Exonucleolytic processing of small nucleolar RNAs from pre-mRNA introns. *Genes Dev* 9: 1411–1424 (1995).
20. Kiss-László Z, Henry Y, Bachelier J-P, Caizergues-Ferrer M, Kiss T: Site-specific ribose methylation of preribosomal RNA: a novel function for small nucleolar RNAs. *Cell* 85: 1077–1088 (1996).
21. Lafontaine D, Tollervey D: Trans-acting factors in yeast pre-rRNA and pre-snoRNA processing. *Biochem Cell Biol* 73: 803–812 (1995).
22. Leader DJ, Clark GP, Watters J, Beven AF, Shaw PJ, Brown JWS: Clusters of multiple different small nucleolar RNA genes in plants are expressed as and processed from polycistronic pre-snoRNAs. *EMBO J* 16: 5742–5751 (1997).
23. Leader DJ, Connelly S, Filipowicz W, Waugh R, Brown JWS: Differential expression of U5snRNA gene variants in maize (*Zea mays*). *Plant Mol Biol* 21: 133–143 (1993).
24. Leader DJ, Sanders JF, Waugh R, Shaw PJ, Brown JWS: Molecular characterisation of plant U14 small nucleolar RNA genes: closely linked genes are transcribed as a polycistronic U14 transcript. *Nucl Acids Res* 22: 5196–5200 (1994).
25. Lygerou Z, Allmang C, Tollervey D, Seraphin B: Accurate processing of a eukaryotic precursor ribosomal RNA by ribonuclease MRP *in vitro*. *Science* 272: 268–270 (1996).
26. Maxwell ES, Fournier MJ: The small nucleolar RNAs. *Annu Rev Biochem* 35: 897–934 (1995).
27. Mishra RK, Eliceiri GL: Three small nucleolar RNAs that are involved in ribosomal RNA precursor processing. *Proc Natl Acad Sci USA* 94: 4972–4977 (1997).
28. Nicoloso M, Qu L-H, Michot B, Bachelier J-P: Intron-encoded, antisense small nucleolar RNAs: the characterisation 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 (1996).
29. Peculis BA, Steitz JA: Sequence and structural elements critical for U8 snRNP function in *Xenopus* oocytes are evolutionarily conserved. *Genes Dev* 8: 2241–2255 (1994).
30. Sollner-Webb B, Tycowski KT, Steitz JA: Ribosomal RNA processing in eukaryotes. In: Zimmerman RA, Dahlberg AE (eds) *Ribosomal RNA*, pp. 469–490. CRC Press, Boca Raton, FL (1995).
31. Tollervey D, Kiss T: Function and synthesis of small nucleolar RNAs. *Curr Opin in Cell Biol* 3: 337–342 (1997).
32. Tyc K, Steitz JA: U3, U8 and U13 comprise a new class of mammalian snRNPs localised in the cell nucleolus. *EMBO J* 8: 3113–3119 (1989).
33. Tycowski KT, Shu M-D, Steitz JA: A small nucleolar RNA is processed from an intron of the human gene encoding ribosomal protein S3. *Genes Dev* 6: 1120–1130 (1993).
34. Tycowski KT, Shu M-D, Steitz JA: A mammalian gene with introns instead of exons generating stable RNA products. *Nature* 379: 464–466 (1996).
35. Tycowski KT, Smith CM, Shu M-D, Steitz JA: A small nucleolar RNA requirement for site-specific ribose methylation of rRNA in *Xenopus*. *Proc Natl Acad Sci USA* 93: 14480–14485 (1996).
36. Venema J, Tollervey D: Processing of pre-ribosomal RNA in *Saccharomyces cerevisiae*. *Yeast* 11: 1629–1650 (1995).
37. Watkins NJ, Leverette RD, Xia L, Andrews MT, Maxwell ES: 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 (1996).
38. Xia L, Watkins NJ, Maxwell ES: Identification of specific nucleotided sequences and structural elements required for intronic U14snoRNA processing. *RNA* 3: 17–26 (1996).
39. Zagorski J, Tollervey D, Fournier MJ: Characterisation of an SNR gene locus in *Saccharomyces cerevisiae* that specifies both dispensible and essential small nuclear RNAs. *Mol Cell Biol* 8: 3282–3290 (1988).