5196–5203 Nucleic Acids Research, 1994, Vol. 22, No. 24

Molecular characterisation of plant U14 small nucleolar RNA genes: closely linked genes are transcribed as polycistronic U14 transcripts

David J.Leader, Jane F.Sanders+, Robbie Waugh, Peter Shaw[1] and John W.S.Brown*
Department of Cell and Molecular Genetics, Scottish Crop Research Institute, Invergowrie, Dundee DD2 5DA and [1]John Innes Institute, Colney Lane, Norwich NR4 7UH, UK

Received September 30, 1994; Accepted October 28, 1994 EMBL accession nos Z35298–Z35303

ABSTRACT

U14snoRNAs are highly conserved eukaryotic nucleolar small RNAs involved in precursor ribosomal RNA processing. In vertebrates, U14snoRNAs and a number of other snoRNAs are transcribed within introns of protein coding genes and are released by processing. We have isolated potato and maize genomic U14 clones using PCR-amplified plant U14 probes. Plant U14s show extensive homology to those from yeast and animals but contain plant-specific sequences. One of the isolated maize clones contains a cluster of four U14 genes in a region of only 761 bp, confirming the close linkage of U14 genes in maize, potato and barley as established by PCR. The absence of known plant promoter elements, the proximity of the genes and the detection of transcripts containing linked U14s by RT–PCR indicates that some plant U14snoRNAs are transcribed as precursor RNAs which are then processed to release individual U14s. Whether plant U14snoRNAs are intron-encoded or transcribed from novel promoter sequences, remains to be established.

INTRODUCTION

The two major classes of eukaryotic small nuclear RNAs are the spliceosomal snRNAs involved in pre-messenger RNA splicing and the small nucleolar RNAs (snoRNAs) involved in processing of pre-ribosomal RNA (pre-rRNA) transcripts and ribosome formation (1). At least 12 snoRNAs have been identified in yeast and evidence supporting their involvement in pre-rRNA processing or ribosome formation is largely based on their nucleolar localisation, propensity for base pairing with rRNA transcripts and their association with nucleolar proteins (2). In animals, the best characterised snoRNAs are U3, U8, U13 and U14 (3, 4) and their properties and roles in pre-rRNA processing have been reviewed recently (5). U14snoRNA is evolutionarily conserved, occurring widely in eukaryotes (6), and is required for a normal growth phenotype and for pre-rRNA transcript processing in yeast (7–9). In addition, certain sequence elements which are required for function and accumulation have been defined for yeast U14. These include the conserved box C and D sequences found in a number of nucleolar snoRNAs, of which box C is required for binding the nucleolar protein, fibrillarin (10), the 5′ and 3′ terminal stem, two conserved sequence elements which are complementary to 18S rRNA, and yeast specific sequences (11–13).

The genomic organisation of animal U14s is of particular significance. In mouse, rat, *Xenopus* and human, three copies of U14 are located in introns 5, 6 and 8 of the constitutively expressed cognate *hsc70* gene (14). In addition, the rainbow trout *hsc70* analogue, *hsc71*, contains U14 sequences in five introns (15). The orientation of the U14 sequences in the coding strand of the *hsc70* introns, the lack of typical UsnRNA promoter sequences (14), and the lack of a cap structure at the 5′ end suggest that these U14 sequences are transcribed as part of the *hsc70* pre-mRNA and processed from the intron sequences containing them (16).

More recently, a number of other snoRNAs have been shown to be encoded within introns of, for the most part, proteins involved in pre-rRNA processing and ribosome assembly, or which are associated with the nucleolus. U15 is located in the first intron of the human ribosomal protein S3 gene (17) and one U16 and four U18s in introns of the L1 ribosomal protein gene of *Xenopus laevis*, *X.tropicalis* and man (18, 19). In addition, two U17 genes (also called E1; 20) are found in introns 1 and 2 of a human cell cycle regulatory protein gene, *RCCI* (21, 22), while six U17s are located in introns of the *X.laevis* ribosomal protein S8 gene (23). Examples of other snoRNAs, U19–U21, Y, E2 and E3 (22), which appear to be or are likely to be encoded in introns of ribosomal or nucleolar protein genes have been recently reviewed (5, 24–26). The widespread occurrence of intron-encoded snoRNA genes raises several important questions such as how this gene arrangement facilitates the regulation and co-ordination of expression of nucleolar or ribosomal components, what are the processing mechanisms and components required for release of the snoRNAs and how conserved are these in evolutionary terms?

*To whom correspondence should be addressed

+Present address: University of Wales, College of Medicine, Heath Park, Cardiff CF4 4XN, UK

In relation to the latter, U14snoRNA in yeast does not appear to be intron-encoded but is likely to be processed from a precursor containing both U14 and a second nucleolar snRNA, snR190 (8, 27). Furthermore, the lack of U17snoRNAs in introns 1 and 2 of *Drosophila* and *Schizosaccharomyces pombe RCC1* analogues originally suggested that such gene arrangements may only be found in higher eukaryotes (21). However, recently intron-encoded snoRNA genes have been found in yeast (A.Balakin, J.Ni and M.Fournier, personal communication), again illustrating the widespread occurrence of this gene organisation and making it likely that such gene arrangements may also be found in plants. Given the lack of information on plant snoRNA genes [with the exception of U3 (28–32) and 7-2/MRP snRNA (33)], and the high degree of homology between U14 sequences from yeast and vertebrates, we have isolated U14snoRNA gene sequences from maize and potato. Evidence is presented which shows that, at least, some plant U14 genes are clustered in the genome. Furthermore, this novel organisation, the absence of promoter elements and the detection of polycistronic transcripts by RT–PCR suggests co-transcription of the genes and processing to release individual U14s.

MATERIALS AND METHODS

Materials

Restriction enzymes and T4 DNA ligase were purchased from Boehringer (Mannheim), Pharmacia and Promega. SP6 RNA polymerase, T7 RNA polymerase and RNase inhibitor were from Pharmacia. Radionucleotides, T4 polynucleotide kinase and Hybond N^+ filters were obtained from Amersham. RNase A, RNase T_1 and *Taq* DNA polymerase were purchased from Boehringer-Mannheim. MMLV reverse transcriptase and RNase-free, DNase RQI were obtained from Gibco-BRL and Sequenase from United States Biochemicals.

RNA and DNA extraction

RNA was isolated from leaf, stem and tuber tissue of the potato (*Solanum tuberosum* L.) cultivars Cara and Desirée, and from leaf, ovule and silk material of the maize (*Zea mays* L.) variety Kelvedon Glory, as described (34). DNA was extracted from leaf tissue of Cara and Kelvedon Glory as described previously (35).

RT–PCR and intergenic PCR

Total RNA preparations for reverse transcriptase–polymerase chain reaction (RT–PCR) were extensively DNased with RQI RNase-free DNase prior to first strand cDNA synthesis with primer II, 5′-CATATGATCA(G/A)ACATCCAAGGAAGG-3′, complementary to a conserved region (element 3 in ref. 12) at the 3′ end of mouse U14 (positions 65–82) and yeast U14 (positions 107–124). PCR reactions were carried out following addition of a second oligonucleotide primer, I, 5′-ACGCGTCG-ACCATTCGNNGTTTCCAC-3′, homologous to element A (12) in mouse U14 (positions 24–41) and yeast U14 (positions 31–46). RT–PCR reactions were carried out as described (36). Amplified products were isolated following acrylamide gel electrophoresis, sub-cloned into pUC13 (Promega) and sequenced. Primers III, 5′-CCCGGGAGCTCAGGCTTGAG-AGCTAGTGC-3′, and IV, 5′-GAATTCTGCAGCANGGCG-AAAGCTCTTAG-3′, were designed on the basis of the maize RT–PCR amplified sequences for use in inverse PCR and were used to amplify intergenic sequences between adjacent, closely linked U14 genes in both PCR and RT–PCR reactions (Fig. 1). Similar primers, V, 5′-CCCGGGAGCTCAGGCTTGAG-AGCATATGC-3′, and VI, 5′-GAATTCTGCAGCGAAGGCG-AAAGCACTTAG-3′, were used in PCR and RT–PCR reactions with potato DNA and RNA.

Cloning of plant U14 genomic sequences

Maize and potato genomic libraries in lambda EMBL4 were screened with inserts from IgPCR clones. Two maize genomic clones, ZmU14.1 and ZmU14.4, and one potato genomic clone, StU14.1, were purified and hybridising fragments identified by Southern analysis (results not shown) were subcloned into pUC13. ZmU14.1 was initially subcloned as four *Pst*I fragments of 1.1, 1.0, 0.7 and 0.07 kbp; ZmU14.4 was subcloned as a 2.2 kbp *Eco*RI fragment and StU14.1 as an 0.9 kbp *Eco*RI–*Pst*I fragment. ZmU14.1 contained four intact U14 gene sequences, ZmU14.1a–d, while ZmU14.4 and StU14.1 contained single genes. Plasmids for *in vitro* transcription for RNase A/T_1 mapping and nucleolar localisations were constructed by subcloning smaller genomic fragments into pGEM3zf+. ZmU14.1a was cloned as a 432 bp *Rsa*I fragment (pgMU14.1a) and contained 276 bp of upstream flanking sequence. ZmU14.1b was isolated as a 266 bp *Rsa*I fragment (pgMU14.1b) containing 35 bp of upstream flanking sequence and 101 bp of downstream sequence including the 36 bp U14 fragment identical to the 3′ end of ZmU14.1b. Plasmid pgMU14.1d contained the whole ZmU14.1d coding sequence, the 3′-most two-thirds of ZmU14.1c and 327 bp of sequence downstream of the U14 cluster, isolated on a 581 bp *Taq* I–*Bgl*II fragment. The second maize genomic clone, MzU14.4, was subcloned as a 254 bp *Nsi*I–*Bfr*I fragment (pgZmU14.4) and the potato genomic clone as a 438 bp *Eco*RI–*Hin*dIII fragment (pgStU14.1) for *in vitro* transcription.

Nucleolar localisation of U14snRNAs in plant tissue

Antisense riboprobes complementary to maize and potato U14s were generated by *in vitro* transcription from linearised plasmids in the presence of digoxygenin–UTP (Boehringer). *In situ* hybridisations on pea root cells were carried out as described previously (37). The preparations were counterstained with DAPI to show the chromatin and imaged by epifluorescence microscopy on a Nikon Microphot microscope.

Primer extension

Mapping of the 5′ end of U14 transcripts by primer extension was carried out on DNase-treated total RNA preparations using ^{32}P-end-labelled oligonucleotide primers: PotPex 1, 5′-TTA-GGCGGCCACTGCGAATG-3′, MzPex 1, 5′-TTAGGCGGC-AACTGCGAATG-3′, and MzPex 2, 5′-TTAG GCGGCT-ACTGCGAATG-3′. The two maize primers differed in a single nucleotide to reflect sequence variation seen in the cloned maize U14 genes. Primers were end-labelled with polynucleotide kinase and gel purified. Primer extension was carried out as in the RT step of RT–PCR (36) and products were separated on DNA sequencing gels.

RNase A/T_1 protection analysis

Preparation of ^{32}P-labelled RNA transcripts complementary to maize and potato U14 genes from linearised plasmids was as described previously (38). RNase A/T_1 protection analyses were performed on 10 μg of total RNA as described (39) using 4×10^5 dpm of labelled probe, and products separated on DNA sequencing gels.

RESULTS

Isolation of U14 sequences from maize and potato by PCR

Expressed plant U14 sequences were amplified by RT−PCR from total RNA preparations of maize, potato and barley using primers I and II based on relatively highly conserved regions of mouse and yeast U14s (Fig. 1a). Amplified products of the expected size of ~80 bp were subcloned and sequenced. Four maize, two potato and two barley sequences showed homology to mouse and yeast U14 sequences, suggesting that putative plant U14 sequences had been amplified (results not shown). Primers III and IV (maize) and V and VI (potato) were designed to amplify in the 5′ and 3′ direction from U14 coding sequences (Fig. 1a) originally for inverse PCR (IPCR) to generate complete U14 gene sequences. IPCR reactions with both potato and maize DNA generated major amplified DNA products of only 150−350 bp. Given the unexpectedly short size of these fragments, we reasoned that these products may represent the amplification of sequences between adjacent, closely linked genes (Fig. 1b), which was confirmed by standard PCR on genomic DNA (results not shown). PCR products were subcloned and the sequences of four maize and four potato clones were obtained (Fig. 2). Each of the clones contained ~45−50 bp of the 3′ end of the upstream gene (including the primer III or V sequence) and 70 bp of the 5′ end of the downstream U14 gene (including the primer IV or VI sequence) and also contained primer I and II sequences respectively. The intergenic regions of the four maize sequences varied in size from 53 to 140 bp (distances were measured between the box D of the upstream gene and box C of the downstream gene) while those of the potato clones ranged from 124 to 156 bp (Fig. 2). Some homology was apparent among intergenic sequences from the same species. For example ZmIgPCR3 and 4 differed by a 16 bp insertion in the latter. In addition, all four potato clones contained a sequence GTT-CTTG(C/T)TTGATGATGAATA (interestingly containing a box C sequence) 41−44 bp downstream of the box D sequence and ~75−100 bp upstream of box C in the adjacent downstream gene, however this sequence was not present in the potato genomic clone.

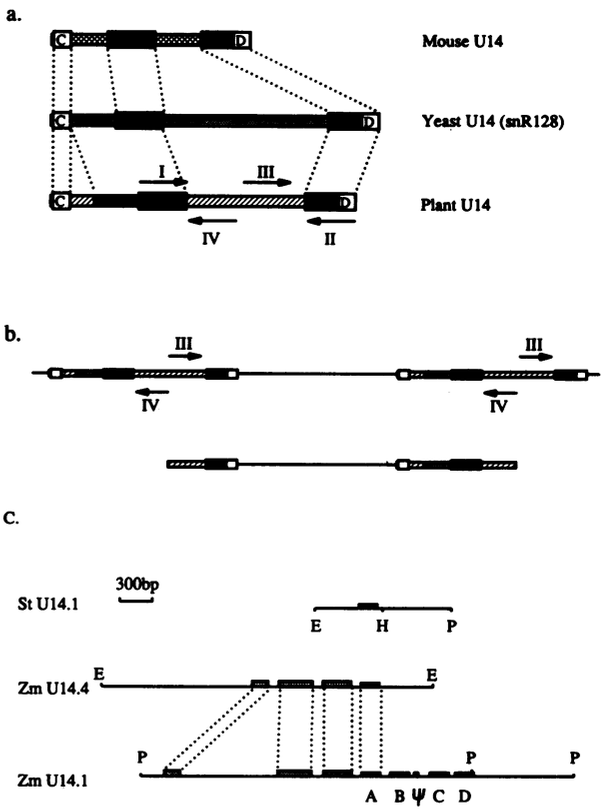


Figure 1. Schematic diagrams of PCR cloning strategy and plant U14 genomic clones. (**a**) Mouse, yeast and plant U14s showing regions of homology (black boxes), boxes C and D and positions of primers (I and II) used in initial RT−PCR reactions and primers (III and IV) used in intergenic PCRs. (**b**) Intergenic PCR showing structure of PCR products amplified by primers III and IV and presented in Fig. 2. (**c**) Regions of potato and maize U14 genomic clones showing positions of coding regions (black boxes) and the positions of three blocks of sequence homology between the 5′ flanking regions of the two maize genomic clones (shaded boxes).

Putative plant U14 gene sequences show extensive homology to vertebrate and yeast U14s but contain plant-specific regions

Screening maize and potato genomic libraries with maize U14 and potato U14 probes generated by intergenic PCR (IgPCR) identified a number of positive clones. Two maize and one potato genomic clone were purified and U14-containing fragments subcloned and sequenced (Fig. 3). Maize U14.1 gave an extremely strong hybridising signal and contained four apparently intact U14 coding sequences (ZmU14.1a−d) and a fragment of a fifth U14 sequence, all contained within a 761 bp region (Figs 1c and 3). ZmU14.4 and StU14.1 each contained a single U14 coding sequence. The sequences of the four apparently full length U14 coding regions of maize U14.1, and the single genes of maize U14.4 and potato U14.1 are aligned with yeast and mouse U14 sequences in Fig. 3. The putative coding regions have been defined by the inverted repeat sequences adjacent to the box C and box D sequences (Fig. 3). All six plant sequences show high homology and extensive regions of identity to one another. When compared to yeast and mouse U14s, sequence homologies ranged from 53 to 63% and 68 to 76% (expressed as the % of the yeast and mouse U14 sequences respectively). Although sequence variation exists among the plant, yeast and animal U14 sequences, they contain two regions of virtual identity: the box C sequence, and an 18 nt region including box D, lying directly upstream of the 3′ terminal inverted repeat. These two regions were called regions 1 and 3 in the comparison of yeast and mouse U14 (12) and the latter contains a sequence (18S-B) complementary to 18S rRNA which is also highly conserved in plants. In addition, a region of extensive similarity was identified in the 5′ halves of the molecules. This region includes region 2 (12), which again has a sequence, 18S-A, with complementarity to 18S rRNAs (Fig. 3). The plant U14 sequence information allows a realignment of the mouse U14 central region which could extend region 2 in the 5′ and 3′ directions for all of the sequences (Fig. 3). In addition, 5′ to this sequence is a 5 bp sequence homology among plant and yeast sequences (Fig. 3). The new alignment potentially gives one of the functionally essential yeast-specific sequences (12) a counterpart in both mouse and plant U14s, while a second essential yeast-specific region is still absent in mouse but overlaps one of the plant-specific sequences (P1−P3). The first of these is an additional 7 or 8 nucleotides between the 5′ extended region 2 sequence and box C, which are found only in plant U14s, while the other two overlap yeast-specific sequences, although their lengths and sequences differ (Fig. 3). The inverted repeats in

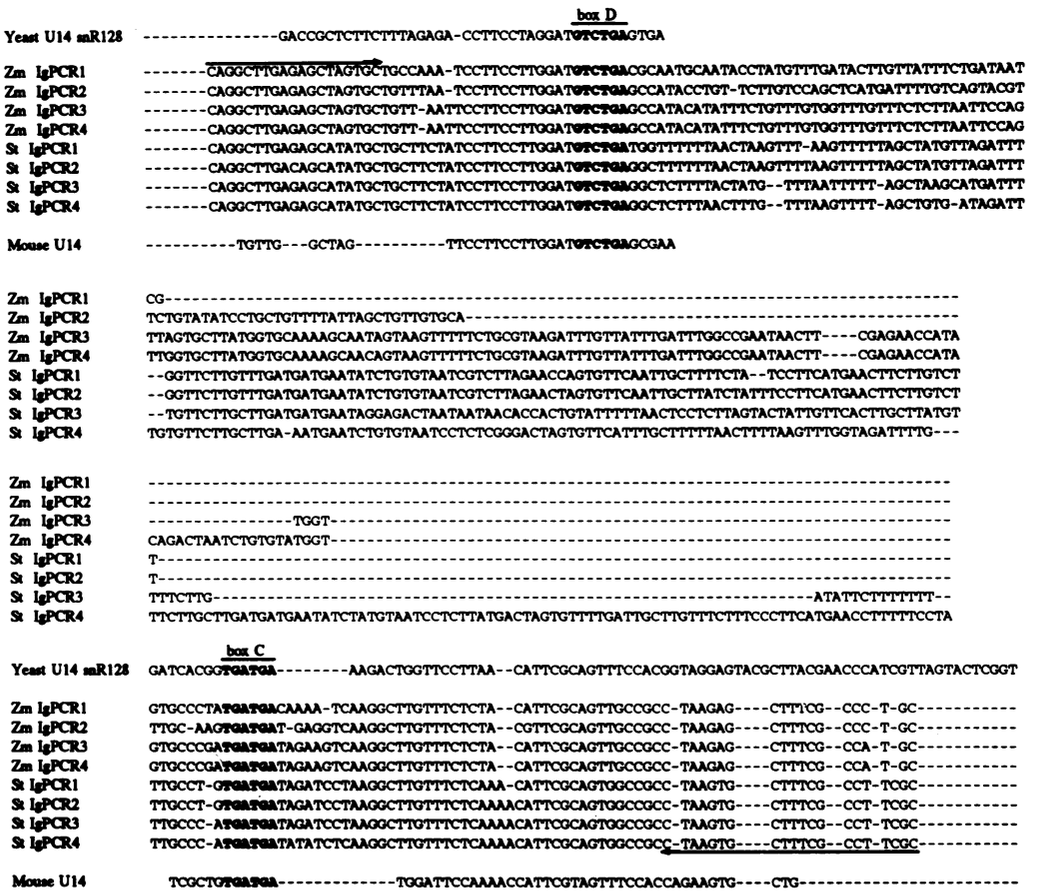


Figure 2. Alignment of sequences of U14 IgPCR products from maize (Zm) and potato (St). Conserved boxes C and D at the 3′ end of the upstream gene and 5′ end of the downstream gene respectively are indicated. The positions of maize primers III and IV (and the corresponding potato primers V and VI) are shown by arrows. Partial yeast and mouse U14 coding sequences are presented.

the plant genes vary in length and sequence and there are 2–3 nt between the 5′ IR and box C sequences, while the 3′ IR is usually directly adjacent to the box D sequence. Predicted sizes of the plant U14s are intermediate between those of mouse and yeast, with maize and potato U14s being ~117–120 nt long.

Of the genomic clones ZmU14.1, ZmU14.4 and StU14.1, 2872, 1280 and 873 bp were sequenced respectively. Notably, the sequence across the junction of the ZmU14.1c and 1d genes was identical to one of the sequences generated by IgPCR (Zm U14 IgPCR1; Fig. 2) and the sequence of ZmU14 IgPCR2 corresponded to the junctions between either ZmU14.1a and 1b or the gene fragment downstream of 1b and gene 1c. These latter sequences are virtually identical to one another as they are contained within a large direct repeat in the gene cluster. The U14 flanking regions were analysed for open reading frames (ORFs) and AU/GC content in an attempt to identify potential protein-encoding exon sequences. Although three regions of nucleic acid homology were observed in the 5′ flanking regions of the two maize genomic clones (Fig. 1c), these regions did not encode similar ORFs and therefore did not appear to correspond to exon sequences. Database searches with ORFs (>50 amino acids) and flanking DNA sequences failed to identify significant homologies with known genes. In addition, no typical plant gene promoter elements were identified in the flanking or coding regions of the U14 genes. A search with the maize U14 coding region identified >95% homology to an 80 bp region of an *Arabidopsis* expressed sequence tag (GenBank accession no. Z26461).

Nucleolar localisation of plant U14snoRNAs

Antisense riboprobes, labelled with digoxygenin–UTP, were generated by *in vitro* transcription of linearised plasmids pgMU14.1a and pgStU14.1. *In situ* hybridisation of these antisense probes complementary to potato and maize U14 sequences showed clear localisation to nucleoli with little or no labelling of the nucleoplasm or cytoplasm (Fig. 4), confirming that the isolated plant sequences encoded a nucleolar snRNA. Control *in situ* hybridisations were carried out using oligonucleotides complementary to plant U1snRNA and U3snoRNA sequences and showed specific labelling to the nucleoplasm and nucleoli respectively (results not shown).

Northern analysis and primer extension mapping of the 5′ end of U14 transcripts

The sizes of maize and potato U14s were estimated from Northern analysis of total RNA hybridised with the 316 bp *Eco*RI–*Hin*dIII insert of pgMU14.1b containing the maize U14.1b gene. A hybridising RNA band of ~120 nt was observed, consistent with

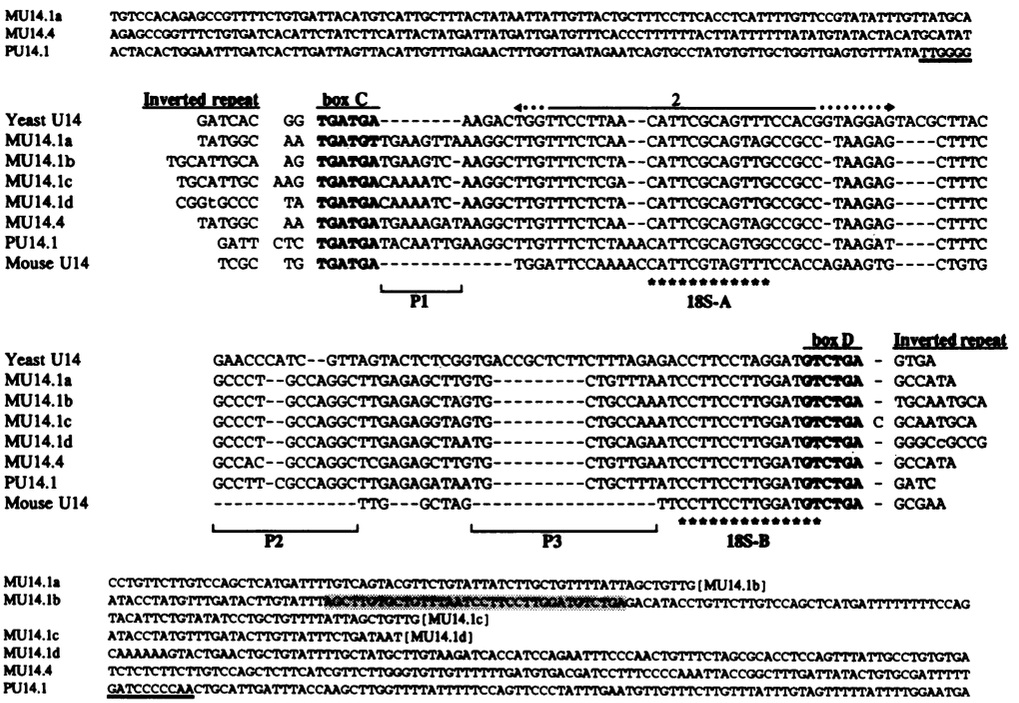


Figure 3. Alignment of plant U14 genomic sequences with mouse and yeast U14 coding regions. Inverted repeats at the 5′ and 3′ ends of the genes are shown; boxes C and D are in bold; 100 bp of 5′ and 3′ flanking sequences of the single genes MU14.4 and PU14.1 and flanking the U14 gene cluster of MU14.1, and the sequences downstream of genes 1a, 1b and 1c are presented; adjoining downstream U14 coding regions in the MU14.1 gene cluster are indicated by brackets. The gene fragment downstream of mU14.1b is shaded. Regions of complementarity to 18S rRNA (18S-A and 18S-B) are indicated by asterisks and conserved region 2 (12) is overlined and possible extensions of this region of homology are indicated by dotted lines. The three plant-specific sequences (P1–P3) are shown by brackets and the putative extended inverted repeats in PU14.1 are underlined in bold.

the sequence data (results not shown). There was no evidence from Northern analyses of larger hybridising RNA species representing unprocessed or partially processed U14 precursor RNAs.

Primer extension analysis of total plant RNA with primers complementary to equivalent maize and potato U14 sequences generated a number of major products. The MzPex 1 and MzPex 2 primers produced major primer extension products with maize ovule RNA in the size range 54–59 and 54–57 nt respectively, of which those of 56 nt were the most intense (Fig. 5a, lanes 2 and 3). All of the products mapped within the 5′ inverted repeat (Fig. 5b). Primer extension with total RNA from potato leaf using primer PotPex 1 gave a range of products from 56 to 60 nt (Fig. 5a, lane 1), which again mapped to the 5′ IR (Fig. 5b). The range of major products was not expected because, from the sequences, there was virtually no length variation in the U14 variants in the region copied by primer extension. The products may therefore reflect multiple 5′ ends resulting from processing, as seen in yeast U14s (27). Besides the major primer extension products described above, other products in the size range ~80–350 nt were observed (Fig. 5a) and were indicative of 5′ extensions to at least some U14 transcripts. In contrast, a control primer extension reaction with total maize ovule RNA and a primer complementary to maize U5snRNA gave a single primer extension product of expected size (results not shown). By comparison to the intensities of the major products, the extended products were low in abundance in the total RNA preparations. They were not derived from residual DNA, as the same RNA preparations were used in RT–PCR reaction analyses where PCR controls gave no DNA-derived products (see below).

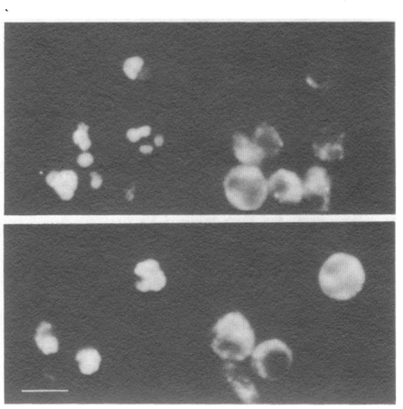
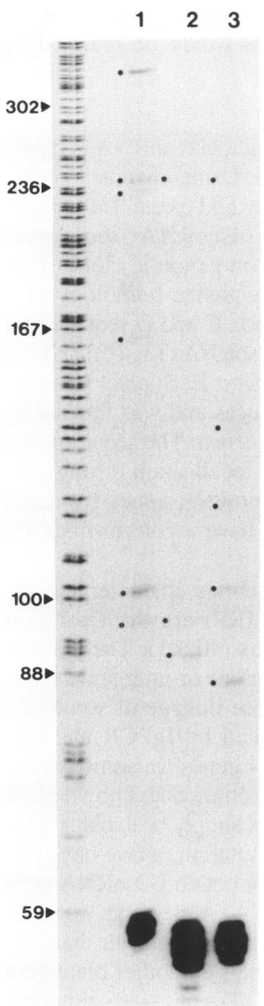


Figure 4. Nucleolar localisation of plant U14s. *In situ* hybridisations of potato U14 antisense RNA probes to potato U14 to pea root cell preparations showed virtually exclusive labelling of the nucleolus with little labelling of the rest of the nucleus. Two fields of cells are shown: left hand image, U14 *in situ* labelling; right hand image, nuclear chromatin stained with 4,6-diamidino-2-phenylindole (DAPI). Nucleoli are visible in the DAPI images as dark regions in the nuclei and are the sole sites of U14 *in situ* labelling. Bar 10 μm.

Expression of U14s by RNase A/T_1 protection mapping

To investigate whether the isolated genes were expressed, RNase A/T_1 protection mapping was carried out on maize ovule and



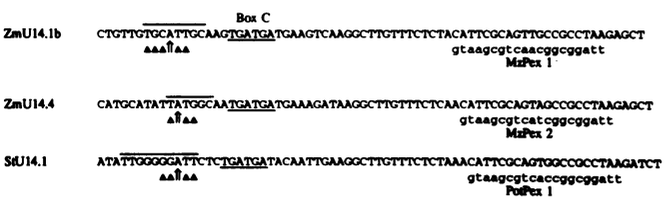


Figure 5. Primer extension analysis. (**a**) Mapping of 5′ ends of U14-containing transcripts in total RNA of maize with primers MzPex 1 and MzPex 2 (lanes 2 and 3 respectively) and potato (lane 1). Major products ranged in size from 54 to 59 nt and larger minor products are indicated by dots (lanes 1–3). (**b**) The major products observed in (a) map to the inverted repeat (IR) regions of plant U14s; arrow, major products; arrowheads, less intense products; box C sequences are underlined and the sequences and position of primers are shown.

potato leaf total RNA with antisense probes complementary to ZmU14.1a, 1b and 1d, ZmU14.4 and StU14.1 (Fig. 6, lanes 1–5 respectively). Full length protected products in the range 118–126 nt were observed. Within the limitations of specificity of RNase A/T_1 protection, these products indicated expression of the above U14 genes. Consistent with the sequence analysis

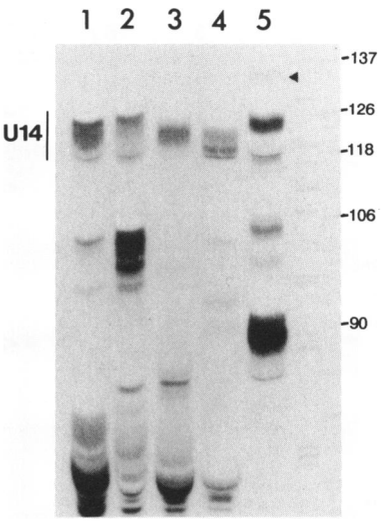


Figure 6. Expression of maize and potato U14s. RNase A/T_1 protection of total maize and potato RNA preparations with antisense RNA probes to ZmU14.1a, 1b and 1d, ZmU14.4 and StU14.1 (lanes 1–5 respectively). Full length protection products are labelled (U14) and a longer fragment protected by the StU14.1 probe is indicated with an arrowhead. Sizes of a DNA sequence ladder marker are given.

(Fig. 3), the potato U14 products were slightly larger than those of maize. The range of sizes of the products may reflect sequence variation in the U14s and/or different 5′ termini, as already observed with primer extension (Fig. 5). Much less intense, larger products were observed higher in the RNase A/T_1 gels (not shown) and could represent precursor U14 transcripts, but their identity has not been investigated further. Mapping with the StU14.1 probe showed a protected fragment of 134–135 nt which may reflect a particularly stable precursor possibly due to formation of a terminal stem involving 10 bp GC-rich inverted repeats flanking this potato U14 gene (Fig. 3). All probes showed smaller protected fragments arising from hybridisation of the probes to transcripts from other U14 gene variants. Control reactions without added plant RNA showed no protected fragments. Taken together, the positions of the 5′ ends mapped by primer extension and the lengths of RNase A/T_1 full length products suggest that the 3′ ends of the U14 transcripts would map to or immediately adjacent to the 3′ inverted repeat.

Detection of polycistronic U14 transcripts by RT–PCR

Intergenic PCR (IgPCR) and the gene arrangement of ZmU14.1 suggested that some U14 genes are clustered in both potato and maize, raising the possibility that these U14 genes are transcribed together as a polycistronic transcript. RT–PCR on maize RNA using primer III for initial cDNA production and primers III and IV in the PCR reaction produced a number of labelled products of 135, 152, 180, 261, 266 and 272 nt (Fig. 7, lane 1). Similar reactions with potato RNA preparations from three different potato organs using primers V and VI gave products of 82, 118, 258, 277, 278 and ~360 nt (Fig. 7, lanes 2–4). Control reactions without reverse transcriptase (Fig. 7, lanes 5–8) gave no products, showing no residual DNA in the RNA preparations and confirming that the RT–PCR products in lanes 1–4 were derived from RNA. From the positions of the primers, the RT–PCR products in Fig. 7 would contain ~120 nt of coding region from the two adjacent genes, leaving intergenic regions

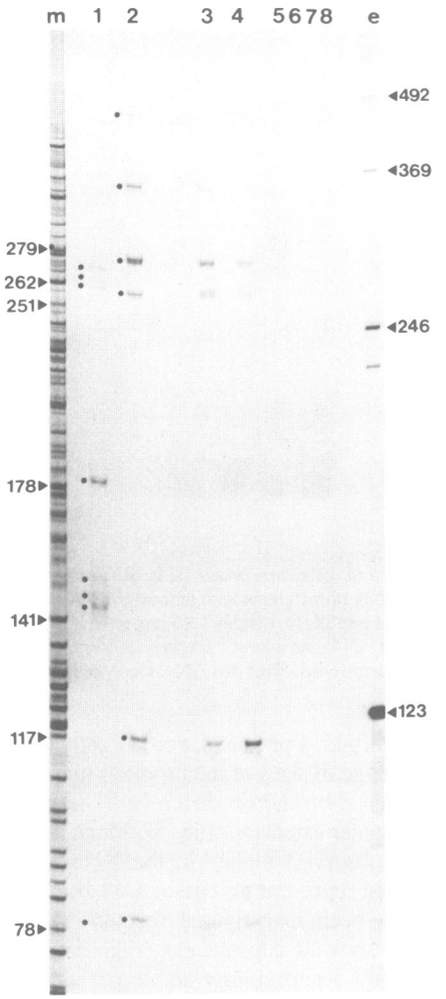


Figure 7. Intergenic RT−PCR of total maize and potato RNA. RT−PCR reactions using primers III and IV (maize) and V and VI (potato) (see Materials and Methods), where primers III and V were end-labelled with ^{32}P were separated on a urea−polyacrylamide gel. Lane 1, maize ovule RNA; lanes 2−4, potato leaf, stem and tuber RNA respectively; lanes 5−8, control reactions without reverse transcriptase for lanes 1−4 respectively. RT−PCR products were separated alongside a DNA sequence marker (lane m) and end-labelled 123 nt ladder marker (lane e).

in the range 15−150 nt for maize and 140−240 nt for potato, assuming that the products contain the arrangement shown in Figs 1b and 2. The identity of the smaller RT−PCR products is unknown but could arise from rearranged U14 genes or gene fragments transcribed as part of a polycistronic U14 transcript. Cloning and sequencing of some IgRT−PCR products confirmed the presence of 3′ and 5′ fragments of adjacent genes separated by an intergenic sequence, as already obtained by IgPCR (Fig. 3). Four sequenced clones were virtually identical to the sequence of StD IgPCR4 (Fig. 3), two were identical to StD IgPCR2 and one was identical to Zm IgPCR1 (Fig. 3). Two further maize-derived clones contained the 3′ and 5′ fragments of adjacent genes separated by only 17 and 18 bp (results not shown), which did not have a counterpart in the IgPCR clones nor in the ZmU14.1 gene cluster. Thus, RT−PCR has detected a number of transcripts containing adjacent U14 genes, suggesting that U14 gene clusters as seen in the genomic ZmU14.1 clone are likely to be transcribed as a single polycistronic transcript and inferring that individual U14s would be released by processing.

DISCUSSION

The essential U14 nucleolar snRNA is highly conserved between yeast and vertebrates. Using oligonucleotide primers to conserved sequences in mouse and yeast U14snoRNAs we have PCR-amplified plant U14snoRNA sequences. Plant U14 gene sequences, isolated on genomic clones from maize and potato, showed extensive sequence homology to vertebrate and yeast U14s in containing box C and D sequences [present in fibrillarin-associated nucleolar snRNAs (3, 10)] and regions complementary to 18S rRNA sequences. Each plant U14 sequence also contained plant-specific sequences and was flanked by inverted repeats of variable size and sequence. The sequence and structural similarity and their nucleolar localisation strongly suggest that we have isolated plant U14snoRNA genes from both potato and maize which are likely to have a role in plant rRNA processing and ribosome assembly.

The promoter elements characteristic of plant snRNA genes (U1−U6 and 7-2/MRP) are the Upstream Sequence Element (USE) and TATA box (40, 41). These elements were not present in the 5′ flanking regions or intergenic sequences of the genomic U14 sequences. Close linkage of some of the plant U14 genes has been demonstrated by IgPCR and the organisation of the cluster of maize U14 genes. In some cases, intergenic distances were too short to accommodate known plant promoter elements. These findings cannot simply be explained by the IgPCR approach used for U14 gene isolation, as we have previously used IgPCR to establish linkage of potato U2snRNA genes. In each case plant UsnRNA USE and TATA elements were observed and intergenic sequences were substantially greater than observed here (42). The absence of these signals and other plant gene promoter elements suggested that the U14 genes were either transcribed from novel promoter sequences or as part of larger precursor RNA molecules. The detection of transcripts containing at least two U14 coding sequences by RT−PCR and the identification of an RNA-derived RT−PCR product with the sequence of part of the ZmU14.1 cluster further suggests that clusters of U14 genes are transcribed as polycistronic transcripts from which individual U14s are processed. On the basis of flanking sequences, we are unable, at present, to establish whether these plant U14 gene sequences are intron-encoded. In animal systems only a single snoRNA gene is contained in any one intron. In the maize U14 cluster, obvious 3′ and 5′ splice sites are not observed in the intergenic regions, suggesting that if these genes are intron-encoded, they would all be contained within the same intron. It is, therefore, possible that the gene clusters are transcribed as a specialised U14 transcript from a single upstream promoter, from novel promoter elements contained within short intergenic or coding sequences, or as part of a pre-mRNA transcript.

The location of snoRNAs in introns of genes encoding proteins involved in pre-rRNA processing or ribosome assembly suggest that their expression is both achieved and regulated by exploiting the active or constitutive expression of the host genes (5, 26). In addition it has been postulated that the relative abundance of intron-encoded snoRNAs is reflected in the number of copies of the genes, such that multiple copies (but only one per intron), as observed with vertebrate U14 and *Xenopus* U17 and U18 sequences, may represent a means of achieving higher levels of snoRNA transcripts (26). The novel polycistronic U14 gene

arrangements in plants may also ensure abundant levels of U14 transcripts regardless of whether the precursor is intronic or transcribed from a novel promoter. Analyses of splicing and processing of U16 and U18 snoRNA-containing intron transcripts have suggested a two step processing pathway where endonucleolytic cleavages generate snoRNAs with terminal extensions which are then rapidly removed by either endo- or exonucleolytic activity (18, 19, 43–45). However, recent *in vivo* and *in vitro* analysis of processing of human U17 and U19 snoRNAs have demonstrated that processing requires splicing of the intron and proceeds by exonucleolytic digestion of flanking regions from the debranched intron lariat (T.Kiss and W.Filipowicz, personal communication). The tight clustering of some plant U14 genes suggests that, in plants, initial endonucleolytic cleavages would be involved in generating individual U14s. The absence of larger transcripts on Northern analysis and the low abundance of potential U14 precursor transcripts in primer extension and RNase A/T_1 mapping would point to rapid processing of plant U14 precursors *in vivo*. The stability of mature, processed vertebrate snoRNAs and certain intermediates of the processing reaction suggest that these molecules are protected from the exonucleolytic activity required to generate the mature snoRNAs, (16, 17, 21). While this may be due to stable stem formation between the 5′ and 3′ termini of, for example, U14–U16snoRNAs and binding of proteins such as fibrillarin (16–19, 27), some processed snoRNAs (U17 and U18) do not contain inverted repeats or box C or D sequences and protection may therefore involve binding of other protein factors (45). Further analyses of genomic flanking sequences and an examination of the transcriptional activity and processing of plant U14 precursors is currently underway to determine the mechanism of plant U14snoRNA gene expression.

ACKNOWLEDGEMENTS

This research was supported by Gene Shears, Pty, Sydney, Australia, the Scottish Office Agriculture and Fisheries Department, and the BBSRC via a grant-in-aid to the John Innes Centre.

REFERENCES

1. Dahlberg J. E. and Lund, E. (1988) In: Structure and Function of Major and Minor Small Ribonucleoprotein Particles (ed. Birnstiel, M.L.). Springer-Verlag, Berlin, pp. 38–70.
2. Fournier, M. J. and Maxwell, E. S. (1993) *Trends Biochem. Sci.* **18**, 131–135.
3. Tyc, K. and Steitz, J. A. (1989) *EMBO J.* **8**, 3113–3119.
4. Maxwell, E. S. and Martin, T. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 7261–7265.
5. Filipowicz, F. and Kiss, T. (1993) *Mol. Biol. Rep.* **18**, 149–159.
6. Trinh-Rohlik, Q. and Maxwell, E. S. (1988) *Nucleic Acids Res.* **16**, 6041–6056.
7. Tollervey, D. (1987) *EMBO J.* **6**, 6149–4175.
8. Zagorski, J., Tollervey, D. and Fournier, M. J. (1988) *Mol. Cell. Biol.* **8**, 3282–3290.
9. Li, H. V., Zagorski, J. and Fournier, M. J. (1990) *Mol. Cell. Biol.* **10**, 1145–1152.
10. Baserga, S. J., Yang, W. and Steitz, J. A. (1991) *EMBO J.* **10**, 2645–2651.
11. Jarmolowski, A., Zagorski, J., Li, H. V. and Fournier, M. J. (1990) *EMBO J.* **9**, 4503–4509.
12. Li, H. V. and Fournier, M.J. (1992) *EMBO J.* **11**, 683–689.
13. Huang, G. M., Jarmolowski, A., Strack, J.C.R., and Fournier, M.J. (1992) *Mol. Cell. Biol.* **12**, 4456–4463.
14. Liu, J. and Maxwell, E. S. (1990) *Nucleic Acids Res.* **18**, 6565–6571.
15. Zafarullah, M., Wisniewski, J., Shworak, N.W., Schieman, S., Misra, S. and Gedamu, L. (1992) *Eur. J. Biochem.* **204**, 893–900.
16. Leverette, R.D., Andrews, M.T. and Maxwell, E.S. (1992) *Cell* **71**, 1215–1221.
17. Tycowski, K.T., Shu, M.-D and Steitz, J.A. (1993) *Genes Dev.* **7**, 1176–1190.
18. Fragapane, P., Prislei, S., Michienzi, A., Caffarelli, E. and Bozzoni, I. (1993) *EMBO J.* **12**, 2921–2928.
19. Prislei, S., Michienzi, A., Presutti, C., Fragapane, P., and Bozzoni, I. (1993) *Nucleic Acids Res.* **21**, 5824–5830.
20. Ruff, E. A., Rimoldi, O.J., Raghu, B. and Eliceiri, G. (1993) *Proc. Natl. Acad. Sci. USA* **90**, 635–638.
21. Kiss, T. and Filipowicz, W. (1993) *EMBO J.* **12**, 2913–2920.
22. Nag, M.K., Thai, T.T., Ruff, E.A., Selvamuragan, N., Kunnimalaiyaan, M., and Eliceiri, G.L. (1993) *Proc. Natl. Acad. Sci. USA* **90**, 9001–9005.
23. Cecconi, F., Mariottini, P., Loreni, F., Pierandrei-Amaldi, P., Campioni, N. and Amaldi F. (1994) *Nucleic Acids Res.* **22**, 732–741.
24. Mount, S. and Henikoff, S. (1993) *Curr. Biol.* **3**, 382–374.
25. Seraphin, B. (1993) *Trends Biochem. Sci.* **18**, 330–331.
26. Sollner–Webb, B. (1993) *Cell* **75**, 403–405.
27. Balakin, A.G., Lempicki, R.A., Huang, G.M. and Fournier, M.J. (1994) *J. Biol. Chem.* **269**, 739–746.
28. Kiss, T. and Solymosy, F. (1990) *Nucleic Acids Res.* **18**, 1941–1949.
29. Marshallsay, C., Kiss, T. and Filipowicz, W. (1990) *Nucleic Acids Res.* **18**, 3459–3466.
30. Marshallsay, C., Connelly, S. and Filipowicz, W. (1992) *Plant Mol. Biol.* **29**, 973–983.
31. Kiss, T., Marshallsay, C. and Filipowicz, W. (1991) *Cell* **65**, 517–526.
32. Leader, D.J., Connelly, S., Filipowicz, W. and Brown, J.W.S. (1994) *Biochim. Biophys. Acta* in press.
33. Kiss, T., Marshallsay, C. and Filipowicz, W. (1992) *EMBO J.* **11**, 3737–3746.
34. Chomczynski, P. and Sacchi, N. (1987) *Anal. Biochem.* **162**, 156–159.
35. van der Ven, W.T.G., Powell, W., Ramsay, G. and Waugh, R. (1990) *Heredity* **65**, 329–342.
36. Simpson, C.G., Sinibaldi, R. and Brown, J.W.S. (1992) *Plant J.* **2**, 835–836.
37. Highett, M.I., Bevan, A.F. and Shaw, P.J. (1993) *J. Cell Sci.* **105**, 1151–1158.
38. Waugh, R., Clark, G., Vaux, P. and Brown, J.W.S. (1991) *Nucleic Acids Res.* **19**, 249–256.
39. Goodall, G.J., Kiss, T. and Filipowicz, W. (1989) *Oxford Surv. Plant Mol. Cell Biol.* **7**, 255–296.
40. Vankan, P. and Filipowicz, W. (1989) *EMBO J.* **8**, 3875–3882.
41. Waibel, F. and Filipowicz, W. (1990) *Nature* **346**, 199–202.
42. Waugh, R., Clark, G. and Brown, J.W.S. (1991) *Gene* **107**, 197–204.
43. Fragapane, P., Caffarelli, E., Lener, M., Prislei, S., Santoro, B. an Bozzoni, I. (1992) *Mol. Cell. Biol.* **12**, 1117–1125.
44. Prislei, S., Sperando, S., Fragapane, P., Caffarelli, E., Presutti, C. and Bozzoni, I. (1992) *Nucleic Acids Res.* **20**, 4473–4479.
45. Caffarelli, E., Arese, M., Santoro, B., Fragapane, P. and Bozzoni, I. (1994) *Mol. Cell. Biol.* **14**, 2966–1974.