Nucleic Acids Research, 1993, Vol. 21, No. 23 5391–5397

SnR31, snR32, and snR33: three novel, non-essential snRNAs from *Saccharomyces cerevisiae*

Andrey G.Balakin, Gregory S.Schneider, Max S.Corbett, Jingwei Ni and Maurille J.Fournier*
Department of Biochemistry and Molecular Biology and Program in Molecular and Cellular Biology, Lederle Graduate Research Center, University of Massachusetts, Amherst, MA 01002, USA

Received August 9, 1993; Revised and Accepted October 6, 1993 GenBank accession nos L22433–L22435 (incl.)

ABSTRACT

Genes for three novel yeast snRNAs have been identified and tested for essentiality. Partial sequence information was developed for RNA extracted from isolated nuclei and the respective gene sequences were discovered by screening a DNA sequence database. The three RNAs contain 222, 188 and 183 nucleotides and are designated snR31, snR32 and snR33, respectively. Each RNA is derived from a single copy gene. The *SNR31* gene is adjacent to a gene for an unnamed protein associated with the cap-binding protein eIF-4E. The *SNR32* gene is next to a gene for ribosomal protein L41 and the gene for *SNR33* is on chromosome III, between two open reading frames with no known function. Genetic disruption analyses showed that none of the three snRNAs is required for growth. The new RNAs bring the number of non-spliceosomal snRNAs characterized thus far in *S.cerevisiae* to 14, of which only three are essential.

INTRODUCTION

The yeast *Saccharomyces cerevisiae* contains upwards of 30 small nuclear RNAs (1,2). Sixteen of these RNAs have been at least partially characterized, including the five snRNAs involved in mRNA splicing and 11 RNAs that appear to be associated with the nucleolus (reviewed in ref.3,4). The nucleolar small RNAs in yeast and other organisms have been implicated in ribosome biogenesis. Several of these species have been shown to be important for processing of ribosomal RNA precursors, based on: i) disruption of rRNA processing in cells or extracts depleted of specific snRNAs; ii) co-purification with precursor rRNAs, and; iii) cross-linking with rRNA (reviewed in refs 3–5). Only three of the 11 non-spliceosomal snRNAs characterized thus far from *S. cerevisiae* are essential for growth, U3, U14 and snR30 (6–8). The other genes are dispensable, a situation that begs explanation.

Each of the three essential nucleolar snRNAs is required for normal processing of rRNA. Loss of any one impairs processing of the primary transcript and leads to a shortage of mature 18S rRNA (9–11). A fourth snRNA, snR10, is not essential for viability, but loss of this species also alters the flux of intermediates in the rRNA processing pathway and leads to a cold-sensitive phenotype (12). Effects nearly identical to those observed with loss of the essential snRNAs have also been observed with depletion of four nucleolar proteins, NOP1, GAR1, and NOP3, which are all essential, and a non-essential protein NSR1 (13–17). GAR1, NOP1 and a protein known as SSB1 have been shown to be associated with snoRNAs (14,18,19). Taken together, these results suggest that processing of rRNA, like that of mRNA, may be mediated by an assembly of snRNA particles in a large 'processome' complex (4).

The pool of uncharacterized snRNAs can be expected to include additional snRNAs that participate in ribosome biogenesis—at any stage—but functions unrelated to the manufacture of ribosome are also possible. The present report describes progress made in an ongoing effort to characterize novel snRNAs in the yeast *Saccharomyces cerevisiae*. Our approach is to clone new snRNA genes and determine if these genes are essential for growth. If a phenotype is observed with an inactive gene, conditional alleles can then be developed to investigate the role of the new RNA. The effectiveness of this strategy was established first for the spliceosomal snRNAs and then extended to the vital non-spliceosomal snRNAs. The properties of three novel snRNAs are described in this communication.

MATERIALS AND METHODS

Strains and media

Yeast nuclear RNA was prepared from strain BJ2407 (a/α, *leu2/leu2, trp1/trp1, ura3-52/ura3-52, prb1-1122/prb1-1122, prc-407/prc-407, pep4-3/pep4-73, gal2/gal2*; 20,21) and genomic DNA from strain GRF167 (α, *his3Δ, ura3-167*; G.Fink, unpublished). Disrupted yeast *SNR* genes were characterized in the His$^-$, Ura$^-$ strain MH2 (a/α, *ade2-101/ade2-101, trp1-Δ901/trp1-Δ901, ura3-52/ura3-52, leu2-3,112/leu2-3,112, his3/his3*; M. Fitzgerald-Hayes, unpublished). Yeast strains were grown on YPD liquid medium/plates (1% yeast extract, 2% peptone and 2% glucose) or selective plates with SD medium (2% glucose and 0.67% yeast nitrogen base) containing 20μg/ml of histidine or uracil at 32°C unless stated otherwise. Alternate

* To whom correspondence should be addressed

carbon sources included 2% galactose, 2% glycerol or 2% potassium acetate.

Yeast were transformed by a lithium acetate procedure (22). Tetrad dissections were carried out essentially as described by Sherman and Hicks (23). Diploid cells were sporulated in 1% potassium acetate, asci were treated with zymolyase (ICN Biochemicals, Inc.) and the spores separated on YPD plates.

Cloning manipulations were carried out with *E.coli* DH5αF' (F'/*end*A1, *hsd*R17(r^-m^+), *supE44, thi-1, gyrA*(Nal^r), *recA1, rel*A1, Δ(*argF-lacZYA*)U169 (*ϕ80d*lacZΔM15) or strain JM56 (*F'lacI^q-proAB, dam3,rpsL*; M.J.Fournier, unpublished) essentially as described by Sambrook *et al* (24).

RNA isolation, 3' end labeling and enzymatic sequencing

Our nuclear RNA isolation protocol was a combination of two protocols. Spheroplasts were prepared with lyticase (Sigma; 25), lysed and the nuclei were purified by centrifugation through a 0.5M sucrose cushion (26). RNA was extracted from washed nuclei with phenol and 3' end labeled by a variation of the method developed by Bruce and Uhlenbeck (27). RNA was incubated with 50 pmols 5'[^{32}P]pCp and 36 units of RNA ligase in 50 μl of solution containing: 50 mM potassium HEPES, pH 7.5, 20 mM $MgCl_2$, 1 mM ATP, 10mM DTT, 10% glycerol, 10% DMSO, 0.1 mg/ml BSA and 40 units RNasin. Individual RNAs were purified by two rounds of electrophoresis: first in a 1 mm thick gel and then in a thin sequencing gel. Electrophoresis in the latter gel was continued until RNAs differing by one nucleotide were resolved. 3' end sequences of pCp labeled RNAs were determined with a RNA sequencing kit (Pharmacia).

Gene disruption

DNA segments containing the new snRNA genes were amplified by PCR using the following primers:
SNR31 gene: 5' side, GTGCGGAATTCGATATCGGGTGC-AAT and 3' side, GGTGCGATATCGAATTCAGAAAAGTG, both with *Eco*RI sites
SNR32 gene: 5' side, GTGCGGAATTCCTCTAGAGAGACGT-CGTC and 3' side, CGTGGGAATTCCCGTGAACG-AGCCGTACA, both with *Eco*RI sites
SNR33 gene: 5' side, GCTTCCTCGAGCACACACTCTCTC-TTATT and 3' side, CAAGGGAATTCCTGATCTCTGGCG-AACGT, with *Xho*I and *Eco*RI sites, respectively.

The PCR products were cloned into shuttle vector pRS314 (*SNR32, SNR33*; 28) or the *E.coli* vector pACYC184 (*SNR31*; 29) using indicated restriction enzymes. Unique restriction sites in the *SNR32* and *SNR33* coding regions were used for marker gene insertions. An *Ase*I site within pACYC184 sequence was eliminated so that an *Ase*I site in the middle of *SNR31* could be used to insert a marker gene. *Bam*HI fragments derived from plasmids YDp-U and YDp-H (30) containing *URA3* and *HIS3* genes were used for the gene disruption analyses.

DNA isolation and analysis

The procedure used to prepare yeast genomic DNA is described in Rose *et al* (31). Southern analyses were carried out according to Sambrook *et al* (*SNR32* and *SNR33*; 24) or under conditions similar to our Northern hybridization protocol (*SNR31*). Radioactive DNA probes were prepared either by nick translation (24) or by using the New England Biolab Random Priming System I kit. A Southern blot of *Eco*RI-digested genomic DNA from nine different organisms ('Zoo-Blot') was obtained from Clontech Labs.,Inc.

Northern hybridization and characterization of terminal nucleotides

Total yeast RNA was isolated by a sodium dodecyl sulfate/hot phenol method (32). Glass beads were used to break the cells. RNA was separated in an 8% polyacrylamide gel, blotted onto Gene Screen membranes and UV cross-linked as recommended by the manufacturer. The membranes were probed with radiolabeled DNA fragments in solution containing 5×SSPE buffer, 50% formamide, 0.2% SDS and 50 μg/ml heparin (33). 20×SSPE contained 3M NaCl, 0.2M Na_2HPO_4 and 20mM Na_2.EDTA.

A Northern blot of anti-TMG immunoprecipitated RNAs from *S.cerevisiae* was kindly provided by R.A.Lempicki.

Primer extension and RNA dideoxy sequencing were carried out essentially as described (34). Three oligonucleotides were used as primers: (5'-TAGGAACTCATGGTGTAATT-3'), complementary to snR31 from position 8 to position 27; (5'-CTGATTCCCAGTGTCTCGTTTC-3'), complementary to snR32 positions 153−174; and, (5'-GATTGTCCACACAC-TTC-3'), complementary to snR33 positions 167−183.

RNA 3' end sequences were determined by a variation of a T4 RNA ligase/PCR-based approach described by Kiss and Filipovicz (35). Instead of oligoribonucleotide, an oligodeoxiribonucleotide of the same sequence (35) was used for ligation with total yeast RNA. PCR products obtained with oligonucleotides specific to snR31, snR32 and snR33 were gel-purified and sequenced directly with the same snRNA-specific oligonucleotides by using the dsDNA Cycle Sequencing System kit from BRL Life Technolgies, Inc. Oligonucleotides were: (5'-GGAAGCCTTTCTCTTTCACC-3'), corresponding to positions 161−170 of snR31; (5'-GATAGATTGAACGTTGC-TGG-3'), corresponding to positions 111−130 of snR32; and, (5'-GGTTTGAGTCGGTTCCTTCG-3'), corresponding to positions 124−143 of snR33.

Computer sequence analysis

Analyses of RNA and DNA sequences were performed with the programs designed by the University of Wisconsin Genetics Computer Group (36).

RESULTS

Isolation and 3' sequencing of new snRNAs

The snRNA genes described here were identified by partial sequencing of purified RNA and searching the GenBank DNA sequence database for related sequences. The RNA sequences were developed in the following manner. Nuclei were purified by centrifugation through a 0.5M sucrose cushion and extracted with spheroplast lysis buffer. The snRNAs resistant to this treatment were then isolated by phenol extraction. The resulting RNAs were labeled at the 3' ends with [^{32}P]-pCp and fractionated by gel electrophoresis (Fig.1). Several RNAs which appeared to be amenable to isolation were purified by two rounds of electrophoresis. Nine novel snRNAs and a number of species already characterized by others were identified by enzymatic sequencing.

Discovery of three gene sequences in a DNA database

Homologous DNA sequences were found in the GenBank database for three of the nine novel RNA sequences (Figs.1 and 2). After this work was completed we learned that the RNA

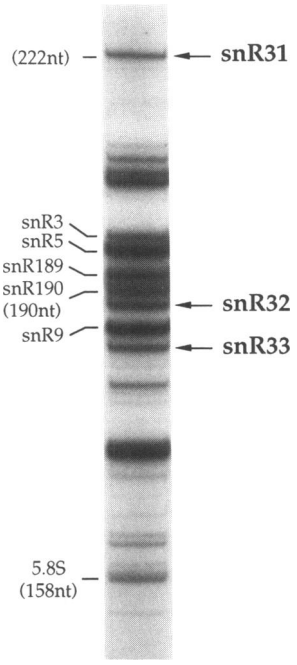

Figure 1. Fractionation of nuclear small RNAs. Small RNA was prepared from isolated nuclei pre-extracted with 25mM K·HEPES pH7.6, 5mM magnesium acetate, 0.25 mM sucrose and 0.1% NP-40. SnRNAs resistant to this treatment were labeled at the 3′ end with [^{32}P] pCp and fractionated by electrophoresis in a 8% polyacrylamide gel. The pattern shown is for snRNAs in the size range of 158−222 nucleotides. The bands corresponding to snRNA species snR31, snR32 and snR33 were purified by a second cycle of electrophoresis and subjected to direct sequence analysis. The identity of the known RNAs marked at the left was established by direct sequencing.

estimated to contain approximately 220 nucleotides corresponds to a species already named snR31 (3 and D.Tollervey, pers. commun.). To avoid confusion we have elected to use the same identifying name for this RNA, and to adopt consecutively numbered designators for the two other RNAs described here, i.e., snR32 and snR33. This system of nomenclature will also be used for several other new snRNAs under study.

Each of the RNA coding sequences found in the database occurs in genomic segments that have been sequenced for unrelated reasons. DNA corresponding to snR31 is in a segment (GenBank accession number X15731) that specifies an unnamed 20 kDa protein believed to be associated with the translation factor eIF-4E (37; Figs.2A and 3). A DNA segment (GenBank accession number M62392) corresponding to snR32 is adjacent to one of two sequenced genes for ribosomal protein L41 (38; Figs.2B and 4). The snR31 and snR32 coding sequence are located 99 bp and 268 bp respectively downstream of the protein genes, in a tail-to-tail orientation.

The segment encoding snR33 was found in the completely sequenced chromosome III (GenBank accession number X59720), between nucleotide positions 141,410 and 142,838 (39; Figs.2C and 5). The snR33 sequence occurs between two uncharacterized open reading frames located 1200 and 200 bp up- and downstream, respectively; the two ORF elements and snR33 coding sequence are oriented in the same direction. A gene for a glycine tRNA (GenBank accession number K02654) and a solo delta element also occur upstream, approximately 160 and 310 nucleotides from the start of the snR33 coding sequence.

SNR31

```
        10         20         30         40         50
GAATTCAGAA AAGTGAAGCA AAATTACACC ATGAGTTCCT ATTAACGTCA  50
GCCTCTTTTG TACTCATAAT GTGGCTTCGG ATGTTTGATG GGCGTTGCTT 100
CAGTGCGTAC GGCTCATGGT AGATTAATTA TTAGAAAGAT GTATCTCCAG 150
CTGTTGATAT TAGAGGGGGA AGCCTTTCTC TTTCACCTCG CCTTTTTAAA 200
CACCTGATAC AGTTGGTCAT GATTCGTTCT ACATTTTAAT TTCTTTATGT 250
AACCACGTTG TATATTTTCA ATCTGGAAAG TAACCTAAAA GAGTAAAAAC 300
TGTTTATTAA AAAAAATGTT ATTCAAAATT TATACATAGT ACACGTACAG 350
CTGCCTGAAG AAAGCATGAG CGAATGACTT ATGCTTCGTC GTCTTCGTCT 400
```

SNR32

```
        10         20         30         40         50
TTCGTGCTAG CCCTACAACG GAACCCAAAA AATCAATCAC CAGGGTAAGC   50
                                 *  *****
GGGAAAAAAA AAAAGTGTAG GAACCGCACT ATAAAAAGTC GAAGGTTTGT  100
ATTAATTTTC CACACCTCTT TAGTTTTGTT ATAATCAAAT TCCTTTTGGT  150
TTTTGTTTTA AAAAGGCCAA CATCATAATT ATATCGCGAA TGTACTACTA  200
GTATATGCAG TTTTCTTAAA ATTGCAGGGA ATGACTAAAA GCCCACGTTT  250
TTCCCACTTT TAGTAGTACT GAAGCGTAGA TAGATTGAAC GTTGCTGGGC  300
GCCTGGTGTT GATCATTTCT GAAATGAGAT ATTGGGAATC AGCTGTATCT  350
ACATTTTATT TATACATATT TATTTTTACT GGTTCTTAGT AACGTACTTT  400
AGTAGCAATG CTGTGTTTTA GTCGAACGCC TACGTGTCTT GTTAACTTAT  450
CATGGTTGCC AACGCGAACG TTTACAACAG GTTCAAAGGC AGGATGAAAA  500
```

SNR33

```
        10         20         30         40         50
                           ******
GCTTGCGCAT CTTTTTTTAC TTTATATACT AGTTTTTTTT TTTTTTTTCT   50
TTTTCCCAAA TTTTTTCATG AAAAATTTGG CGGAACGGTA CATAAGAATA  100
GAAGAGATTC ATTATGAAAA TTTTCTGCTC TCTTCCACAT TTTTTTTTCA  150
TAAGAATTAA AAAAAATAAG CTTAATGCCC TCTTTGTACG ATGGTGTCAC  200
TCTAAAATTA TTTAGGTAAT TTAATTGGAT GTGATATCAA GCACTAGTTT  250
TGTTTGGAGG AGCAGAGATT GAAAGCCTAG CTTTTACACC GGTTTGAGTC  300
GGTTCCTTCG TTTTCTTAGT TGGAATTGAT ATAGAAGTGT GTGGACAATC  350
GATTTGCAAA ACTTTTTATT TTATGTATTA CTACTGAGCT TGTTTATCTG  400
TATACTTTTT TCTTATTATT ATGTAGTTAG AGGTTAGTTC TTATCGTTTT  450
CTTTAAAAGC GACCTTTCTT CGCACAAAGC AATGTTGACA ATCGAGCCAA  500
```

Figure 2. Sequences of the gene regions encoding snR31, snR32 and snR33. The snRNA genes were identified by searching the GenBank DNA sequence database for related sequences. The snRNA sequences are highlighted by italics and underlines. The nearest candidate TATA elements are marked with overhead asterisks (*). The 5′ and 3′ ends of the coding sequences were determined by primer extension and a T4 RNA ligase/PCR-based sequencing strategy (see Materials and Methods).

Sequence and termini of new snRNAs

The RNA sequences were initially deduced from: the 3′ end partial sequences, the sizes of the RNAs (estimated electrophoretically), and the sequences from the DNA database. The actual boundaries of the coding sequences (Fig.2) were later determined by precise mapping of the 5′ and 3′ ends of the RNAs.

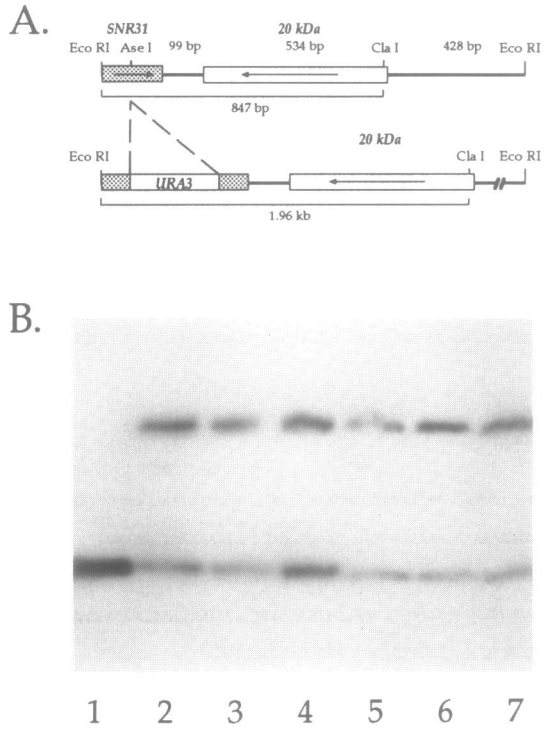

Figure 3. Cloning and disruption of the *SNR31* gene region. **A:** A 1.3 kb fragment of genomic DNA was PCR amplified, cloned and disrupted with the *URA3* gene. The snR31 coding sequence is located 99 bp downstream of a gene for a 20 kDa protein believed to be associated with the m^7G cap binding protein eIF-4E. The *URA3* gene fragment was inserted at an *AseI* site in the midst of the *SNR31* gene. **B:** Southern analysis of DNA from URA^+ transformants. DNA digested with *Eco*RI and *Cla*I was probed with the 847 bp *Eco*RI−*Cla*I fragment radiolabeled by nick-translation. Lane 1, DNA from homozygous wild type cells; lanes 2−7, DNA from heterozygotes containing wild type and null alleles of *SNR31*. The wild type DNA fragment is 860 bp and the disrupted gene fragment is 1.9 kb.

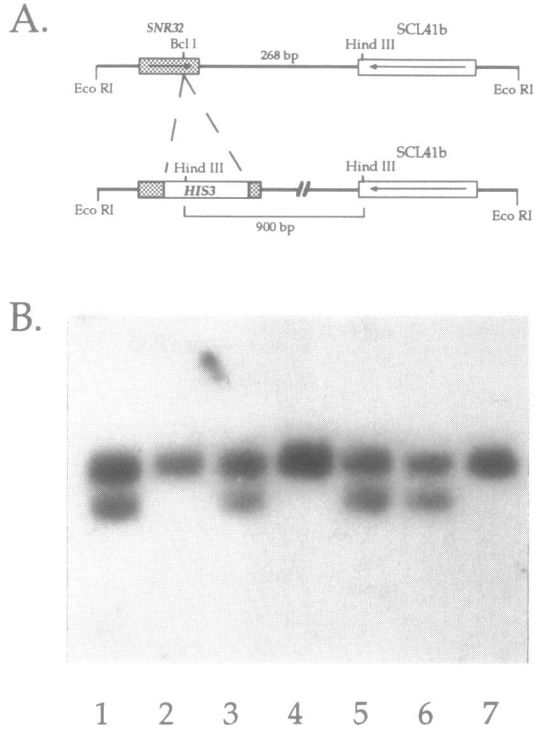

Figure 4. Cloning and disruption analysis of the *SNR32* gene locus. **A:** The snR32 coding segment is located 268 bp downstream of the gene for ribosomal protein L41b (*SCL41b*), in the opposite orientation. A 1.8 kb DNA fragment containing both genes was cloned by PCR amplification and the *SNR32* gene was disrupted with a *HIS3* gene fragment inserted at a *Bcl*I site. The cloned region extends from nucleotide −378 upstream of the snRNA segment to a position 1253 bp beyond the snR32 coding sequence; the latter site is 211 bp upstream of *SCL41b*. The structures of the wild type and *snr185::HIS3* alleles are shown. **B:** Southern blot analysis of DNA from HIS^+ transformants and wild type cells. Genomic DNA was digested with *Hind*III and probed with the 525 bp *Eco*RI−*Bcl*I fragment radiolabeled by random priming. Lanes 1,3,5 and 6, DNAs from heterozygous isolates; lanes 2,4 and 7, DNA from cells homozygous for *SNR32* (wild type). The allele disrupted with *HIS3* yields a fragment 900bp smaller than the fragment derived from the wild type allele.

The snR32 and snR33 sequences were verified by reverse transcriptase dideoxy sequencing and the 5′ ends were established at the same time. The 5′ end of snR31 was mapped by primer extension analysis only.

The 3′ terminal nucleotides were identified by a clever approach described by Kiss and Filipovicz (35), applied with some modifications. In the original procedure an RNA oligonucleotide is first ligated to the 3′ end of the RNA and cDNA synthesis is primed with a DNA oligonucleotide complementary to this segment. The cDNA is then amplified by PCR using a second DNA oligonucleotide corresponding to an internal RNA sequence. The PCR product is then subcloned into a plasmid and sequenced. We determined that more readily available DNA oligonucleotides can be used in the initial ligation with RNA and that the sequence of the joint RNA−DNA molecule can be successfully amplified by PCR. We omitted the subcloning and sequenced the PCR products directly after gel-purification, as the 3′ ends of all three RNAs had been determined to be homogeneous from enzymatic sequencing.

SnR31, snR32 and snR33 were determined to have a 5′ trimethylguanosine cap (TMG), based on hybridization probing of electrophoretically fractionated snRNA prepared by precipitation with TMG antibodies (data not shown).

All three snRNAs are encoded by single copy, non-essential genes

The availability of the genomic sequence information made it possible to clone each gene by a PCR amplification strategy (Figs.3−5). Cloned DNAs were used to prepare hybridization probes and to construct null alleles to test the essentiality of the genes. The gene copy number was determined by Southern analysis of genomic DNA digested with a variety of restriction enzymes. The data showed each snRNA to be encoded by a single copy gene.

The essentiality of the snRNA genes was tested by replacing the wild type *SNR* genes with alleles disrupted by marker genes. *URA3* and *HIS3* were simply inserted in the midst of *SNR31* and *SNR32* genes respectively (Figs.3A, 4A). A null allele of *SNR33* was prepared by replacing the 3′ half of the snRNA coding sequence with *URA3* (Fig.5A). The disrupted alleles were transformed into Ura^- or His^- test strains and transformants heterozygous for the *SNR* loci were identified by growth phenotype and Southern blot analysis (Figs.3B, 4B, 5B).

Tetrad analysis showed the genes for snR31, snR32 and snR33 to be dispensable. Northern assays carried out with cells from

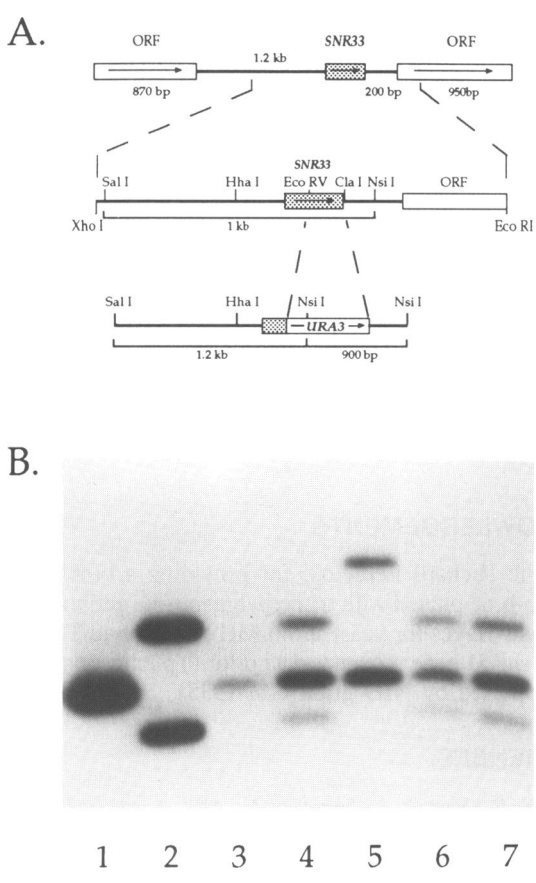

Figure 5. Cloning and disruption analysis of the *SNR33* gene region. **A:** A 1.4 kb DNA fragment containing the snR33 gene was cloned by PCR amplification and the 115 bp *Eco*RV−*Cla*I from the gene was replaced with a *URA3* marker gene. The structures of the wild type and disrupted alleles (*snr33::URA3*) are shown. The cloned region extends from nucleotides 141,410 to 142,838 on chromosome III. **B:** DNA from URA$^+$ transformants was screened by Southern blot analysis, by probing *Nsi*I/*Sal*I digests with the 500 bp *Nsi*I−*Hha*I fragment radiolabeled by random priming. Lanes 1 and 3, DNA from a homozygous wild type strain and plasmid containing intact *SNR33* (1kb fragment); lane 2, plasmid with a disrupted allele (0.9 kb and 1.2 kb fragments); lanes 4,6, and 7, genomic DNA from transformants heterozygous for the wild type and disrupted alleles; lane 5, DNA from a heterozygote containing the disrupted allele with the *URA3* gene inserted in the opposite orientation.

individual spores confirmed the absence of the test RNAs in haploid cells containing the disrupted allele (Fig.6). No phenotypic difference was evident for the haploid cells at 16°C, 25°C, 32°C or 38°C. Cells lacking snR33 exhibited similar growth rates on media containing 2% glucose, 2% galactose, 2% glycerol or 2% potassium acetate; the effect of carbon source was not tested for cells lacking the other two snRNAs. The fact that each of the RNAs is dispensable precludes the possibility of using conditional mutants to study function as has been done for several essential snRNAs.

Relatedness of snR31, snR32 and snR33 to other RNAs

Computer searches were carried out with each RNA sequence to determine: i) if the snRNAs are evolutionarily conserved; ii) if sequence motifs conserved in other snRNAs are present, such as the box A−D elements in several nucleolar snRNAs; and, iii) if the new RNAs contain segments of complementarity with rRNA that might suggest roles in rRNA synthesis.

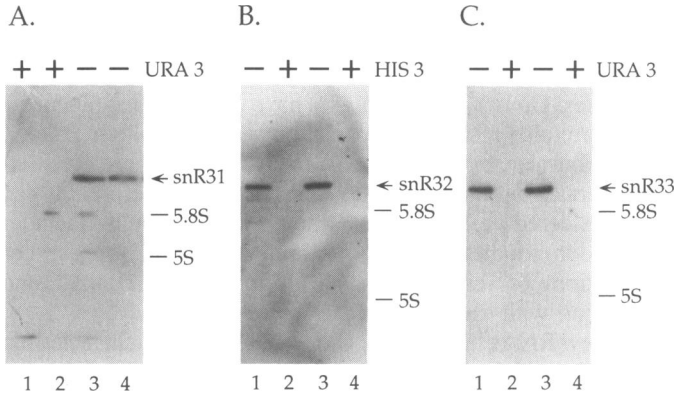

Figure 6. Northern analysis of cells containing disrupted *SNR* genes.Haploid cells containing single disrupted alleles for snR31, snR32 and snR33 were analyzed by Northern blotting (panels A−C). Signals corresponding to the test RNAs were absent in each case. Each series includes RNA isolated from the four spores of a single tetrad (lanes 1−4). (−) or (+) corresponds to spores containing wild type alleles or alleles disrupted with marker genes (*URA3* or *HIS3*). A 844 bp *Hind*III−*Eco*RI fragment (see Fig.4) labeled by random priming was used as a probe for snR32. Probes for snR31 and snR33 are identified in Figs.3 and 5. The positions of 5S and 5.8S RNAs are marked.

No impressive relationship to other sequenced snRNAs was found. The search for snRNA-conserved elements was also negative, with no recognizable element in any of the three species. Conservation of the snR32 and snR33 sequences was also examined experimentally by Southern analysis. No signals were detected with genomic DNAs from human, monkey, rat, mouse, dog, cow, rabbit and chicken (data not shown).

As expected, many segments of complementarity exist between the three snRNAs and pre-rRNA, because of the large size of the rDNA coding unit. While none of these may be relevant the largest regions will be described since they have the highest potential for direct binding; of course, important interactions could also occur through pairing of short regions. The longest complementary regions in rRNA were: 1) a 12 base segment within the 5′ external transcribed sequence, at nucleotides 567−578 from the 5′ end of pre-rRNA (snR32); 2) a 10 base segment within internal transcribed sequence 1, at nucleotides 2447−2456 (snR33); and, 3) a 19 base segment in a region that includes an enhancer element and the termination signal T2, at nucleotides 6843−6861 (snR31; 40). Although each of these snRNA-related segments can be imagined to have a role in rRNA synthesis, it is not clear that any of the complementarities are statistically significant. A similar search using spliceosomal snRNA sequences as controls yielded complementarities upwards of 10 contiguous bases. Many of the test and control complementarities involve stretches of adenines and uridines and the high AT content of the yeast genome argues that these relationships may not be important statistically.

DISCUSSION

The three genes identified bring the number of *S.cerevisiae* snRNA genes analyzed to date to 19, corresponding to 5 spliceosomal and 14 non-spliceosomal RNAs (reviewed in 3,4). The novel snRNAs described here have extraction properties like those of known nucleolar snRNAs (7,12, data not shown), however, it is premature to conclude that these RNAs are actually

nucleolar. Resistance to extraction at modest salt concentration is believed to be a property for nucleolar snRNAs (12), but this behavior could also apply to other RNAs associated with stable complexes. Consistent with this we have observed that precursor tRNAs are also resistant to extraction with low salt (A.Balakin and M.Fournier, in prep.). The snRNAs identified here may well be nucleolar, but more direct evidence is needed before this can be considered established. Other criteria could include: localization to the nucleolus with hybridization probes, co-purification or cross-linking with precursor rRNA, and association with known nucleolar proteins.

The new RNAs lack the box C or box D elements characteristic of several nucleolar RNAs from yeast and other sources (41–44), but this does not argue against a nucleolar location. A number of other snRNAs exhibit nucleolar localization properties, but lack box C and D. These include the yeast snRNA species snR3, -4, -5, -8, -9 and -10 and human HeLa snRNAs U17 (E1), E2, and E3 (12,34,45). The yeast snR31, snR32 and snR33 species could belong to this latter group of snRNAs.

Several snRNAs participating in processing of mRNA or rRNA are believed to function through complementary interactions with substrate molecules and each other (4,46). The yeast nucleolar snRNAs U3 and U14, for example, contain rRNA-related sequences which have been implicated in rRNA processing experimentally. Yeast U3 has been cross-linked to rRNA at complementary sites (47) and mutational analysis has shown that U14 activity depends on one of two segments complementary to 18S rRNA (48). At this point the only link between the snRNA species described here and rRNA synthesis is the tenuous nucleolar localization data and complementarity with several regions of ribosomal DNA. It remains to be seen if any of the new snRNAs has any role in ribosome biogenesis.

The finding that each of the the novel snRNAs is derived from a single gene was not surprising. All but one of the *S.cerevisiae* snRNAs characterized thus far are encoded by single-copy DNA (49). The exception is U3 which is specified by two genes, both containing introns (6,50). No other snRNA genes from *S.cerevisiae* are known to contain introns, including the three described in the present study.

The proximity of the new snRNA genes with protein genes or ORF elements at the same locus raises the possibility of a functional relationship between the RNA and protein genes. The involvement of snRNAs in ribosome biogenesis makes the linkage between the snRNA genes and genes for a putative mRNA cap binding protein (snR31; 37) and ribosomal protein (snR32; 38) especially intriguing. A similar situation also exists for the yeast snRNA species snR189 which is located 300 bp from a gene for ribosomal protein rp59 (*CRY1*; 2). The *SNR31* and *SNR32* genes are in a tail-to-tail arrangement with the protein genes, whereas *SNR189* and *CRY1* are in a tail-to-head arrangement. The closeness of these genes suggests a possible regulatory link between ribosome biogenesis and translation. This notion is strengthened by recent discoveries showing that several small nucleolar RNAs in higher eukaryotes are encoded within introns of ribosomal protein genes and genes for a heat shock protein (41, 43, 51).

Of the 14 non-spliceosomal snRNA genes analyzed in *S.cerevisiae* thus far, only four have been shown to influence growth and one of these, snR10, is non-essential (6–8,52). All of the non-spliceosomal snRNAs characterized are believed to be nucleolar and, for this reason predicted to be involved in ribosome biogenesis. If the present species should prove to be nucleolar as well, the number of dispensable small RNAs associated with this complex increases to 11, corresponding to 80% of the known nucleolar small RNAs.

As the number of non-essential species expands it seems increasingly unlikely that these RNAs are simply redundant, with no distinguishing function. More plausible is the possibility that these RNAs provide overlapping or even unique advantages that are not essential. Dispensable snRNAs could, for example, influence the kinetics or accuracy of rRNA processing or the rate of assembly of ribosomal RNP intermediates with little or no effect on growth rate. Unless ribosome content becomes limiting growth phenotypes would not be observed with loss of non-essential snRNAs. The contributions of such RNAs would only be apparent with quantitative biochemical assays or discovery of genetic conditions where deleterious effects of snRNA loss are amplified.

ACKNOWLEDGEMENTS

We thank Richard Lempicki for providing a blot of snRNAs immuno-precipitated with trimethylguanosine antibodies, Molly Fitzgerald-Hayes for yeast strain MH2 and David Tollervey for sharing snR31 gene sequence data prior to publication. This work was supported by NIH grant GM-19351.

REFERENCES

1. Riedel, N., Wise, J. A., Swerdlow, H., Mak, A. and Guthrie, C. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 8097–8101.
2. Thompson, J.R., Zagorski, J., Woolford, J.L. and Fournier, M.J. (1988) *Nucleic Acids Res.* **16**, 5587–5601.
3. Mattaj, I.W., Tollervey, D. and Seraphin, B. (1993) *FASEB J.* **7**, 47–53.
4. Fournier, M.J. and Maxwell, E.S. (1993) *Trends Biochem. Sci.* **18**, 131-135.
5. Sollner-Webb, B., Tyc, K. and Steitz, J. (1993) in *Ribosomal RNA: Structure, Evolution, Gene Expression and Function in Protein Synthesis* (Zimmermann, R.A. and Dahlberg, A.E., eds), CRC Press (in press).
6. Hughes, J.M.X., Konings, D.A.M. and Cesareni, G. (1987) *EMBO J.* **6**, 2145–2155.
7. Zagorski, J., Tollervey, D. and Fournier, M.J. (1988) *Mol. Cell. Biol,* **8**, 3282–3290.
8. Bally, M., Hughes, J. and Cesareni, G. (1988) *Nucleic Acids Res.* **16**, 5291–5303.
9. Hughes, J.M.X. and Ares, Jr., M. (1991) *EMBO J.* **10**, 4231–4239.
10. Li, H.V., Zagorski, J. and Fournier, M. (1990) *Mol. Cell. Biol.* **10**, 1145–1152.
11. Morrissey, J.P. and Tollervey, D. (1993) *Mol. Cell. Biol.* **13**, 2469–2477.
12. Tollervey, D. (1987) *EMBO J.* **6**, 4169–4175.
13. Tollervey, D., Lehtonen, H., Carmo-Fonseca, M. and Hurt, E.C. (1991) *EMBO J.* **10**, 573–583.
14. Girard, J.-P., Lehtonen, H., Caizergues-Ferrer, M., Amalric, F., Tollervey, D. and Lapeyre, B. (1992) *EMBO J.* **11**, 673–682.
15. Russell, I.D. and Tollervey, D. (1992) *J. Cell Biol.* **119**, 737–747.
16. Lee, W-C., Zabetakis, D. and Melese, T. (1992) *Mol. Cell. Biol.* **12**, 3865–3871.
17. Kondo, K., Kowalski, L.R. and Inouye, M. (1992) *J. Biol. Chem.* **267**, 16259–16265.
18. Schimmang, T., Tollervey, D., Kern, H., Frank, R., and Hurt, E.C. (1989) *EMBO J.* **8**, 4015–4024.
19. Clark, M.W., Yip, M.L., Campbell, J. and Abelson, J. (1990) *J. Cell Biol.* **111**, 1745–1751.
20. Jones, E.W. (1977) *Genetics* **85**, 23–33.
21. Zubenko, G.S., Mitchell, A.P., and Jones, E.W. (1980) *Genetics* **96**, 137–146.
22. Rose, M.D., Winston, F. and Hieter, P. (1990) *Methods in Yeast Genetics: A Laboratory Course Manual.* Cold Spring Harbor Laboratory Press, Cold Spring Harbor Laboratory pp. 122–123.
23. Sherman, F. and Hicks, J.B. (1991) *Methods Enzymol.* **194**, 21–37.
24. Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual.* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

25. Aris, J.P., and Blobel, G. (1991) *Methods Enzymol.* **194**, 735–749.
26. Wise, J.A. (1991) *Methods Enzymol.* **194**, 405–415.
27. Bruce, A.G., and Uhlenbeck, O.C. (1978) *Nucleic Acids Res.* **5**,3665–3677.
28. Sikorski, R.S. and Hieter, P. (1989) *Genetics* **122**, 19–27.
29. Chang,A.C.Y., and Cohen,S.N. (1978) *J. Bacteriol.* **134**, 1141–1156.
30. Berben, G., Dumont, J., Gilliquet, V., Bolle, P.-A. and Hilger, F. (1991) *Yeast* **7**, 475–477.
31. Rose, M.D., Winston, F. and Hieter, P. (1990) *Methods in Yeast Genetics: A Laboratory Course Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor Laboratory pp. 130–132.
32. Kohrer, K., and Domdey, H. (1991) *Methods Enzymol.* **194**, 398–405.
33. Singh, L., and Jones, K.W. (1984) *Nucleic Acids Res.* **12**, 5627–5638.
34. Shelness, G.S., and Williams, D.L. (1985) *J. Biol. Chem.* **260**, 8637–8646.
35. Kiss, T. and Filipovicz, W. (1993) *EMBO J.* **12**, 2913–2920.
36. Devereux, J., Haeberli, P. and Smithies, O. (1984) *Nucleic Acids Res.* **12**, 387–395.
37. Altmann, M., Krieger, M. and Trachsel, H. (1989) *Nucleic Acids Res.* **17**, 7520.
38. Kawai, S., Murao, S., Mochizuki, M., Shibuya, I., Yano, K. and Takagi, M. (1992) *J. Bacteriol.* **174**, 254–262.
39. Oliver, S.G., van der Aart, Q.J.M., Agostoni-Carbone, M. L., Aigle, M., Alberghina, L., Alexandraki, D., Antoine, G., Anwar, R., Ballesta, J.P.G., Benit, P., Berben, G., Bergantino, E., Biteau, N., Bolle, P. A., Bolotin-Fukuhara, M., Brown, A., Brown, A.J.P., Buhler, J.M.,Carcano, C., Carignani, G., Cederberg, H., Chanet, R., Contreras, R., Crouzet, M., Daignan-Fornier, B., Defoor, E., Delgado, M., Demolder, J., Doira, C., Dubois, E., Dujon, B., Dusterhoft, A., Erdmann, D., Esteban, M., Fabre, F., Fairhead, C., Faye, G., Feldmann, H., Fiers, W., Francingues-Gailard, M.C., Franco, L., Frontali, L., Fukuhara, H., Fuller, L.J., Galland, P., Gent, M.E., Gigot, D., Gilliquet, V., Glansdorff, N., Goffeau, A., Grenson, M., Grisanti, P., Grivell, L.A., de Haan, M., Haasemann, M., Hatat, D., Hoenicka, J., Hegemann, J., Herbert, C.J., Hilger, F., Hohmann, S., Hollenberg, C.P., Huse, K., Iborra, F., Indge, K.J., Isono, K., Jacq, C., Jacquet, M., James, C.M., Jauniaux, J.C., Jia, Y., Jimenez, A., Kelly, A., Kleinhans, U., Kreisl, P., Lanfranchi, G., Lewis, C., van der Linden, C.G., Lucchini, G., Lutzenkirchen, K., Maat, M.J., Mallet, L., Mannhaupt, G., Martegani, E., Mathieu, A., Maurer, C.T.C., McConnell, D., McKee, R.A., Messenguy, F., Mewes, H.W., Molemans, F., Montague, M.A., Muzi Falconi, M., Navas, L., Newlon, C.S., Noone, D., Pallier, C., Panzeri, L., Pearson, B.M., Perea, J., Philippsen, P., Pierard, A., Planta, R.J., Plevani, P., Poetsch, B., Pohl, F., Purnelle, B., Ramezani Rad, M., Rasmussen, S.W., Raynal, A., Remacha, M., Richterich, P., Roberts, A.B., Rodriguez, F., Sanz, E., Schaaff-Gerstenschlager, I., Scherens, B., Schweitzer, B., Shu, Y., Skala, J., Slonimski, P.P., Sor, F., Soustelle, C., Spiegelberg, R., Stateva, L.I., Steensma, H.Y., Steiner, S., Thierry, A., Thireos, G., Tzermia, M., Urrestarazu, L.A., Valle, G., Vetter, I., van Vliet-Reedijk, J.C., Voet, M., Volckaert, G., Vreken, P., Wang, H., Warmington, J.R., von Wettstein, D., Wicksteed. B.L., Wilson, C., Wurst, H., Xu, G., Yoshikawa, A., Zimmermann, F.K. and Sgouros, J.G. (1992) *Nature* **357**, 38–46.
40. Raue, H.A. and Planta, R.J. (1991) *Prog. Nucl. Acid Res. Mol. Biol.* **41**, 89–129.
41. Tyc, K. and Steitz, J.A. (1989) *EMBO J.* **8**, 3113–3119.
42. Huang, G.M., Jarmolowski, A., Struck, J.C.R., and Fournier, M. (1992) *Mol. Cell. Biol.* **12**, 4456–4463.
43. Tycowski, T.K., Shu, M-D. and Steitz, J.A. (1993) *Genes Devel.* **7**, 1176–1190.
44. Fragapane, P., Prislei, S., Michienzi, A., Caffarelli, E. and Bozzoni, I. (1993) *EMBO J.* **12**, 2921–2928.
45. Ruff, E.A., Rimoldi, O.J. and Eliceiri, G.L. (1993) *Proc. Natl. Acad. Sci. USA* **90**, 635–638.
46. Steitz, J. (1992) *Science* **257**, 888–889.
47. Beltrame, M. and Tollervey, D. (1992) *EMBO J.* **11**, 1531–1542.
48. Jarmolowski, A., Zagorski, J., Li, H.V. and Fournier, M.J. (1990) *EMBO J.* **9**, 4503–4509.
49. Wise, J.A., Tollervey, D., Maloney, D., Swerdlow, H., Dunn, E. and Guthrie, C. (1983) *Cell* **35**, 743–751.
50. Myslinski, E., Segault, V. and Branlant, C. (1990) *Science* **247**, 1213–1216.
51. Liu, J. and Maxwell, E.S. (1990) *Nucleic Acids Res.* **18**, 6565–6571.
52. Tollervey, D. and Guthrie, C. (1985) *EMBO J.* **4**, 3873–3878.