Identification and characterization of two novel box H/ACA snoRNAs from *Schizosaccharomyces pombe* *

LI Siguang, ZHOU Hui, ZHANG Peng, LUO Yuping and QU Lianghu **

(Key Laboratory of Gene Engineering of the Ministry of Education, Biotechnology Research Center, Zhongshan University, Guangzhou 510275, China)

Received September 16, 2004

Abstract Through constructing a specialized cDNA library based on small RNAs isolated from partial purified nuclei of *Schizosaccharomyces pombe*, two novel noncoding RNAs, termed Sp15-70 and Sp18-61, have been identified. Bioinformatics analysis reveals that both the novel RNAs possess a typical secondary structure of box H/ACA snoRNA and antisense elements to rRNAs. According to the relationship between the structure and function of box H/ACA snoRNA, Sp15-70 was predicted to direct pseudouridylation in 25S rRNA at U2401 and U2298; Sp18-61 was predicted to direct pseudouridylation in 18S rRNA at U208 and 25S rRNA at U2341. The four predicted pseudouridylation sites were all verified experimentally by the CMC-primer extension analysis. Both Sp15-70 and Sp18-61 were encoded by single copies which were located in the intergenic regions between the CDS of two protein-coding genes on chromosome Ⅰ and Ⅲ of *S. pombe*, respectively. Putative TATA-like elements can be found upstream from the 5′ end of these snoRNA genes, suggesting that they could be transcribed from their own promoters. Comparison of the two snoRNAs and their functional homologues in diverse organisms reveals that extensive recombinations among different snoRNAs have occurred during the evolution from their primitive progenitors.

Keywords: box H/ACA snoRNA, RNA pseudouridylation, gene evolution, *Schizosaccharomyces pombe*.

In eukaryotes, a large number of small nucleolar RNAs (snoRNAs) accumulated within the nucleolus play important roles in precursor ribosomal RNA (pre-RNA) processing and maturation[1]. All snoRNAs, with the exception of RNase MRP, can be broadly divided into two expending groups, box C/D and H/ACA snoRNAs, based on conserved sequence elements[2]. Box C/D snoRNAs contain two conserved short sequence motifs, box C (UGAUGA) and box D (CUGA), located only a few nucleotides away from the 5′ and 3′ ends, respectively, generally as part of a typical 5′-3′ terminal stem-box structure. Guiding of 2′-O-methlation involves base-pairing of 10 to 21 nt-long sequence positioned upstream of box D (or D′) to the target RNA, with the nucleotide positioned 5 base pairs upstream of the D/D′ box being selected for methylation[3,4]. Box H/ACA snoRNAs are characterized by the presence of a "hairpin-hinge-hairpin-tail" secondary structure and two conserved sequence motifs, the H (ANANNA) and ACA boxes, located in the single-stranded hinge and tail regions, respectively. Most of known box H/ACA snoRNAs are rRNA pseudouridylation guides and possess bipartite guide sequences in the internal loop of one or both of the two hairpin domains. The guide sequences complementary to target RNAs can form a 9—13 bp duplex that defines the site of pseudouridylation[5,6]. Recent studies showing the diversity and complexity of two snoRNA guide families continue to expand with the discovery of snoRNAs possessing novel structure and function. A novel type of RNAs, exhibiting structural features of both box C/D and H/ACA snoRNAs, have been demonstrated in mammal to function as a guide to RNAs in site-specific synthesis of 2′-O-ribosome-methylated nucleotides and pseudouridines in spliceosomal small nuclear RNAs (snRNAs) transcribed by the RNA polymerase Ⅱ[7,8]. However, these guide RNAs exclusively accumulate within the cajal bodies rather than nucleolus and hence are termed scaRNAs (small cajal body-specific RNAs). Homologues of the two families of snoRNAs have been recently identified in Archaea[9—11] and even a box C/D snoRNA homolog from Archaea plays a role in tRNA methylation[12]. In addition, an increasing number of novel members of "orphan snoRNAs"[11,13—16], imprinted snoRNAs and tissue-specific snoRNAs have been identified and characterized[17,18]. It is certain that the structure and function of snoRNAs are far more abundant and complex than what have been anticipated.

* Supported by National Natural Science Foundation of China (Grant Nos. 30230200 and 30170216)
** To whom correspondence should be addressed. E-mail: lsbrc04@zsu.edu.cn
?1994-2018 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

Schizosaccharomyces pombe is a single-celled free-living organism. Although both *S. pombe* and *S. cerevisiae* belong to Ascomycetes, many findings show that *S. pombe* differs from *S. cerevisiae* but resembles higher eukaryotes in some aspects such as pre-snRNA processing and modification[19,20]. So it has been served as an excellent model organism for the study of pre-RNA processing and modification. Identification of *S. pombe* snoRNAs and study on their structure, function and expression are significant in revealing the mechanisms of modification and maturation of RNAs in eukaryotes. To date, 21 box H/ACA snoRNAs have been characterized in the budding yeast *S. cerevisiae*[2,21], while only 3 box H/ACA snoRNAs were identified in *S. pombe* and one of them was verified to be served as a guide for pseudouridylation in rRNA[22]. Obviously, the box H/AC snoRNAs, especially the guides for pseudouridylation, have been underestimated in the fission yeast genome. With the goal of finding and identifying novel box H/ACA snoRNAs, we have constructed a specialized cDNA library with small RNAs from *S. pombe* nucleoli. Through screening of the cDNA library and DNA sequencing, six sequences exhibiting the conserved motifs of box H/ACA snoRNA were obtained and two of these were confirmed to be box H/ACA snoRNAs which were expressed stably in cells. The structures, functions and genomic organizations of these two snoRNAs were further analyzed and an evolutionary mechanism for the snoRNAs was discussed.

1 Materials and methods

1.1 Strains and media

The *S. pombe* wild-type haploid strain sp972 was used for construction of the cDNA library and all RNA analyses. This strain was grown in rich YPD medium (1% yeast extract, 1% peptone, 2% glucose) at 30 ℃. *Escherichia coli* strain TG1 [F′/ *supE*, *hsd* △5, *thi* △(*lac-proAB*)] grown on 2× YT(1.6% Bacto tryptone, 1% Bacto yeast extract, 0.5% NaCl) liquid or solid medium was used for all cloning procedures.

1.2 Nuclear RNAs isolation and cDNA library construction

Preparation and purification of nuclei from *S. pombe* were essentially performed as described previously[23]. Briefly, after the digestion of the cell by lywallzyme, spheroplasts were lysed using a homogenizer in the F-buffer containing 18% Ficoll 400, 20 mmol/L PIPES, 0.5 mmol/L $MgCl_2$ and 1 mmol/L PMSF. The lysate was then transferred to 50%, 40% and 30% Ficoll step gradients and centrifuged at 58000 g for 60 min at 2 ℃. The nuclei occurred on the 40%—50% interface. Nuclear RNAs were then isolated and purified from the nucleus pellet by guanidine thiocyanate/phenol-chloroform extraction[24].

For cDNA library construction, 8 μg of nuclear RNAs were tailed with ATP by using the poly (A) polymerase (Takara). A-tailed RNA was subsequently reversely transcribed into cDNA by using the primer poly (T) and 10 U of AMV reverse transcriptase (Promega). The reaction mixture was size fractionated on 8% polyacrylamide-8 mol/L urea gels. Subsequently, cDNAs ranging in size from 50 to 500 nt were excised and eluted from the gels. The cDNA was tailed at the 3′ end with dGTP by using 7 U of terminal deoxynucleotidyl transferase (TDT, Takara). G-tailed cDNAs were then amplified by PCR with a forward primer *Hin*dⅢ-(T)16 having a *Hin*dⅢ restriction site and a reverse primer *Bam*HⅠ-(C)16 carrying a *Bam*HⅠ restriction site. After digestion with *Hin*dⅢ and *Bam*HⅠ, the amplified fragment was cloned into the corresponding restriction site of pTZ18 and transformed into the strain TG1 of *E. coli*.

1.3 cDNA library screen and DNA sequence analyses

To exclude clones carrying rRNA fragments and known small RNAs, the library was screened by 5′-^{32}P-end labeled oligonucleotides antisense to known small RNAs and random-priming probes derived from 18S and 28S rRNA genes. Duplicate membranes were hybridized with the labeled probes and then exposed to a phosphor screen and analyzed by Typhoon 8600 variable mode imager. Only the clones showing low hybridization signals were selected for DNA sequencing. Sequencing of the clones was performed with an automatic DNA sequencer (ABI Prism 377) by using the BigDye terminator cycle sequencing kit (PE Applied Biosystems) and the P47 and P48 universal primers.

Genomic locations of the sequences from the cDNA library were analyzed by using BLAST program from GenBank (http://www.ncbi.nlm.nih.gov/

?1994-2018 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

BLAST/). The secondary structures of the novel box H/ACA snoRNAs were analyzed by the mfold program (http://www.bioinfo.rpi.edu/applications/mfold/old/rna/form1.cgi).

1.4 RNA analyses

Total cellular RNA was extracted from exponentially growing cultures of *S. pombe* using a hot phenol method as described previously[25]. For Northern analysis, 15 μg total cellular RNA was separated by electrophoresis on 8% acrylamide-8 mol/L urea gels and then electrotransferred onto nylon membrane (Hybond-N+, Amersham), followed by UV light irradiation for 4 min. Subsequently, hybridization with 5′ end labeled probes was performed[19]. The membranes were exposed to a phosphor screen and analyzed by the Typhoon 8600 variable mode imager. Reverse transcription was carried out in 20 μL reaction mixture containing 15 μg of total cellular RNA, 8 ng 5′ end labeled primer and 500 μmol/L dNTPs. After being denatured at 65 ℃ for 5 min and then cooled to 42 ℃, 200 units of M-MLV reverse transcriptase (Promega) were added and the extension was carried out at 42 ℃ for 50 min. The cDNAs were separated on 8% acrylamide-8 mol/L urea gels, and then analyzed by the imager.

1.5 rRNA pseudouridylated site mapping

The mapping of rRNA pseudouridylated sites was performed essentially according to the CMC (*N*-cyclohexyl-*N*′-β-(4-methyl morpholinium) ethylcarbodiimide *p*-tosylate)-primer extension method[26]. 50 μg of total cellular RNA was treated with 100 μL of 0.17 mmol/L CMC at 37 ℃ for 20 min. Subsequently, the CMC-treated RNA was subjected to alkali hydrolysis in the presence of 50 mmol/L Na_2CO_3 (pH 10.4) at 37 ℃ for 4 h. As a control, another 50 μg of total RNA, untreated with CMC, was subjected to alkali hydrolysis under the same conditions. Reverse transcriptions were carried out in the presence of 3—5 μg CMC-treated or -untreated RNA and an appropriate primer labeled at the 5′-end with ^{32}P. The rDNA fragments of *S. pombe* 18S and 25S rRNAs were amplified by PCR with the primer pairs Sp-18F/Sp-18R and Sp-25F/Sp-25-R, respectively, and then cloned into the pMD18-T vector. The plasmid DNA insert was directly sequenced with the same primer as used for rRNA pseudouridylation mapping and run in parallel with the reverse transcription reactions on 8% polyacrylamide denaturing gels. In the reaction of RNA with CMC, although CMC links to U, G and pseudouridine, alkali can remove the adduct from U and G but not from the N_3 of pseudouridine. The CMC group at the N_3 position of pseudouridine efficiently blocks reverse transcription, resulting in a stop band one residue 3′ to the pseudouridine on sequencing gels[26]. So the pseudouridines in RNA molecules can be mapped precisely.

1.6 Oligonucleotides

The oligonucleotides listed below were synthesized and purified by Sangon Co. (Shanghai, China). They were 5′-^{32}P-end labeled according to a standard protocol and directly used as probes for Northern hybridization or submitted to a prior purification by electrophoresis on 8% acrylamide-8 mol/L urea gels before utilization as primer for reverse transcription and rDNA sequencing.

The oligonucleotides used for construction of the cDNA library were as follows: *Hin*d Ⅲ-(T)16 (5′-CCCCAAAAGCTTTTTTTTTTTTTTTT-3′), *Bam*H Ⅰ-(C)16 (5′-GGAATTCGGATCCCCCCCCCCCCCCCC-3′). Those used for Northern blotting and reverse transcription were as follows: Psp15-70 (5′-CAGGCAACGAAAAATCAAACTAA-3′), Psp18-61 (5′-GACACATAGAGGGTTTTCCATAGA-3′). The oligonucleotides used for mapping of ribose pseudouridines were: 18S-251 (5′-ATTCGAAAAGTTATTATGAC-3′), 25S-2371 (5′-CACTAGTTAGATGACGAGG-3′), and 25S-2425 (5′-TCGCTAGATAGTAGATAGG-3′). The oligonucleotides used for PCR amplification of *S. pombe* 18S and 25S rRNA gene fragments were: Sp-18F (5′-TACCTGGTTGATCCTGCCAG-3′), Sp-18R (5′-TGATCCTTCCGCAGGTTCACC-3′), Sp-25F (5′-CCTCAAATCAGGTAGGACTAC-3′), and Sp-25R (5′-GGCTTAATCTCAGCAGATCG-3′).

2 Results

2.1 Identification of two novel box H/ACA snoRNAs from *S. pombe*

A cDNA library was generated with small RNAs isolated from purified nuclei of *S. pombe*. By screening of the cDNA library and DNA sequencing, we obtained 6 sequences exhibiting the conserved motifs ANANNA and ACA in the single-stranded hinge and tail regions, respectively, and sized from 140 to 210 nt (Table 1). These molecules were further analyzed

?1994-2018 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

Table 1. The sequences of six box H/ACA snoRNA candidates determined from the *S. pombe* cDNA library

snoRNA	Sequence[a)]
Sp15-70	TTTGCACTATTAATCAGAAATAGTTCACAGTTTA- TGAGCGTATTTCCTTCATGCAGTGCAT AGAATAA- TGTTTTCATACTCACTTTACCCTCAGGTGCTTCA- AACTAACTTAGTTTGATTTTTCGTTGCCTGCGGG- CCATACATGTATGAT ATA TTTC
Sp18-61	ttattgcaaccttctaataactttacatcttcttatccaaagaagaagggttagaatg- taccgtgttttacggaagttccattcccgtagggttacaatgtgttgct agaata- aaccgtaagtcatagtacatatagaggccaattctgtaaactttTTTAAAG- TCTATGGAAAACCCTCTATGTGTCTCCAATAGA- GAACGG ACA TCTC
Sp12-35	AATTTGAAGGTATTGAGAAACTCAGCTATGTTGA- ATTAATGCAATCGTATACGAAATTACTTCAATGT- GTATATTATTTTGCCATC ACAAAT ACTCCTTT- GCAATTCATGTTAAAAAATAGCACTACGTTCA- TGTGAAAAATTATAACAGAATTATAA ACA TAT
Sp17-2	GGAATGTCTCCCTTGCCAGTACTGCTAGGGTTTTT- CTTTCAAACTATGGAAGCCCATTCAAGCTGC AT- ATTACGATTTTGTTTTTCGCTTTTAGAAAGTGG- TTTAGATGAGATAATAGAAAAATTCTTGATCTCCG- ACA GAT
Sp17-10	ATGAATCTGAAGTTACATTGTTCAGATTCTTTCG- AGAATCTTGCGTGTAACATGGTTGGCGTGTAGAG- GTTG ACAAG AGTACAGACAGGGCGGCGATTGT- CCTTGCAGTTTCGACTGCAGGTTGAGAGTTGGCG- TCTTGCCATGAAAATTGACTAAAGGATGCTGCGT- AGCCTAGCCA ACA GATT
Sp17-17	ACACAACTAAAATCAAATGATGAACGAGTTTTTAA- ATTGAGTTTGAGCGCCCTAATCAAGTGCTCTAACT- CGCTCCATGGAGTCTTTGGATTTTATGGATGTGGG- GCATTG AGAAGA CGACCTCGCTCTTTCACCACC- TTTTTGGGGTCGAATGTTGTTCTG AGA TTT

a) All the sequences determined from cDNA are in uppercase letters. Sp18-61 has partial cDNA sequence, but its intact sequence has been derived from genomic sequence (in lowercase letters) according to our Northern blot and reverse transcription analyses. Box H and ACA motifs in the sequences are shaded.

by Northern blot hybridization with antisense probes that were synthesized according to the cDNA sequences. Sp15-70 and Sp18-61 were positively detected (Fig. 1(a)), suggesting that they are snoRNAs expressed stably in cells of *S. pombe*. The size of Sp15-70 revealed by Northern blot is 158 nt, agreeing well with that of the cDNA sequence, while the size of Sp18-61 revealed is larger than its cDNA size. Reverse transcription was employed to determine the 5′ terminal of Sp18-61 (Fig. 1(b)) and its full sequence was obtained from genomic sequence database of *S. pombe* according to the results of Northern blot and reverse transcription (Table 1). As compared with known snoRNAs, these two snoRNAs were identified for the first time experimentally from *S. pombe*. The remaining four snoRNA candidates have no hybridization signals in the Northern analysis. It is not clear if the transcriptions of these candidates were too low to be detected by the Northern blot or the sequences were not accumulated stably in the cells.

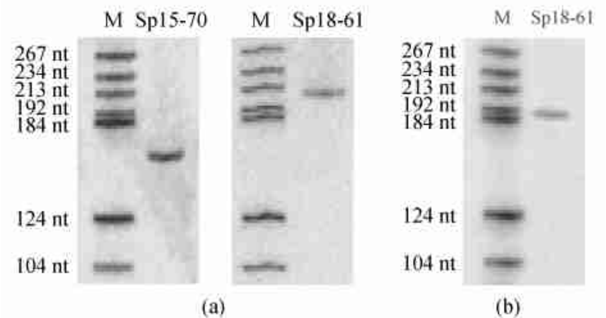

Fig. 1. Positive detection of the box H/ACA snoRNAs. (a) Northern hybridization; (b) reverse transcription. Lane M, molecular weight markers (pBR322 digested with *Hae* III and 5′-end labeled with [γ-^{32}P] ATP).

2.2 Structure and function of the two novel snoRNAs

Analysis of the secondary structures of Sp15-70 and Sp18-61 showed that both of them share the canonical structural features of box H/ACA snoRNA family, i.e. "hairpin-hinge-hairpin-tail" secondary structure and two conserved sequence motifs, the H (ANANNA) and ACA boxes, located in the single-stranded hinge and tail regions, respectively (Fig. 2(a)). All of the "pseudouridylation pockets" (first loop) on every hairpin-like domain of these two snoRNAs contain two short (8—12 nt) single-stranded sequences. Comparison of these single sequences with any portion of the *S. pombe* 18S and 25S rRNAs revealed that every pseudouridine pocket possesses a region complementary to rRNAs and the bipartite antisense sequences are 3—9 nt in length (Fig. 2(b)). According to the relationship between the structure and function of box H/ACA snoRNA, Sp15-70 was predicted to direct the pseudouridylation of U2401 and U2298 in 25S rRNA, and Sp18-61 was predicted to direct the pseudouridylation of U208 in 18S rRNA and U2341 in 25S rRNA. Notably, the spacing between the selected pseudouridine and the H or ACA box of the snoRNAs was remarkably conserved in about 14—16 nt, which is a critical determinant of pseudouridylation site.

?1994-2018 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

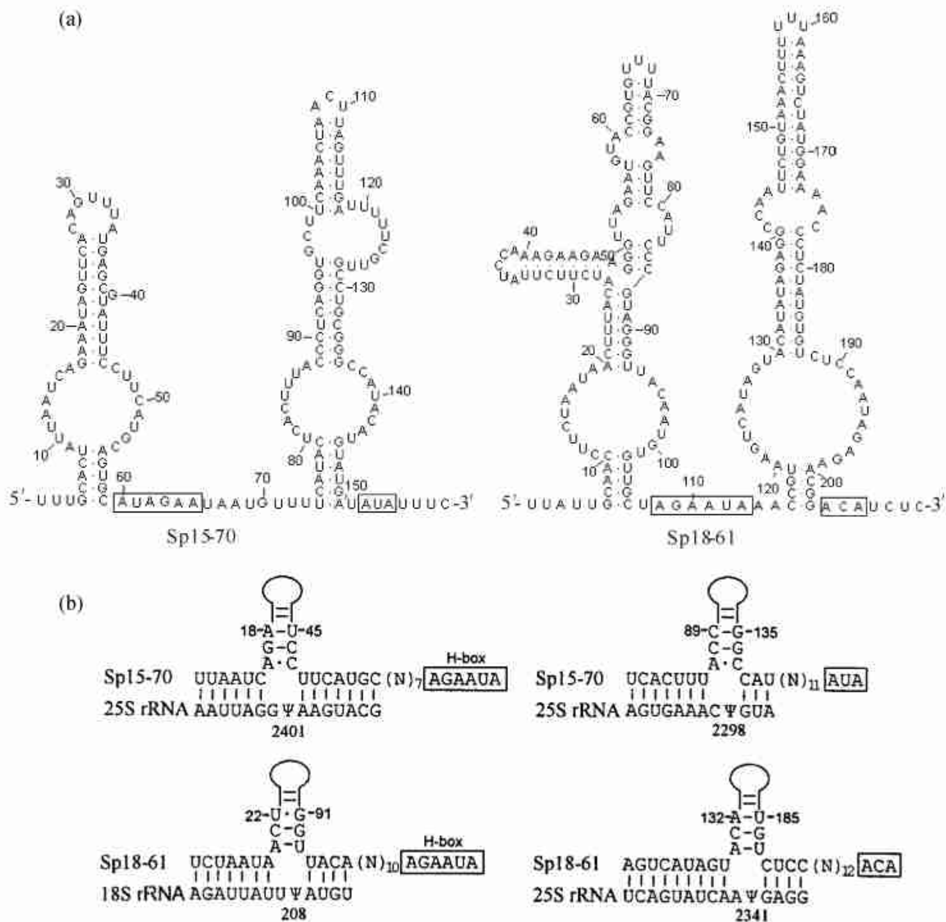

Fig. 2. Secondary structural and functional analyses of snoRNAs. (a) Secondary structures of snoRNAs. Box H and ACA motifs are boxed. (b) Potential base-pairing interactions between snoRNAs and rRNAs. The snoRNA sequences in a 5′-3′ orientation are shown in the upper strands, and hairpin domains are schematized by a solid line. rRNA sequences in a 3′-5′ orientation are shown in the lower strands, in which predicted pseudouridylation sites are denoted by "Ψ".

2.3 Determination of the rRNA pseudouridylation sites by the novel snoRNAs

The four pseudouridylation sites predicted by Sp15-70 and Sp18-61 were verified by CMC-primer extension analysis. As shown in Fig. 3, when reverse transcription analyses were carried out using the CMC-treated total cellular RNA as templates, the pauses related to pseudouridylated nucleotide on 18S rRNA at U208 and 25S rRNA at U2298, U2341 and U2401 were precisely determined, suggesting that they are all indeed pseudouridylated nucleotides. In addition, in the assay to determine the Ψ2341 in 25S rRNA, three closely clustered pseudouridylation sites, Ψ2345, Ψ2351 and Ψ2353, were also mapped precisely (Fig. 3). Comparison of rRNA pseudouridylation sites among yeast, plant and animal shows these clustered pseudouridylation sites are most highly conserved in evolution[27, 28], while their cognate snoRNAs in *S. pombe* have not been found.

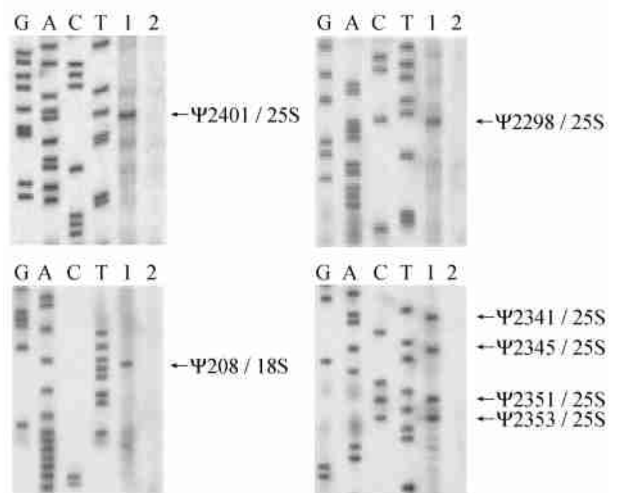

Fig. 3. Determination of the rRNA pseudouridylation sites by CMC-primer extension method. Lane 1, reverse transcription with CMC-treated total RNA; lane 2, reverse transcription reaction with CMC-untreated total RNA; lanes A, C, G and T, the rDNA sequence ladder; positions of pseudouridine residues are indicated by arrows.

?1994-2018 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

2.4 Genomic organization of the novel snoRNA genes

The genomic locations of these two box H/ACA snoRNAs were confirmed by searching genomic sequence of *S. pombe* with BLAST program. Both Sp17-70 and Sp18-61 were encoded by a single copy gene and located in the intergenic regions, sized 1247 and 720 nt, between the CDS of two protein-coding genes in *S. pombe* chromosomes I and III, respectively (Fig. 4). Putative TATA-like elements can be found upstream of the 5′ end of these snoRNA genes, suggesting that they could be transcribed independently from their own promoters.

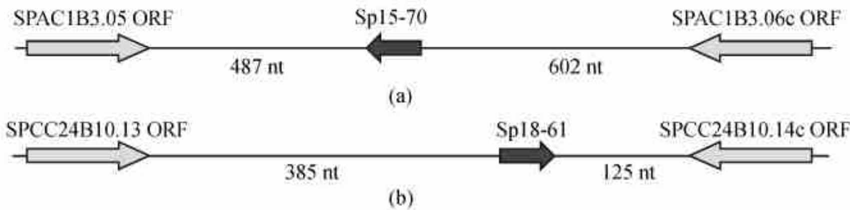

Fig. 4. Locations of the two novel snoRNA genes. (a) and (b) represent the locations of Sp15-70 and Sp18-61 genes respectively. SnoRNA genes are shown by black arrows and protein-coding genes by gray arrows (not drawn to scale). Numbers under the lines indicate the sizes of the intergenic regions between snoRNA genes and protein-coding genes.

3 Discussion

To date, two main methods have been developed for finding and identifying novel non-coding RNA genes. Computational RNomics is largely based on the search of conserved sequences and structure elements of novel non-coding RNAs in gemonic database[14, 29, 30]; whereas experimental RNomics depends on the construction and sequencing of specialized cDNA libraries enriched with small RNA[11, 13—16]. The presence of hallmark boxes and long complementarity to the site of ribose methylation in box C/D snoRNAs allowed many of them to be identified by computational researches of completely sequenced genomes. A large number of novel box C/D snoRNAs have been identified so far through computational RNomics from *S. cerivisera*[31], *A. thaliana*[32] and *O. sativa*[33]. In contrast, no effective computational genomic search of box H/ACA snoRNAs has been reported yet, mainly due to shorter box motif sequences and shorter bipartite guide sequences. Recently, a new algorithm taking the folding parameters in RNA molecule into account has been developed to screen yeast genome for novel box H/ACA snoRNAs[34]. However, the computational method is insufficient for identifying novel box H/ACA snoRNAs with diverse sequence motifs and structure, especially for the orphan snoRNAs. In fact, to date box H/ACA snoRNAs have been exclusively discovered by experimental approaches. In this study, we obtained six sequences with the property of conserved motifs of box H/ACA snoRNAs through generating a specialized cDNA library based on small nuclear RNA of *S. pombe* and cDNA sequencing, and two of these, Sp15-70 and Sp18-61, were confirmed to be novel snoRNAs and expressed stably in cells. The four pseudouridylation sites directed by Sp15-70 and Sp18-61 are phylogenetically conserved. However, it is unlikely to identify the homologues of Sp15-70 and Sp18-61 in other organisms by BLASTN search in GenBank database. This once more emphasizes the importance of experimental RNomics for finding novel box H/ACA snoRNAs.

Many box H/ACA snoRNAs have two functional elements that guide respectively two pseudouridylation sites phylogenetically conserved. Comparison of the bipartite sequence complementarities to rRNA and the corresponding pseudouridylation target sites among different organisms reveals some homologous snoRNAs in distant organisms are partially homologous. As shown in Fig. 5(a), *S. pombe* Sp15-70 shares the guide sequence in pseudouridylation pocket 1 with pocket 1 of *Arabidopsis* snoR83 and pocket 2 of *Drosophila melanogaster* Dm-3, and shares the other guide sequence with *Arabidopsis* snoR99. While guide sequence in pocket 1 of Dm-3 is equivalent to those in pocket 2 of *S. pombe* Sp18-61, *Arabidopsis* snoR92 and MBI-1 and so on. This complicated pattern of snoRNA homologues suggests that extensive recombinations (two guide sequence correspond to different snoRNAs, respectively, in different organism) and transposition (equivalent guide sequences in counterparts located in different hairpin domains) between different snoRNAs have occurred during the evolution from their primitive progenitors. Thus this large snoRNA gene family is an ideal model for investigating the mechanisms of gene evolution.

?1994-2018 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

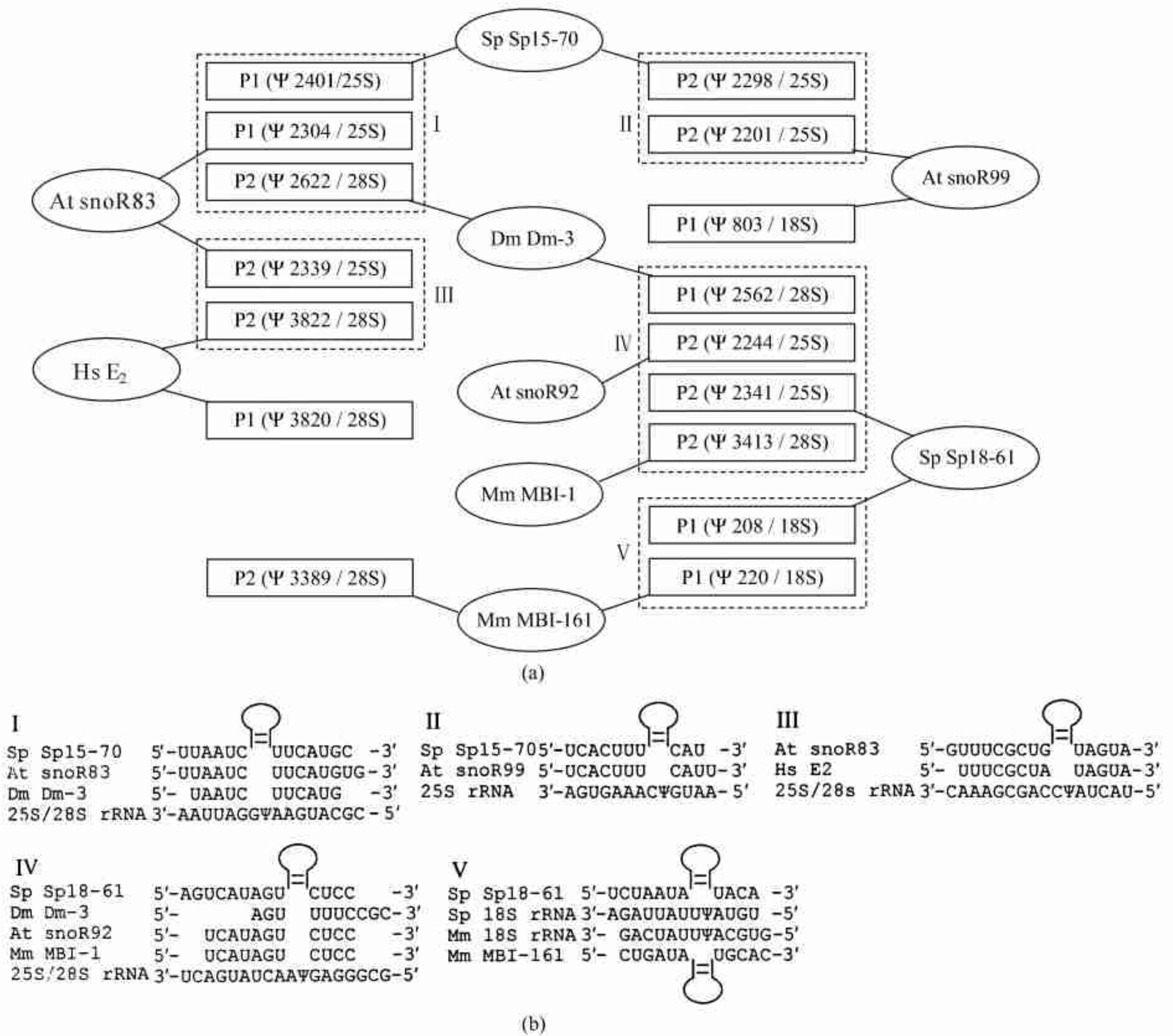

Fig. 5. Analysis of box H/ACA snoRNA homologues among distant organisms. (a) Schematic diagrams of correlation of some box H/ACA snoRNAs among diverse organisms. SnoRNAs are represented by ovals. Pseudouridylation pockets of each box H/ACA snoRNA are boxed by rectangles. P1 and P2 represent pseudouridylation pockets 1 and 2, respectively. The pockets in broken-line boxes possess homologous guide sequences that direct a phylogenetically conserved site of rRNA pseudouridylation in different organisms. (b) I-V show the sequences of each snoRNA—rRNA duplexes, respectively.

Genomic organization and expression of snoRNA genes exhibit a striking diversity in different organisms[35]. Almost all guide snoRNAs in vertebrates are nested within introns of host genes[1, 36] and at most one snoRNA is encoded in any particular intron, which is processed from the pre-mRNA by exonucleolytic digestion of the debranched lariat[1]. In contrast, most plant snoRNAs are transcribed from gene clusters as polycistronic transcripts which are processed by endo- and exonucleases to release mature snoRNAs[33, 35, 37]. In *S. cerevisiae*, a majority of snoRNA genes are transcribed independently and only a few of them have been found in gene clusters or encoded in the intron of protein genes[31]. Both the two novel box H/ACA snoRNAs identified in this study are located in intergenic regions of protein-coding genes, and express as monocistronic RNA transcripts. The other three box H/ACA snoRNAs identified so far in *S. pombe* are also encoded by genes expressed independently[22], suggesting that monocistron is likely to be a predominant gemonic organization of box H/ACA snoRNA genes in *S. pombe*.

Acknowledgement We are grateful to Chen Xiaohong for her technical assistance.

References

1 Bachellerie J. P., Cavaille J. and Huttenhofer A. The expanding snoRNA world. Biochimie, 2002, 84(8): 775—790.

?1994-2018 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

2 Balakin A. G., Smith L. and Fournier M. J. The RNA world of the nucleolus: two major families of small RNAs defined by different box elements with related functions. Cell, 1996, 86(5): 823—834.

3 Kiss-Laszlo Z., Henry Y., Bachellerie J. P. et al. Site-specific ribose methylation of preribosomal RNA: a novel function for small nucleolar RNAs. Cell, 1996, 85(7): 1077—1088.

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

5 Ganot P., Bortolin M. L. and Kiss T. Site-specific pseudouridine formation in preribosomal RNA is guided by small nucleolar RNAs. Cell, 1997, 89(5): 799—809.

6 Ni J., Tien A. L. and Fournier M. J. Small nuclear RNAs direct site-specific synthesis of pseudouridine in ribosomal RNA. Cell, 1997, 89(4): 565—573.

7 Jady B. E. and Kiss T. A small nucleolar guide RNA functions both in 2′-O-ribose methylation and pseudouridylation of the U5 spliceosomal RNA. EMBO J., 2001, 20(3): 541—551.

8 Darzacq X., Jady B. E., Verheggen C. et al. Cajal body-specific small nuclear RNAs: a novel class of 2′-O-methylation and pseudouridylation guide RNAs. EMBO J., 2002, 21(11): 2746—2756.

9 Gaspin C., Cavaille J., Erauso G. et al. Archaeal homologs of eukaryotic methylation guide small nucleolar RNAs: lessons from the Pyrococcus genomes. J. Mol. Biol., 2000, 297(4): 895—906.

10 Omer A. D., Lowe T. M., Russell A. G. et al. Homologs of small nucleolar RNAs in Archaea. Science, 2000, 288(5465): 517—522.

11 Tang T. H., Bachellerie J. P., Rozhdestvensky T. et al. Identification of 86 candidates for small non-messenger RNAs from the archaeon Archaeoglobus fulgidus. Proc. Natl. Acad. Sci. USA, 2002, 99(11): 7536—7541.

12 Clouet d'Orval B., Bortolin M. L., Gaspin C. et al. Box C/D RNA guides for the ribose methylation of archaeal tRNAs. The tRNATrp intron guides the formation of two ribose-methylated nucleosides in the mature tRNATrp. Nucleic Acids Res., 2001, 29(22): 4518—4529.

13 Huttenhofer A., Kiefmann M., Meier-Ewert S. et al. RNomics: an experimental approach that identifies 201 candidates for novel, small, non-messenger RNAs in mouse. EMBO J., 2001, 20(11), 2943—2953.

14 Huttenhofer A., Brosius J. and Bachellerie J. P. RNomics: identification and function of small, non-messenger RNAs. Curr. Opin. Chem. Biol., 2002, 6(6): 835—843.

15 Marker C., Zemann A., Terhorst T. et al. Experimental RNomics: identification of 140 candidates for small non-messenger RNAs in the plant Arabidopsis thaliana. Curr. Biol., 2002, 12(23): 2002—2013.

16 Yuan G., Klambt C., Bachellerie J. P. et al. RNomics in Drosophila melanogaster: identification of 66 candidates for novel non-messenger RNAs. Nucleic Acids Res., 2003, 31(10): 2495—2507.

17 Cavaille J., Buiting K., Kiefmann M. et al. Identification of brain-specific and imprinted small nucleolar RNA genes exhibiting an unusual genomic organization. Proc. Natl. Acad. Sci. USA, 2000, 97(26): 14311—14316.

18 Filipowicz W. Imprinted expression of small nucleolar RNAs in brain: time for RNomics. Proc. Natl. Acad. Sci. USA, 2000, 97(26): 14035—14037.

19 Zhou H., Chen Y. Q., Du Y. P. et al. The Schizosaccharomyces pombe mgU6-47 gene is required for 2′-O-methylation of U6 snRNA at A41. Nucleic Acids Res., 2002, 30(4): 894—902.

20 Massenet S., Mougin A. and Branlant C. Posttranscriptional modifications in the U small nuclear RNAs. In: Modification and Editing of RNA. Washington DC: ASM Press, 1998, 201—227.

21 Badis G., Fromont-Racine M. and Jacquier A. A snoRNA that guides the two most conserved pseudouridine modifications within rRNA confers a growth advantage in yeast. RNA. 2003, 9(7): 771—779.

22 Qu G. S., Zhou H., Yu C. H. et al. Characterization of three novel snoRNAs from Schizosaccharomyces pombe implies the structural and functional diversity of box H/ACA snoRNA family. Chinese Science Bulletin, 2003, 48(21): 2320—2327.

23 Dove J. E., Brockenbrough J. S. and Aris J. P. Isolation of nuclei and nucleoli from the yeast Saccharomyces cerevisiae. Methods Cell Biol., 1998, 53: 33—46.

24 Chomczynski P. and Sacchi N. Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. Anal. Biochem., 1987, 162(1): 156—159.

25 Schmitt M. E., Brown T. A. and Trumpower B. L. A rapid and simple method for preparation of RNA from Saccharomyces cerevisiae. Nucleic Acids Res., 1990, 18(10): 3091—3092.

26 Bakin A. and Ofengand J. Four newly located pseudouridylate residues in Escherichia coli 23S ribosomal RNA are all at the peptidyltransferase center: analysis by the application of a new sequencing technique. Biochemistry, 1993, 32(37): 9754—9762.

27 Ofengand J. and Bakin A. Mapping to nucleotide resolution of pseudouridine residues in large subunit ribosomal RNAs from representative eukaryotes, prokaryotes, archaebacteria, mitochondria and chloroplasts. J. Mol. Biol., 1997, 266(2): 246—268.

28 Ofengand J. Ribosomal RNA pseudouridines and pseudouridine synthases. FEBS Lett., 2002, 514(1): 17—25.

29 Rivas E. and Eddy S. R. Noncoding RNA gene detection using comparative sequence analysis. BMC Bioinformatics, 2001, 2(1): 8.

30 Rivas E., Klein R. J., Jones T. A. et al. Computational identification of noncoding RNAs in E. coli by comparative genomics. Curr. Biol., 2001, 11(17): 1369—1373.

31 Lowe T. M. and Eddy S. R. A computational screen for methylation guide snoRNAs in yeast. Science, 1999, 283(5405): 1168—1171.

32 Barneche F., Gaspin C., Guyot R. et al. Identification of 66 box C/D snoRNAs in Arabidopsis thaliana: extensive gene duplications generated multiple isoforms predicting new ribosomal RNA 2′-O-methylation sites. J. Mol. Biol., 2001, 311(1): 57—73.

33 Chen C. L., Liang D., Zhou H. et al. The high diversity of snoRNAs in plants: identification and comparative study of 120 snoRNA genes from Oryza sativa. Nucleic Acids Res., 2003, 31(10): 2601—2613.

34 Edvardsson S., Gardner P. P., Poole A. M. et al. A search for H/ACA snoRNAs in yeast using MFE secondary structure prediction. Bioinformatics, 2003, 19(7): 865—873.

35 Brown J. W., Echeverria M., Qu L. H. et al. Plant snoRNAs: functional evolution and new modes of gene expression. Trends Plant Sci., 2003, 8(1): 42—49.

36 Maxwell E. S. and Fournier M. J. The small nucleolar RNAs. Annu. Rev. Biochem., 1995, 64: 897—934.

37 Brown J. W., Clark G. P., Leader D. J. et al. Multiple snoRNA gene clusters from Arabidopsis. RNA. 2001, 7(12): 1817—1832.

?1994-2018 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net