Volume 12 Number 13 1984 Nucleic Acids Research

The complete nucleotide sequence of a rice 17S rRNA gene

Fumio Takaiwa, Kiyoharu Oono* and Masahiro Sugiura[1]

National Institute of Agrobiological Resources, Tsukuba Science City, Yatabe Ibaraki 305, and [1]Department of Biology, Faculty of Science, Nagoya University, Chikusa, Nagoya 464, Japan

Received 17 April 1984; Revised and Accepted 25 May 1984

ABSTRACT

The complete nucleotide sequence of a rice nuclear 17S rRNA gene (rDNA) has been determined. The rice rDNA is 1812 bp long and its G + C content is 51.3%. This nucleotide sequence shows 79%, 80% and 80% homology to those of yeast, Xenopus laevis and rat 18S rDNAs, respectively. Divergency of nucleotide sequences is largely attributed to five blocks of highly variable regions, where eukaryotic specific sequences can be observed.

INTRODUCTION

Eukaryotic ribosomes contain four species of rRNAs. The 25-28S, 5.8S and 5S rRNAs are associated with the 60S ribosomal subunit and the 17-18S rRNA is associated with the 40S subunit. Determination of primary structures of eukaryotic 17-18S rRNAs is essential for understanding the subunit association, ribosomal assembly and interaction with mRNA, 5S rRNA and tRNA. Furthermore, comparative sequence analysis of rDNAs from a variety of organisms is regarded as one of the effective methods for deducing a secondary structure as well as for providing some insight into evolutional trends. However, the complete nucleotide sequences so far determined are limited to yeast (1), Xenopus laevis (2), rat (3) and Dictyostelium discoideum (4).

We have recently cloned a major class (7.7 kbp, formerly designated as 5.0 Md) of rice rDNA repeating units and determined the entire nucleotide sequence of a spacer region between 17S and 25S rDNAs (5). In the present study, we present the complete sequence of a rice 17S rDNA and compared it with corresponding sequences of yeast, X. laevis and rat. This is -to our knowledge- the first complete nucleotide sequence of a 17S rDNA from

higher plants. Its length and sequence is highly conserved among eukaryotic organisms.

MATERIALS AND METHODS

Recombinant plasmid pRR217 contains whole rRNA gene sequences in its inserted 7.7 kbp EcoRI fragment (5). The plasmid DNA was cleaved with BamHI + HincII, BamHI + BstEII or SacII. The digested DNA fragments were fractionated in 4.5% polyacrylamide gels. The 2.4 kbp BamHI - BstEII fragment, 0.78 kbp BamHI - HincII fragment, 0.29 kbp and 1 kbp HincII fragments and 0.8 kbp SacII fragment that encode 17S rDNA were recovered from the excised gel pieces and sequenced by the chemical method as described (5).

RESULTS AND DISCUSSION

Recombinant plasmid pRR217 contains a major class (7.7 kbp) of rice rDNA repeating unit (5). The DNA fragments encoding the 17S rDNA were isolated from pRR217 DNA and used for sequencing. A detailed physical map and the sequencing strategy are shown in Fig. 1. Most of the regions were confirmed by sequencing in opposite directions and more than twice in the same directions. Fig. 2 shows the complete nucleotide sequence (RNA-like strand) of the 17S rDNA. The 5' end of rice 17S rDNA was deduced by comparison with 18S rRNA sequences from other organisms. The 3' end was assigned based on the 3' terminal sequences of wheat 17S rRNA (6). The T residue at the 5' end is found to be identical among eukaryotic organisms sequenced so far, while the G residue at the 3' end seems to be unique to higher plants. The coding region of rice 17S rDNA is estimated to be 1812 bp. The rice 17S rDNA is 23 bp longer than the yeast 18S rDNA (1789 bp) and 57 bp shorter than the rat 18S rDNA (1869 bp). This is mainly due to two insertions (5 bp and 10 bp) at positions 651-655 and 995-1004 and two long deletions (11 bp and 26 bp) at positions 191/192 and 225/226 (Fig. 2). It has recently been presented that *D. discoideum* 18S rDNA is 1872 bp long (4) and is the longest of all species reported so far. This is attributed to an unique insertion (79 bp) at positions 1366/1367. Eukaryotic 17-18S rRNAs are similar in size to each other, which

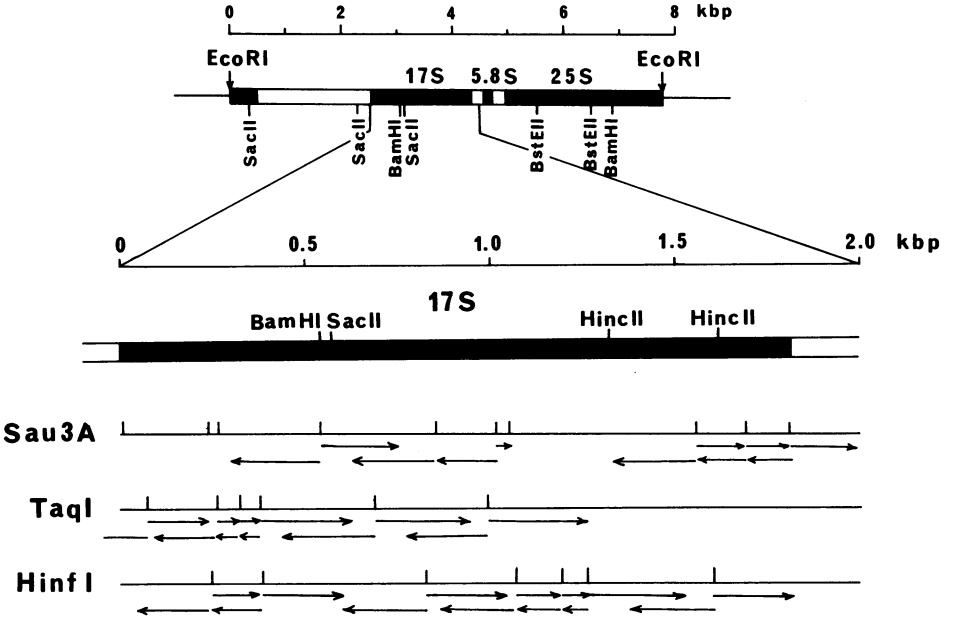
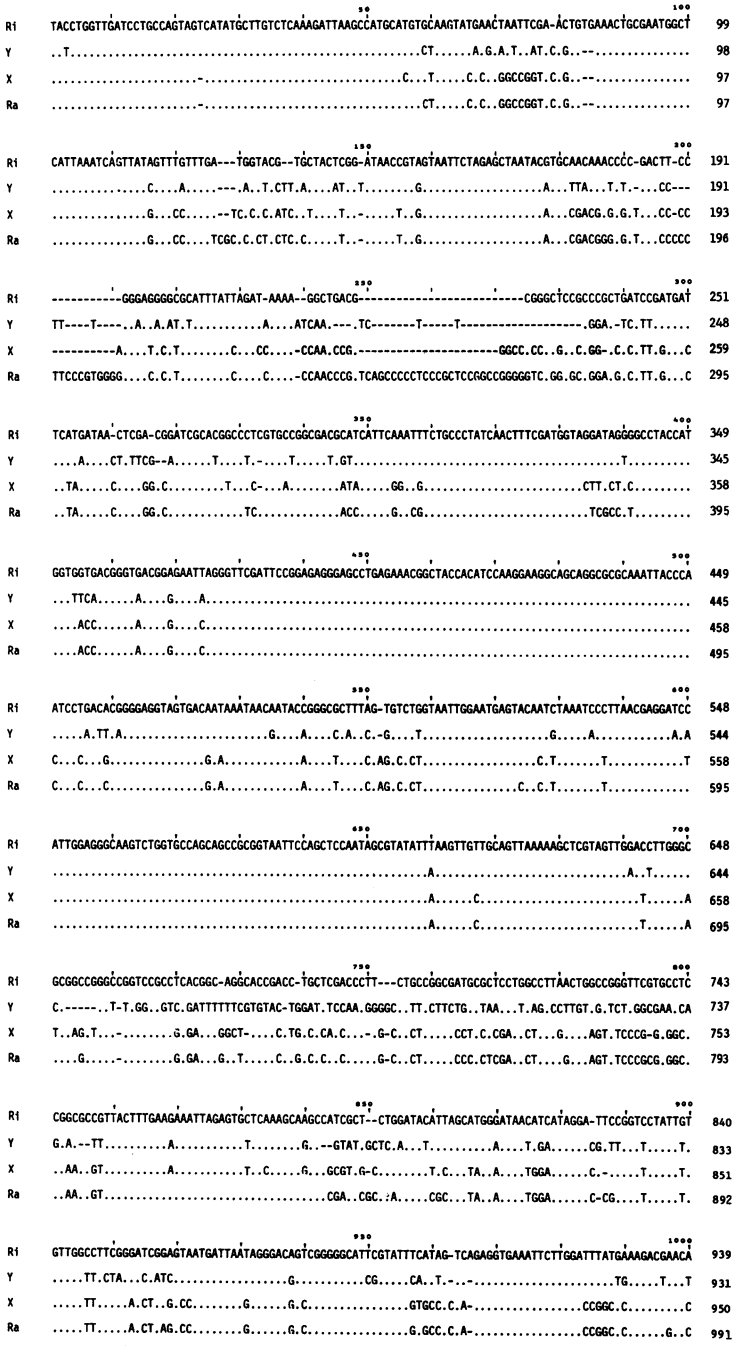


Fig. 1. Physical map of the cloned 7.7 kbp rDNA from rice and the sequence strategy of 17S rDNA. Thick lines represent the inserted rice DNA fragments and single lines show parts of pBR325. Coding regions are shown in black. Horizontal arrows indicate the direction and extent of the DNA regions sequenced.

is in remarkable constract to what is observed in 25-28S rRNAs. The 17-18S rRNA range in size from 1789 bp to 1872 bp, whereas the 25-28S rRNAs from 3380 bp (rice) to 4718 bp (rat) (7). Such a conservation in the size of 17-18S rRNA may be related to important function in protein synthesis.

When the nucleotide sequence of rice 17S rDNA is aligned with those of yeast, X. laevis and rat 18S rDNA so as to maximize sequence homology, numbers of identical residues are 1434, 1449 and 1446, respecively (Fig. 2). We have recently presented that the rice 5.8S rDNA sequence shows 80%, 71% homologies to those of yeast and X. laevis (5). Therefore, the degree of homology found in 17-18S rDNA is slightly higher than that of 5.8S rDNAs.

Conserved regions are observed all over the 17-18S rDNAs (Fig. 2). Consecutive identical sequences longer than 40 bp are detected in positions 373-449, 549-608 and 1766-1811 (3' terminal



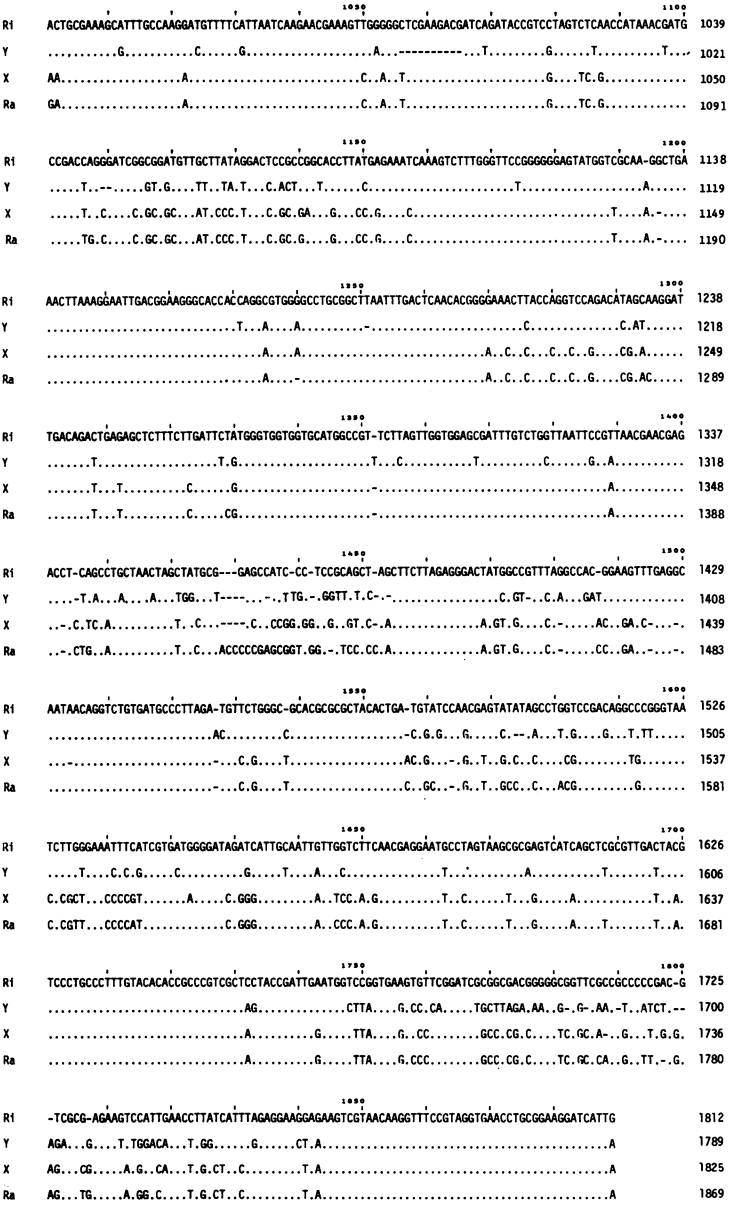


Fig. 2. Nucleotide sequence of the rice 17S rDNA (Ri). The non-coding (RNA-like) strand is shown. Sequences of yeast (Y), X. laevis (X) and rat (Ra) are presented for comparison. Dots indicate nucleotides identical with the rice sequence. Bars present gaps introduced to maximize homology. Numbers in the right-hand margin refer to actual nucleotide residues from 5' end.

Table 1 G + C content value

	Complete	Highly variable regions					Rest of regions
		I	II	III	IV	V	
Rice	51.3% 1812bp	55.4% (175-268)	76.0% (648-751)	57.8% (1045-1089)	60.0% (1340-1384)	61.3% (1674-1766)	48.1%
Yeast	45.0% 1789bp	29.2% (177-265)	55.4% (644-743)	44.2% (1027-1069)	39.5% (1321-1365)	48.4% (1653-1743)	45.3%
X. laevis	53.8% 1825bp	65.1% (176-278)	73.3% (658-761)	63.1% (1059-1100)	63.0% (1350-1396)	61.1% (1685-1779)	50.5%
Rat	55.6% 1869bp	69.3% (178-314)	76.9% (695-800)	68.9% (1097-1141)	64.0% (1391-1440)	62.1% (1729-1823)	51.6%

Numbers in parentheses indicate positions of variable regions in rice, yeast, X. laevis and rat.

regions), suggesting that these regions are important for ribosome function. The 3' terminal regions of 17-18S rRNAs are well characterized among eukaryotes as well as prokaryotes (6,8). It has been proposed that the 3' terminal regions are involved in interaction with the 5' terminal regions of mRNAs or the 3' terminal regions of 5S rRNAs (8). The function of two other conserved regions is as yet unknown.

Highly variable regions (longer than 40 bp) which show less than 45% homology are found in positions 175-268, 648-751, 1045-1089, 1340-1384 and 1674-1766. These sequences found in rice, X. laevis and rat are 7-28% more rich in G + C content than the rest of regions, whereas the corresponding regions of yeast are similar or less except for region II (Table 1). The G + C content of 17-18S rRNA progressively increases from 45% (yeast) to 56% (rat) during evolution. As conserved and semiconserved regions of 17-18S rRNAs are similar in G + C content among organisms, overall increases of the G + C content in the rRNAs are mainly due to increase in the G + C content found in highly variable regions. It is known that increase in size of 17-18S rRNA during evolution from prokaryotes to eukaryotes is the result of the four insertions (9,10). These inserts can be detected at the highly variable regions. Therefore, it is

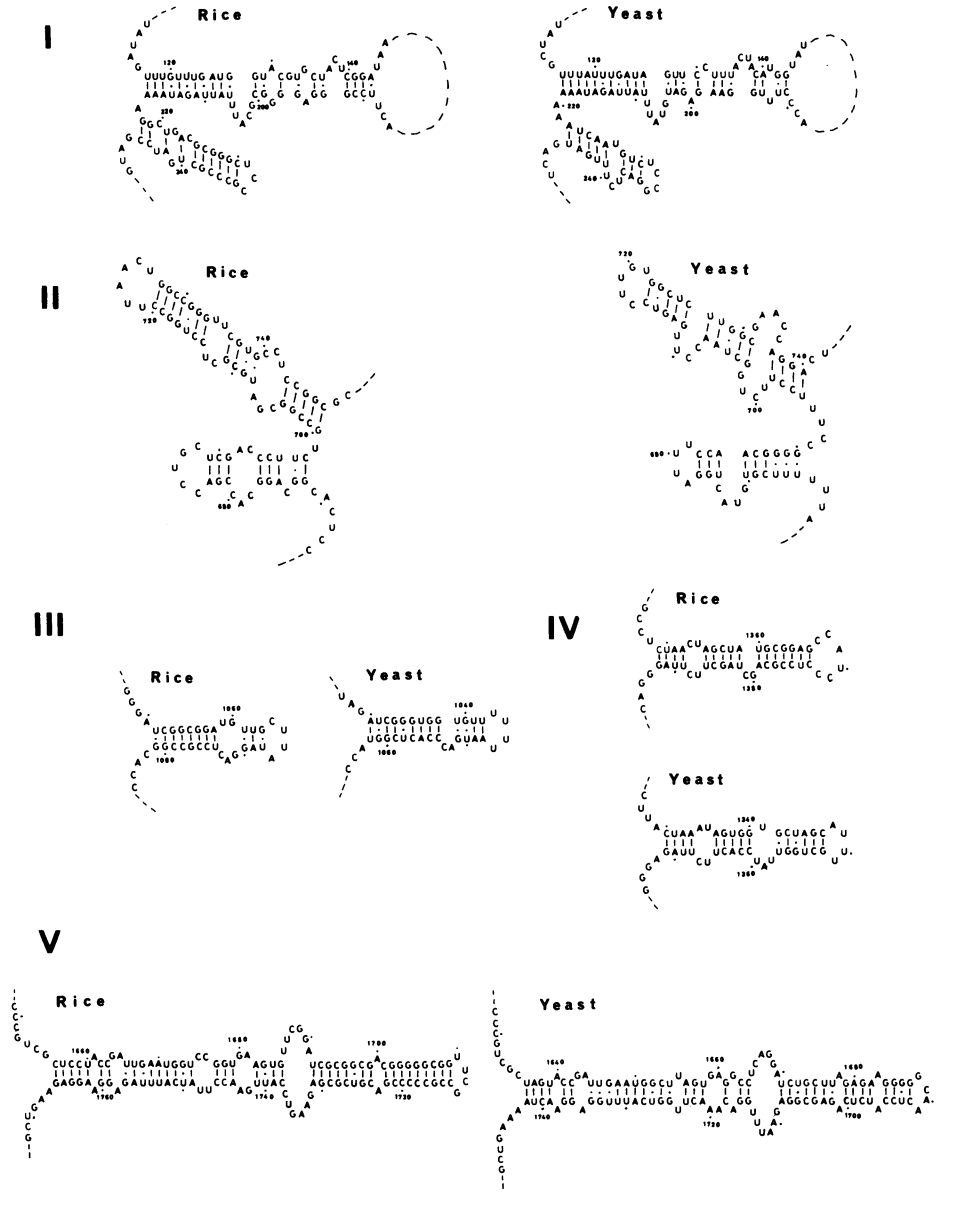


Fig. 3. Possible secondary structures for five highly variable regions of rice and yeast 17-18S rRNAs deduced from the DNA sequences.

suggested that highly variable regions were originated by insertions of eukaryotic specific sequences into non-essential regions during evolution.

Several secondary structure models have been proposed for eukaryotic 17-18S rRNAs (9-12). When the secondary structure of rice 17S rRNA was constructed according to the model of Atmadja et al. (12), it can be folded by base pairing in a similar way as shown in yeast and X. laevis. Most of the herical regions are maintained by many compensating base substitutions. It is noteworthy that highly variable regions also show roughly similar secondary structures to yeast and X. laevis (Fig. 3). Therefore, higher plant 17-18S rRNA is similar in primary and secondary structures to other eukaryotic rRNAs.

ACKNOWLEDGEMENTS

We thank Mrs. A. Yamada for preparation of this manuscript and Dr. S. Suzuki and Mr. M. Oshima for computer programming.

* To whom correspondence should be addressed.

REFERENCES

1. Rubtsov, P.M., Musakhanov, M.M., Zakhayev, V.M., Krayev, A.S., Skryabin, K.G. and Bayev, A.A. (1980) Nucleic Acids Res. 8, 5779-5794.
2. Salim, M. and Maden, B.E.H. (1981) Nature 291, 205-208.
3. Torczynski, R., Bollon, A. and Fuke, M. (1983) Nucleic Acids Res. 11, 4879-4890.
4. McCarroll, R., Olsen, G.J., Stahl, Y.D., Woese, C.R. and Sogin, M.L. (1983) Biochemistry 22, 5858-5868.
5. Takaiwa, F., Oono, K. and Sugiura, M. (1984) in preparation
6. Hagenbuchele, O., Santer, M. and Steiz, J.A. and Mans, R.J. Cell 13, 551-563.
7. Chan, Y-L., Olvera, J. and Wool, I.G. (1983) Nucleic Acids Res. 11, 7819-7831.
8. Azad, A.A. and Deacon, N.J. (1980) Nucleic Acids Res. 8, 4365-4376.
9. Zwieb, C., Glotz, C. and Brimacombe, R. (1981) Nucleic Acids Res. 9, 3621-3640.
10. Olsen, G.J., McCarroll, R. and Sogin, M.L. (1983) Nucleic Acids Res. 11, 8037-8049.
11. Chan, Y-L., Gutell, R., Noller, H.F. and Wool, I.G. (1984) J. Biol. Chem. 259, 224-230.
12. Atmadja, J., Brimacombe, R. and Maden, B.E.H. (1984) Nucleic Acids Res. 12, 2649-2667.