# Sequence analysis of *Citrus limon* DNA coding for 26 S rRNA

## Evidence of heterogeneity in the 3′-region

V.O. Kolosha, V.M. Kryukov and I. Fodor

*Institute of Biochemistry and Physiology of Microorganisms, Pushchino, Moscow Region 142292, USSR*

Received 9 December 1985

Abstract not received

*rRNA* *DNA sequence* (Citrus limon) *Heterogeneity*

## 1. INTRODUCTION

The structural organization of rDNA cistrons is evolutionarily conserved in various higher eukaryotes. Comparative analysis of the rDNA structure provides some insights into evolutionary trends. Higher plants and several lower eukaryotes are characterized by the small size of the cytoplasmic large rRNA subunit. Furthermore, considerable heterogeneity in nucleotide sequence and in length has been observed in some plants which is linked mainly with the large nontranscribed spacer of rDNA [1-3].

We have sequenced and characterized the DNA coding for 26 S rRNA from *Citrus limon*, a representative of the subclass Dicotyledoneae. Having compared the structure of the 500 bp region isolated from two independent plasmid clones pCIRS and pCIRW, we have found heterogeneity in the coding sequences of plant rDNA.

## 2. MATERIALS AND METHODS

Plasmid DNA was isolated by alkaline extraction [4] followed by CsCl equilibrium ultracentrifugation. The gel filtration step was included in the isolation of plasmid DNA when dephosphorylation by calf intestinal phosphatase (Boehringer Mannheim) and (5′-$^{32}$P)-end-labelling of DNA fragments were needed. 5′- or 3′-labelled ends were separated by secondary restriction digestion. DNA sequencing was carried out using the method of Maxam and Gilbert [5] with slight modifications [6].

The homology was calculated aftr the alignment procedure [7]. Insertions or deletions longer than 16 bp were neglected, those shorter than 16 bp being considered as mismatches.

## 3. RESULTS AND DISCUSSION

The restriction map of *C. limon* rDNA as well as the cloning of rDNA fragments have been described in [8]. The recombinant plasmids pCIB33, pCIRS and pCIRW contained the 3′-portions of 26 S rRNA genes [9]. To facilitate sequence analysis and obtain a detailed restriction map of this rDNA region, the 1.4 kb *Bam*HI fragment was isolated from PCIB33 and its *Sau*3A subfragments cloned in pUR222 [10]. *Pst-Eco*RI fragments were isolated from pCIRS and pCIRW.

Our sequencing strategy is shown in fig.1. We determined the 1437 bp sequence of pCIB33 comprising the 3′-end of 26 S rRNA (fig.2); the 38 bp coding sequence located to the right of the *Bam*HI (*Sau*3A) site (fig.1) was sequenced using the

*Published by Elsevier Science Publishers B.V. (Biomedical Division)*
00145793/86/$3.50 

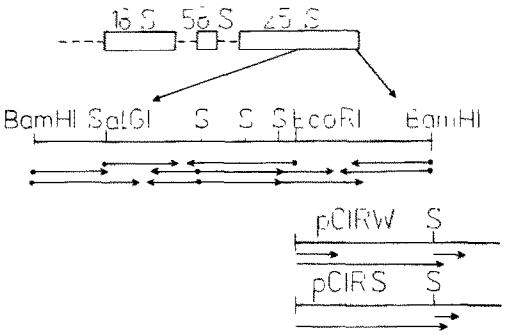


Fig.1. Restriction map and sequencing strategy used for *C. limon* part 26 S rDNA. Simplified restriction map. S, *Sau*3A site.

plasmids PCIRW and pCIRS. The position of the 3′-end of 26 S rRNA cistron was determined by comparison with other analogous eukaryotic structures [11,12].

The sequenced portion of *C. limon* 26 S rDNA was found to have 79% homology with *Saccharomyces* 26 S rRNA, which is greater than that with *Physarum polycephalum* 26 S rRNA and *E. coli* 23 S rRNA – 63% [12–14]. We compared in more detail the homologous 3′-portions of 26 S rRNA gene from yeast and lemon spanning 1483 and 1475 bp, respectively. Four areas in yeast and lemon rDNA fragments having high homology and lacking any deletions or insertions were identified

```
GGATCCGTAA CTTCGGGAAA AGGATTGGCT CTGAGGGCTG GGCACGGGGG TCCCAGTCCC GAACCCGTCG   70
GCTGTCGGCG GACTGCTCGA GCTGCCACCG CGGCGAGAGC GGGTCGCCGC GTGCCGGCCG GGGGACGGAC  140
TGGGAACGGT CCTTTCGGGG GCCTTCCCCG GGCGTCGAAC AGTCGACTCA GAACTGGTAC GGACAAGGGG  210
AATCCGACTG TTTAATTAAA ACAAAGCATT GCGATGGTCC CTGCGGATGC TCACGCAATG TGATTTCTGC  280
CCAGTGCTCT GAATGTCAAA GTGAAGAAAT TCAACCAAGC GCGGGTAAAC GGCGGGAGTA ACTATGACTC  350
TCTTAAGGTA GCCAAATGCC TCGTCATCTA ATTAGTGACG CGCATGAATG GATTAACGAG ATTCCCACTG  420
TCCCTGTCTA CTATCCAGCG AAACCACAGC CAAGGGAACG GGCTTGGCAG AATCAGCGGG GAAAGAAGAC  490
CCTGTTGAGC TTGACTCTAG TCCGACTTTG TGAAATGACT TGAGAGGTGT AGTATAAGTG GGAGCCGGAA  560
ACGGCGAAAG TGAAATACCA CTACTTTTAA CGTTATTTTA CTTATTCCGT GAATCGGAGG CGGGGCACTG  630
CCCCTCTTTT TGGACCCAGC CGCCCTCACG GGGGCCGATC CGGGCGGAAG ACATTGTCAG GTGGGGAGTT  700
TGGCTGGGGC GGCACATCTC TTAAAAGATA ACGCAGGTGT CCTAAGATGA GCTCAACGAG AACAGAAATC  770
TCGTGTGGAA CAAAAGGGTA AAAGCTCGTT TGATTCTGAT TTTCAGTACG AATACGAACC GTGAAAGCGT  840
GGCCTATCGA TCCTTTAGAC CTTCGGAATT TGAAGCTAGA GGTGTCAGAA AAGTTACCAC AGGGATAACT  910
GGCTTGTGGC AGCCAAGCGT TCATAGCGAC GTTGCTTTTT GATCCTTCGA TGTCGGCTCT TCCTATCATT  980
GTGAAGCAGA ATTCACCAAG TGTTGGATTG TTCACCCACC AATAGGGAAC GTGAGCTGGG TTTAGACCGT 1050
CGTGAGACAG GTTAGTTTTA CCCTACTGAT GACTGCGTCG TAATAGTAAT TCAACCTAGT ACGAGAGGAA 1120
CCGTTGATTC GCACAATTGG TCATCGCGCT TGGTTGAAAA GCCAGTGGCG CGAAGCTACC GTGCGCTGGA 1190
TTATGACTGA ACGCCTCTAA GTCAGAATCC GGGCTAGACC GACGCGTGCG CCCGCCGCCC GTTTGCCGAC 1260
CTGCAGTAGG GACCTCCCGG TCCCCAGAGG CACGTGTCGT AGGCCAAGCC CCCGCGGCGG AAGGGCCGCG 1330
GCGGCCGCCT TGAATCGTAA TTCCCATCGA GCGGCGGGTA GAATCCTTTG CAGACGACTT AAATACGCGA 1400
CGGGGTATTG TAAGTGGCAG AGTGGCCTTG CTGCCACGAT CCACTGAGAT TCAGCCCTGT GTCGCTCAGA 1470
TTCGT                                                                    1475
```

Fig.2. Nucleotide sequence of 1475 bp fragment which spans the *C. limon* large rRNA subunit from 1919 to 3393. The region 1–1436 was determined from the *Bam*HI fragment of pCIB 33. The nucleotide sequence of the region from 1133 to 1436 is the same in plasmids pCIB33, pCIRW and pCIRS. The sequence from 1436 to 1475 bp is in the *Eco*RI-*Pst* fragment of pCIRS.

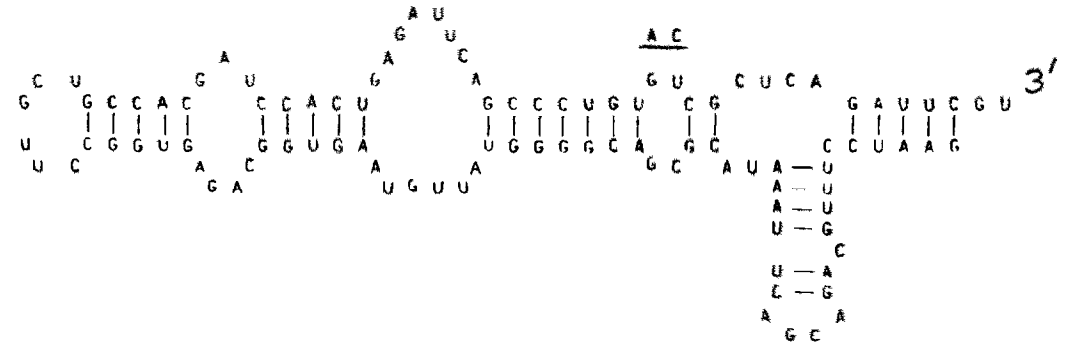


Fig.3. Putative secondary structure of the 3'-region of lemon 26 S rRNA. Dinucleotide substitution GU-AC is underlined.

at positions 1909–1947, 2177–2426, 2572–2700 and 2740–2993 of *Saccharomyces* 26 S rRNA [12]. It is noteworthy that even in the 3 variable regions 1948–2090, 2426–2571 and 3137–3233 with 52, 64 and 45% homology, respectively, insertions or deletions longer than 4 bp were not observed. The high sequence homology between large rRNA subunits in yeast and higher plants is in good agreement with their phylogenetic relationship. To investigate further the length heterogeneity found in *C. limon* rDNA, the *Eco*RI fragments were cloned in pBR325. The plasmids carrying *Eco*RI fragments from two main ribosomal length repeats present in *C. limon* rDNA in equal ratios were designated as pCIRS and pCIRW [8,9]. Certain differences in the restriction maps of these two plasmids were found in the large spacer but not in the coding region [9]. When comparing 500 bp sequences from the 3'-region of the 26 S rRNA gene in pCIRS and pCIRW, one dinucleotide substitution was detected: GT in pCIRS was replaced by AC in pCIRW, in the position 13 bp from the 3'-end of the 26 S rRNA cistron. It is unlikely that the substitution took place during the cloning procedure. The dinucleotide is localised in the evolutionarily variable region and the substitution apparently has no effect on the secondary structure suggested in fig.3. Nucleotide polymorphism in large rRNA cistrons of two main ribosomal repeats in a particular plant apparently accounts for these differences. Another type of heterogeneity was also observed in this part of the 26 S rRNA cistron: a *Bam*HI heterogeneity involves the presence of a *Bam*HI site in pCIB33 and its substitution to *Sau*3A in pCIRS and pCIRW.

The heterogeneity of *Bam*HI and *Bgl*II sites in 26 S rRNA genes in higher plants described by others may reflect plant methylation specificity [1,15,16].

There are several examples of polymorphism observed in the rDNA coding region in various organisms. Two main ribosomal repeats were found in the nematode *Ascaris lumbricoides*. They differ in point substitution in the 5'-coding region of 26 S rRNA [17]. The *Hin*dII polymorphism involves the presence of a *Hin*dII site in the 3'-end proximal third of the 28 S rRNA gene of some repeats in primate species and humans [18]. It was also shown that cloned *Xenopus laevis* DNA fragments containing the 5'-region of 28 S rRNA gene differ at one position [19].

Another example of heterogeneity was found in *Drosophila melanogaster* where slightly different 18 S rRNA is transcribed from X and Y chromosomes [20]. It is of interest to determine whether the heterogeneity of lemon ribosomal cistrons described above can be detected at the rRNA level.

## REFERENCES

[1] Appels, R., Gerlach, W.L., Dennis, E.S., Swift, H. and Peacock, W.J. (1980) Chromosoma 78, 293–311.
[2] Delseny, M., Aspart, L., Cooke, R., Grellet, F. and Penon, P. (1979) Biochem. Biophys. Res. Commun. 91, 540–547.
[3] Yakura, K., Kato, A. and Tanifuji, S. (1984) Mol. Gen. Genet. 193, 400–405.
[4] Birnboim, H.C. and Doly, J. (1979) Nucleic Acids Res. 7, 1513–1523.

[5] Maxam, A.M. and Gilbert, W. (1980) Methods Enzymol. 65, 499-560.
[6] Voleckaert, G., Winter, G. and Gaillard, C. (1984) in: Advanced Molecular Genetics (Pühler, A. and Timmis, K.N. eds).
[7] Roytberg, M.A. (1984) The Algorithm for Sequence Homology Determination, Pushchino, Preprint.
[8] Fodor, I. and Beridze, T. (1980) Biochem. Int. 1, 493-501.
[9] Kolosha, V. and Fodor, I. (1986) Mol. Biol. (USSR), in press.
[10] Rüther, U., Koenen, M., Otto, K. and Müler-Hill, B. (1981) Nucleic Acids Res. 9, 4087-4098.
[11] Kominami, R., Mishima, Y., Urano, Y., Sakai, M. and Muramutsu, M. (1982) Nucleic Acids. Res. 10, 1963-1979.
[12] Georgiev, O.I., Nikolaev, N., Hadjiolov, A.A., Skryabin, K.G., Zakharyev, V.M. and Bayev, A.A. (1981) Nucleic Acids Res. 9, 6953-6958.
[13] Otsuka, T., Nomiyama, H., Yoshida, H., Kukita, T., Kuhara, S. and Sakaki, Y. (1983) Proc. Natl. Acad. Sci. USA 80, 3163-3167.
[14] Brosius, J., Dull, T.J. and Noller, H.F. (1980) Proc. Natl. Acad. Sci. USA 77, 201-204.
[15] Rafalshi, J.A., Wiewiorowski, M. and Söll, D. (1983) FEBS Lett. 152, 241-246.
[16] Varsanyi-Breiner, A., Gusella, J.F., Keys, C., Housman, D.E., Sallivan, D., Brisson, N. and Verma, D.P.S. (1979) Gene 7, 317-334.
[17] Back, E., Müller, F. and Tobler, H. (1984) Nucleic Acids Res. 12, 1313-1332.
[18] Arnheim, N., Krystal, M., Schmickel, R., Wilson, G., Ryder, O. and Zimmer, E. (1980) Proc. Natl. Acad. Sci. USA 77, 7323-7327.
[19] Stewart, M.A., Hall, L.M.C. and Maden, B.E.H. (1983) Nucleic Acids Res. 11, 629-646.
[20] Yagura, T., Yagura, M. and Muramatsu, M. (1979) J. Mol. Biol. 133, 533-547.