Volume 11 Number 1 1983 Nucleic Acids Research

# Collection of published 5S and 5.8S ribosomal RNA sequences

Volker A.Erdmann, Erik Huysmans*, Antoon Vandenberghe* and Rupert De Wachter*

Institut für Biochemie, Fachbereich Chemie, Freie Universität Berlin, Thielallee 69-73, 1000 Berlin 33 (Dahlem), FRG and *Department Celbiologie, Universiteit Antwerpen, B-2610 Wilrijk, Belgium

The 1981 collection (1) of mature 5S and 5.8S RNA sequences is updated. The sequences reported are based on actual sequence determinations and not on the basis of oligonucleotide catalogues. Some of the earlier published oligonucleotide catalogues are given in reference 2.

For a summary of the possible structure and function of the ribosomal 5S and 5.8S RNAs the reader is refered to references 3-8.

For this year the collection of 5S RNA sequences will be presented according to the recently proposed structure model of De Wachter, Chen and Vandenberghe (8), even though we realize that these structures, being deduced by a comparative approach, have not been experimentally proven to be correct. It may well be that next years sequences will be presented in the previous (1) form or in alignment with another secondary structural model.

We would also like to thank those colleagues who have sent us their pre- or reprints on small ribosomal RNA sequences in 1982, and who have pointed out several mistakes in our previous collection.

## SEQUENCE ALIGNMENT

The procedure followed to align the 5S RNA has been described previously (8). It is based on the presence of marker residues and sequences that are conserved, with very few exceptions, in all 5S RNAs. Examples of such markers are $YRC_{35}$, $CRU_{48}$, $AARC_{74}$,

0305-1048/83/1101-r105$2.00/0

$RRU_{95}$, $RAR_{121}$, and $YGY_{140}$, to cite only those consisting of at least 3 adjacent conserved residues (Y, pyrimidine; R, purine). The number of positions needed to accomodate all presently known 5S RNA sequences is 150, not including an exceptionally long insertion of 108 nucleotides in Halococcus morrhuae, between alignment positions 129 and 130.

## SECONDARY STRUCTURE MODEL

The topological rules followed in the derivation of a uniform secondary structure model applicable to all presently known 5S RNA sequences is fully documented elsewhere (8). A brief description of these requires the following definitions:

- Double-stranded area: an area consisting of a single helix segment or of several helix segments connected by bulges or, more exceptionally, by small interior loops.
- Helix segment: a stretch of double helix uninterrupted by bulges or interior loops.
- Standard base pair: one of the base pairs G•C, A•U and G•U.
- Non-standard or odd base pair: one of the 7 other conceivable base pairs.
  Such pairs are occasionally encountered in the stems of a tRNA cloverleaf (9,10) and are assumed not to disturb the helix structure under the conditions detailed below.

The topological rules can then be summarized as follows. A helix segment should consist of at least two standard base pairs. If it contains odd base pairs these should be intercalated between two standard base pairs and separated from the segment ends by at least one G•C or two other standard base pairs. A hairpin loop should contain at least 3 bases.

For all 5S RNA sequences it is possible to distinguish 5 double-stranded areas, occurring in corresponding positions of the sequence alignment. These areas are labeled A-A' to E-E' in their order of appearance 5'- to 3'-end. The extreme positions of the double-stranded areas in the alignment, and their usual segment composition, are summarized in Table 1. The general shape of the secondary structure model is represented in Figure 1. The double-stranded areas are separated from each other by multibranched loop M and by interior loops $I_1$ and $I_2$ Areas C-C' and E-E' are closed by hairpin loops $H_1$ and $H_2$.

Table 1. Positions and segment composition of double-stranded areas

| | | usual composition (a) | | |
|---|---|---|---|---|
| double stranded area | maximum extension in alignment (positions) | number of helix segments | discontinuity between segments | non-standard base pairs |
| A<br>A' | 2- 14<br>135-147 | 1 | - | absent |
| B<br>B' | 19- 30<br>73- 82 | 2 | bulge in strand B' | absent |
| C<br>C' | 30- 41<br>55- 72 | 2 | bulge in strand C' | absent |
| D<br>D' | 83- 93<br>121-135 | 1 | - | present in plants & eubacteria |
| E<br>E' | 95-106<br>111-120 | 1 or 2 | bulge in strand E if 2 segments | frequently present if single segment |

(a) There are several exceptions. As an example, the two segments of area C-C' are separated by an internal loop in yeasts and in cyanobacteria, by a 5'-proximal bulge (in C) in chloroplasts. Similar exceptions occur in other areas.

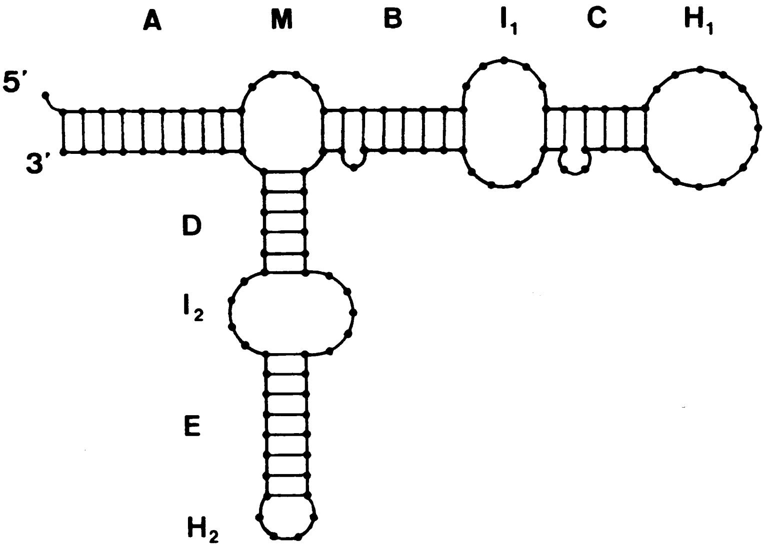


Fig. 1. General shape of secondary structure model. A to E, double-stranded areas. M, multibranched loop, and $I_1$, $I_2$, interior loops connecting the 5 areas. $H_1$, $H_2$, hairpin loops. A slightly different alternative base pairing scheme in area C is possible for all sequences (8).

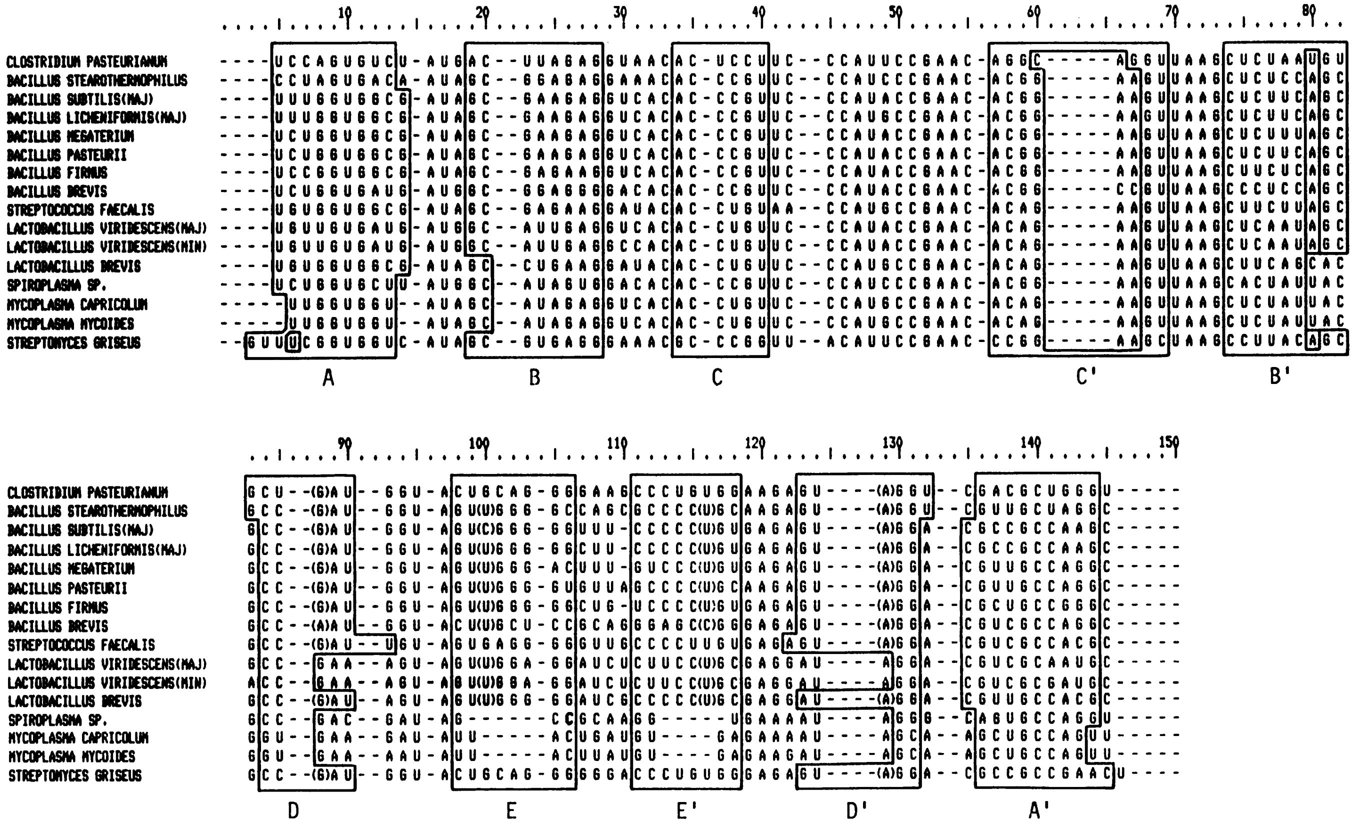
## ABBREVIATIONS AND CONVENTIONS USED IN THE 5S RNA ALIGNMENTS

The double-stranded areas A-A', B-B', C-C', D-D' and E-E' are boxed. Bulges or small interior loops connecting helix segments within a double-stranded area are indicated by nested boxes. Bases belonging to odd base pairs within a helix segment are put in parentheses. Terminal residues found in submolar amounts because of length heterogeneity in the 5S RNA preparations are written in lower case characters.

For some species, more than one sequence has been published. The following cases may be distinguished:

- A published sequence has been reexamined later by a more reliable method and corrected: in this case only the corrected sequence appears in the alignment, but the reference of the original paper is mentioned, followed by the reference of the correction.
- Different sequences have been published and it is not known which is the correct one, or different strains have been examined: all reported sequences are aligned and numbered sequentially.
- A single species contains more than one 5S RNA variant due to gene polymorphism: if the variants have been isolated and separately sequenced, each variant sequence is aligned. If heterogeneity was detected during sequencing of a mixture, a single sequence is aligned and heterogeneous positions are mentioned in a footnote on the reference page. The position numbers refer to the alignment and not to the distance from the 5'-terminus.

## Eubacterial 5 S RNA Sequences



```
CLOSTRIDIUM PASTEURIANUM           ----UCCAGUGUCU-AUGAC--UUAGAGGUAACAC-UCCUUC--CCAUUCCGAAC-AGGC------AGGUUAAGCUCUAAUGU
BACILLUS STEAROTHERMOPHILUS        ----CCUAGUGACA-AUAGC--GGAGAGGAAACAC-CCGUUC--CCAUCCCGAAC-ACGG------AAGUUAAGCUCUCCAGC
BACILLUS SUBTILIS(MAJ)             ----UUUGGUGGCG-AUAGC--GAAGAGGUCACAC-CCGUUC--CCAUACCGAAC-ACGG------AAGUUAAGCUCUUCAGC
BACILLUS LICHENIFORMIS(MAJ)        ----UUUGGUGGCG-AUAGC--GAAGAGGUCACAC-CCGUUC--CCAUGCCGAAC-ACGG------AAGUUAAGCUCUUCAGC
BACILLUS MEGATERIUM                ----UCUGGUGGCG-AUAGC--GAAGAGGUCACAC-CCGUUC--CCAUACCGAAC-ACGG------AAGUUAAGCUCUUUAGC
BACILLUS PASTEURII                 ----UCUGGUGGCG-AUAGC--GAAGAGGUCACAC-CCGUUC--CCAUACCGAAC-ACGG------AAGUUAAGCUCUUCAGC
BACILLUS FIRMUS                    ----UCCGGUGGCG-AUAGC--GAGAAGGUCACAC-CCGUUC--CCAUACCGAAC-ACGG------AAGUUAAGCUUCUCAGC
BACILLUS BREVIS                    ----UCUGGUGAUG-AUGGC--GGAGGGGACACAC-CCGUUC--CCAUACCGAAC-ACGG------CCGUUAAGCCCUCCAGC
STREPTOCOCCUS FAECALIS             ----UGUGGUGGCG-AUAGC--GAGAAGGAUACAC-CUGUAA--CCAUGCCGAAC-ACAG------AAGUUAAGCUUCUUAGC
LACTOBACILLUS VIRIDESCENS(MAJ)     ----UGUUGUGAUG-AUGGC--AUUGAGGUCACAC-CUGUUC--CCAUACCGAAC-ACAG------AAGUUAAGCUCAAUAGC
LACTOBACILLUS VIRIDESCENS(MIN)     ----UGUUGUGAUG-AUGGC--AUUGAGGCCACAC-CUGUUC--CCAUACCGAAC-ACAG------AAGUUAAGCUCAAUAGC
LACTOBACILLUS BREVIS               ----UGUGGUGGCG-AUAGC--CUGAAGGAUACAC-CUGUUC--CCAUGCCGAAC-ACAG------AAGUUAAGCUUCAGCAC
SPIROPLASMA SP.                    ----UCUGGUGCUU-AUGGC--AUAGUGGACACAC-CCGUUC--CCAUCCCGAAC-ACGG------AAGUUAAGCACUAUUAC
MYCOPLASMA CAPRICOLUM              -----UUGGUGGU--AUAGC--AUAGAGGUCACAC-CUGUUC--CCAUGCCGAAC-ACAG------AAGUUAAGCUCUAUUAC
MYCOPLASMA MYCOIDES                -----UUGGUGGU--AUAGC--AUAGAGGUCACAC-CUGUUC--CCAUGCCGAAC-ACAG------AAGUUAAGCUCUAUUAC
STREPTOMYCES GRISEUS               --GUUUCGGUGGUC-AUAGC--GUGAGGGAAACGC-CCGGUU--ACAUUCCGAAC-CCGG------AAGCUAAGCCUUACAGC
```

```
CLOSTRIDIUM PASTEURIANUM           GCU--(G)AU--GGU-ACUGCAG-GGGAAGCCCUGUGGAAGAGU----(A)GGU--CGACGCUGGGU-----
BACILLUS STEAROTHERMOPHILUS        GCC--(G)AU--GGU-AGU(U)GGG-GCCAGCGCCCC(U)GCAAGAGU----(A)GGU--CGUUGCUAGGC-----
BACILLUS SUBTILIS(MAJ)             GCC--(G)AU--GGU-AGU(C)GGG-GGUUU-CCCCC(U)GUGAGAGU----(A)GGA--CGCCGCCAAGC-----
BACILLUS LICHENIFORMIS(MAJ)        GCC--(G)AU--GGU-AGU(U)GGG-GGCUU-CCCCC(U)GUGAGAGU----(A)GGA--CGCCGCCAAGC-----
BACILLUS MEGATERIUM                GCC--(G)AU--GGU-AGU(U)GGG-ACUUU-GUCCC(U)GUGAGAGU----(A)GGA--CGUUGCCAGGC-----
BACILLUS PASTEURII                 GCC--(G)AU--GGU-AGU(U)GGG-GUGUUAGCCCC(U)GCAAGAGU----(A)GGA--CGUUGCCAGGC-----
BACILLUS FIRMUS                    GCC--(G)AU--GGU-AGU(U)GGG-GGCUG-UCCCC(U)GUGAGAGU----(A)GGA--CGCUGCCGGGC-----
BACILLUS BREVIS                    GCC--(A)AU--GGU-ACU(U)GCU-CCGCAGGGAGC(C)GGGAGAGU----(A)GGA--CGUCGCCAGGC-----
STREPTOCOCCUS FAECALIS             GCC--(G)AU--UGU-AGUGAGG-GGGUUGCCCCUUGUGAGAGU----(A)GGA--CGUCGCCACGC-----
LACTOBACILLUS VIRIDESCENS(MAJ)     GCC--GAA--AGU-AGU(U)GGA-GGAUCUCUUCC(U)GCGAGGAU----AGGA--CGUCGCAAUGC-----
LACTOBACILLUS VIRIDESCENS(MIN)     ACC--GAA--AGU-AGU(U)GGA-GGAUCUCUUCC(U)GCGAGGAU----AGGA--CGUCGCGAUGC-----
LACTOBACILLUS BREVIS               GCC--(G)AU--AGU-AGU(U)GGG-GGAUCGCCCCC(U)GCGAGGAU----(A)GGA--CGUUGCCACGC-----
SPIROPLASMA SP.                    GCC--GAC--GAU-AG-------CCGCAAGG------UGAAAAU----AGGG--CAGUGCCAGGU-----
MYCOPLASMA CAPRICOLUM              GGU--GAA--GAU-AUU------ACUGAUGU----GAGAAAAU----AGCA--AGCUGCCAGUU-----
MYCOPLASMA MYCOIDES                GGU--GAA--AAU-AUU------ACUUAUGU----GAGAAGAU----AGCA--AGCUGCCAGUU-----
STREPTOMYCES GRISEUS               GCC--(G)AU--GGU-ACUGCAG-GGGGGACCCUGUGGGAGAGU----(A)GGA--CGCCGCCGAACU----
```

## Eubacterial 5 S RNA Sequences

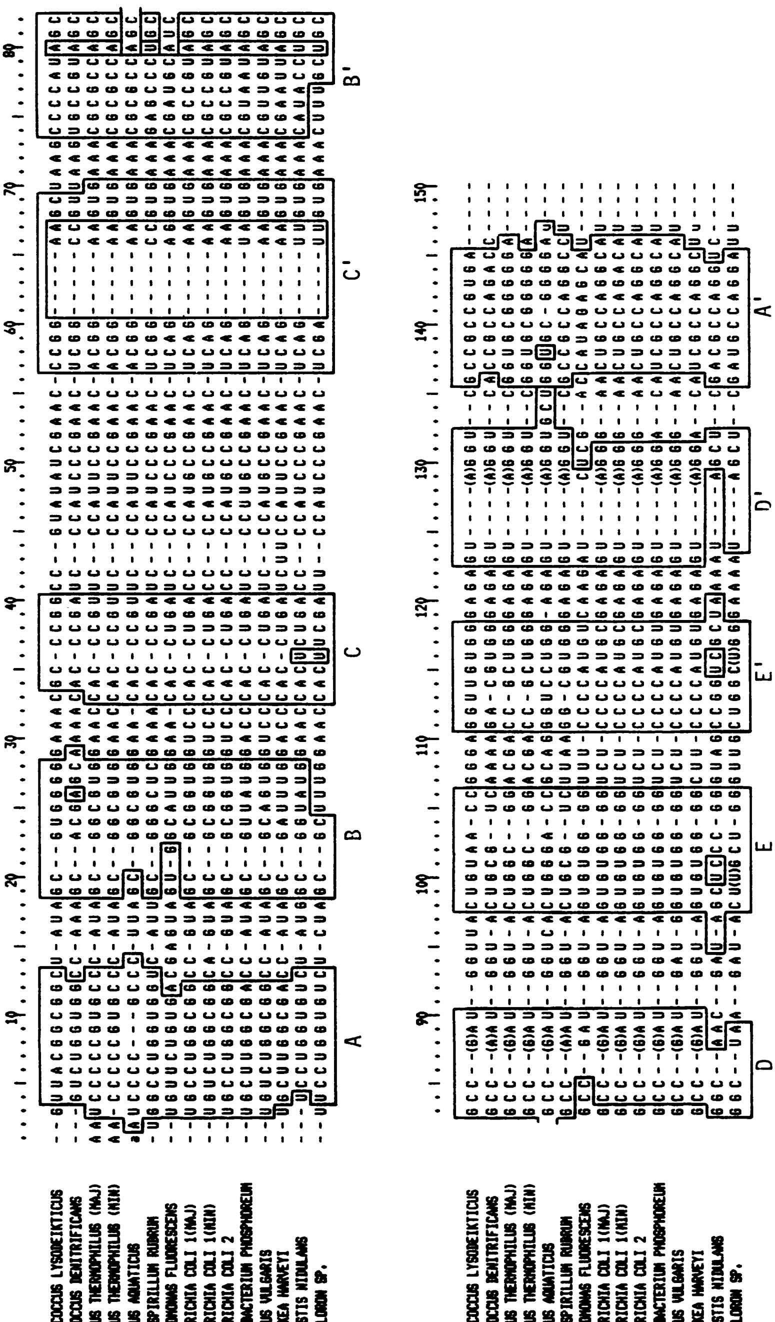
| Species | Specification | Reference Number |
|---|---|---|
| *Clostridium pasteurianum* | ATCC 6013 | 11 |
| *Bacillus stearothermophilus* | Strain 1430FV | 12 |
| *Bacillus subtilis* (major) | Strain BD 170 | 12 |
| *Bacillus licheniformis* (major) | Strain S244 | 13 |
| *Bacillus megaterium* [a] | Strain KM | 14 |
| *Bacillus pasteurii* [b] | ATCC 11859 | 14 |
| *Bacillus firmus* [c] | ATCC 14575 | 14 |
| *Bacillus brevis* | ATCC 8185 | 14 |
| *Streptococcus faecalis* | — | 14 |
| *Lactobacillus viridescens* (major) | ATCC 12706 | 15 |
| *Lactobacillus viridescens* (minor) | ATCC 12706 | 16 |
| *Lactobacillus brevis* | Strain X-2 | 14 |
| *Spiroplasma sp.* | Strain BC3 | 17 |
| *Mycoplasma capricolum* | ATCC 27343 | 18 |
| *Mycoplasma mycoides* | ssp, capri, strain PG3 | 17 |
| *Streptomyces griseus* | Strain 45-H | 19 |

(a) Sequence heterogeneities : $G/A_{8}$, $G/A_{88}$, $U/G_{118}$, $C/U_{141}$.

(b) Sequence heterogeneities : $C/U_{6}$, $G/A_{143}$.

(c) Sequence heterogeneities : $C/U_{7}$, $G/A_{101}$, $G/A_{142}$.

## Eubacterial 5 S RNA Sequences



Eubacterial 5 S RNA Sequences

| Species | Specification | Reference Number |
|---|---|---|
| *Micrococcus lysodeikticus* | ATCC 4698 | 20 |
| *Paracoccus denitrificans* | ATCC 13543 | 21 |
| *Thermus thermophilus* (major) | Strain HB 8 | 22 |
| *Thermus thermophilus* (minor) | Strain HB 8 | 22 |
| *Thermus aquaticus* | ATCC 25104 | 23 |
| *Rhodospirillum rubrum* | Strain S1 | 24 |
| *Pseudomonas fluorescens* [a] | ATCC 13430 | 25 |
| *Escherichia coli* 1 (major) [b] | MRE 600 | 26 |
| *Escherichia coli* 1 (minor) | MRE 600 | 27 |
| *Escherichia coli* 2 [c] | CA 265 | 26 |
| *Photobacterium phosphoreum* | Strain 8265 | 28 |
| *Proteus vulgaris* | — | 27 |
| *Beneckea harveyi* | Strain 392 | 29 |
| *Anacystis nidulans* | Strain 1405/1 Kratz/Allen | 30 |
| *Prochloron sp.* | Lissoclinum patella host | 21 |

---

(a) Sequence heterogeneities : $G/U_9$, $A/G_{22}$.

(b) Sequence heterogeneity : $G/U_{16}$.

(c) Sequence heterogeneity : $C/A_{14}$.

## Chloroplast and Mitochondrial 5 S RNA Sequences

```
                            ....|....10....|....20....|....30....|....40....|....50....|....60....|....70....|....80..
DRYOPTERIS ACUMINATA CHL.   -uAUUCUGGUGUCC-CAGGC--GUAGAGGAACCACACCGAUC---CAUCUCGAAC-UUGG-----UGGUGAAACUCUGCCGC
LEMNA MINOR CHL.            -uAUUCUGGUGUCC-UAGGC--GUAGAGGAACCACACCAAUC---CAUCCCGAAC-UUGG-----UGGUUAAACUCUACUGC
NICOTIANA TABACUM CHL.      -uAUUCUGGUGUCC-UAGGC--GUAGAGGAACCACACCAAUC---CAUCCCGAAC-UUGG-----UGGUUAAACUCUACUGC
SPINACIA OLERACEA CHL.      -UAUUCUGGUGUCC-UAGGC--GUAGAGGAACCACACCAAUC---CAUCCCGAAC-UUGG-----UGGUUAAACUCUACUGC

                                  A              B              C                         C'             B'

TRITICUM AESTIVUM MIT.      -aAACCGGGC-ACU-ACGGUGAGACGUGAAAACAC-CCGAUC--CCAUUCCGACC-UCGAUAUAUAUGUGGAAUCGUCUUGC


                            ..|....90....|....100....|....110....|....120....|....130....|....140....|....150
DRYOPTERIS ACUMINATA CHL.   GGU--AAC--CAA-UACUCGG-GGGGGGCCCUGCGGAAAAAU----AGCU--CGAUGCCAGGAua---
LEMNA MINOR CHL.            GGU--GAC--GAU-ACUGUAG-GGGAGGUCCUGCGGAAAAAU----AGCU--CGACGCCAGAAu----
NICOTIANA TABACUM CHL.      GGU--GAC--GAU-ACUGUAG-GGGAGGUCCUGCGGAAAAAU----AGCU--CGACGCCAGGAu----
SPINACIA OLERACEA CHL.      GGU--GAC--GAU-ACUGUAG-GGGAGGUCCUGCGGAAAAAU----AGCU--CGACGCCAGGAUG---

                              D                E           E'              D'          A'

TRITICUM AESTIVUM MIT.      GCCA-UAU---GU-ACUG----AAAUUGUU---CGGGAGACA----UGGU--CAAAGCCCGGAAa---
```

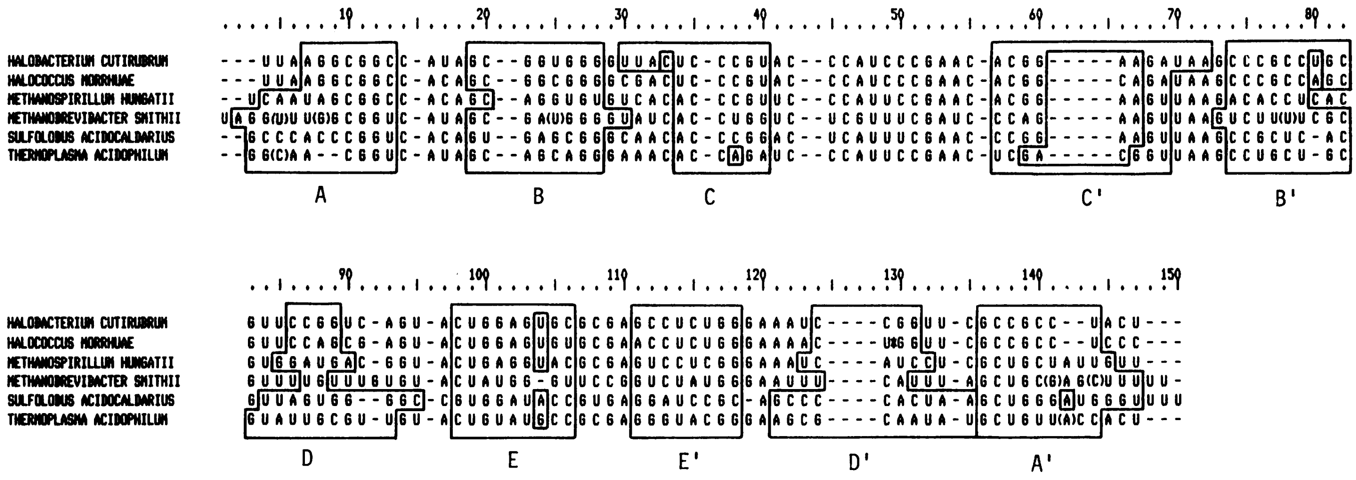
## Chloroplast 5 S RNA Sequences

| Species | Specification | Reference Number |
|---|---|---|
| *Dryopteris acuminata* chl. | — | 31 |
| *Lemna minor* chl. | — | 32 |
| *Nicotiana tabacum* chl. | var. Bright Yellow 4 | 33 |
| *Spinacia oleracea* chl. | — | 34 |

## Mitochondrial 5 S RNA Sequence

| Species | Specification | Reference Number |
|---|---|---|
| *Triticum aestivum* mit. | var. Thatcher | 35 |

## Archaebacterial 5 S RNA Sequences



```
                                     ....|....10....|....20....|....30....|....40....|....50....|....60....|....70....|....80..
HALOBACTERIUM CUTIRUBRUM             - - - U U A A G G C G G C C - A U A G C - - G G U G G G G U U A C U C - C C G U A C - - C C A U C C C G A A C - A C G G - - - - - - A A G A U A A G C C C G C C U G C
HALOCOCCUS MORRHUAE                  - - - U U A A G G C G G C C - A C A G C - - G G C G G G G C G A C U C - C C G U A C - - C C A U C C C G A A C - A C G G - - - - - - C A G A U A A G C C C G C C A G C
METHANOSPIRILLUM HUNGATII            - - U C A A U A G C G G C C - A C A G C - - A G G U G U G U C A C A C - C C G U U C - - C C A U U C C G A A C - A C G G - - - - - - A A G U U A A G A C A C C U C A C
METHANOBREVIBACTER SMITHII           U A G G(U)U U(G)G C G G U C - A U A G C - - G A(U)G G G G U A U C A C - C U G G U C - - U C G U U U C G A U C - C C A G - - - - - - A A G U U A A G U C U U(U)U C G C
SULFOLOBUS ACIDOCALDARIUS            - - G C C C A C C C G G U C - A C A G U - - G A G C G G G C A A C A C - C C G G A C - - U C A U U U C G A A C - C C G G - - - - - - A A G U U A A G C C G C U C - A C
THERMOPLASMA ACIDOPHILUM             - - G G(C)A A - - C G G U C - A U A G C - - A G C A G G G A A A C A C - C A G A U C - - C C A U U C C G A A C - U C G A - - - - - - C G G U U A A G C C U G C U - G C
                                               A                 B               C                                      C'                   B'

                                     ..|....90....|....100....|....110....|....120....|....130....|....140....|....150
HALOBACTERIUM CUTIRUBRUM             G U U C C G G U C - A G U - A C U G G A G U G C G C G A G C C U C U G G G A A A U C - - - - C G G U U - C G C C G C C - - U A C U - - -
HALOCOCCUS MORRHUAE                  G U U C C A G C G - A G U - A C U G G A G U G U G C G A A C C U C U G G G A A A A C - - - - U*G G U U - C G C C G C C - - U C C C - - -
METHANOSPIRILLUM HUNGATII            G U G G A U G A C - G G U - A C U G A G G U A C G C G A G U C C U C G G G A A A U C - - - - A U C C U - C G C U G C U A U U G U U - - -
METHANOBREVIBACTER SMITHII           G U U U U G U U U G U G U - A C U A U G G - G U U C C G G U C U A U G G G A A U U U - - - - C A U U U - A G C U G C(G)A G(C)U U U U U -
SULFOLOBUS ACIDOCALDARIUS            G U U A G U G G - - G G C - C G U G G A U A C C G U G A G G A U C C G C - A G C C C - - - - C A C U A - A G C U G G G A U G G G U U U U
THERMOPLASMA ACIDOPHILUM             G U A U U G C G U - U G U - A C U G U A U G C C G C G A G G U A C G G G A A G C G - - - - C A A U A - U G C U G U U(A)C C A C U - - -
                                             D                 E             E'               D'              A'
```

Archaebacterial 5 S RNA Sequences

| Species | Specification | Reference number |
|---|---|---|
| *Halobacterium cutirubrum* | NRC 34001 | 36,37 |
| *Halococcus morrhuae* [a] | ATCC 17082 | 38 |
| *Methanospirillum hungatii* | — | 37 |
| *Methanobrevibacter smithii* | Strain PS | 37 |
| *Sulfolobus acidocaldarius* [b] | — | 39 |
| *Thermoplasma acidophilum* | Strain 122-1B2 or 122-1B3 | 40 |

---

(a) The chain length of the complete 5 S RNA is 231 nucleotides. The following sequence has to be inserted between positions 129 and 130 : GCCUCGCAAGAGGGGCCAAGUGUGAGCCCAGGAGCGCAAUCGGUAGUACACGCGGAGUCCGUCCCGGUGGAACCCGAGACCGCUCCGUUACAACGGCUCAUAACUAGU

(b) The C residue at position 38 has an unidentified modification.

## Eukaryotic 5 S RNA Sequences

```
                              .........10........20........30........40........50........60........70........80..
HOMO SAPIENS                  ---GUCU-ACGGCC-AUACC--ACCCUGAACGCGC-CCGAUC--UCGU-CUGAUC-UCGG------AAGCUAAGCAGGGUCGG
RATTUS RATTUS                 ---GUCU-ACGGCC-AUACC--ACCCUGAACGCGC-CCGAUC--UCGU-CUGAUC-UCGG------AAGCUAAGCAGGGUCGG
XENOPUS LAEVIS(O)             ---GCCU-ACGGCC-ACACC--ACCCUGAAAGUGC-CUGAUC--UCGU-CUGAUC-UCAG------AAGCGAUACAGGGUCGG
NOTOPHTALMUS VIRIDESCENS      ---GCUU-ACGGCC-AUACC--ACCCUGAAAGCAC-CUGAUC--UCGU-CUGAUC-UCAG------AAGUUAAGCAGGGUCGG
MISGURNUS FOSSILIS(S)         ---GCUU-ACGGCC-AUACC--ACCCUGAGCACGC-CCGAUC--UCGU-CCGAUC-UCGG------AAGCUAAGCAGGGUCGG
MISGURNUS FOSSILIS(O)         ---GCUU-ACGGCC-ACACC--A(A)CCUGAGCAAGC-CCGAUC--UCGU-CUGAUC-UCGG------AAGCCAAGCAGG(G)UUGG
ACYRTHOSIPHON MAGNOLIAE       ---G(G)CA-ACG(A)CC-AUACC--ACGUUGAAUACAC-CAGUUC--UCGU-CCGAUC-ACUG------AAGUUAAGCAACGUCGG
DROSOPHILA MELANOGASTER       ---GCCA-ACGACC-AUACC--ACGCUGAAUACAU-CGGUUC--UCGU-CCGAUC-ACCG------AAAUUAAGCAGCGUCGG
PHILOSAMIA CYNTHIA RICINI     ---GCCA-ACG(U)CC-AUACC--ACGUUGAAAACAC-CGGUUC--UCGU-CCGAUC-ACCG------AAGUUAAGCAACGUCGG
BOMBYX MORI                   ---GCCA-ACGUCC-AUACC--AUGUUGAAUACAC-CGGUUC--UCGU-CCGAUC-ACCG------AAGUCAAGCAACAUCGG
ARTEMIA SALINA                ---ACCA-ACGGCC-AUACC--ACGUUGAAAGUAC-CCAGUC--UCGU-CAGAUC-CUGG------AAGUCACACAACGUCGG
LYTECHINUS VARIEGATUS         ---GCCU-ACGACC-AUACC--AUGCUGAAUAUAC-CGGUUC--UCGU-CCGAUC-ACCG------AAGUCAAGCAGCAUAGG
LINGULA ANATINA               ---GUCU-ACG(A)CC-AUACC--ACGUUGAAAGCAC-CGGUUC--UCGU-CCGAUC-ACCG------AAGUUAAGCAACGUCGG
HELIX POMATIA                 ---GUCU-ACGGCC-AUACC--ACGUUGAAAAUAC-CGGUUC--UCGU-CCGAUC-ACCG------AAGUCAAGCAACGUCGG
ARION RUFUS                   ---GUCU-ACGGCC-AUACC--ACGUUGAAAAUAC-CGGUUC--UCGU-CCGAUC-ACCG------AAGUCAAGCAACGUCGG
MYTILUS EDULIS                ---GUCU-ACGACC-AUAUC--ACGUUGAAAACAC-CGGUUC--UCGU-CCGAUC-ACCG------AAGUUAAGCAACGUCGA
ANTHOPLEURA JAPONICA          ---GUCU-ACGGCC-AUACC--ACCGGGAAAAAAC-CGGUUC--UCGU-CCGAUC-ACCG------AAGUCAAGCCCGGUAGG
                                     A         B         C                       C'           B'

                              ..|.....90........100.......110.......120.......130.......140.......150
HOMO SAPIENS                  GCCU-GGUU-AGU-ACUUGG(A)UGGGAGACCG(C)CUGGGAAUACC---GGGU---GCUGU-AGGCUu---
RATTUS RATTUS                 GCCU-GGUU-AGU-ACUUGG(A)UGGGAGACCG(C)CUGGGAAUACC---GGGU---GCUGU-AGGCUUU--
XENOPUS LAEVIS(O)             GCCU-GGUU-AGU-ACCUGG(A)UGGGAGACCG(C)CUGGGAAUACC---AGGU---GUCGU-AGGCUU---
NOTOPHTALMUS VIRIDESCENS      GCCU-GGUU-AGU-ACC(U)GG(A)UGGGAGACCG(C)CU(U)GGAAUACC---AGGU---GCUGU-AGGCUU---
MISGURNUS FOSSILIS(S)         GCCU-GGUU-AGU-ACUUGG(A)UGGGAGACUG(C)CUGGGAAUACC---AGGU---GUUGU-AAGCUu---
MISGURNUS FOSSILIS(O)         GCCU-GGUU-AGU-ACUUGG(A)UGGGAGACUG(C)CUGGGAAUACC---AGGU---GUUGU-AAGCUu---
ACYRTHOSIPHON MAGNOLIAE       GCGU-AGUU-AGU-ACUUGG(A)UGGGUGACCG(C)UUGGGAACACU---ACGU---G(C)CGU-UG(G)CAU---
DROSOPHILA MELANOGASTER       GCGC-GGUU-AGU-ACUUAG(A)UGGGGGACCG(C)UUGGGAACACC---GCGU---GUUGU-UGGCCU---
PHILOSAMIA CYNTHIA RICINI     GCGC-GGUC-AGU-ACUUGG(A)UGGGUGACCA(C)CUGGGAACACC---GCGU---G(C)CGU-UGGCU----
BOMBYX MORI                   GCGU-GGUC-AGU-ACUUGG(A)UGGGUGACCG(C)CUGGGAACACC---ACGU---GAUGU-UGGCUU---
ARTEMIA SALINA                GCCC-GGUC-AGU-ACUUGG(A)UGGGUGACCG(C)CUGGGAACACC---GGGU---GCUGU-UGGCAU---
LYTECHINUS VARIEGATUS         GCUC-GGUU-AGU-ACUUGG(A)UGGGAGACCG(C)CUGGGAAUACC---GGGU---GUUGU-AGGCUU---
LINGULA ANATINA               GCCA-GGUU-AGU-ACUUGG(A)UGGGUGACCG(C)CUGGGAAUACC---UGGU---G(C)CGU-AGACA----
HELIX POMATIA                 GCGU-GGUU-AGU-ACUUAG(A)UGGGUGACCG(C)UUGGGAACACC---ACGU---GUUGU-AGAUA----
ARION RUFUS                   GCGU-GGUU-AGU-ACUUAG(A)UGGGUGACCG(C)UUGGGAACACC---ACGU---GUUGU-AGACA----
MYTILUS EDULIS                GCCC-GGUU-AGU-ACUUGG(A)UGGGUGACCG(C)CUGGGAAUACC---GGGU---GUUGU-AGACA----
ANTHOPLEURA JAPONICA          GCCA-GGUU-AGU-ACUUGG(A)UGGGUGACCG(C)CUGGGAAUACC---UGGU---GCUGU-AGACUU---
                                 D              E              E'             D'         A'
```

Eukaryotic 5 S RNA Sequences

| Species | Specification | Reference Number |
|---|---|---|
| *Homo sapiens* | KB cells | 41 |
| *Rattus rattus* | Liver | 42 |
| *Xenopus laevis* (oocytes) | Oocytes | 43 |
| *Notophthalmus viridescens* | — | 44 |
| *Misgurnus fossilis* (somatic) | Liver | 45 |
| *Misgurnus fossilis* (oocytes) | Oocytes | 45 |
| *Acyrthosiphon magnoliae* | Whole tissue | 46 |
| *Drosophila melanogaster*[a] | — | 47,48 |
| *Philosamia cynthia ricini* | Posterior silkgland | 49 |
| *Bombyx mori* | Posterior silkgland | 50 |
| *Artemia salina* | Cysts | 51 |
| *Lytechinus variegatus* | — | 52 |
| *Lingula anatina*[b] | Whole tissue | 53 |
| *Helix pomatia* | Whole tissue | 54 |
| *Arion rufus* | Whole tissue | 54 |
| *Mytilus edulis* | Whole tissue | 54 |
| *Anthopleura japonica* | Whole tissue | 55 |

---

(a) Sequence heterogeneity : $G/A_{106}$

(b) Sequence heterogeneity : $G/A_{32}$

# Eukaryotic 5 S RNA Sequences

```
                              ....|....10....|....20....|....30....|....40....|....50....|....60....|....70....|....80..
HALICLONA OCULATA             ---GCCU-GCGGCC-AUACC--ACGUUGAAUGCAC-CGGUUC--CCAU-CUGAAC-ACCG------AAGUUAAGCAACGUCGG
HALICHONDRIA PANICEA          ---GCCU-ACGGCC-AUACC--ACGUUGAAAACAC-CGGUUC--UCGU-CUGAUC-ACCG------AAGUUAAGCAACGUAGG
HYMENIACIDON SANGUINEA        ---GUCU-ACGACC-AUACC--ACGUUGAAAACAC-CGGUUC--UCGU-CCGAUC-ACCG------AAGUUAAGCAACGUAGG
CAENORHABDITIS ELEGANS        ---GCUU-ACGACC-AUAUC--ACGUUGAAUGCAC-GCCAUC--CCGU-CCGAUC-UGGC------AAGUUAAGCAACGUUGA
CAENORHABDITIS BRIGGSIAE      ---GCUU-ACGACC-AUAUC--ACGUUGAAUGCAC-GCCAUC--CCGU-CCGAUC-UGGC------AAGUUAAGCAACGUUGA
EUGLENA GRACILIS 1            --GGCGU-ACGGCC-AUACU--ACCGGGAAUACAC-CUGAAC--CCGU-UCGAUU-UCAG------AAGUUAAGCCUGGUCAG
EUGLENA GRACILIS 2            ---GCGU-ACGGCC-AUACU--ACC(G)GGAAUACAC-CUGAAC--CCGU-UCGAUU-UCAG------AAGUUAAGCC(G)GGUUAG
CRITHIDIA FASCICULATA         ---GAGU-ACG(A)CC-AUACU--UG(A)GUGAAAACAC-CAUAUC--CCGU-CCGAUU-UGUG------AAGUUAAGCAC(C)CACAG
TETRAHYMENA THERMOPHILA       ---GCUA-UCGGCC-AUACU--AAGGUGAAAACAC-CGGAUC--CCAU-UCGAAC-UCCG------AAGUUAAGCGCCUUAAG
BLEPHARISMA JAPONICUM         ---GUUG-UCGGCC-AUACU--AUGCCUAACGCAC-CAGAUC--CCAU-CCGAAC-UCUG------AAGUUAAGCGGCAUAAG
PARAMECIUM TETRAURELIA        ---GUUG-(G)UGGCC-AUACU--AAGCCUAAAGCAC-CGGAUC--CCAU-UCGAAC-UCCG------AAGUUAAGCGGCUUAAG
ACANTHAMOEBA CASTELLANII      ---GGAU-ACGGCC-AUACU--GCGCAGAAAGCAC-CGCUUC--CCAU-CCGAAC-AGCG------AAGUUAAGCUGCGCCAG
CRYPTHECODINIUM COHNII        ---GCUG-ACGGCC-AUACC--GUGUCGAAUGCAC-CGGAUC--UCUU-CUGACC-UCCG------AAGUUAAGCGGCACAGG
PHYSARUM POLYCEPHALUM         ---GGAU-GCGGCC-AUACU--AAGGAGAAAGCAC-CUCAUC--CCGU-CCGAUC-UGAG------AAGUUAAGCUCCUUCAG
DICTYOSTELIUM DISCOIDEUM      ---GUAU-ACGGCC-AUACU--AGGUUGGAAACAC-AUCAUC--CCGU-UCGAUC-UGAU------AAGUAAAUCGACCUCAG
SAPROLEGNIA FERAX             ---GGAA-UCGGCC-AUAUU--AGCCUGACUAGAC-UGCAUC--CCGU-CCGCUC-UGCA------AAGUUAACCAGGCUCAA
PYTHIUM HYDNOSPORUM           ---GUAG-ACGGCC-AUCUU--AGGCUGAGAACAC-CGUAUC--CCGU-UCGCUC-UGCG------AAGUUAAGCAGCCUCAA
                                     A                 B            C                          C'             B'

                              ..|....90....|....100....|....110....|....120....|....130....|....140....|....150
HALICLONA OCULATA             GCCA-GCUU-AGU-ACCUGG(A)UGGGUGACCG(C)CUGGGAAUCGC---UGGU---GCUGC-AGGCUU---
HALICHONDRIA PANICEA          GCCU-GCCC-AGU-ACUUGG(A)UGGGUGACCG(C)CUGGGAACAGC---AGGU---GUUGU-AGGCUU---
HYMENIACIDON SANGUINEA        GCCU-GCCC-AGU-ACUUGG(A)UGGGUGACCG(C)CUGGGAACAGC---AGGU---GUUGU-AGACUU---
CAENORHABDITIS ELEGANS        GUCC-AGUU-AGU-ACUUGG(A)UCGGAGACGG(C)CUGGGAAUCCU---GGAU---GUUGU-AAGCU----
CAENORHABDITIS BRIGGSIAE      GUCC-AGUU-AGU-ACUUGG(A)UCGGAGACGG(C)CUGGGAAUCCU---GGAU---GUUGU-AAGCU----
EUGLENA GRACILIS 1            GCCC-AGUU-AGU-ACUGAGGUGGGCGACCACUUGGGAACACU---GGGU---GCUGU-ACGCUU---
EUGLENA GRACILIS 2            GCCC-AGUU-AGU-ACUGAG-UGGGCGACCACUUGGGAACACU---GGGU---GCUGU-ACGCUU---
CRITHIDIA FASCICULATA         GCUU-AGUU-AGU-ACUGAGGUCAGUGAUGACUCGGGAACCCU---GAGU---G(C)CGU-ACUCCC---
TETRAHYMENA THERMOPHILA       GCUG-GGUU-AGU-ACUAAGGUGGGGGACCGCUUGGGAAGUCC---CAGU---GUCGA-UAGCCU---
BLEPHARISMA JAPONICUM         GCGA-GGUU-AGU-ACUUGGGUGGGGGACCGCCAGGGAAGCCC---UCGU---GCUGA-CAGCUA---
PARAMECIUM TETRAURELIA        GCGA-GGUU-AGU-ACUAAGGUGGGGGACCGCUUGGGAAGUCC---UCGU---GUUG(A)-CAACCC---
ACANTHAMOEBA CASTELLANII      GCGG-UGUU-AGU-ACUGGGGUGGGCGACCACCCGGGAAUCCA---CCGU---GCCGU-AUCCU----
CRYPTHECODINIUM COHNII        GCCC-GGAU-AGU-ACUGGGGUGGGGGACCGCCCGGGAAGUCCUUAGGGU---GCUGU-CAGCU----
PHYSARUM POLYCEPHALUM         GCGU-GGUU-AGU-ACUGGGGUGGGGGACCACCUGGGAAUCCC---ACGU---GCUGC-AUUCUU---
DICTYOSTELIUM DISCOIDEUM      GCCU-UCCA-AGU-ACUCUGGUUGGAGACAACAGGGGAACAUA---GGGU---GCUGU-AUACU----
SAPROLEGNIA FERAX             GCGU-GGGU-AGU-ACUCAGGUGGGUGACCACUGGGGAAGUCC---ACGU---GCUG--AUUCU----
PYTHIUM HYDNOSPORUM           GCUC-GGGU-AGU-ACUCGGGUGGGUGACCACCGGGGAAGUCC---GAGU---GCUGU-CUAC-----
                                D               E              E'         D'            A'
```

## Eukaryotic 5 S RNA Sequences

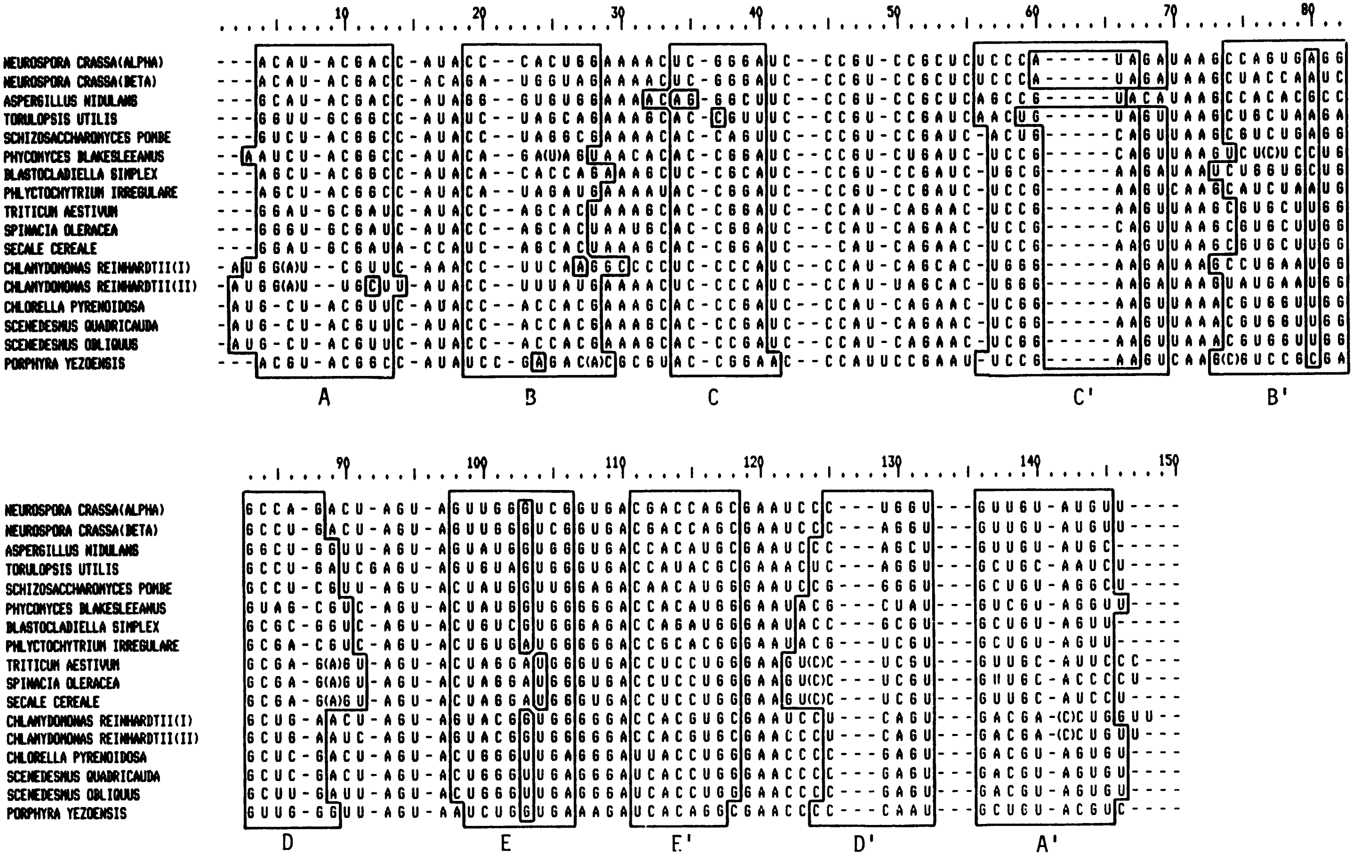
| Species | Specification | Reference Number |
|---|---|---|
| *Haliclona oculata* | Whole tissue | 56 |
| *Halichondria panicea* | Whole tissue | 56 |
| *Hymeniacidon sanguinea* | Whole tissue | 56 |
| *Caenorhabditis elegans* | var. Bristol strain N2 | 57 |
| *Caenorhabditis briggsae* | — | 57 |
| *Euglena gracilis* 1 | Strain Z | 58 |
| *Euglena gracilis* 2[a] | Strain Z, CM 500 | 59 |
| *Crithidia fasciculata* | — | 60 |
| *Tetrahymena thermophila*[b] | Strain B | 61 |
| *Blepharisma japonicum* | R 13 | 62 |
| *Paramecium tetraurelia* | Mating type VIII | 62 |
| *Acanthamoeba castellanii* | ATCC 30010 | 63 |
| *Crypthecodinium cohnii* | — | 64 |
| *Physarum polycephalum* | — | 65 |
| *Dictyostelium discoideum* | Strain NC-4 | 66 |
| *Saprolegnia ferax* | ATCC 26116 | 67 |
| *Pythium hydnosporum* | ATCC 26929 | 67 |

---

(a) Sequence heterogeneity : $C/A_5$ . $U_{48}$ stands for pseudouridine.

(b) Sequence heterogeneities : $C/U_5$, $A/G_7$, $U/C_{142}$, $G/A_{144}$.

## Eukaryotic 5 S RNA Sequences



```
                                10        20        30        40        50        60        70        80
NEUROSPORA CRASSA(ALPHA)        ---ACAU-ACGACC-AUACC--CACUGGAAAACUC-GGGAUC--CCGU-CCGCUCUCCCA------UAGAUAAGCCAGUGAGG
NEUROSPORA CRASSA(BETA)         ---ACAU-ACGACC-ACAGA--UGGUAGAAAACUC-GGGAUC--CCGU-CCGCUCUCCCA------UAGAUAAGCUACCAAUC
ASPERGILLUS NIDULANS            ---GCAU-ACGACC-AUAGG--GUGUGGAAAACAG-GGCUUC--CCGU-CCGCUCAGCCG------UACAUAAGCCACACGCC
TORULOPSIS UTILIS               ---GGUU-GCGGCC-AUAUC--UAGCAGAAAGCAC-CGUUUC--CCGU-CCGAUCAACUG------UAGUUAAGCUGCUAAGA
SCHIZOSACCHAROMYCES POMBE       ---GUCU-ACGGCC-AUACC--UAGGCGAAAACAC-CAGUUC--CCGU-CCGAUC-ACUG------CAGUUAAGCGUCUGAGG
PHYCOMYCES BLAKESLEEANUS        --AAUCU-ACGGCC-AUACA--GA(U)AGUAACACAC-CGGAUC--CCGU-CUGAUC-UCCG------CAGUUAAGUCU(C)UCCUG
BLASTOCLADIELLA SIMPLEX         ---AGCU-ACGGCC-AUACA--CACCAGAAAGCUC-CGCAUC--CCGU-CCGAUC-UGCG------AAGAUAAUCUGGUGCUG
PHLYCTOCHYTRIUM IRREGULARE      ---AGCU-ACGGCC-AUACA--UAGAUGAAAAUAC-CGGAUC--CCGU-CCGAUC-UCCG------AAGUCAAGCAUCUAAUG
TRITICUM AESTIVUM               ---GGAU-GCGAUC-AUACC--AGCACUAAAGCAC-CGGAUC--CCAU-CAGAAC-UCCG------AAGUUAAGCGUGCUUGG
SPINACIA OLERACEA               ---GGGU-GCGAUC-AUACC--AGCACUAAUGCAC-CGGAUC--CCAU-CAGAAC-UCCG------CAGUUAAGCGUGCUUGG
SECALE CEREALE                  ---GGAU-GCGAUA-CCAUC--AGCACUAAAGCAC-CGGAUC---CAU-CAGAAC-UCCG------AAGUUAAGCGUGCUUGG
CHLAMYDOMONAS REINHARDTII(I)    -AUGG(A)U--CGUUC-AAACC--UUCAAGGCCCCUC-CCCAUC--CCAU-CAGCAC-UGGG------AAGAUAAGCCUGAAUGG
CHLAMYDOMONAS REINHARDTII(II)   -AUGG(A)U--UGCUU-AUACC--UUUAUGAAAACUC-CCCAUC--CCAU-UAGCAC-UGGG------AAGAUAAGUAUGAAUGG
CHLORELLA PYRENOIDOSA           -AUG-CU-ACGUUC-AUACC--ACCACGAAAGCAC-CCGAUC--CCAU-CAGAAC-UCGG------AAGUUAAACGUGGUUGG
SCENEDESMUS QUADRICAUDA         -AUG-CU-ACGUUC-AUACC--ACCACGAAAGCAC-CCGAUC--CCAU-CAGAAC-UCGG------AAGUUAAACGUGGUUGG
SCENEDESMUS OBLIQUUS            -AUG-CU-ACGUUC-AUACC--ACCACGAAAGCAC-CCGAUC--CCAU-CAGAAC-UCGG------AAGUUAAACGUGGUUGG
PORPHYRA YEZOENSIS              ---ACGU-ACGGCC-AUAUCC-GAGAC(A)CGCGUAC-CGGAAC--CCAUUCCGAAU-UCCG------AAGUCAAG(C)GUCCGCGA
                                      A             B            C                       C'             B'

                                  90        100       110       120       130       140       150
NEUROSPORA CRASSA(ALPHA)        GCCA-GACU-AGU-AGUUGGGUCGGUGACGACCAGCGAAUCCC---UGGU---GUUGU-AUGUU----
NEUROSPORA CRASSA(BETA)         GCCU-GACU-AGU-AGUUGGGUCGGUGACGACCAGCGAAUCCC---AGGU---GUUGU-AUGUU----
ASPERGILLUS NIDULANS            GGCU-GGUU-AGU-AGUAUGGUGGGUGACCACAUGCGAAUCCC---AGCU---GUUGU-AUGC-----
TORULOPSIS UTILIS               GCCU-GAUCGAGU-AGUGUAGUGGGUGACCAUACGCGAAACUC---AGGU---GCUGC-AAUCU----
SCHIZOSACCHAROMYCES POMBE       GCCU-CGUU-AGU-ACUAUGGUUGGAGACAACAUGGGAAUCCG---GGGU---GCUGU-AGGCU----
PHYCOMYCES BLAKESLEEANUS        GUAG-CGUC-AGU-ACUAUGGUGGGGGACCACAUGGGAAUACG---CUAU---GUCGU-AGGUU----
BLASTOCLADIELLA SIMPLEX         GCGC-GGUC-AGU-ACUGUCGUGGGAGACCAGAUGGGAAUACC---GCGU---GCUGU-AGUU-----
PHLYCTOCHYTRIUM IRREGULARE      GCGA-CGUC-AGU-ACUGUGAUGGGGGACCGCACGGGAAUACG---UCGU---GCUGU-AGUU-----
TRITICUM AESTIVUM               GCGA-G(A)GU-AGU-ACUAGGAUGGGUGACCUCCUGGGAAGU(C)C---UCGU---GUUGC-AUUCCC---
SPINACIA OLERACEA               GCGA-G(A)GU-AGU-ACUAGGAUGGGUGACCUCCUGGGAAGU(C)C---UCGU---GUUGC-ACCCCU---
SECALE CEREALE                  GCGA-G(A)GU-AGU-ACUAGGAUGGGUGACCUCCUGGGAAGU(C)C---UCGU---GUUGC-AUCCU----
CHLAMYDOMONAS REINHARDTII(I)    GCUG-AACU-AGU-AGUACGGUGGGGGACCACGUGCGAAUCCU---CAGU---GACGA-(C)CUGGUU--
CHLAMYDOMONAS REINHARDTII(II)   GCUG-AAUC-AGU-AGUACGGUGGGGGACCACGUGCGAACCCU---CAGU---GACGA-(C)CUGUU---
CHLORELLA PYRENOIDOSA           GCUC-GACU-AGU-ACUGGGUUGAGGGAUUACCUGGGAACCCC---GAGU---GACGU-AGUGU----
SCENEDESMUS QUADRICAUDA         GCUC-GACU-AGU-ACUGGGUUGAGGGAUCACCUGGGAACCCC---GAGU---GACGU-AGUGU----
SCENEDESMUS OBLIQUUS            GCUU-GAUU-AGU-ACUGGGUUGAGGGAUCACCUGGGAACCCC---GAGU---GACGU-AGUGU----
PORPHYRA YEZOENSIS              GUUG-GGUU-AGU-AAUCUGGGUGAAAGAUCACAGGCGAACCCC---CAAU---GCUGU-ACGUC----
                                 D              E          E'          D'          A'
```

Eukaryotic 5 S RNA Sequences

| Species | Specification | Reference Number |
|---|---|---|
| *Neurospora crassa* (alpha) | — | 68,69 |
| *Neurospora crassa* (beta) | — | 69 |
| *Aspergillus nidulans* | Strain paba A1, bi A1 | 68 |
| *Torulopsis utilis* | — | 70,71 |
| *Schizosaccharomyces pombe* | IFO 0345 | 72 |
| *Phycomyces blakesleeanus* | NRRL 1555 | 73 |
| *Blastocladiella simplex* | ATCC 24579 | 67 |
| *Phlyctochytrium irregulare* | ATCC 32066 | 67 |
| *Triticum aestivum* | Embryo | 74 |
| *Spinacia oleracea* | var. 424 | 75 |
| *Secale cereale* [a] | cv. Lavoszpatonai | 76 |
| *Chlamydomonas reinhardtii* (I) | Cell wall mutant CW 15 | 77 |
| *Chlamydomonas reinhardtii* (II) | Cell wall mutant CW 15 | 77 |
| *Chlorella pyrenoidosa* | ATCC 11469 | 78,79 |
| *Scenedesmus quadricauda* | ATCC 11460 | 79 |
| *Scenedesmus obliquus* [b] | — | 80 |
| *Porphyra yezoensis* | — | 81 |

---

(a) The sequence in positions 14 to 19, 45, and 142 to 147 is questioned by McKay et al.[74].

(b) The residue in position 144 was not identified but in the alignment it is assumed to be U as in *Scenedesmus quadricauda*.

## Eukaryotic 5.8S RNA Sequences

| | 1 10 20 30 40 50 60 70 80 90 100 |
|---|---|
| A.C. | pAACΨCCUAACAAC$\underset{m}{G}$GAΨAUCUUGGUUCUCGCGAGGAUGAAGA$\underset{m}{A}$CGCAGCGAAAUGCGAUACGUAGUGUGAAUCGCAGGGAUCAGUGAA$\underset{m}{U}$CAΨCGAAUCUΨ |
| A.S. | pAGAAUGACCCUGGAGGAUGGAUCACUUGGCUCACAUUACGAAGA$\underset{m}{C}$G$\underset{m}{A}$ACGCAGCUAGACGCGUGAUUCCAUGCGAACUGCAGGACACAUGGAACGUCUAU |
| C. | p$\underline{\underline{C}}$AACUCUUAGCGG$\underset{m}{U}$GGAUCACUCGGCUCGUGGGUCGAUGAAGAACGCAGCGCUAGC$\underset{U}{\Psi}$GCGAGAAUUAAUGΨGAAUU$\underset{m}{G}$CAGGACACAUUGAUCAUCGACAC |
| C.C. | pACAACUUUCAGCAGUUGAUUCCUUGGUUCAGACCUCGAUGAAGGGCACUGCGAAAGUGAAUGGCAUGUGAAUGCAGGCAUCCGGGAAUUGAGAGCUUCUU |
| C.F. | pAACGUGU$\underset{m}{C}$GCGAUGGAUGACUUGGCUUCCUAUCUCGUUGAAGA$\underset{m}{A}$CGCAGUAAAGUGCGAUAAGUGGUAΨCAAUUG$\underset{m}{C}$AGAAUCAUUCAAUUACCGAAUCUU |
| C.R. | pAACUCUCAACAACGGAUAUCUUGGCUCUCGGAUCGAUGAAGGACGCAGCGAAAUGCGAUACGUAGUGUGAACUGCAGAAAUACGUGAACUAUCGAAUCCC |
| D.D. | pUUAAGCAUAAACGGUGAAUACCUCGACUCCCAAAUUGAUGAAGACCGUAGCAAACUGCGAUAAUUCACUUGAAUUGCAGCCUACUGGGAUAGUUGAAAUG |

| | 101 110 120 130 140 150 160 |
|---|---|
| A.C. | UGAACGCAAGUUGCGCUCUCGUGGUUUAACCCCCCGGGAGCACGUUCGCUUGAGUGCCGCU $\underline{\underline{U}}$ $_{OH}$ |
| A.S. | AUUUΨGAACGCAAAUUGCAGUCCAGCCUUUGAGGUUGGACUACGUCUGGCUGAGAGACGGAU$_{OH}$ |
| C. | UUCGAACGCACUUGCGGCCCCGGGUUCCUCCCGGGGCUACGCCUGCCUGAGCGUCGCUU$_{OH}$ |
| C.C. | GAAUGCAUACUGCUCCAGCUGACUUGUCAGUUGGAUAAUCUCCCUCAGGGACUA$_{OH}$ |
| C.F. | UGAACGCAAACGGCGCAUGGGAGAAGCUCUUUUGAGUCAUCCCCGUGCAUGCCAUAUUCUCCA$\underset{m}{G}$UGUCGAA(C)$_{OH}$ |
| C.R. | UGAACGUAUACUGCGCCCGAGGCCCCGGUAGAGCAUGUCUGCCUUAGUGCUGGGUU$_{OH}$ |
| D.D. | UUGAACGCACAUGAUGACAUCGGUCCUUUCGGAUUAGGUGUUAUACUUGGGUGAGAGUGGUC |

---

The double underlined nucleotides occur in less than one mole per mole 5.8S RNA. Nucleotide written directly under another nucleotide in the sequence indicates that it may also be found in this position. m indicates that the nucleotide is methylated.

## Eukaryotic 5.8S RNA Sequences

| Abbreviation | RNA Source | Reference Number |
|---|---|---|
| A.C. | *Acanthamoeba castellani* | 82 |
| A.S. | *Artemia salina* | 83 |
| C. | Chicken (embryonic cells) | 84 |
| C.C. | *Cryothecodinium cohnii* | 85 |
| C.F. | *Crithidia fasciculata* | 86 |
| C.R. | Chlamydomonas reinhardii | 87 |
| D.D. | *Dictyostelium discoideum* | 88 |

## Eukaryotic 5.8S RNA Sequences

|  | 1 . 10 . 20 . 30 . 40 . 50 . 60 . 70 . 80 . 90 . 100 |
|---|---|
| H.L. | pCGACUCUUAGCGGUGGAUCACUCGGCUCGUGCGUCGAUGAAGAACGCAGCGCUAGCΨGCGAGAAUUAAUGΨGAAUUGCAGGACACAUUGAUCAUCGACAC |
| M. | pCGACUCUUAGCGGUGGAUCACUCGGCUCGUGCGUCGAUGAAGAACGCAGCGCUAGCΨGCGAGAAUUAAUGΨGAAUUGCAGGACACAUUGAUCAUCGACAC |
| N.C. | pAAACUUUCAACAACGGAUCUCUUGGUUCUGGCAUCGAUGAAGAACGCAGCGAAAUGCGAUAGGUAAUGUGAAUUGCAGAAUUCAGUGAAUCAUCGAAUCU |
| N.H. | pCGACUCUUAGCGGUGGAUCACUCGGCUCGUGCGUCGAUGAAGAACGCAGCGCUAGCΨGCGAGAAUUAAUGΨGAAUUGCAGGACACAUUGAUCAUCGACAC |
| R.T. | pCAACUCUUAGCGGUGGAUCACUCGGCUCGUGCGUCGAUGAAGAACGCAGCGCUAGCΨGCGAGAACUAAUGΨGAAUUGCAGGACACAUUGAUCAUUGACAC |
| S.P. | pAAACUUUCAGCAACGGAUCUCUUGGCUCUCGCAUCGAUGAAGAACGCAGCGAAAUGCGAUACGUAAUGUGAAUUGCAGAAUUCCGUGAAUCAUCGAAUCU |
| T. | pCAACUCUUAACGGUGGAUCACUCGGCUCGUGCGUCGAUGAAGAACGCAGCGCUAGCUGCGAGAAUUAAUGΨGAAUUGCAGGACACAUUGAUCAUCGACAC |

|  | 101 . 110 . 120 . 130 . 140 . 150 . |
|---|---|
| H.L. | UUCGAACGCACUUGCGGCCCCGGGUUCCUCCCGGGGCUACGCCUGUCUGAGCGUCGCUU$_{OH}$ |
| M. | UUCGAACGCACUUGCGGCCCCGGGUUCCUCCCGGGGCUACGCCUGUCUGAGCGUCGCUU$_{OH}$ |
| N.C. | UUGAACGCACAUUGCGCUCGCCAGUAUUCUGGCGAGCAUGCCUGUUCGAGCGUCAUUU$_{OH}$ |
| N.H. | UUCGAACGCACUUGCGGCCCCGGGUUCCUCCCGGGGCUACGCCUGUCUGAGCGUCGCU$_{OH}$ |
| R.T. | UUCGAACGCACUUUGCGGCCCCAGGGGUUCCUCCUGGGGCUACGCCUGUCUGAGUGUCGCUU$_{OH}$ |
| S.P. | UUGAACGCACAUUGCGCCUUUGGGUUCUACCAAAGGCAUGCCUGUUUGAGUGUCAUU |
| T. | UUCGAACGCACUUGCGGCCCCGGGUUCCUCCCGGGGCUACGCCUGUCUGAGCGUCGCUU$_{OH}$ |

---

The double underlined nucleotides occur in less than one mole per mole 5.8S RNA. Nucleotide written directly under another nucleotide in the sequence indicates that it may also be found in this position. m indicates that the nucleotide is methylated. N.C. sequence has been derived from rDNA.

## Eukaryotic 5.8S RNA Sequences

| Abbreviation | RNA Source | Reference Number |
|---|---|---|
| H.L. | HeLa cells | 84,89 |
| M. | Mouse (MPC-11 cells) | 89 |
| N.C. | *Neurospora crassa* | 90 |
| N.H. | Novikoff hepatoma ascites cells | 91 |
| R.T. | Rainbow trout (*Salmo gairdneri*, RTG-2) | 92 |
| S.P. | Schizosaccharomyces pombe | 93 |
| T. | Turtle (heart cells CCL 50) | 94 |

## Eukaryotic 5.8S RNA Sequences

```
         1       10        20        30        40        50        60        70        80        90       100
V.F.    pAGAAUGACUCUCGGCAACGGAΨAUCUAGGCUCUUGCAUCGAUGAAGAACGUAGCGAAAUGCGAUACUUGGUGUGAAUΨGCAGAAUCCCGUGAACCAUCGA
                                                       m                               m
W.E.    pCACACGACUCUCGGCAACGGAΨAUCUCGGCUCUCGCAUCGAUGAAGAACGUAGCGAAAUGCGAUACCUGGUGUGAAUΨGCAGAAUCCCGCGAACCAUCGA
         =                                             m                               m
X.B.    pCGACUCUUAGCGGUGGAUCACUCGGCUCGUGCGUCGAUGAAGAACGCAGCGCUAGCΨGCGAGAAUUAGUGΨGAAUUGCAGGACACAUUGAUCAUCGACAC
         =                                                                           m
X.L.    pCGACUCUUAGCGGUGGAUCACUCGGCUCGUGCGUCGAUGAAGAACGCAGCGCUAGCΨGCGAGAAUUAGUGΨGAAUUGCAGGACACAUUGAUCAUCGACAC
         =                                                       a                   m
Y.S.Ce. pAAACUUUCAACAACGGAUCUCUUGGUUCUCGCAUCGAUGAAGAACGCAGCGAAAUGCGAUACGUAAUGUGAAΨUGCAGAAUUCCGUGAAUCAUCGAAUCU

D.M.*   pAACUCUAAGCGGUGGAUCACUGGGCUCAUGGGUCGAUGAAGAACGCAGCAAACUGUGCGUCAUCGUGUGAACUGCAGGACACAUGAACAUCGACAUUUUG
                                                                                  m


        101    110       120       130       140       150
V.F.    GUCUUUGAACGCAAGUUGCCCCGAUGCCAUUAGGUUGAGGGCACGUCUGCCUGGGUGUCACAU OH
W.E.    GUCUUUGAACGCAAGUUGCGCCCGAGCCACUCGCCGAGGGCACGCCUGCCUGGGCGUCACGC OH
X.B.    UUCGAACGCACCUUGCGGCCCCGGGUUCCUCCCGGGGCUACGCCUGUCUGAGGGUCGCUC OH
                                                                   =
X.L.    UUCGAACGCACCUUGCGGCCCCGGGUUCCUCCCGGGGCCACGCCUGUCUGAGGGUCGCUCC OH
                                                                    =
Y.S.Ce. UUGAACGCACAUUGCGCCCCUUGGUAUUCCAGGGGGCAUGCCUGUUUGAGCGUCAUUU OH
D.M.*   AACGCAUAUCGCAGUCCAUGCUGuuauguacuuuaauuaauuuuauagugcUGCUUGGACUACAUAUGGUUGAGGGUUGUA OH
```

---

The double underlined nucleotides occur in less than one mole per mole 5.8S RNA. Nucleotides written directly under another nucleotide in the sequence indicates that it may also be found in this position. m indicates that the nucleotide is methylated. N.C. sequence has been derived from rDNA.* Nucleotides 1-123 correspond to m 5.8S and nucleotides 152-181 to 2S RNA; nucleotides 124-151 are part of the precursor and removed during processing. For abbreviation of organisms and literature references see opposite page.

## Eukaryotic 5.8S RNA Sequences

| Abbreviation | RNA Source | Reference Number |
|---|---|---|
| V.F. | *Vicia faba* (broad bean) | 95 |
| W.E. | Wheat embryo | 96 |
| X.B. | *Xenopus borealis* (somatic) | 97 |
| X.L. | *Xenopus leavis* (somatic | 84,97 |
| Y.S.Ce. | Yeast (*Saccharomyces cerevisiae* A364A gal-1 ade-1 ade-2 ura-1 his-7 lys-2 try-1 (ATCC 22 244)) | 98 |
| D.M.* | *Drosophila melanogaster* | 99 |

## REFERENCES

1. Erdmann, V.A. (1981) Nucleic Acids Res. 9, r25-r42.
2. Erdmann, V.A. (1978) Nucleic Acids Res. 5, r1-r13.
3. Monier, R. (1974) in Ribosomes, Nomura, M., Tissières, A. and Lengyel, P., Eds., pp. 141-168, Cold Spring Harbor Laboratory, New York.
4. Erdmann, V.A. (1976) in Progress in Nucleic Acids Research and Molecular Biology, Cohn, W.E., Ed., Vol. 18, pp 45-90, Academic Press, New York.
5. Studnicka, G.M., Eiserling, F.A., Lake, J.A. (1981) Nucleic Acids Res. 9, 1885-1904.
6. Böhm, S., Fabian, H., Welfle, H. (1982) Acta biol. med. germ. 41, 1-16.
7. Delihas, N., Andersen, J. (1982) Nucleic Acids Res. , in press.
8. De Wachter, R., Chen, M.-W., Vandenberghe, A. (1982) Biochimie 64, 311-329.
9. Sprinzl, M., Gauss, D.H. (1982) Nucleic Acids Res. 10, r1-r55.
10. Ninio, J. (1979) Biochimie 61, 1133-1150.
11. Pribula, C.D., Fox, G.E., Woese, C.R. (1976) FEBS Lett. 64, 350-352.
12. Marotta, C.A., Varricchio, F., Smith, I., Weissman, S.M., Sogin, M.L., Pace, N.R. (1976) J. Biol. Chem. 251, 3122-3127.
13. Raué, H.A., Stoof, T.J., Planta, R.J. (1975) Eur. J. Biochem. 59, 35-42.
14. Woese, C.R., Luehrsen, K.R., Pribula, C.D., Fox, G.E. (1976) J. Mol. Evol. 8, 143-153.
15. Alexander, L.J., Stewart, T.S. (1980) Nucleic Acids Res. 8, 979-987.
16. Alexander, L.J., Stewart, T.S. (1981) FEBS Lett. 135, 151-154.
17. Walker, R.T., Chelton, E.T.J., Kilpatrick, M.W., Rogers, M.J., Simmons, J. (1982) Nucleic Acids Res. 10, 6363-6367.
18. Hori, H., Sawada, M., Osawa, S., Murao, K., Ishikura, H. (1981) Nucleic Acids Res. 9, 5407-5410.
19. Simoncsits, A. (1980) Nucleic Acids Res. 8, 4111-4124.
20. Hori, H., Osawa, S., Murao, K., Ishikura, H. (1980) Nucleic Acids Res. 8, 5423-5426.
21. MacKay, R.M., Salgado, D., Bonen, L., Stackebrandt, E., Doolittle, W.F. (1982) Nucleic Acids Res. 10, 2963-2970.
22. Kumagai, I., Digweed, M., Erdmann, V.A., Watanabe, K., Oshima, T. (1981) Nucleic Acids Res. 9, 5159-5162.
23. Nazar, R.N., Matheson, A.T. (1977) J. Biol. Chem. 252, 4256-4261.
24. Newhouse, N., Nicoghosian, K., Cedergren, R.J. (1982) Can. J. Biochem. 59, 921-932.
25. DuBuy, B., Weissman, S.M. (1971) J. Biol. Chem. 246, 747-761.
26. Brownlee, G.G., Sanger, F., Barrell, B.G. (1968) J. Mol. Biol. 34, 379-412.
27. Fischel, J.-L., Ebel, J.-P. (1975) Biochimie 57, 899-904.
28. Woese, C.R., Pribula, C.D., Fox, G.E., Zablen, L.B. (1975) J. Mol. Evol. 5, 35-46.
29. Luehrsen, K.R., Fox, G.E. (1981) J. Mol. Evol. 17, 52-55.
30. Corry, M.J., Payne, P.I., Dyer, T.A. (1974) FEBS Lett. 46, 63-66.

31. Takaiwa, F., Sugiura, M. (1982) Nucleic Acids Res. 10, 5369-5373.
32. Dyer, T.A., Bowman, C.M. (1979) Biochem. J. 183, 595-604.
33. Takaiwa, F., Sugiura, M. (1981) Mol. Gen. Genet. 182, 385-389.
34. Delihas, N., Andersen, J., Sprouse, H.M., Dudock, B. (1981) Nucleic Acids Res. 9, 2801-2805.
35. Spencer, D.F., Bonen, L., Gray, M.W. (1981) Biochemistry 20, 4022-4029.
36. Nazar, R., Wildeman, A.T. (1978) J. Biol. Chem. 253, 5464-5469.
37. Fox, G.E., Luehrsen, K.R., Woese, C.R. (1982) Zbl. Bakt. Hyg. 1 Abt. Orig. C3, 330-345.
38. Luehrsen, K.R., Nicholson, D.E., Eubanks, D.C., Fox, G.E. (1981) Nature 293, 755-756.
39. Stahl, D.A., Luehrsen, K.R., Woese, C.R., Pace, N.R. (1981) Nucleic Acids Res. 9, 6129-6137.
40. Luehrsen, K.R., Fox, G.E., Kilpatrick, M.W., Walker, R.T., Domdey, H., Krupp, G., Gross, H.J. (1981) Nucleic Acids Res. 9, 965-970.
41. Forget, B.G., Weissman, S.M. (1969) J. Mol. Biol. 244, 3148-3165.
42. Aoyama, K., Hidaka, S., Tanaka, T., Ishikawa, K. (1982) J. Biochem. 91, 363-367.
43. Federoff, N.V., Brown, R.D. (1977) Cold Spring Harbor Symp. Quant. Biol. 42, 1195-1200.
44. Kay, B.K., Gall, J.G. (1981) Nucleic Acids Res. 9, 6457-6469.
45. Mashkova, T.D., Serenkova, T.I., Mazo, A.M., Avdonina, T.A., Timofeyeva, M.Ya., Kisselev, L.L. (1981) Nucleic Acids Res. 9, 2141-2151.
46. Kawata, Y., Ishikawa, H. (1982) Nucleic Acids Res. 10, 1833-1840.
47. Benhamou, J., Jourdan, R., Jordan, B.R. (1977) J. Mol. Evol. 9, 279-298.
48. Tschudi, C., Pirrotta, V. (1980) Nucleic Acids Res. 8, 441-451.
49. Gu, X.-R., Nicoghosian, K., Cedergren, R.J. (1982) Nucleic Acids Res. 10, 5711-5716.
50. Komiya, H., Kawakami, M., Shimizu, N., Takemura, S. (1980) Nucleic Acids Res. Symposium Series No. 8, S119-S122.
51. Diels, L., De Baere, R., Vandenberghe, A., De Wachter, R. (1981) Nucleic Acids Res. 9, 5141-5144.
52. Lu, A-L., Steege, D.A., Stafford, D.W. (1980) Nucleic Acids Res. 8, 1839-1853.
53. Komiya, H., Shimizu, N., Kawakami, M., Takemura, S. (1980) J. Biochem. 88, 1449-1456.
54. Fang, B.-L., De Baere, R., Vandenberghe, A., De Wachter, R. (1982) Nucleic Acids Res. 10, 4679-4685.
55. Kumazaki, T., Hori, H., Osawa, S. (1982) FEBS Lett. 146, 307-310.
56. Dams, E., Vandenberghe, A., De Wachter, R. (1982) Nucleic Acids Res. 10, 5297-5301.
57. Butler, M.H., Wall, S.M., Luehrsen, K.R., Fox, G.E., Hecht, R.M. (1981) J. Mol. Evol. 18, 18-23.
58. Delihas, N., Andersen, J., Andresini, W., Kaufman, L., Lyman, H. (1981) Nucleic Acids Res. 9, 6627-6633.

59. Kumazaki, T., Hori, H., Osawa, S. (1982) J. Mol. Evol. 18, 293-296.
60. MacKay, R.M., Gray, M.W., Doolittle, W.F. (1980) Nucleic Acids Res. 8, 4911-4917.
61. Luehrsen, K.R., Fox, G.E., Woese, C.R. (1980) Curr. Microbiol. 4, 123-126.
62. Kumazaki, T., Hori, H., Osawa, S., Mita, T., Higashinakagawa, T. (1982) Nucleic Acids Res. 10, 4409-4412.
63. MacKay, R.M., Doolittle, W.F. (1981) Nucleic Acids Res. 9, 3321-3334.
64. Hinnebusch, A.G., Klotz, L.C., Blanken, R.L., Loeblich, A.R. (1981) J. Mol. Evol. 17, 334-347.
65. Komiya, H., Takemura, S. (1981) J. Biochem. 90, 1577-1581.
66. Hori, H., Osawa, S., Iwabuchi, M. (1980) Nucleic Acids Res. 8, 5535-5539.
67. Walker, W.F., Doolittle, W.F. (1982) Nucleic Acids Res. 10, 5717-5721.
68. Piechulla, B., Hahn, U., McLaughlin, L.W., Küntzel, H. (1981) Nucleic Acids Res. 9, 1445-1450.
69. Selker, E.U., Yanofsky, C., Driftmier, K., Metzenberg, R.L., Alzner-DeWeerd, B., RajBhandary, U.L. (1981) Cell 24, 819-828.
70. Nishikawa, K., Takemura, S. (1974) J. Biochem. 76, 935-947.
71. Tabata, S. (1980) Eur. J. Biochem. 110, 107-114.
72. Komiya, H., Miyazaki, M., Takemura, S. (1981) J. Biochem. 89, 1663-1666.
73. Andersen, J., Andresini, W., Delihas, N. (1982) J. Biol. Chem. 257, 9114-9118.
74. MacKay, R.M., Spencer, D.F., Doolittle, W.F., Gray, M.W. (1980) Eur. J. Biochem. 112, 561-576.
75. Delihas, N., Andersen, J., Sprouse, H.M., Kashdan, M., Dudock, B. (1981) J. Biol. Chem. 256, 7515-7517.
76. Payne, P.I., Dyer, T.A. (1976) Eur. J. Biochem. 71, 33-38.
77. Darlix, J.-L., Rochaix, J.-D. (1981) Nucleic Acids Res. 9, 1291-1299.
78. Jordan, B.R., Galling, G., Jourdan, R. (1974) J. Mol. Biol. 87, 205-225.
79. Luehrsen, K.R., Fox, G.E. (1981) Proc. Natl. Acad. Sci. USA 78, 2150-2154.
80. Green, G.A., McCoy, J.M., Jones, D.S. (1982) Nucleic Acids Res. 10, 6389-6392.
81. Takaiwa, F., Kusuda, M., Saga, N., Sugiura, M. (1982) Nucleic Acids Res. 10, 6037-6040.
82. Mackay, R.M., Doolittle, W.F. (1981) Nucleic Acids Res. 9, 3321-3334.
83. Ursi, D., Vandenberghe, A., De Wachter, R. (1982) Nucleic Acids Res. 10, 3517-3530.
84. Khan, N.S.N., Maden, B.E.H. (1977) Nucleic Acids Res. 4, 2495-2505.
85. Hinnebusch, A.G., Klotz, L.C., Blanken, R.I., Loeblich, A.R. (1981) J. Mol. Evol. 17, 334-347.
86. Schnare, M.N., Gray, M.W. (1982) Nucleic Acids Res. 10, 2085-2092
87. Darlix, J.-L., Rochaix, J.-D. (1981) Nucleic Acids Res. 9, 1291-1299.
88. Olsen, G.J., Sogin, M.L. (1982) Biochem. 21, 2335-2343.

89. Nazar, R.N., Sitz, T.O., Busch, H. (1976) Biochem. 15, 505-508.
90. Selker, E., Yanofsky, C. (1979) Nucleic Acids Res. 6. 2461-2567.
91. Nazar, R.N., Sitz, T.O., Busch, H. (1975) J. Biol. Chem. 250, 8591-8597.
92. Nazar, R.N., Roy, K.L. (1978) J. Biol. Chem. 253, 395-399.
93. Schaak, J., Mao, J., Söll, D. (1982) Nucleic Acids Res. 9, 2851-2864.
94. Nazar, R.N., Roy, K.L. (1976) FEBS Lett. 72, 111-116.
95. Tanaka, Y., Dyer, T.A., Brownlee, G.G. (1980) Nucleic Acids Res. 8, 1259-1272.
96. Barber, C., Nichols, J.L. (1978) Can. J. Biochem. 56, 357-364.
97. Ford, P.Y., Mathieson, T. (1978) Eur. J. Biochem. 87, 199-214.
98. Rubin, G.M. (1973) J. Biol. Chem. 248, 3860-3875.
99. Parlakis, G.N., Jordan, B.R., Wurst, R.M., Vournakis, J.N. (1979) Nucleic Acids Res. 7, 2213-2238.