Nucleic Acids Research, 1993, Vol. 21, No. 13 3055–3074

# A compilation of large subunit (23S and 23S-like) ribosomal RNA structures: 1993

Robin R.Gutell*, Michael W.Gray[1]* and Murray N.Schnare[1]
Molecular, Cellular, and Developmental Biology, Campus Box 347, University of Colorado, Boulder, CO 80309-0347, USA and [1]Department of Biochemistry, Dalhousie University, Halifax, Nova Scotia B3H 4H7, Canada

## INTRODUCTION

This compilation is part of an on-going effort to maintain a comprehensive and continually updated collection of large subunit (LSU; 23S and 23S-like) rRNA secondary structures and associated sequence and citation information. Table 1 gives a breakdown of the number and phylogenetic distribution of sequences currently in this LSU rRNA database. A listing of accession numbers for all complete LSU rRNA sequences now in the public domain is presented in Table 2. The reference list is an update of the information provided in the 1992 compilation [1]; here, we include only revised or new citations to LSU rRNA sequences. The complete bibliography of LSU rRNA sequences is available electronically, as is the complete listing of accession numbers presented in Table 1 (see below for instructions on how to obtain this information).

## GROWTH OF THE DATABASE

The past year has seen the largest annual increase in the number of LSU rRNA sequences published and/or released through the EMBL, GenBank and DDBJ databanks (Table 1). This trend has also been apparent in each of the previous years we have monitored, emphasizing the steady increase in the rate at which LSU rRNA sequences are being determined. The rate of growth within each of the phylogenetic categories does vary, however. In the period covered by the present compilation, more than half (34/66) of the new sequences that appeared were (eu)bacterial ones. The second largest increase occurred in the Plastids category, due primarily to the release of 15 new *Chlamydomonas* chloroplast sequences. In the previous year, the greatest increases were in mitochondrial and eucaryotic (nuclear) sequences. Mitochondrial sequences still constitute the largest single category in the LSU rRNA database, as they have in every previous compilation.

## MODELS OF HIGHER-ORDER STRUCTURE

As in previous compilations [1–3], each LSU rRNA primary sequence is presented in a two-dimensional format (the secondary structure) that underpins the biologically relevant conformation of an rRNA molecule. When LSU rRNA sequences are configured in this manner, phylogenetic information becomes readily apparent, as do patterns of variability and conservation in sequence and structure. At the same time, the secondary structure serves as an effective template for relating form to function. Principles for deducing higher-order structure have been enunciated in previous compilations [1–3] (see also [4–6]).

Three phylogenetically and structurally distinct examples of current LSU rRNA secondary structures are presented in this communication. They are: (i) *Escherichia coli*, a typical (eu)bacterial structure (our reference standard, to which other LSU rRNA structures are compared and against which they are modelled); (ii) *Saccharomyces cerevisiae* (yeast), a representative eucaryotic (nucleocytoplasmic) structure; and (iii) a 'minimalist' structure, typified by the unusually small mitochondrial LSU rRNA from the nematode worm, *Caenorhabditis elegans*. While the latter two structures exemplify the type of size variation found among LSU rRNAs, an even smaller and quite peculiar structure has been described for the mitochondrial LSU rRNA of trypanosomatid protozoa (*Trypanosoma*, *Leishmania*, *Crithidia*). At the other extreme, LSU rRNA structures larger than the *S.cerevisiae* nuclear one can be found; these are primarily mammalian nucleocytoplasmic ones, which contain large insertions at the positions of some of the variable regions of their *S.cerevisiae* counterpart.

Due to the extensive sequence and size differences within some of the mitochondrial and eucaryotic variable regions, and the lack of a sufficient number of phylogenetically close examples from which to deduce significant primary and secondary structure homology within these variable regions, some of them had been left unstructured in previous compilations. However, with the recent large increase in the number of available LSU rRNA sequences, we have now deduced sequence and structural homology for many of these variable regions. Because we have a high degree of confidence in these newly solved structures, structural variation and evolution in these mitochondrial and nucleocytoplasmic LSU rRNAs can now be examined critically and in depth. The results of these analyses will be published in due course, along with other LSU rRNA structural characteristics that distinguish the different phylogenetic assemblages.

At the same time that they permit the elucidation of secondary structure pairings, comparative methods allow one to infer tertiary interactions. The three LSU rRNAs presented in this compendium display all of the currently known secondary and tertiary structure constraints. A small but growing number of these newer and more

---

* To whom correspondence should be addressed

complex structural elements have been experimentally verified [4,5], thereby justifying the application of these comparative methods. New and more powerful correlation analysis algorithms are now being developed [6,7; R.R. Gutell, unpublished] and promise to extend our appreciation of the structural (and other) constraints acting on these LSU rRNAs.

In the immediate future we anticipate a continuing steady increase in the number of available LSU rRNA sequences (and thus their structures), in parallel with additional refinements in each of these structures. The information provided should lead us to higher-resolution structures and new insights into how they evolved.

## ACCURACY OF THE DATA

As noted previously [1], independently determined versions of the same primary sequence exist for several LSU rRNAs (see Table 2 and the List of References). Usually, these alternative versions differ from one another at a number of positions, and at least some of these differences are likely to be sequencing errors. Often, the secondary structure predicts which version of a sequence is likely to be correct at a particular position. On the other hand, it is probable that some of the variation actually reflects genuine inter- or intra-strain sequence heterogeneity, particularly if the differences occur within variable regions [8]. We have also noted that a few published primary sequences differ at one or more positions from their GenBank/EMBL/DDBJ listings, without an indication that the database entry represents a subsequently revised version of the original published sequence. Again, secondary structure modelling may or may not indicate which version is likely to be correct. This year, the List of References contains brief [NOTES] on individual sequences, where discrepancies between different versions of the same sequence or between published and database entries of the same sequence are noted. In cases of non-identical versions of the same sequence, the reader is referred to these notes for information on which sequence was used for the secondary structures we present.

## AVAILABILITY AND STATUS OF STRUCTURE FIGURES

As in [1] and [2], the actual 23S and 23S-like rRNA secondary structures are not published here, except for the three examples shown. The comprehensive set of LSU rRNA secondary structures may be obtained in one of two ways. **Hardcopy printouts of this set are available directly from us (inquiries should be referred to M.W.G. at the address listed below).** Release 1.0, which comprised 88 structures, was made available for distribution in July, 1991, and copies may still be obtained. Newly modeled and refined LSU rRNA secondary structures are included in Release 2.0 (January, 1993), which contains 51 structures (13 archaebacterial, 18 eubacterial, 13 plastid, 3 mitochondrial and 4 eucaryotic nuclear). Those structures available in Release 1.0 and 2.0 are noted in Table 2. The next release (2.1) will comprise considerably revised and updated eucaryotic nuclear structures, whereas subsequent releases will contain revised mitochondrial structures (Release 2.2) and models of new prokaryotic (archaebacterial and eubacterial) and plastid sequences (Release 2.3). **Requests for hardcopy printouts, which will be sent out as soon as they become available, should state explicitly which Release (1.0, 2.1–2.3) is being sought. Anyone who is on our current mailing list should already have received Release 2.0 and will automatically receive future releases.**

Alternatively, individuals with access to the Internet telecommunications network and a laser printer capable of processing PostScript™ files may elect to obtain files of LSU rRNA secondary structures by anonymous file transfer protocol (ftp). **These files are deposited with the Ribosomal RNA Database Project** [9] on the RDP computer at the Argonne National Laboratory, and may be accessed as indicated below. Inquiries concerning this on-line service may be directed to R.R.G. (e-mail and postal addresses as noted below).

As time and facilities permit, the rRNA information available on-line will be increased. Currently it includes the set of LSU rRNA secondary structures (in PostScript™ format) that are available in hard copy in Release 2.0, the table of LSU rRNA sequences and their accession numbers, and the associated publication reference list. Refinements to existing secondary structures as well as newly modeled secondary structures will be released on-line as soon as we have completed our own analysis. From time to time, we will also update the associated table and reference list. Currently we include only those LSU rRNA sequences that are complete (or nearly so).

Finally, we invite comments from readers, and welcome suggested revisions to and/or alternative interpretations of our proposed secondary structures, as well as suggestions for improvement in the content and/or form of the database. We also solicit newly determined LSU rRNA sequences in advance of publication, in order to accelerate their inclusion in the compendium.

## ACKNOWLEDGEMENTS

We gratefully acknowledge the computer programming expertise of Bryn Weiser and Tom Macke in this endeavor. Work on this compendium has been supported by an operating grant (MT-11212) from the Medical Research Council of Canada to M.W.G. and by NIH grant GM48207 to R.R.G., who also acknowledges a generous donation of computer equipment from SUN Microsystems. We thank the W.M. Keck Foundation for their generous support of RNA science on the Boulder campus, and the Canadian Institute for Advanced Research (CIAR) for providing funds in support of the production and distribution of secondary structure figures. M.W.G. is a Fellow and R.R.G. is an Associate in the Program in Evolutionary Biology of the CIAR.

## REFERENCES

1. Gutell,R.R., Schnare,M.N. and Gray,M.W. (1992) *Nucleic Acids Res.* **20**, **Supplement**, 2095–2109.
2. Gutell,R.R., Schnare,M.N. and Gray,M.W. (1990) *Nucleic Acids Res.* **18**, **Supplement**, 2319–2330.
3. Gutell,R.R. and Fox,G.E. (1988) *Nucleic Acids Res.* **16**, **Supplement**, r175–r269.
4. Gutell,R.R. In Nierhaus,K.H., Subramanian,A.R., Erdmann,V.A., Franceschi,F., and Wittman-Liebold,B. (eds.). *The Translational Apparatus*. Plenum, in press.
5. Gutell,R.R., Larsen,N. and Woese,C.R. In Zimmermann,R.A. and Dahlberg,A.E. (eds.). *Ribosomal RNA: Structure, Evolution, Gene Expression and Function in Protein Synthesis*. CRC Press, Boca Raton, FL, in press.
6. Gutell,R.R., Power,A., Hertz,G.Z., Putz,E.J. and Stormo,G.D. (1992) *Nucleic Acids Res.* **20**, 5785–57–95.
7. Gutell,R.R. (1993) *Curr. Opin. Struct. Biol.* **3**, in press.

8. Gonzalez,I.L., Gorski,J.L., Campen,T.J., Dorney,D.J., Erickson,J.M., Sylvester,J.E. and Schmickel,R.D. (1985) *Proc. Natl. Acad. Sci. U.S.A.* **82**, 7666–7670.
9. Olsen,G.J., Overbeek,R., Larsen,N., Marsh,T.L., McCaughey,M.J., Maciukenas,M.A., Kuan,W.-M., Macke,T.J., Xing,Y. and Woese,C.R. (1992) *Nucleic Acids Res.* **20**, **Supplement**, 2199–2200.

## REQUESTS

**For hard copies of LSU rRNA secondary structures, either the complete compilation or selected portions thereof (please specify Release 1.0, 2.0, 2.1, 2.2 or 2.3):**

M.W. Gray,
Department of Biochemistry,
Sir Charles Tupper Medical Building,
Dalhousie University,
Halifax, Nova Scotia B3H 4H7, Canada.
FAX: (902)494-1355
e-mail: mgray@ac.dal.ca (via Bitnet)

**For information about on-line availability of LSU rRNA secondary structure files and other information contained in this compendium:**

R.R. Gutell,
Molecular, Cellular, and Developmental Biology,
Campus Box 347,
University of Colorado,
Boulder, Colorado 80309-0347, USA.
FAX: (303)492-7744
e-mail: rgutell@boulder.colorado.edu

**To obtain PostScript™ files of LSU rRNA secondary structures via anonymous ftp on the Internet telecommunications network:**

ftp site: info.mcs.anl.gov
directory: /pub/RDP/LSU_rRNA/sec_struct

The present publication should be quoted as reference for secondary structure data obtained either as hard copies from the authors or from the electronically accessible compilation.

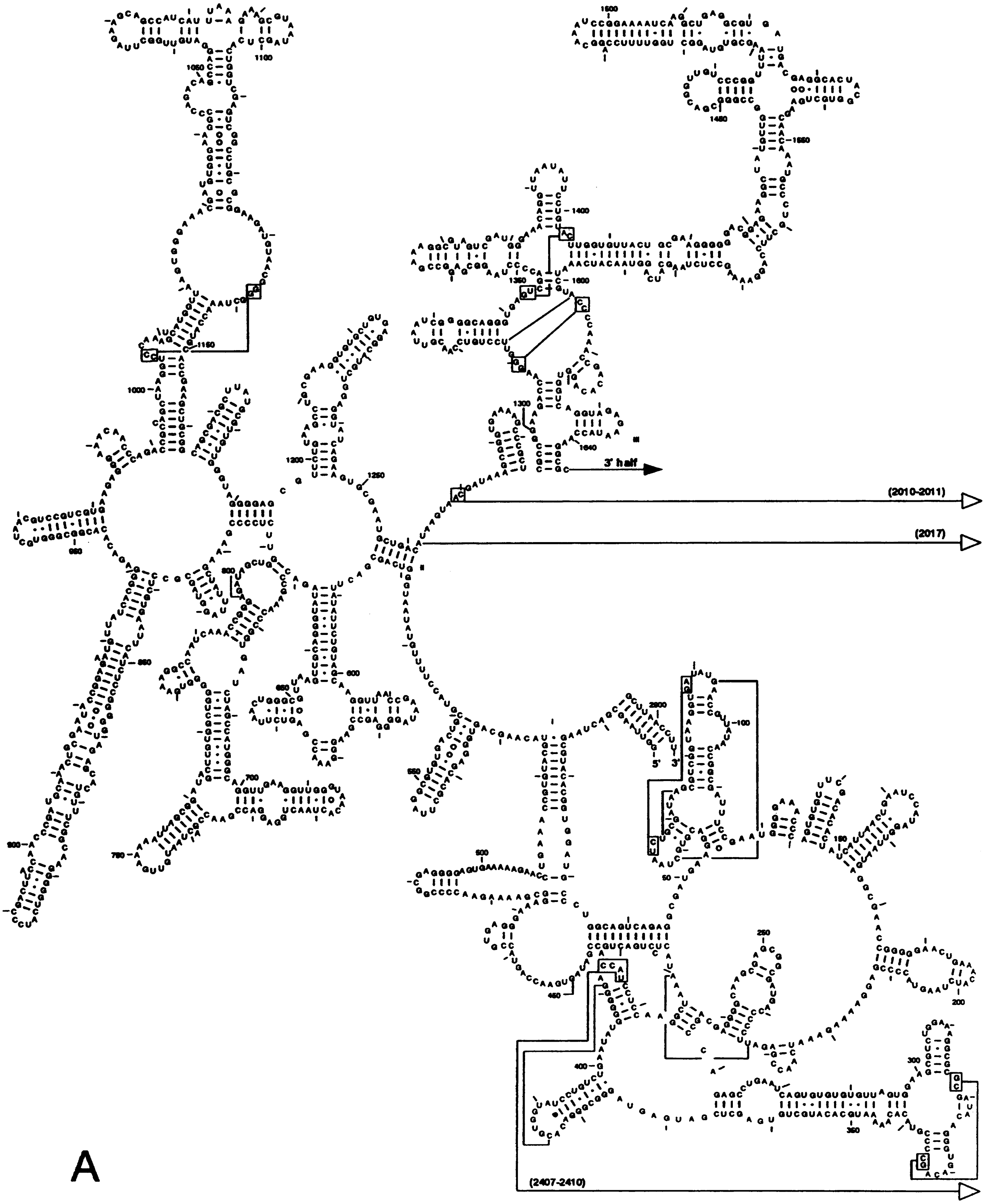
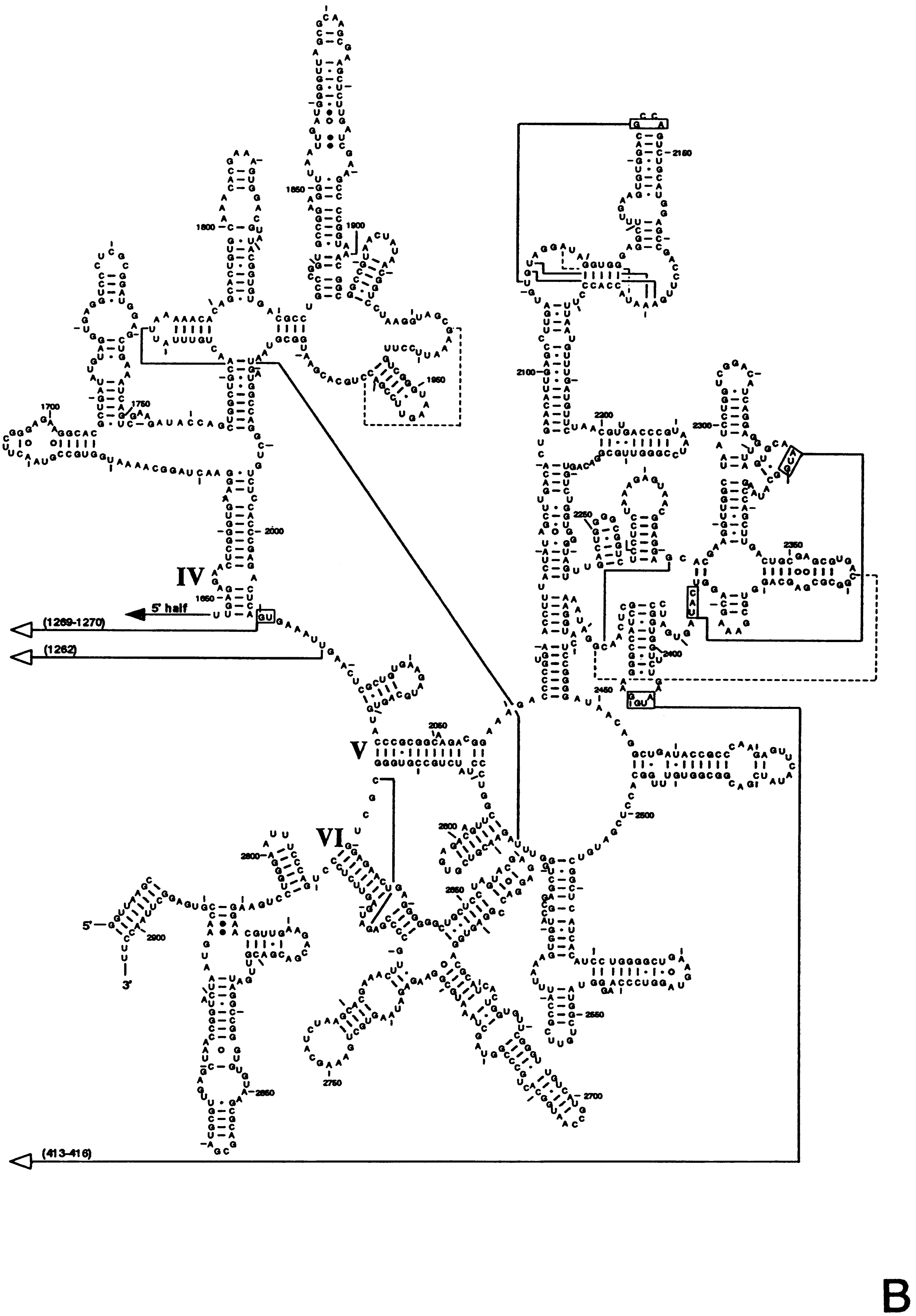


**Figure 1.** Higher-order structure model for a typical (eu)bacterial LSU rRNA (*Escherichia coli* 23S; accession no. J01695). **A**, 5′-half; **B**, 3′-half. C:G and U:A base pairs are connected by lines, G:U pairs by dots, A:G pairs by open circles, and other, non-canonical pairings by closed circles. Proposed tertiary interactions connected with thicker, continuous lines are characterized by a sufficient number of compensatory base changes to make their identification firm; more tentative asssignments are connected by thinner, dashed lines. Every 10th position in the sequence is indicated by a tick mark, while every 50th position is numbered.


IV
V
VI
5' half
(1269-1270)
(1262)
(413-416)
5'
3'
1650
1700
1750
1800
1850
1900
1950
2000
2050
2100
2150
2200
2250
2300
2350
2400
2450
2500
2550
2600
2650
2700
2750
2800
2850
2900
B

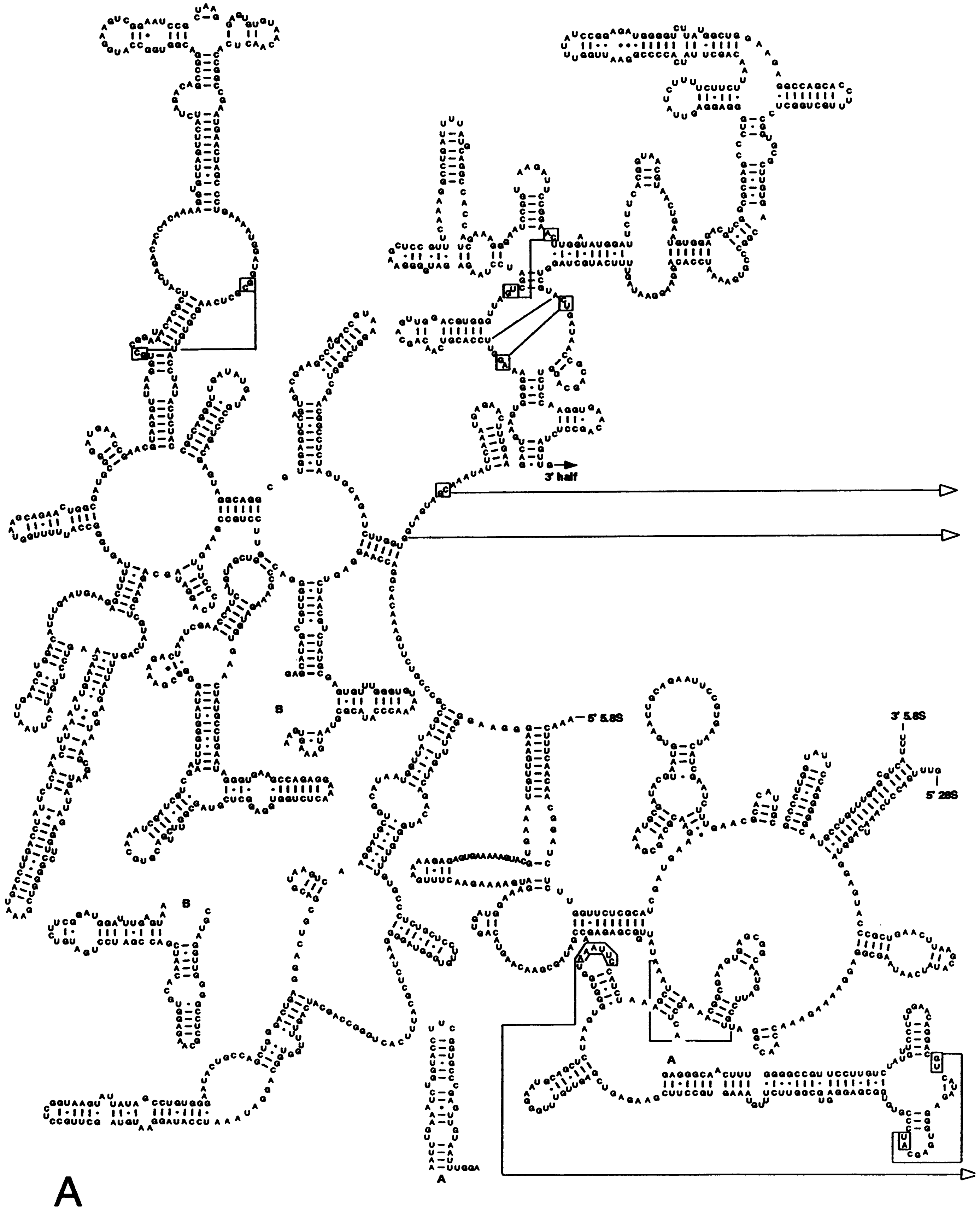
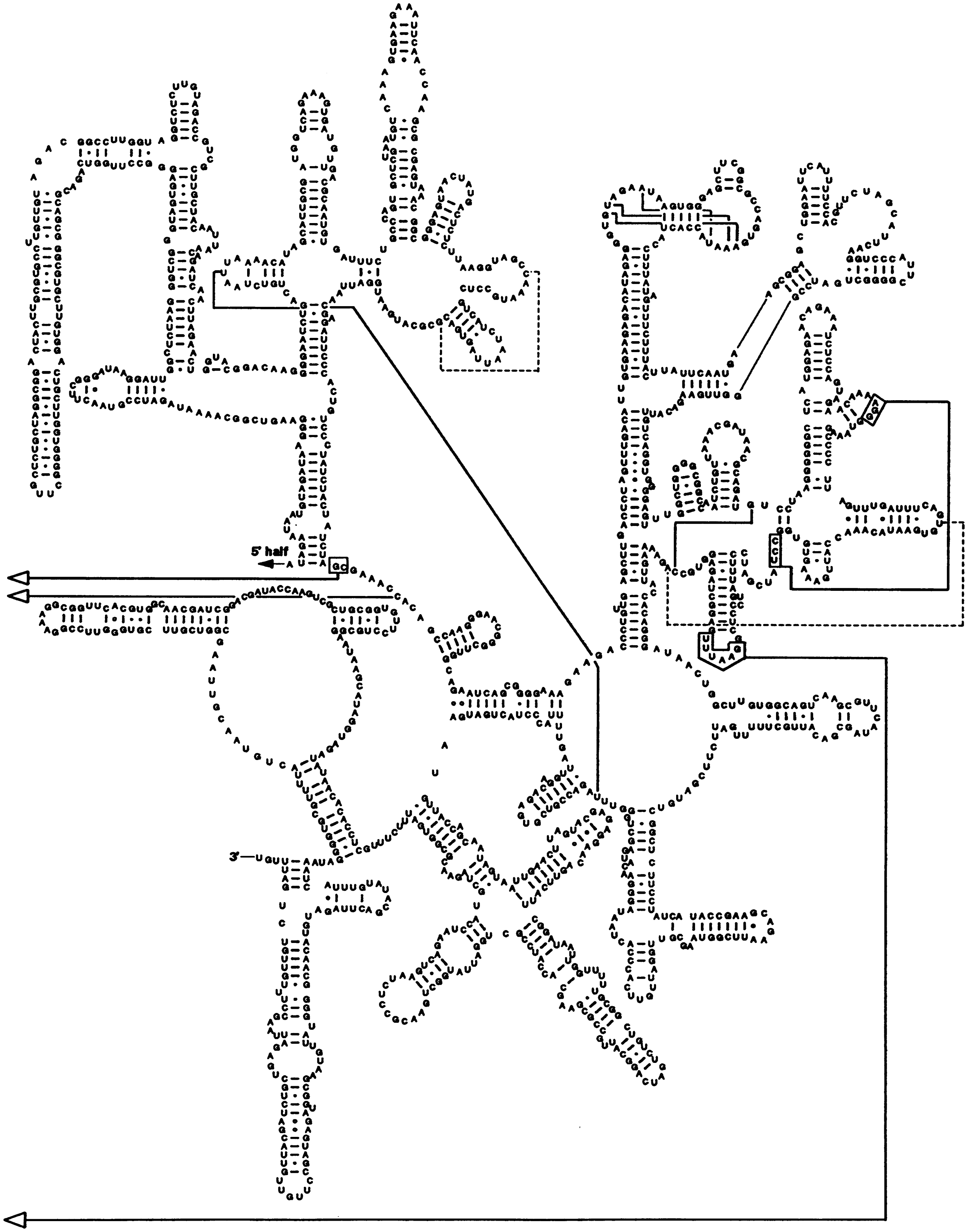


**Figure 2.** Higher-order structure model for a representative eucaryotic nucleocytoplasmic LSU rRNA (*Saccharomyces cerevisiae* 5.8S + 26S; accession nos. J01355, K01048). **A**, 5'-half; **B**, 3'-half. Secondary and tertiary interactions are denoted as in Fig. 1.



B

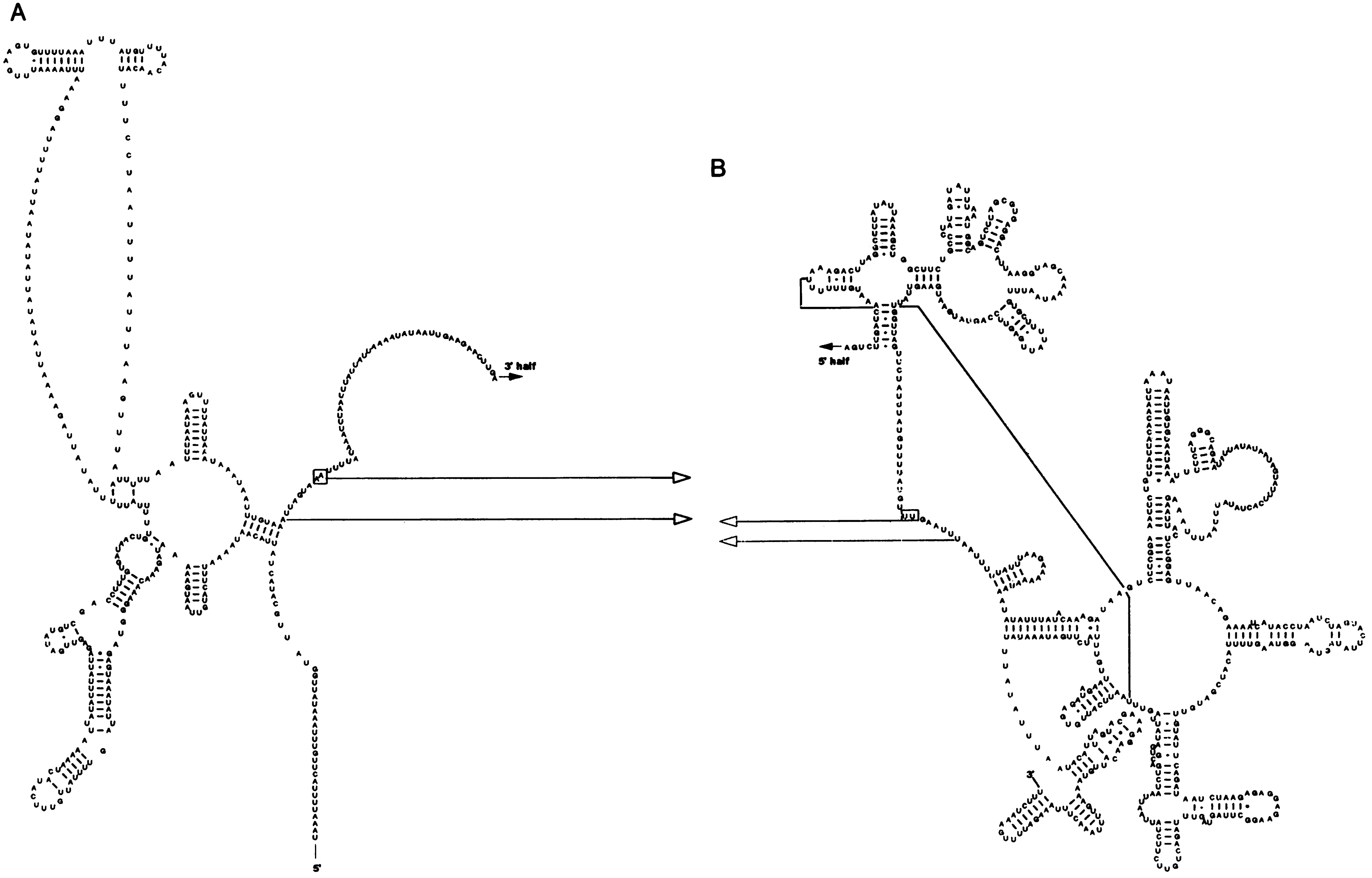


**Figure 3.** Higher-order structure model for the highly reduced mitochondrial LSU rRNA of the nematode worm, *Caenorhabditis elegans* (accession no. X54252). **A**, 5'-half; **B**, 3'-half. Secondary and tertiary interactions are denoted as in Fig. 1.

**Table 1.** Growth of the LSU rRNA Database

| | 1988 | 1990 | 1992 | 1993 |
|---|---|---|---|---|
| Archaea | 6 | 7 | 12 | 13 |
| Bacteria | 5 | 13 | 21 | 55 |
| Plastid | 4 | 8 | 13 | 31 |
| Mitochondria | 17 | 27 | 54 | 61 |
| Eucarya | 10 | 16 | 28 | 34 |
| **All** | **42** | **71** | **128** | **194** |

**Table 2.** List of complete LSU rRNA sequences

| | Organism name[1] | Status[2] | Accession Number |
|---|---|---|---|
| **ARCHAEA**[3] (ARCHAEBACTERIA) | | | |
| **Euryarchaeota**[3] | | | |
| | *Archaeoglobus fulgidus* | 2.0 | M64487 |
| | *Halobacterium halobium* | 2.0 | X00872;X03407 |
| | *Halobacterium marismortui* | 2.0 | X13738 |
| | *Halococcus morrhuae* | 2.0 | X05481 |
| | *Methanobacterium thermoautotrophicum* | 2.0 | X15364 |
| | *Methanococcus vannielii* | 2.0 | X02729 |
| | *Methanospirillum hungatei* | 2.0 | M61738;M81323 |
| | *Thermococcus celer* | 2.0 * | M67497 |
| | *Thermoplasma acidophilum* | 2.0 | M20822;M32298 |
| **Crenarchaeota**[3] | | | |
| | *Desulfurococcus mobilis* | 2.0 | X05480 |
| | *Sulfolobus solfataricus* | 2.0 | M67495 |
| | *Thermofilum pendens* | 2.0 | X14835 |
| | *Thermoproteus tenax* | 2.0 | |
| **BACTERIA**[3] (EUBACTERIA) | | | |
| **Thermotogales**[4] | | | |
| | *Thermotoga maritima* | 2.0 * | M67498 |
| **Radio-resistant micrococci and relatives**[4] | | | |
| | *Thermus thermophilus* | 2.0 | X12612 |
| **Spirochetes and relatives**[4] | | | |
| | *Borrelia burgdorferi* | * | M88330 |
| | *Leptospira interrogans* | 2.0 | X14249 |
| **Planctomyces and relatives**[4] | | | |
| | *Pirellula marina* | 1.0 | X07408 |
| **Green sulfur bacteria**[4] | | | |
| | *Chlorobium limicola* | 1.0 | M31904;M62805 |
| **Flexibacter-cytophaga-bacteroides phylum**[4] | | | |
| | *Flavobacterium odoratum* | 1.0 | M32365;M62807 |
| | *Flexibacter flexilis* | 1.0 | M31904;M62806 |
| **Purple bacteria**[4] | | | |
| Alpha subdivision | | | |
| | *Rhodobacter capsulatus* | 2.0 | X06485 |
| | *Rhodobacter sphaeroides* | 2.0 | X53853–55 |
| Beta subdivision | | | |
| | *Bordetella avium* | * | X70370 |
| | *Bordetella bronchiseptica* | * | X70371 |
| | *Bordetella parapertussis* | * | X68368 |
| | *Bordetella pertussis* | * | X68323 |
| | *Neisseria gonorrhoeae* | * | X67293 |
| | *Neisseria meningitidis* | * | X67300 |
| | *Pseudomonas cepacia* | 2.0 | X16368 |
| Gamma subdivision | | | |
| | *Aeromonas hydrophila* | * | X67946 |
| | *Escherichia coli* | 2.0 | J01695 |
| | *Plesiomonas shigelloides* | 2.0 * | X65487 |
| | *Pseudomonas aeruginosa* | 2.0 | Y00432 |
| | *Pseudomonas perfectomarina* | * | L03788 |
| | *Ruminobacter amylophilus* | 1.0 | X06765 |
| **Cyanobacteria**[4] | | | |
| | *Synechococcus* sp. 6301 (*Anacystis nidulans*) | 2.0 | X00343;X00512 |
| **Gram-positive phylum**[4] | | | |
| High G+C subdivision | | | |
| | *Frankia* sp. #1 | 2.0 | M55343 |
| | *Frankia* sp. #2 | * | M88466 |

**Table 2.** *continued*

| | Organism name[1] | Status[2] | Accession Number |
|---|---|---|---|
| | *Micrococcus luteus* | 2.0 | X06484 |
| | *Mycobacterium kansasii* | * | Z17212 |
| | *Mycobacterium leprae* | 2.0 | X56657 |
| | *Streptomyces ambofaciens* | 2.0 | M27245 |
| | *Streptomyces griseus* | 2.0 | M76388;X55435;X61478 |
| Clostridial subdivision | | | |
| | *Clostridium botulinum* | * | X65602 |
| | *Clostridium tyrobutyricum* | * | L08062 |
| | *Pectinatus frisingensis* | * | X68423 |
| | *Peptococcus niger* | * | X68428 |
| Mycoplasmas and relatives | | | |
| | *Mycoplasma hyopneumoniae* | * | X68421 |
| | *Mycoplasma pneumoniae* | * | X68422 |
| Lactobacillus-bacillus-streptococcus phylum | | | |
| | *Bacillus* sp. (strain PS3) | 2.0 | X60981 |
| | *Bacillus alcalophilus* | * | D11459 |
| | *Bacillus anthracis* | * | X64645 |
| | *Bacillus cereus* | * | S43429 |
| | *Bacillus globisporus* | * | X68424 |
| | *Bacillus licheniformis* | * | X68433 |
| | *Bacillus stearothermophilus* | 2.0 # | K02663;S43424 |
| | *Bacillus subtilis* (*rrnB*) | 2.0 | K00637;M10606;X00007 |
| | *Bacillus subtilis* (*rrnI*) | D11460 | |
| | *Lactobacillus delbrueckii* | * | X68426 |
| | *Lactococcus lactis* subsp. *cremoris* | * | X68434 |
| | *Lactococcus lactis* subsp. *lactis* (strain IL1403) | * | X64887 |
| | *Lactococcus lactis* subsp. *lactis* (strain DSM 20481$^{T}$) | * | X68435 |
| | *Listeria monocytogenes* | * | S46732;X64533;X68420 |
| | *Staphylococcus aureus* | * | X68425 |
| | *Staphylococcus carnosus* | * | X68419 |
| | *Streptococcus oralis* | * | X68427 |
| | *Streptococcus thermophilus* | * | X68429 |
| **PLASTIDS** | | | |
| **Protoctista**[5] | | | |
| Apicomplexa[6] | | | |
| | *Plasmodium falciparum*[6] | | X61660 |
| Chlorophyta (green algae) | | | |
| | *Chlamydomonas* 66.72 | * | X68923–26 |
| | *Chlamydomonas eugametos* | 2.0 # | Z17234 |
| | *Chlamydomonas frankii* | * | X68905–09 |
| | *Chlamydomonas gelatinosa* | * | Z15151 |
| | *Chlamydomonas geitlerii* | * | X68891–92 |
| | *Chlamydomonas humicola* | * | X68921–22 |
| | *Chlamydomonas indica* | * | X68893–98 |
| | *Chlamydomonas iyengarii* | * | X68885–86 |
| | *Chlamydomonas komma* | * | X68927–29 |
| | *Chlamydomonas mexicana* | * | X68910–12 |
| | *Chlamydomonas moewusii* | * | X68913–18 |
| | *Chlamydomonas pallidostigmatica* | * | X68899–904 |
| | *Chlamydomonas peterfii* | * | X68887–88 |
| | *Chlamydomonas pitschmanni* | * | Z15152 |
| | *Chlamydomonas reinhardtii* | 2.0 | X15727;X16686–87 |
| | *Chlamydomonas starii* | * | X68889–90 |
| | *Chlamydomonas zebra* | * | X68919–20 |
| | *Chlorella ellipsoidea* | 2.0 | M36158 |
| Euglenophyta (euglenoid flagellates) | | | |
| | *Astasia longa* | 2.0 | X14386 |
| | *Euglena gracilis* | 2.0 # | X12890 |
| Phaeophyta (brown algae) | | | |
| | *Pylaiella littoralis* | * | X61179 |
| Rhodophyta (red algae) | | | |
| | *Palmaria palmata* | * | Z18289 |
| **Plantae** | | | |
| Angiospermophyta (flowering plants) | | | |
| | *Alnus incana* (alder) | 2.0 | M75719;M75722;M76448 |
| | *Conopholis americana* | 2.0 | X59768 |
| | *Epifagus virgiana* (beechdrops) | 2.0 * | M81884;X62099 |
| | *Nicotiana tabacum* (tobacco) | 2.0 | J01446;Z00044 |

| | Organism name[1] | Status[2] | Accession Number |
|---|---|---|---|
| | *Oryza sativa* (rice) | 2.0 | X15901 |
| | *Pisum sativum* (pea) | 2.0 | M37430 |
| | *Zea mays* (maize) | 2.0 | X01365 |
| Bryophyta | | | |
| | *Marchantia polymorpha* (liverwort) | 2.0 | X01647;X04465 |
| **MITOCHONDRIA** | | | |
| **Protoctista**[5] | | | |
| Apicomplexa | | | |
| | *Plasmodium falciparum* | * | M76611;M99416; S87790,/792,/801,/807,/809–10, /833–34,/836–40 |
| | *Plasmodium yoelii* | | M29000 |
| Chlorophyta (unicellular green algae) | | | |
| | *Chlamydomonas reinhardtii* | 1.0 | M22649;M25123–25130;X54860 |
| Ciliophora (ciliates) | | | |
| | *Paramecium primaurelia* | 1.0 | K00634 |
| | *Paramecium tetraurelia* | 1.0 | K01749;X15917 |
| | *Tetrahymena pyriformis* | 1.0 | M58010;M58011 |
| Zoomastigina (zooflagellates) | | | |
| | *Crithidia fasciculata* | 1.0 | X02548 |
| | *Crithidia oncopelti* | 1.0 | X51736 |
| | *Leishmania tarentolae* | 1.0 | X02354 |
| | *Leptomonas* sp. | 1.0 | J03814 |
| | *Trypanosoma brucei* | 1.0 | X02547 |
| **Plantae** | | | |
| Angiospermophyta (flowering plants) | | | |
| | *Oenothera berteriana* (primrose) | 1.0 | X02559 |
| | *Triticum aestivum* (wheat) | 1.0 # | M37274;Z11889 |
| | *Zea mays* (maize) | 2.0 | K01868 |
| Bryophyta | | | |
| | *Marchantia polymorpha* (liverwort) | | M68929 |
| **Fungi** | | | |
| Ascomycota | | | |
| | *Aspergillus nidulans* | 1.0 | X06961 |
| | *Neurospora crassa* | 1.0 | X55443 |
| | *Penicillium chrysogenum* | * | D13859 |
| | *Podospora anserina* | 1.0 | M30937;X14735 |
| | *Saccharomyces cerevisiae* | 1.0 | J01527 |
| | *Schizosaccharomyces pombe* | 1.0 | X06597 |
| **Animalia** | | | |
| Arthropoda | | | |
| | *Aedes albopictus* (mosquito) | 1.0 | X01078 |
| | *Apis mellifera* (honeybee) | 1.0 # | L06178;X05011 |
| | *Artemia salina* (brine shrimp) | 1.0 | M21833;X12965 |
| | *Drosophila melanogaster* (fruit fly) | 1.0 | X53506 |
| | *Drosophila yakuba* (fruit fly) | 1.0 | X03240 |
| | *Locusta migratoria* (locust) | 1.0 | X05287 |
| | *Spodoptera frugiperda* | | M76713 |
| Chordata | | | |
| | *Aepyceros melampus* | | M86496 |
| | *Antilocapra americana* | 1.0 | M55540 |
| | *Balaenoptera physalus* (fin whale) | | X61145 |
| | *Bos taurus* (ox) | 1.0 | J01394 |
| | *Boselaphus tragocamelus* | | M86494 |
| | *Capra hircus* (goat) | 1.0 | M55541 |
| | *Carassius auratus* | | |
| | *Cephalophus maxwelli* | | M86498 |
| | *Cervus unicolor* (sambar deer) | 1.0 | M35875 |
| | *Crossostoma lacustre* (freshwater loach) | * | M91245 |
| | *Cyprinus carpio* (carp) | | X61010 |
| | *Damaliscus dorcas* | | M86499 |
| | *Gallus gallus domesticus* (chicken) | 1.0 | X52392 |
| | *Gazella thomsoni* | | M86501 |
| | *Homo sapiens* (human) | 1.0 | J01415;M12548;V00710 |
| | *Hydropotes inermis* (Chinese water deer) | 1.0 | M35876 |
| | *Kobus ellipsiprymnus* | | M86497 |
| | *Madoqua kirki* | | M86495 |
| | *Muntiacus reevesi* (Reeves' muntjac) | 1.0 | M35877 |
| | *Mus musculus* (mouse) | 1.0 | J01420 |
| | *Mytilus edulis* (blue mussel) | * | M83756 |

**Table 2.** *continued*

| | Organism name[1] | Status[2] | Accession Number |
|---|---|---|---|
| | *Odocoileus virginianus* (white-tailed deer) | 1.0 | M35874 |
| | *Oryx gazella* | | M86500 |
| | *Phoca vitulina* (harbor seal) | * | S37044;X63726 |
| | *Rana catesbeiana* (frog) | 1.0 | X12841 |
| | *Rattus norvegicus* (rat) | 1.0 | J01438 |
| | *Tragelaphus imberbis* | | M86493 |
| | *Tragulus napu* | 1.0 | M55539 |
| | *Xenopus laevis* (toad) | 1.0 | M10217 |
| Echinodermata (sea urchins) | | | |
| | *Paracentrotus lividus* | 1.0 | J04815 |
| | *Strongylocentrotus purpuratus* | 1.0 | X12631 |
| Nematoda | | | |
| | *Ascaris suum* | 2.0 | X54253 |
| | *Caenorhabditis elegans* | 2.0 | X54252 |
| **EUCARYA**[3] (EUCARYOTES, NUCLEOCYTOPLASMIC) | | | |
| **Archezoa**[7] | | | |
| | *Giardia ardeae* | 2.0 | X58290 |
| | *Giardia intestinalis* | 2.0 | X52949 |
| | *Giardia muris* | 2.0 * | X65063 |
| **Protoctista**[5] | | | |
| Acrasiomycota (cellular slime molds) | | | |
| | *Dictyostelium discoideum* | | X00601 |
| Ciliophora (ciliates) | | | |
| | *Tetrahymena pyriformis* | | M10752;X54004 |
| | *Tetrahymena thermophila* | 1.0 | X54512 |
| Dinoflagellata (dinoflagellates) | | | |
| | *Prorocentrum micans* | | X15973;X16108 |
| Euglenophyta (euglenoid flagellates) | | | |
| | *Euglena gracilis* | 1.0 | X53361 |
| Myxomycota (plasmodial slime molds) | | | |
| | *Didymium iridis* | | X60210 |
| | *Physarum polycephalum* | 1.0 | V01159 |
| Rhizopoda (amastigote amoebas) | | | |
| | *Entamoeba histolytica* | * | X65163 |
| Zoomastigina (zooflagellates) | | | |
| | *Crithidia fasciculata* | 1.0 | Y00055 |
| | *Trypanosoma brucei* | | X14553 |
| **Fungi** | | | |
| Ascomycota | | | |
| | *Candida albicans* | * | L07796;X70659 |
| | *Pneumocystis carinii*[8] | * | S41246;M86760 |
| | *Saccharomyces carlsbergensis* | | J01352;V01285;V01325 |
| | *Saccharomyces cerevisiae* | 2.0 | J01355;K01048 |
| | *Schizosaccharomyces pombe* | * | J01359;V01361;Z19136 |
| Zygomycota | | | |
| | *Mucor racemosus* | | M26190 |
| **Plantae** | | | |
| Angiospermophyta (flowering plants) | | | |
| | *Arabidopsis thaliana* | | X52320 |
| | *Brassica napus* (rapeseed) | * | D10840;S39880 |
| | *Citrus limon* (lemon) | | X05910 |
| | *Fragaria ananassa* (strawberry) | | X15589;X58118 |
| | *Lycopersicon esculentum* (tomato) | | X07889;X13557 |
| | *Oryza sativa* (rice) | 1.0 | M11585 |
| | *Sinapis alba* (mustard) | | X15915;X57137 |
| **Animalia** | | | |
| Arthropoda | | | |
| | *Drosophila melanogaster* | 1.0 | M21017 |
| Chordata | | | |
| | *Herdmania momus* | | X53538 |
| | *Homo sapiens* (human) | 1.0 | J01866;M11167 |
| | *Mus musculus* (mouse) | 1.0 | J01871;X00525 |
| | *Rattus norvegicus* (rat) | 1.0 | J01881;K01591;X00521;X01069 |
| | *Xenopus borealis* (toad) | | X59733 |
| | *Xenopus laevis* (toad) | 1.0 | K01376;X00136;X59734 |
| Nematoda | | | |
| | *Caenorhabditis elegans* | 1.0 | X03680 |

---

[1] Eucaryotic organisms are classified according to the scheme of Margulis,L. and Schwartz,K.V. (1988) [*Five Kingdoms. An Illustrated Guide to the Phyla of Life on Earth* (2nd edition). W.H. Freeman and Co., New York].

[2] New listings that did not appear in the previous edition [Gutell,R.R., Schnare,M.N. and Gray,M.W. (1992) *Nucleic Acids Res.* **20, Supplement**, 2095–2109]

are denoted by *, whereas previous listings containing revised or additional accession number(s) are indicated by #. The most recently updated secondary structure now available in hardcopy form is listed as 1.0 (Release 1.0; July, 1991) or 2.0 (Release 2.0; January. 1993). Underlining of organism names in the above table indicates that an associated [NOTE] concerning the sequence appears with that name in the List of References.

[3] See Woese,C.R., Kandler,O. and Wheelis,M.L. (1990) Towards a natural system of organisms: Proposal for the domains Archaea, Bacteria, and Eucarya. *Proc. Natl. Acad. Sci. U.S.A.* **87**, 4576–4579.

[4] The classification used here follows the taxonomically ordered listing for SSU rRNA sequences on the RDP anonymous server. See Woese,C.R., Stackebrandt,E., Macke,T.J. and Fox, G.E. (1985) A phylogenetic definition of the major eubacterial taxa. *System. Appl. Microbiol.* **6**, 143–151; and Woese,C.R. (1991) The use of ribosomal RNA in reconstructing evolutionary relationships among bacteria. In *Evolution at the Molecular Level* (Selander,R.K., Clark,A.G. and Whittman,T.S., eds.), pp. 1–24. Sinauer, Sunderland, MA.

[5] The term 'Protoctista', as defined and used by Margulis and Schwartz (1988), includes but is not limited to the 'Protista' (unicellular eucaryotes, or protists).

[6] See [NOTE] in List of References, below.

[7] See Cavalier-Smith,T. (1987) Eukaryotes with no mitochondria. *Nature* **326**, 332–333.

[8] Placement of *Pneumocystis carinii* within the Ascomycota is probable but not yet proven; see Bruns,T.D., Vilgalys,R., Barns,S.M., Gonzalez,D., Hibbett,D.S., Lane,D.J., Simon,L., Stickel,S., Szaro,T.M., Weisburg,W.G. and Sogin, M.L. (1992) Evolutionary relationships within the fungi: analyses of nuclear small subunit rRNA sequences. *Mol. Phylogent. Evol.* **1**, 231–241.

## List of references to LSU rRNA Sequences (new or revised from 1992 edition)

### ARCHAEA (ARCHAEBACTERIA)

***Halobacterium marismortui***
Brombach,M., Specht,T., Erdmann,V.A. and Ulbrich,N. (1989) *Nucleic Acids Res.* **17**, 3293.
Complete nucleotide sequence of a 23S ribosomal RNA gene from *Halobacterium marismortui.*

[NOTE: We used the published sequence, which differs from the database entry (X13738) at *E.coli* positions 2506 (U vs. A) and 2523 (G vs. deletion).]

***Sulfolobus solfataricus***
Woese,C.R. (1991) Unpublished.

***Thermococcus celer***
Woese,C.R. and Achenbach,L. (1991) Unpublished.
The sequence of *Thermococcus celer* 23S rRNA.

### BACTERIA (EUBACTERIA)

***Aeromonas hydrophila***
East,A.K. and Collins,M.D. (1993) *FEMS Microbiol. Lett.* **106**, 129–134.
Molecular characterization of DNA encoding 23S rRNA and 16S-23S rRNA intergenic spacer regions of *Aeromonas hydrophila.*

***Bacillus alcalophilus***
Sutherland,K.J., Kudo,T. and Horikoshi,K. (1992) Unpublished.
*B. alcalophilus* 23S ribosomal RNA, partial sequence.

***Bacillus anthracis***
Ash,C. and Collins,M.D. (1992) *FEMS Microbiol. Lett.* **94**, 75–80.
Comparative analysis of 23S ribosomal RNA gene sequences of *Bacillus anthracis* and emetic *Bacillus cereus* determined by PCR-direct sequencing.

***Bacillus cereus*** **(see *Bacillus anthracis*)**

***Bacillus globisporus***
Ludwig,W., Kirchhof,G., Klugbauer,N., Weizenegger,M., Betzl,D., Ehrmann,M., Hertel,C., Jilg,S., Tatzel,R., Zitzelsberger,H., Liebl,S., Hochberger,M., Shah,J., Lane,D., Wallnöfer,P.R. and Schleifer,K.H. (1992) *System. Appl. Microbiol.* **15**, 487–501
Complete 23S ribosomal RNA sequences of gram-positive bacteria with a low DNA G+C content.

***Bacillus licheniformis*** **(see *Bacillus globisporus*)**

***Bacillus stearothermophilus*** **(see also *Bacillus anthracis*)**
Kop,J., Wheaton,V., Gupta,R., Woese,C.R. and Noller,H.F. (1984) *DNA* 3, 347–357.
Complete nucleotide sequence of a 23S ribosomal RNA gene from *Bacillus stearothermophilus.*

[NOTE: We used the primary sequence presented in Kop *et al.* (1984). The secondary structure in this paper has an extra G between *E.coli* positions 11 and 12. This G is not present in the sequence newly determined by Ash and Collins (1992) (see ***Bacillus anthracis***) which, however, differs at two other positions from that of Kop *et al.* (1984).]

***Bacillus subtilis*** **(*rrnB*)**
Green,C.J., Stewart,G.C., Hollis,M.A., Vold,B.S. and Bott,K.F. (1985) *Gene* **37**, 261–266.
Nucleotide sequence of the *Bacillus subtilis* ribosomal RNA operon, *rrnB.*

***Bacillus subtilis*** **(*rrnI*)**
Sutherland,K.J., Kudo,T. and Horikoshi,K. (1992) Unpublished.

***Bordetella avium***
Mueller,M. (1993) Unpublished.
*B. avium* gene for 23S rRNA.

***Bordetella bronchiseptica***
Mueller,M. (1993) Unpublished.
*B. bronchiseptica* gene for 23S rRNA.

***Bordetella parapertussis***
Mueller,M. (1992) Unpublished.
*B. parapertussis* gene for 23S ribosomal RNA.

***Bordetella pertussis***
Mueller,M. (1992) Unpublished.
*B. pertussis* gene for 23S ribosomal RNA.

***Borrelia burgdorferi***
Schwartz,I., Gazumyan,A. and Schwartz,J.J. (1992) Unpublished.
Sequence and structure of the ribosomal RNA genes of *Borrelia burgdorferi*.

***Clostridium botulinum***
East,A.K., Thompson,D.E. and Collins, M.D. (1992) *J. Bacteriol.* **174**, 8158–8162.
Analysis of operons encoding 23S rRNA of *Clostridium botulinum* type A.

***Clostridium tyrobutyricum***
van der Meer,J.R., Ludwig,W. and de Vos,W.M. (1992) Unpublished.
Characterization of a ribosomal RNA gene cluster from *Clostridium tyrobutyricum*: phylogenetic positioning based on the 16S and 23S nucleotide sequences.

***Frankia*** sp.
Normand,P., Cournoyer,B., Simonet,P. and Nazaret,S. (1992) *Gene* **111**, 119–124.
Analysis of a ribosomal RNA operon in the actinomycete *Frankia*.
Johnson,D.A. (1992) Thesis, Biology, University of Ottawa.
Analysis of an rRNA operon from *Frankia* strain AcN14a.

[NOTE: The two *Frankia* sequences have many differences and so are listed as separate sequences (#1 and #2) in Table 1.]

***Lactobacillus delbrueckii*** subsp. ***bulgaricus*** (see ***Bacillus globisporus***)

***Lactococcus lactis*** subsp. *cremoris* (see *Bacillus globisporus*)

***Lactococcus lactis*** subsp. ***lactis*** (strain IL1403)
Chiaruttini,C. and Milet,M. (1993) *J. Mol. Biol.* **230**, 57–76.
Gene organization, primary structure and RNA processing analysis of a ribosomal RNA operon in *Lactococcus lactis*.

***Lactococcus lactis*** subsp. *lactis* (strain DSM 20481$^T$) (see ***Bacillus globisporus***)

[NOTE: Although the three *Lactococcus* sequences are very similar, strain DSM 20481$^T$ has a different hairpin sequence at *E.coli* positions ~1170–1180.]

***Listeria monocytogenes*** (see also ***Bacillus globisporus***)
Thompson,D.E., Balsdon,J.T., Cai,J. and Collins,M.D. (1992) *FEMS Microbiol. Lett.* **96**, 219–224.
Studies on the ribosomal RNA operons of *Listeria monocytogenes*.

***Micrococcus luteus***
Regensburger,A., Ludwig,W., Frank,R., Blöcker,H. and Schleifer,K.H. (1988) *Nucleic Acids Res.* **16**, 2344.
Complete nucleotide sequence of a 23S ribosomal RNA gene from *Micrococcus luteus*.

[NOTE: We used the published sequence, which differs from the database entry (X06484) at the following *E.coli* positions (published sequence first, database entry second): 112, A vs. U; between 174 and 175, G vs. U.]

***Mycobacterium kansasii***
Fleurbaaij,G.A., McNally,M. and Marich,J.E. (1992) Unpublished.
The complete nucleotide sequence of *Mycobacterium kansasii* 23S ribosomal RNA.

***Mycoplasma hyopneumoniae*** (see ***Bacillus globisporus***)

***Mycoplasma pneumoniae*** (see ***Bacillus globisporus***)

***Neisseria gonorrhoeae***
Wolff,K., Sperka,S. and Stern,A. (1992) *Nucleic Acids Res.* **20**, 4657.
Phylogeny and nucleotide sequence of a 23S rRNA gene from *Neisseria gonorrhoeae* and *Neiserria meningitidis*.

***Neisseria meningitidis*** (see ***Neisseria gonorrhoeae***)

***Pectinatus frisigensis*** (see ***Bacillus globisporus***)

***Peptococcus niger*** (see ***Bacillus globisporus***)

***Plesiomonas shigelloides***
East, A.K., Allaway,D. and Collins, M.D. (1992) *FEMS Microbiol. Lett.* **95**, 57–62.
Analysis of DNA encoding 23S rRNA and 16S-23S rRNA intergenic spacer regions from *Plesiomonas shigelloides*.

***Pseudomonas cepacia***
Höpfl,P., Ludwig,W., Schleifer,K.H. and Larsen,N. (1989) *Eur. J. Biochem.* **185**, 355–364.
The 23S ribosomal RNA higher-order structure of *Pseudomonas cepacia* and other prokaryotes.

[NOTE: The position of the 3′-terminus is unclear in Höpfl *et al.* (1989) (there is a discrepancy between the number given in text and the secondary structure figure). We have assigned the 3′-terminus by comparison with *E.coli* 23S rRNA.]

***Pseudomonas perfectomarina*** (= *P. stutzeri*)
Kerkhof,L.J. (1992) Unpublished.
A species-specific probe for the marine bacterium, *Pseudomonas perfectomarina* (*P. stutzeri* strain Zobell)

***Rhodobacter sphaeroides***
Dryden,S.C. and Kaplan,S. (1990) *Nucleic Acids Res.* **18**, 7267–7277.
Localization and structural analysis of the ribosomal RNA operons of *Rhodobacter sphaeroides*.

[NOTE: Minor sequence heterogeneity has been noted in the three operons that have been sequenced.]

***Staphylococcus aureus*** (see ***Bacillus globisporus***)

***Staphylococcus carnosus*** (see ***Bacillus globisporus***)

***Streptococcus oralis*** (see ***Bacillus globisporus***)

***Streptococcus thermophilus*** (see ***Bacillus globisporus***)

***Synechococcus*** sp. 6301 (formerly ***Anacystis nidulans***)
Kumano,M., Tomioka,N. and Sugiura,M. (1983) *Gene* **24**, 219–225.
The complete nucleotide sequence of a 23S rRNA gene from a blue-green alga, *Anacystis nidulans*.

Douglas,S.E. and Doolittle,W.F. (1984) *Nucleic Acids Res.* **12**, 3373–3386.
Complete nucleotide sequence of the 23S rRNA gene of the Cyanobacterium, *Anacystis nidulans*.

[NOTE: The Douglas and Doolittle (1984) sequence differs from that of Kumano *et al.* (1983) at a number of positions. We used the former sequence, which is better supported by secondary structure comparisons.]

***Thermotoga maritima***
Achenbach,L. and Woese,C.R. (1991) Unpublished.
Sequence of the ribosomal operon of *Thermotoga maritima*.

---

## PLASTIDS

***Alnus incana***
Lévesque,M. and Johnson,D.A. (1992). *Plant Mol. Biol.* **18**, 601–602.
Nucleotide sequence of the chloroplast 23S rRNA gene from alder (*Alnus incana*).

Imbeault,J.-C. and Johnson,D.A. (1991) Unpublished.
Analysis of the 4.5S-5S rRNA intergenic region in rDNA from alder (*Alnus incana*) chloroplast.

***Chlamydomonas*** sp. 66.72 (see ***Chlamydomonas frankii***)

***Chlamydomonas eugametos***
Turmel,M., Boulanger,J., Schnare,M.N., Gray,M.W. and Lemieux,C. (1991) *J. Mol. Biol.* **218**, 293–311.
Six group I introns and three internal transcribed spacers in the chloroplast large subunit ribosomal RNA gene of the green alga *Chlamydomonas eugametos*.

***Chlamydomonas frankii***
Turmel,M., Gutell,R.R., Mercier,J.-P., Otis,C. and Lemieux,C. (1993) *J. Mol. Biol.*, in press.
Analysis of the chloroplast large subunit ribosomal RNA gene from 17 *Chlamydomonas* taxa. Three internal transcribed spacers and 12 group I intron insertion sites.

***Chlamydomonas geitleri*** (see ***Chlamydomonas frankii***)

***Chlamydomonas gelatinosa*** (see ***Chlamydomonas frankii***)

***Chlamydomonas humicola*** (see ***Chlamydomonas frankii***)

***Chlamydomonas indica*** (see ***Chlamydomonas frankii***)

***Chlamydomonas iyengarii*** (see ***Chlamydomonas frankii***)

***Chlamydomonas komma*** (see ***Chlamydomonas frankii***)

***Chlamydomonas mexicana*** (see ***Chlamydomonas frankii***)

***Chlamydomonas moewusii*** (see ***Chlamydomonas frankii***)

***Chlamydomonas pallidostigmatica*** (see ***Chlamydomonas frankii***)

***Chlamydomonas peterfii*** (see ***Chlamydomonas frankii***)

***Chlamydomonas pitschmannii*** (see ***Chlamydomonas frankii***)

***Chlamydomonas reinhardtii***
Lemieux,C., Boulanger,J., Otis,C, and Turmel,M. (1989) *Nucleic Acids Res.* **17**, 7997.
Nucleotide sequence of the chloroplast large subunit rRNA gene from *Chlamydomonas reinhardtii.*

[NOTE: The 5′-end of the 7S rRNA and the 3′-end of the 23S rRNA have been changed to agree with those of the corresponding species in *C. eugametos* chloroplast; see ***Chlamydomonas eugametos***.)]

***Chlamydomonas starrii*** (see ***Chlamydomonas frankii***)

***Chlamydomonas zebra*** (see ***Chlamydomonas frankii***)

***Chlorella ellipsoidea***
Yamada,T. and Shimaji,M. (1987) *Curr. Genet.* **11**, 347–352.
An intron in the 23S rRNA gene of the *Chlorella* chloroplasts: Complete nucleotide sequence of the 23S rRNA gene.

[NOTE: The putative intron described by Yamada and Shimaji (1987) is probably an internal transcribed spacer that is excised during post-transcriptional processing; see Turmel *et al.* (1991), ***Chlamydomonas eugametos***).

***Epifagus virginiana*** (beechdrops)
Wolfe,K.H., Katz-Downie,D.S., Morden,C.W. and Palmer,J.D. (1992) *Plant Mol. Biol.* **18**, 1037–1048.
Evolution of the plastid ribosomal RNA operon in a nongreen parasitic plant: Accelerated sequence evolution, altered promoter structure, and tRNA pseudogenes.

***Euglena gracilis***
Yepiz-Plascencia,G.M., Jenkins,M.E. and Hallick,R.B. (1988) *Nucleic Acids Res.* **16**, 9340.
Nucleotide sequence of the *Euglena gracilis* chloroplast 23S rRNA gene of the *rrnC* operon.

Schnare,M.N., Yepiz-Plascencia,G.M., Copertino,D.W., Hallick,R.B. and Gray,M.W. (1992) *Nucleic Acids Res.* **20**, 1421.
5′- and 3′-terminal sequences of the chloroplast 16S and 23S ribosomal RNAs of *Euglena gracilis.*

[NOTE: The sequence in Schnare *et al.* (1992) (X12890) revises the original one (X13310) presented in Yepiz-Plascencia *et al.* (1988).]

***Nicotiana tabacum***
Takaiwa,F. and Sugiura,M. (1982) *Eur. J. Biochem.* **124**, 13–19.
The complete nucleotide sequence of a 23-S rRNA gene from tobacco chloroplasts.

Shinozaki,K., Ohme,M., Tanaka,M., Wakasugi,T., Hayashida,N., Matsubayashi,T., Zaita,N., Chunwongse,J., Obokata,J., Yamaguchi-Shinozaki,K., Ohto,C., Torazawa,K., Meng,B.Y., Sugita,M., Deno,H., Kamogashira,T., Yamada,K., Kusuda,J., Takaiwa,F., Kato,A., Tohdoh,N., Shimada,H. and Sugiura,M. (1986) *Plant Mol. Biol. Reporter* 4, 111–147.
The complete nucleotide sequence of the tobacco chloroplast genome.
[See also: *EMBO J.* 5, 2043–2049 (1986). The complete nucleotide sequence of the tobacco chloroplast genome: its gene organization and expression.]

[NOTE: The original sequence of Takaiwa and Sugiura (1982) differs from that presented later in Shinozaki *et al.* (1986). We used the more recent sequence because it agrees with other plant chloroplast sequences and is supported by the secondary structure.]

***Palmaria palmata***
Singh,R.K. (1993) Unpublished.
The organization and primary structure of the plastid rRNA operon from the red alga *Palmaria palmata.*

***Plasmodium falciparum***
Gardner,M.J., Feagin,J.E., Moore,D.J., Rangachari,K., Williamson,D.H. and Wilson,R.J.M. (1993) *Nucleic Acids Res.* **21**, 1067–1071.
Sequence and organization of large subunit rRNA genes from the extrachromosomal 35 kb circular DNA of the malaria parasite *Plasmodium falciparum.*

[NOTE: Phylogenetic analysis by Gardner *et al.* (1993) suggests that the *P.falciparum* 35-kbp DNA may be a plastid DNA remnant; see also Gardner,M.J., Feagin,J.E., Moore,D.J., Spencer,D.F., Gray,M.W., Williamson,D.H. and Wilson,R.J.M. (1991) *Mol. Biochem. Parasitol.* **48**, 77–88.]

***Pylaiella littoralis***
Somerville,C.C., Jouannic,S., Martin,W.F., Kloareg,B. and Loiseaux-de Goër,S. (1993) *Plant Mol. Biol.* **21**, 779–787.
Secondary structure and phylogeny of the chloroplast 23S rRNA gene from the brown alga *Pylaiella littoralis.*

***Zea mays***
Edwards,K. and Kössel,H. (1981) *Nucleic Acids Res.* 9, 2853–2869.
The rRNA operon from Zea mays chloroplasts: nucleotide sequence of 23S rDNA and its homology with E.coli 23S rDNA.

[NOTE: The 23S rRNA is thought to be fragmented at two positions as a result of excision of short internal transcribed spacers; see Kössel,H., Natt,E., Strittmatter,G., Fritzche,E., Gozdzicka-Jozefiak,A. and Przybyl,D. (1985) Structure and expression of rRNA operons from plastids of higher plants. In *Molecular Form and Function of the Plant Genome* (van Vloten-Doting,L., Groot,G.S.P. and Hall,T.C., eds.), NATO ASI Series (Series A: Life Sciences), Vol. 83, pp. 183–198, Plenum, New York.]

---

## MITOCHONDRIA

***Aepyceros melampus***
Allard,M.W., Miyamoto,M.M., Jarecki,L., Kraus,F. and Tennant,M.R. (1992) *Proc. Natl. Acad. Sci. U.S.A.* **89**, 3972–3976.
DNA systematics and evolution of the artiodactyl family Bovidae.

***Apis mellifera***
Vlasak,I., Burgschwaiger,S. and Kreil,G. (1987) *Nucleic Acids Res.* **15**, 2388.
Nucleotide sequence of the large ribosomal RNA of honeybee mitochondria.

Crozier,R.H. and Crozier,Y.C. (1993) *Genetics* **133**, 97–117.
The mitochondrial genome of the honeybee *Apis mellifera*: complete sequence and genome organization.

***Artemia salina***
Palmero,I., Renart,J. and Sastre,L. (1988) *Gene* **68**, 239–248.
Isolation of cDNA clones coding for mitochondrial 16S ribosomal RNA from the crustacean *Artemia*.

[NOTE: The partial cDNA sequence of Palmero *et al.* (1988) shows many differences from the database entry (complete sequence). We used the sequence from the database.]

***Ascaris suum***
Okimoto,R., Macfarlane,J.L., Clary,D.O. and Wolsteholme,D.R. (1992) *Genetics* **130**, 471–498.
The mitochondrial genomes of two nematodes, *Caenorhabditis elegans* and *Ascaris suum*.

***Caenorhabditis elegans*** (see ***Ascaris suum***)

***Carassius auratus*** (goldfish)
Peng,G., Taylor,J.D. and Tchen,T.T. (1992) *Biochem. Biophys. Res. Commun.* **189**, 445–449.
Increased mitochondrial activities in pigmented (melanized) fish cells and nucleotide sequence of mitochondrial large rRNA.

***Crossostoma lacustre*** (freshwater loach)
Tzeng,C.-S., Hui,C.-F., Shen,S.-C. and Huang,P.C. (1992) *Nucleic Acids Res.* **20**, 4853–4858.
The complete nucleotide sequence of the *Crossostoma lacustre* mitochondrial genome: conservation and variations among vertebrates.

***Locusta migratoria***
Uhlenbusch,I., McCracken,A. and Gellissen,G. (1987) *Curr. Genet.* **11**, 631–638.
The gene for the large (16S) ribosomal RNA from the *Locusta migratoria* mitochondrial genome.

[NOTE: The primary sequence figure in Uhlenbusch *et al.* (1987) has the sequence presented backwards and a few positions are not clear. These positions (corresponding to *E.coli* 1933 and 2659) are clearly G in the secondary structure presented in the same paper; however, they are C in the database entry (X05287), which also displays the sequence backwards.]

***Marchantia polymorpha***
Oda,K., Yamato,K., Ohta,E., Nakamura,Y., Takemura,M., Nozato,N., Akashi,K., Kanegae,T., Ogura,Y., Kohchi,T. and Ohyama,K. (1992) *Plant Mol. Biol. Reporter* **10**, 105–163.
Complete nucleotide sequence of the mitochondrial DNA from a liverwort, *Marchantia polymorpha*.
[#See also: *J. Mol. Biol.* **223**, 1–7 (1992). Gene organization deduced from the complete sequence of liverwort *Marchantia polymorpha* mitochondrial DNA: a primitive form of plant mitochondrial genome.]

***Mytilus edulis***
Hoffmann,R.J., Boore,J.L. and Brown,W.M. (1992) *Genetics* **131**, 397–412.
A novel mitochondrial genome organization for the blue mussel, *Mytilus edulis*.

***Paramecium primaurelia***
Seilhamer,J.J. and Cummings,D.J. (1981) *Nucleic Acids Res.* 9, 6391–6406.
Structure and sequence of the mitochondrial 20S rRNA and tRNA tyr gene of Paramecium primaurelia.

Seilhamer,J.J., Gutell,R.R. and Cummings,D.J. (1984) *J. Biol. Chem.* **259**, 5173–5181.
*Paramecium* mitochondrial genes. II. Large subunit rRNA gene sequence and microevolution.

[NOTE: The sequence in Seilhamer and Cummings (1981) differs from that in Seilhamer *et al.* (1984). We used the latter (more recent) version.]

***Penicillium chrysogenum***
Naruse,A., Yamamoto,H. and Sekiguchi,J. (1993) *Biochim. Biophys. Acta* **1172**, 353–356.
Nucleotide sequence of the large mitochondrial rRNA gene of *Penicillium chrysogenum*.

***Phoca vitulina***
Árnason,U. and Johnsson,E. (1992) *J. Mol. Evol.* **34**, 493–505.
The complete mitochondrial DNA sequence of the harbor seal, *Phoca vitulina*.

***Plasmodium falciparum***
Feagin,J.E., Werner,E., Gardner,M.J., Williamson,D.H. and Wilson,R.J.M. (1992) *Nucleic Acids Res.* **20**, 879–887.
Homologies between the contiguous and fragmented rRNAs of two *Plasmodium falciparum* extrachromosomal DNAs are limited to core sequences.

Vaidya,A.B., Lashgari,M.S., Pologe,L.G. and Morrisey,J. (1993) Structural features of *Plasmodium* cytochrome *b* that may underlie susceptibility to 8-aminoquinolines and hydroxynaphthoquinones. *Mol. Biochem. Parasitol.* **58**, 33–42.

***Plasmodium yoelii***
Vaidya,A.B., Akella,R. and Suplick,K. (1989) *Mol. Biochem. Parasitol.* **35**, 97–108.
Sequences similar to genes for two mitochondrial proteins and portions of ribosomal RNA in tandemly arrayed 6-kilobase-pair DNA of a malarial parasite.

[NOTE: A corrigendum by the same authors (*Mol. Biochem. Parasitol.* **39**, 295–296 (1990)) contains corrections to their original sequence.]
Suplick,K., Morrisey,J. and Vaidya,A.B. (1990) *Mol. Cell. Biol.* **10**, 6381–6388.
Complex transcription from the extrachromosomal DNA encoding mitochondrial functions of *Plasmodium yoelii*.

***Spodoptera frugiperda***
Pashley,D.P. and Ke,L.D. (1992) *Mol. Biol. Evol.* **9**, 1061–1075.
Sequence evolution in mitochondrial ribosomal and ND-1 genes in Lepidoptera: implications for phylogenetic analyses.

***Triticum aestivum***
Falconet,D., Sevignac,M. and Quétier,F. (1988) *Curr. Genet.* **13**, 75–82.
Nucleotide sequence and determination of the extremities of the 26S ribosomal RNA gene in wheat mitochondria: evidence for sequence rearrangements in the ribosomal genes of higher plants.

Spencer,D.F., Schnare,M.N., Coulthart,M.B. and Gray,M.W. (1992) *Plant Mol. Biol.* **20**, 347–352.
Sequence and organization of a 7.2 kb region of wheat mitochondrial DNA containing the large subunit (26S) rRNA gene.

---

## EUCARYA (EUCARYOTES, NUCLEOCYTOPLASMIC)

***Brassica napus***
Okumura,S. and Shimada,H. (1992) *Nucleic Acids Res.* **20**, 3510.
Nucleotide sequence of genes for a 5.8S and 25S rRNA from rape seed (*Brassica napus*).

***Candida albicans***
Mercure,S., Rougeau,N., Montplaisir,S. and Lemay,G. (1993) *Nucleic Acids Res.* **21**, 1490.
The nucleotide sequence of the 25S rRNA-encoding gene from *Candida albicans*.

Lott,T.J. (1992) Unpublished.
Nucleotide sequence of the 5.8S rDNA from *Candida albicans*.

***Didymium iridis***
Johansen,S., Johansen,T. and Haugli,F. (1992) *Curr. Genet.* **22**, 305–312.
Extrachromosomal ribosomal DNA of *Didymium iridis*: sequence analysis of the large subunit ribosomal RNA gene and sub-telomeric region.

***Drosophila melanogaster***
Tautz,D., Hancock,J.M., Webb,D.A., Tautz,C. and Dover,G.A. (1988) *Mol. Biol. Evol.* **5**, 366–376.
Complete sequences of the rRNA genes of *Drosophila melanogaster*.

Linares,A.R., Hancock,J.M. and Dover,G.A. (1991) *J. Mol. Biol.* **219**, 381–390.
Secondary structure constraints on the evolution of *Drosophila* 28 S ribosomal RNA expansion segments

Rousset,F., Pélandakis,M. and Solignac,M. (1991) *Proc. Natl. Acad. Sci. U.S.A.* **88**, 10032–10036.
Evolution of compensatory substitutions through GU intermediate state in *Drosophila* rRNA.

[NOTE: Minor corrections are presented in Linares *et al.* (1991) and Rousset *et al.* (1991). The latter contains a sequence of 25 nucleotides absent in Tautz *et al.* (1988).]

***Entamoeba histolytica***
Ramachandran,S., Bhattacharya,A. and Bhattacharya,S. (1993) *Nucleic Acids Res.* **21**, 2011.
Nucleotide sequence analysis of the rRNA transcription unit of a pathogenic *Entamoeba histolytica* strain HM-EIMSS.

***Giardia ardeae***
van Keulen,H., Horvat,S., Erlandsen,S.L. and Jarroll,E.L. (1991) *Nucleic Acids Res.* **19**, 6050.
Nucleotide sequence of the 5.8S and large subunit rRNA genes and the internal transcribed spacer and part of the external spacer from *Giardia ardeae*.

[NOTE: See van Keulen *et al.* (1992) (***Giardia muris***) for primary sequence. The *G. ardeae* sequence in this paper has several differences from the database entry (X58290). We used the *G. ardeae* sequence presented in the *G. muris* paper, which fits the secondary structure better (the V3 structure shown in the *G. ardeae* paper is missing a G).]

***Giardia intestinalis* (= *G. duodenalis, G. lamblia*)**
Healey,A., Mitchell,R., Upcroft,J.A., Boreham,P.F.L. and Upcroft,P. (1990) *Nucleic Acids Res.* **18**, 4006.
Complete nucleotide sequence of the ribosomal RNA tandem repeat unit from *Giardia intestinalis*.

[NOTE: A minor correction (insertion of a C) is presented in the ***G. muris*** reference. The two ambiguous positions are G according to Edlind,T.D., Sharetzsky,C. and Cha, M.E. (1990) Ribosomal RNA of the primitive eukaryote *Giardia lamblia*: large subunit domain I and potential processing signals. *Gene* **96**, 289–293.]

***Giardia muris***
van Keulen,H., Gutell,R.R., Campbell,S.R., Erlandsen,S.L. and Jarrell,E.L. (1992) *J. Mol. Evol.* **35**, 318–328.
The nucleotide sequence of the entire ribosomal DNA operon and the structure of the large subunit rRNA of *Giardia muris*.

***Homo sapiens***
Maden,B.E.H., Dent,C.L., Farrell,T.E., Garde,J., McCallum,F.S. and Wakeman,J.A. (1987) *Biochem J.* **246**, 519–527.
Clones of human ribosomal DNA containing the complete 18 S-rRNA and 28 S-rRNA genes. Characterization, a detailed map of the human ribosomal transcription unit and diversity among clones.

[NOTE: A correction was made in Maden *et al.* (1987) to eliminate an extra GC (following position 50) in the original 5.8S rRNA sequence of Khan,M.S.N and Maden,B.E.H. (1977) *Nucleic Acids Res.* 4, 2495–2505. Nucleotide sequence relationships between vertebrate 5.8 S ribosomal RNAs.]

Gonzalez,I.L., Gorski,J.L., Campen,T.J., Dorney,D.J., Erickson,J.M., Sylvester,J.E. and Schmickel,R.D. (1985) *Proc. Natl. Acad. Sci. U.S.A.* **82**, 7666–7670.
Variation among human 28S ribosomal RNA genes.

***Pneumocystis carinii***
Liu,Y., Rocourt,M., Pan,S., Liu,C. and Leibowitz,M.J.(1992) *Nucleic Acids Res.* **20**, 3763–3772.
Sequence and variability of the 5.8S and 26S rRNA genes of *Pneumocystis carinii*.

***Rattus norvegicus***
Subrahmanyam,C.S., Cassidy,B., Busch,H. and Rothblum,L.I. (1982) *Nucleic Acids Res.* **10**, 3667–3680.
Nucleotide sequence of the region between the 18S rRNA sequence and the 28S rRNA sequence of rat ribosomal DNA.

Chan,Y.-L., Olvera,J. and Wool,I.G. (1983) *Nucleic Acids Res.* **11**, 7819–7831.
The structure of rat 28S ribosomal ribonucleic acid inferred from the sequence of nucleotides in a gene.

Hadjiolov,A.A., Georgiev,O.I., Nosikov,V.V. and Yavachev,L.P. (1984) *Nucleic Acids Res.* **12**, 3677–3693.
Primary and secondary structure of rat 28 S ribosomal RNA.

Wool,I.G. (1986) In Hardesty,B. and Kramer,G. (eds.). *Structure, Function, and Genetics of Ribosomes.* Springer-Verlag, New York, pp. 391–411.
Studies on the structure of eukaryotic (mammalian) ribosomes.

[NOTE: We used the most recent (revised) sequence given in Wool (1986).]

***Saccharomyces carlsbergensis***
Veldman,G.M., Brand, R.C., Klootwijk,J. and Planta,R.J. (1980) *Nucleic Acids Res.* **8**, 2907–2920.
Some characteristics of processing sites in ribosomal precursor RNA of yeast.

Veldman,G.M., Klootwijk,J., van Heerikhuizen,H. and Planta,R.J. (1981) *Nucleic Acids Res.* **9**, 4847–4862.
The nucleotide sequence of the intergenic region between the 5.8S and 26S rRNA genes of the yeast ribosomal RNA operon. Possible implications for the interaction between 5.8S and 26S rRNA and the processing of the primary transcript.

Veldman,G.M., Klootwijk,J., de Regt,V.C.H.F., Planta,R.J., Branlant,C., Krol,A. and Ebel,J.-P. (1981) *Nucleic Acids Res.* 9, 6935–6952.
The primary and secondary structure of yeast 26S rRNA.

***Saccharomyces cerevisiae***
Rubin,G.M. (1973) *J. Biol. Chem.* **248**, 3860–3875.
The nucleotide sequence of *Saccharomyces cerevisiae* 5.8 S ribosomal ribonucleic acid.

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.
The structure of the yeast ribosomal RNA genes. 4. Complete sequence of the 25 S rRNA gene from Saccharomyces cerevisiae.

[NOTE: The *S. carlsbergensis* and *S.cerevisiae* sequences are very similar; most of the differences likely reflect errors in the latter sequence. However, we used the *S.cerevisiae* sequence because the *S. carlsbergensis* 5.8S rRNA sequence is not complete.]

***Schizosaccharomyces pombe***
Schaak,J., Mao,J.I. and Söll,D.G. (1982) *Nucleic Acids Res.* **10**, 2851–2864.
The 5.8S RNA gene sequence and the ribosomal repeat of *Schizosaccharomyces pombe*.

Lapeyre,B.B. and Feliu, J.J. (1992) Unpublished.
Nucleotide sequence of the 25S ribosomal RNA from *Schizosaccharomyces pombe*.

***Sinapis alba***
Capesius,I. (1991) *Plant Mol. Biol.* **16**, 1093–1094.
Nucleotide sequence of a 25S rRNA gene from mustard (*Sinapis alba*).

Rathgeber,J. and Capesius,I. (1989) *Nucleic Acids Res.* **17**, 7522.
Nucleotide sequence of the 18S-25S spacer region from mustard DNA.

***Tetrahymena thermophila***
Engberg,J., Nielsen,H., Lenaers,G., Murayama,O., Fujitani,H. and Higashinakagawa,T. (1990) *J. Mol. Evol.* **30**, 514–521.
Comparison of primary and secondary 26S rRNA structures in two *Tetrahymena* species: Evidence for a strong evolutionary and structural constraint in expansion segments.

Engberg,J. and Nielsen,H. (1990) *Nucleic Acids Res.* **18**, 6915–6919.
Complete sequence of the extrachromosomal rDNA molecule from the ciliate *Tetrahymena thermophila* strain B1868VII.

[NOTE: The sequence shown in Engberg and Nielsen (1990) is missing positions 6601–6700 (corresponding to *E.coli* positions 1599–1698.]

***Xenopus laevis***
Hall,L.M.C. and Maden,B.E.H. (1980) *Nucleic Acids Res.* 8, 5993–6005.
Nucleotide sequence through the 18S-28S intergene region of a vertebrate ribosomal transcription unit.
Ware,V.C., Tague,B.W., Clark,C.G., Gourse,R.L., Brand,R.C.and Gerbi,S.A. (1983) *Nucleic Acids Res.* **11**, 7795–7817.
Sequence analysis of 28S ribosomal DNA from the amphibian *Xenopus laevis*.

Stebbins-Boaz,B. and Gerbi,S.A. (1991) *J. Mol. Biol.* **217**, 93–112.
Structural analysis of the peptidyl transferase region in ribosomal RNA

Ajuh,P.M., Heeney,P.A. and Maden,B.E.H. (1991) *Proc. R. Soc. Lond., B, Biol Sci.* **245**, 65–71.
*Xenopus borealis* and *Xenopus laevis* 28S ribosomal DNA and the complete 40S ribosomal precursor RNA coding units of both species.

Schnare,M.N. and Gray,M.W. (1992) *Nucleic Acids Res.* **20**, 608.
A new 3′-terminus for *Xenopus laevis* 28S ribosomal RNA.

[NOTE: Corrections to the sequence of Domain V in Ware *et al.* (1983) may be found in Stebbins-Boaz and Gerbi (1991). The most recent revisions (Ajuh et al. (1991); Schnare and Gray (1992)) are being incorporated into the current secondary structure model.]