Compilation of small ribosomal subunit RNA sequences

Erik Huysmans and Rupert De Wachter*

Departement Biochemie, Universiteit Antwerpen (UIA), Universiteitsplein 1, B-2610 Antwerpen, Belgium

INTRODUCTION

In the present compilation of small ribosomal subunit RNA sequences (also named 16 S-like ribosomal RNA; further abbreviated srRNA) an attempt is made to list all such sequences published up to and including 1985. The sequences are presented in the form of an alignment based on primary structure homology and containing information on the presumed secondary structure. The 57 sequences included are listed in Table 1 and belong to 14 eukaryotes, 9 archaeobacteria, 13 eubacteria, 4 plastids and 16 mitochondria. Only complete or virtually complete sequences were included. Since papers on srRNA structure are appearing at an increasingly rapid pace, some recent papers may have escaped our attention or reached us too late to meet the publication deadline, for which we apologize to their authors.

CONVENTIONS AND SYMBOLS USED IN THE SEQUENCE ALIGNMENT

The 57 sequences appear in the alignment in the same order, and bearing the same number, as in Table 1. In addition to its number, each sequence is identified by the initials of the species name. The numbering at the top and the bottom applies to alignment positions, not to the nucleotides of an individual sequence. The sequences are divided in 5 blocks, comprising srRNAs from eukaryotes, archaeobacteria, eubacteria, plastids and mitochondria.

All the sequences are presented as RNA chains, i.e. using the symbols A, C, G and U, although most of them have been analyzed on the DNA level. A set of symbols for incompletely specified residues, listed in Table 2, has been recommended by the IUB nomenclature committee (63,64). These symbols are only used to denote sequence heterogeneity detected at certain positions of srRNAs analyzed by RNA sequencing, in other words they denote heterogeneity due to gene polymorphism. In contrast, nucleotides incompletely identified during sequencing of a cloned gene are denoted by the symbol X.

Table 1. srRNA sequences listed in the alignment

Footnotes are listed on the page following this Table

Nr (a)	Species	Specification (b)	Taxonomic position (c)		S- value	Length (d)	Remark
	Eukaryotes						
1	Homo sapiens		Chordata (Ph.)	Mammalia (Cl.)	18S	1868	
2	Mus musculus		Chordata (Ph.)	Mammalia (Cl.)	18S	1869	
3	Rattus norvegicus		Chordata (Ph.)	Mammalia (Cl.)	18S	1869	e
4	Rattus norvegicus		Chordata (Ph.)	Mammalia (Cl.)	18S	1874	e
5	Oryctolagus caniculus	reticulocytes	Chordata (Ph.)	Mammalia (Cl.)	18S	1868	f,g
6	Xenopus laevis		Chordata (Ph.)	Amphibia (Cl.)	18S	1826	h
7	Artemia salina	cysts	Arthropoda (Ph.)	Crustacea (Cl.)	18S	1810	
8	Saccharomyces cerevisiae	strain A364A	Ascomycota (Ph.)	Hemiascomycotina (SPh.)	18S	1802	i
9	Dictyostelium discoideum	strain Ax3	Myxomycota (Ph.)	Acrascomycetes (Cl.)	17S	1872	g
10	Oxytricha nova		Ciliata (Ph.)	Hypotricha (Cl.)	1771	1771	
11	Stylonychia pustulata		Ciliata (Ph.)	Hypotricha (Cl.)	1771	1771	
12	Tetrahymena thermophila	strain C3-368	Ciliata (Ph.)	Holotricha (Cl.)	17S	1753	
13	Oryza sativa		Spermatophyta (Ph.)	Liliopsida (Cl.)	17S	1812	
14	Zea mays	cv. Black Mexican; leaves	Spermatophyta (Ph.)	Liliopsida (Cl.)	17S	1809	j
15	Glycine max	var. Wayne	Spermatophyta (Ph.)	Liliopsida (Cl.)	18S	1807	
	Archaea						
16	Halobacterium cutirubrum		Halobacteriales (O.)		16S	1472	
17	Halobacterium halobium	strain RI	Halobacteriales (O.)		16S	1473	
18	Halobacterium volcanii		Halobacteriales (O.)		16S	1472	
19	Halococcus morrhuae	ATCC 17082	Halobacteriales (O.)		16S	1475	
20	Methanospirillum hungatei	strain JF1	Methanomicrobiales (O.)		16S	1464	
21	Methanococcus vannielii	DSM 1224	Methanococcales (O.)		16S	1466	
22	Methanobacterium formicicum	DSM 1312	Methanobacteriales (O.)		16S	1476	
23	Sulfolobus solfataricus	strain P1	Sulfolobales (O.)		16S	1493	g
24	Thermoproteus tenax		Thermoproteales (O.)		16S	1504	

Table 1, continued

25	Eubacteria						
26	<i>Agrobacterium tumefaciens</i>	DSM 30105	Purple & relatives	alpha	16S	1489	k f, g l
27	<i>Pseudomonas testosteroni</i>	ATCC 11996	Purple & relatives	beta	16S	1536	
28	<i>Escherichia coli</i>	rrnB	Purple & relatives	gamma	16S	1542	
29	<i>Proteus vulgaris</i>		Purple & relatives	gamma	16S	1544	
30	<i>Bacillus subtilis</i>	rrnB	Gram + & relatives	Clostridia & rel.	16S	1550	
31	<i>Helicobacterium chlorum</i>		Gram + & relatives		16S	1551	j
32	<i>Mycoplasma capricolum</i>	rrnB	Gram + & relatives	Mycoplasmata	16S	1521	
33	<i>Mycoplasma sp.</i>	strain PG 50; rrnA	Gram + & relatives	Mycoplasmata	16S	1523	
34	<i>Bacteroides fragilis</i>	ATCC 25285	Bacteroides & relatives	Bacteroides	16S	1533	
35	<i>Flavobacterium heparinum</i>	IFO 12017 (ATCC 13125)	Bacteroides & relatives	Flavobacteria	16S	1528	
36	<i>Desulfovibrio desulfuricans</i>	ATCC 27774	S-dependent & relatives	S-dependent	16S	1551	n o
37	<i>Mycococcus xanthus</i>	strain M0207	S-dependent & relatives	Myxobacteria	16S	1537	
	<i>Anacystis nidulans</i>	strain 6301 (Berkeley)	Cyanobacteria		16S	1487	
	Chloroplasts						
38	<i>Euglena gracilis</i>	strain Z	Euglenophyta (Ph.)		16S	1491	
39	<i>Chlamydomonas reinhardtii</i>		Chlorophyta (Ph.)		16S	1475	j
40	<i>Zea mays</i>		Spermatophyta (Ph.)	Liliopsida (Cl.)	16S	1490	
41	<i>Nicotiana tabacum</i>	var. Bright Yellow 4	Spermatophyta (Ph.)	Magnoliopsida (Cl.)	16S	1486	
	Mitochondria						
42	<i>Homo sapiens</i>	placenta	Chordata (Ph.)	Mammalia (Cl.)	12S	954	n o
43	<i>Mus musculus</i>	L-cell, line C2-1	Chordata (Ph.)	Mammalia (Cl.)	12S	956	
44	<i>Rattus norvegicus</i>	liver	Chordata (Ph.)	Mammalia (Cl.)	12S	953	
45	<i>Bos taurus</i>	heart	Chordata (Ph.)	Mammalia (Cl.)	12S	955	
46	<i>Xenopus laevis</i>		Chordata (Ph.)	Amphibia (Cl.)	12S	819	
47	<i>Drosophila yakuba</i>	stock 2371.6, Ivory coast	Arthropoda (Ph.)	Insecta (Cl.)	12S	789	n o
48	<i>Aspergillus nidulans</i>	strain paba A1, bi A1	Ascomycota (Ph.)	Plectomycetes (Cl.)	15S	1437	
49	<i>Saccharomyces cerevisiae</i>	strain D273-108	Ascomycota (Ph.)	Hemiascomycotina (SPh.)	15S	1686	
50	<i>Paramecium primaurelia</i>	stock 513 (Edinburgh)	Ciliata (Ph.)	Holotricha (Cl.)	1680	1680	
51	<i>Paramecium tetraurelia</i>	stock 51 (Edinburgh)	Ciliata (Ph.)	Holotricha (Cl.)	1680	1680	
52	<i>Crithidia fasciculata</i>		Flagellata (Ph.)		9S	612	n o
53	<i>Trypanosoma brucei</i>		Flagellata (Ph.)		9S	618	
54	<i>Leishmania tarentolae</i>	strain 427	Flagellata (Ph.)		9S	609	
55	<i>Zea mays</i>		Spermatophyta (Ph.)	Liliopsida (Cl.)	18S	1968	
56	<i>Triticum aestivum</i>		Spermatophyta (Ph.)	Liliopsida (Cl.)	18S	1955	
57	<i>Oenothera sp.</i>		Spermatophyta (Ph.)	Magnoliopsida (Cl.)	18S	1912	

Footnotes to Table 1

- a) This number precedes and follows the sequence on each alignment page and is also the reference number of the paper from which the sequence was taken.
- b) This column contains the following data, if specified by the authors.
 - Strain name for laboratory animals, (cultivated) variety for plants, culture collection and strain number in the case of microorganisms.
 - Tissue from which the DNA used for clone construction, or the RNA used for sequencing, was extracted in the case of differentiated organisms.
 - Ribosomal RNA operon to which belongs the cloned srRNA gene in the case of bacteria.
- c) The taxonomic position is described according to the following references : (68) for metazoa and protozoa, (69) for Ascomycota, (70) for algae and phytoflagellates, (71) for higher plants, (72) and (73) for archaeobacteria, (74) for eubacteria. Taxon designations corresponding to an established taxonomic level are followed by the abbreviation Ph. (phylum), Sph. (sub-phylum), Cl. (classis), or O. (ordo).
- d) The chain length is not mentioned in case the location of the srRNA termini on the sequenced DNA is uncertain.
- e) Two different sequences have been reported for rat srRNA, both are listed.
- f) The primary structure was derived by RNA sequencing.
- g) Partial data on modified nucleosides are mentioned in the paper reporting this sequence or other papers cited therein.
- h) A sequence missing one nucleotide (A₆₅) was previously reported in (58), where complete data on nucleoside modification can be found.
- i) This paper contains a sequence corrected with respect to the original report (59).
- j) The termini were allocated by comparison with a sequence from a related species.
- k) A slightly different structure was previously determined by RNA sequencing (60) and a sequence missing one nucleotide (U₆₉) of the rrnB srRNA (61) was corrected in (27). Complete data on nucleoside modification appear in (60) and (61).
- l) Approximately 15 nucleotides adjacent to the 5'-terminus are missing because they were not comprised in the cloned sequence.
- m) A somewhat different sequence was reported previously (62) without localization of the termini.
- n) The location of the termini is uncertain.
- o) The location of the termini is not according to the original paper (53) but according to (54).

Table 2. Symbols denoting sequence heterogeneity

<u>Symbol</u>	<u>Nucleotides present</u>	<u>Symbol</u>	<u>Nucleotides present</u>
R	A and G	H	A, U, and C
Y	U and C	B	G, U, and C
W	A and U	V	A, G, and C
S	G and C	D	A, G, and U
M	A and C	N	A, G, U, and C
K	G and U		

These symbols are used only to denote positional heterogeneity found by RNA sequencing (see text).

Posttranscriptional nucleoside modifications are not mentioned in the alignment, mainly because there exist no conventional one-letter abbreviations for all the modified residues encountered in srRNA. Moreover, data on modified residues exist only for a few sequences. References to papers containing such data are mentioned in Table 1.

The alignment of sequences for optimal homology has been obtained without the aid of a computer program. It is based on the existence of conserved sequence blocks, and on homology in secondary structure (see below). Certain sections of several mitochondrial srRNA sequences are printed in more condensed format, without a space between the characters. These are sections where it is especially difficult to discern homology with srRNAs from other species. It has been attempted, however, to optimize homology between corresponding condensed-format sections in species that are adjacent in the alignment. There are other areas where sequence alignment is rather arbitrary, often corresponding to those loops in the secondary structure models (see below) for which loop length is variable from one species to the other. In general, sequence alignment is much more questionable in the variable areas than in the conserved areas of the secondary structure models. The areas of dubious homology may shrink in the future by the application of alignment algorithms or/and availability of additional sequences for comparison. For the moment being, their existence should be taken into account in applications of the data for studies in molecular evolution.

SECONDARY STRUCTURE

The boxes superimposed on the sequence alignment enclose segments presumably involved in base pairing. The boxes enclosing the complementary sequences forming helix 1 are numbered 1 and 1', and so on. Internal loops and bulges are indicated by nested boxes. Bases that may belong to pairs other than G·C, A·U or G·U (non-standard-, non-canonical or odd base pairs) are put in parentheses. Two sets of helix numbers appear in the alignment. One set is listed under the uppermost block consisting of eukaryotic srRNA sequences, and applies to these structures. A second set of numbers is listed between the eubacterial and the plastidial srRNA sequence blocks and applies to all prokaryotic structures.

In the case of prokaryotic srRNAs, the model proposed by Noller, Woese and collaborators (65,66) is followed. Bacterial and chloroplast srRNAs fit well into this model, but much variability is observed in the structure of mitochondrial srRNAs, which show deletions as well as insertions of secondary

structure areas with respect to the bacterial model. Insertions prevail in plant mitochondrial srRNAs (56) which have chain lengths approaching 2000 nucleotides, while deletions prevail in animal mitochondrial srRNAs (46,47) which have chain lengths of less than 1000 nucleotides. Among the srRNAs of fungal and protist mitochondria, both insertions and deletions account for structural variability with respect to the bacterial secondary structure model. Eukaryotic (cytoplasmic) srRNAs are fitted into a variant model (7) related to those proposed for Xenopus laevis (6) and Dictyostelium discoideum (67) srRNAs. This model was shown to be applicable to srRNAs from 6 eukaryotes in a previous compilation (7) of srRNA sequences, and proved to remain applicable to all of the 14 eukaryotic structures presently published.

The prokaryotic and the eukaryotic srRNA models applied to the sequence alignment are illustrated in Fig. 1 and 2. The two models are for a large part homologous, but there are two kinds of differences, both localized in some of the areas of variable primary and secondary structure that are labeled V1 to V8. The first kind of difference is found in areas V2 and V4, where eukaryotic srRNAs possess extra helices with respect to prokaryotic srRNAs. The second kind of differences resides in area V5, which shows variability of a different type. Here, it is not the absence or presence of a helix that makes the difference, but the topology in which the helices are folded, in other words the order of occurrence of complementary sequences when the chain is scanned from 5'- to 3'-end. It should be noted that eukaryotic srRNA sequences can also be fitted into a model having the prokaryotic topology in area V5, although at the cost of forming a less stable structure. A recent review on srRNA structure (66) lists such models for five eukaryotic srRNAs, but base pairing in the helix corresponding to P18 (our numbering system) is questionable. Conversely, certain prokaryotic srRNAs can be fitted into a eukaryotic-type model, which again does not seem to fit them as nicely as the prokaryotic model, but we have not investigated this possibility for all presently known prokaryotic srRNAs. The possibility that area V5 may undergo a structural switch, or may be a case of evolutionary divergence in secondary structure, has been discussed (7).

The helix numbering system (7) adopted in the alignment and in Figures 1 and 2 attributes a separate number to each helical area, which may consist of an uninterrupted helix, or a set of helix segments connected by bulges or interior loops. Helices (or helical areas) bear a different number only when separated by a multibranch loop (e.g. helices 7 and 8), or when separated by a single-stranded segment which does not form a loop (e.g. helices 23 and 39).

Numbers are attributed in the order of occurrence of the 5'-proximal strand of the helix in the sequence. Helices that are common to the prokaryotic and eukaryotic models bear a single number from 1 to 40. In areas where the two models differ in pairing scheme, helices are numbered in the format Pa or Pa-b (prokaryotic srRNAs) and Ea or Ea-b (eukaryotic srRNAs) where a is the number of the preceding common helix and b is a serial number. As an example, E9-1 and E9-2 are eukaryote-specific helices encountered after common helix 9 when scanning sequences from the 5'-end. The common helices are not universal in the sense that some are missing in certain mitochondrial srRNAs, but they are present in all hitherto examined eukaryotic, archaeobacterial, eubacterial and plastidial srRNAs. Future refinements of the secondary structure models may bring to light additional common helices and necessitate a revision of the numbering system.

In sorting out the base pairing schemes superimposed on the sequence alignment, we were often aided by the availability of secondary structure models proposed in the paper reporting the srRNA sequence or in (66). However, we did not necessarily adopt such models entirely. There are two main differences between the secondary structures proposed here and a set of 20 srRNA secondary structure models published by Gutell et al. (66). In area V5, we assume a different topology in eukaryotes and prokaryotes whereas they assume the same folding. Conversely, in area V3 we assume the presence of the same topology in eukaryotes and prokaryotes while they assume a different one. In the case of trypanosome mitochondrial srRNAs (52-54), it is especially difficult to propose a meaningful alignment and secondary structure model. Not only are the sequences exceptionally short, but they are also very A-U rich and hence very monotonous. Only a limited set of segments of these sequences is aligned with other srRNA sequences according to the proposal of Sloof et al. (52), the remainder being printed in condensed format.

AVAILABILITY AND ACCURACY OF THE DATA

The sequences presented here were entered in our database and checked to the best of our ability for errors. They were not checked against a central data base such as EMBL, one reason being that many of the sequences are published very recently and there is some delay before they appear in central data bases. We intend to carry out such checks in the future and to update the compilation regularly. Authors are invited to inform us about errors and omissions of published sequences, and to send reprints of their newly published sequences.

Fig. 1. Secondary structure model applied to prokaryotic srRNAs. Helices, defined as described in the text, are numbered in the order of occurrence from the 5'- to the 3'-terminus, symbolized respectively by a dot and an arrowhead. Forty helices present in eukaryotic as well as prokaryotic srRNAs (with the exception of certain mitochondrial srRNAs) are hitherto distinguished. Additional helices specific for prokaryotic srRNAs bear a number preceded by P. A probable interaction between the sequence preceding helix 1 and a segment of the multibranched loop separating helices 1 and 2 is not numbered. Bold lines are used for areas relatively conserved in primary and secondary structure. Thin lines are used for variable areas, numbered V1 to V8.

Fig. 2. Secondary structure model applied to eukaryotic srRNAs. Conventions and numbering system are as in Fig. 1, except that helices specific for eukaryotic srRNAs bear a number preceded by E.

The srRNA sequence compilation is available to interested scientists in computer-readable form, either on magnetic tape or on floppy disks. An updated version of a file containing 5 S rRNA sequences, in the format used in previous publications (75) is also available. It is our hope that these compilations will be an aid for studies on rRNA structure and molecular evolution.

[illegible]

	229	230	240	250	260	270	280	
GGUCG	CUCGCUCCUCUCCUACU	UG(G)-AU	ACUGUG	UAAU-U	CUAGAGCUAAU(A)-	CAUGCCBA	CGGGC	H.s. 1
GGUCG	CUCGCUCCUCUCCUACU	UG(G)-AU	AACUGUG	UAAU-U	CUAGAGCUAAU(A)-	CAUGCCBA	CGGGC	N.h. 2
G-UCG	CUCGCUCCUCUCCUACU	UG(G)-AU	AACUGUG	UAAU-U	CUAGAGCUAAU(A)-	CAUGCCBA	CGGGC	R.n. 3
GGUCG	CUCGCUCCUCUCCUACU	UG(G)-AU	AACUGUG	UAAU-U	CUAGAGCUAAU(A)-	CAUGCCBA	CGGGC	R.n. 4
GGUCG	CUCGCUCCUCUCCUACU	UG(G)-AU	AACUGUG	UAAU-U	CUAGAGCUAAU(A)-	CAUGCCBA	CGGGC	C.d. 5
GAUCG	CUCC--AUCUUGUACU	UG(G)-AU	AACUGUG	UAAU-U	CUAGAGCUAAU(A)-	CAUGCCBA	CGGGC	X.l. 6
GAUCG	CUCC--AUAUCCUACU	UG(G)-AU	AACUGUG	UAAU-U	CUAGAGCUAAU(A)-	CAUGCCBA	CGGGC	A.s. 7
GAUCG	CUCC--UUUACUAC	UG(G)-AU	AACUGUG	UAAU-U	CUAGAGCUAAU(A)-	CAUGCCBA	CGGGC	S.c. 8
GACUUG	CGG--GUUUUACCUU	UG(G)-AU	ACCGCAGUAAU	U-CCG	GGCUAAU(A)-	CAUGCCBA	CGGGC	B.d. 9
GAUAAUCAA	--AUUUUAC	UG(G)-AU	ACCGCAGUAAU	U-CCG	GGCUAAU(A)-	CAUGCCBA	CGGGC	O.n. 10
GAUAAUCAA	--AUUUUAC	UG(G)-AU	ACCGCAGUAAU	U-CCG	GGCUAAU(A)-	CAUGCCBA	CGGGC	S.p. 11
GAUAAUCAA	--AUUUUAC	UG(G)-AU	ACCGCAGUAAU	U-CCG	GGCUAAU(A)-	CAUGCCBA	CGGGC	T.t. 12
GAUGGUAACG	--UGCUUAC	UG(G)-AU	ACCGCAGUAAU	U-CCG	GGCUAAU(A)-	CAUGCCBA	CGGGC	O.s. 13
GAUGGUAACG	--UGCUUAC	UG(G)-AU	ACCGCAGUAAU	U-CCG	GGCUAAU(A)-	CAUGCCBA	CGGGC	Z.o. 14
GAUGGUAACG	--UGCUUAC	UG(G)-AU	ACCGCAGUAAU	U-CCG	GGCUAAU(A)-	CAUGCCBA	CGGGC	G.n. 15
		8		8'		9		
UGUGG	-----A	CGG(G)AUA	UCUCGG	GAUA-U	CUAGAGCUAAU(C)CG	AUAAC	CGUUUGC	H.c. 16
UGUGG	-----A	CGG(G)AUA	UCUCGG	GAUA-U	CUAGAGCUAAU(C)CG	AUAAC	CGUUUGC	H.h. 17
UACAG	-----A	GAAC(C)UA	CUUCGG	GAUA-U	CUAGAGCUAAU(C)CG	AUAAC	CGUUUGC	H.h. 18
UUGG	-----A	CC(C)GGAUA	CUUCGG	GAUA-U	CUAGAGCUAAU(C)CG	AUAAC	CGUUUGC	H.h. 19
UGAAG	-----A	GGGAGGAUA	UCCCGGG	GAUA-U	CUAGAGCUAAU(C)CG	AUAAC	CGUUUGC	H.h. 20
U(CA)GG	-----A	GGGAGGAUA	UCCCGGG	GAUA-U	CUAGAGCUAAU(C)CG	AUAAC	CGUUUGC	H.h. 21
UAGG	-----A	UGG(G)UA	CCUUGGG	GAUA-U	CUAGAGCUAAU(C)CG	AUAAC	CGUUUGC	N.v. 22
UCGG	-----A	CGGGGAUA	CCCGGG	GAUA-U	CUAGAGCUAAU(C)CG	AUAAC	CGUUUGC	S.s. 23
UCGG	-----A	CGGGGAUA	CCCGGG	GAUA-U	CUAGAGCUAAU(C)CG	AUAAC	CGUUUGC	T.t. 24
		8		8'		9		
UGCCU	-----G	GGG(G)UA	GCUCGG	GAUA-U	CUAGAGCUAAU(C)CG	AUAAC	CGUUUGC	A.t. 25
AGUAU	-----G	GGG(G)UA	GCUCGG	GAUA-U	CUAGAGCUAAU(C)CG	AUAAC	CGUUUGC	P.t. 26
GAUGA	-----G	GGG(G)UA	GCUCGG	GAUA-U	CUAGAGCUAAU(C)CG	AUAAC	CGUUUGC	E.c. 27
GAUGA	-----G	GGG(G)UA	GCUCGG	GAUA-U	CUAGAGCUAAU(C)CG	AUAAC	CGUUUGC	P.v. 28
GAUGA	-----G	GGG(G)UA	GCUCGG	GAUA-U	CUAGAGCUAAU(C)CG	AUAAC	CGUUUGC	S.p. 29
GAUGA	-----G	GGG(G)UA	GCUCGG	GAUA-U	CUAGAGCUAAU(C)CG	AUAAC	CGUUUGC	H.c. 30
GAUGA	-----G	GGG(G)UA	GCUCGG	GAUA-U	CUAGAGCUAAU(C)CG	AUAAC	CGUUUGC	H.c. 31
GAUGA	-----G	GGG(G)UA	GCUCGG	GAUA-U	CUAGAGCUAAU(C)CG	AUAAC	CGUUUGC	H.c. 32
GAUGA	-----G	GGG(G)UA	GCUCGG	GAUA-U	CUAGAGCUAAU(C)CG	AUAAC	CGUUUGC	B.f. 33
GAUGA	-----G	GGG(G)UA	GCUCGG	GAUA-U	CUAGAGCUAAU(C)CG	AUAAC	CGUUUGC	F.h. 34
GAUGA	-----G	GGG(G)UA	GCUCGG	GAUA-U	CUAGAGCUAAU(C)CG	AUAAC	CGUUUGC	B.d. 35
GAUGA	-----G	GGG(G)UA	GCUCGG	GAUA-U	CUAGAGCUAAU(C)CG	AUAAC	CGUUUGC	N.x. 36
ACAGG	-----G	GGG(G)UA	GCUCGG	GAUA-U	CUAGAGCUAAU(C)CG	AUAAC	CGUUUGC	A.n. 37
		8		8'		9		
UGGGG	-----G	GGG(G)UA	CAAGG	GAUA-U	CGUUUGCUAAU(C)CG	AUAAC	CGUUUGC	E.s. 38
AUCGG	-----G	GGG(G)UA	CAAGG	GAUA-U	CGUUUGCUAAU(C)CG	AUAAC	CGUUUGC	C.r. 39
UGGGG	-----G	GGG(G)UA	CAAGG	GAUA-U	CGUUUGCUAAU(C)CG	AUAAC	CGUUUGC	Z.n. 40
UGGGG	-----G	GGG(G)UA	CAAGG	GAUA-U	CGUUUGCUAAU(C)CG	AUAAC	CGUUUGC	N.t. 41
		8		8'		9		
								H.s. 42
								N.h. 43
								R

290 300 310 320 330 340 350

1 H.s. G C U G A C C C C C U U - - C G C G G G G G G G A U G C G U G C A U U U A U C A G A U C A A A - A C C A A C C C G - G U C A G C C C C C U C
2 H.m. G C U G A C C C C C C U U C C C G - G G G G G G G A U G C G U G C A U U U A U C A G A U C A A A - A C C A A C C C G - G U G A G C C C C C U C
3 R.n. G C U G A C C C C C C U U C C C G U G G G G G G G A C G C G U G C A U U U A U C A G A U C A A A - A C C A A C C C G - G U C A G C C C C C U C
4 R.n. G C U G A C C C C C C U U C C C G U G G G G G G A C G C G U G C A U U U A U C A G A U C A A A - A C C A A C C C G - G U C A G C C C C C U C
5 O.c. G C U G A C C C C C U U - - G U G C G G G A U G C G U G C A U U U A U C A G A U C A A A - A C C A A C C C G C G U C A G C C C C C U C
6 X.l. G C U G A C C C C C U U - - C C A G G G A U G C G U G C A U U U A U C A G A U C A A A - A C C A A C C C G - G G G C C C C C -
7 A.s. C C C A A C U - - U C A C G G A A G G G G U G C U U U U A U A G A U C A A G - A C C A A U C - G - G G G C U U C -
8 S.c. C U C G A C C C - - U U U G G A A G G A G A U G U A U U U A U A G A U A A A A A A U C A A U G C U U C -
9 D.d. G G U G A C U G - - G C A A C G G A A G C U C A G C G A U U A U U A G C A U U C U - A C C A A U G C U U C -
10 O.n. C U G A C U U - - U U A G G A A G G G C U G U A U U U A U A G A U A A C A A A U C A A U A U C C U U -
11 S.p. C C C A C U U - - U G A G G A A G G G C U G U A U U U A U A G A U A A C A A A U C A A U A U C C U U -
12 T.t. U C C G U G U C - - C U G C G A C C G G A A - C G U A U U U A U A G A U A A U U A G A C C A A U C G A G A -
13 D.s. C C C G A C U U C - - G G G A G G A G C G A U U U A U A G A U A A A A G G C U G A C G C G - G G C U C C -
14 Z.m. C C C A C U U C - - G G G A G G G G C G C A U U U A U A G A U A A A A G G C U G A C G C G - G G C U C U -
15 G.m. C C C G A C U U C - - G G A A G G G A U G C A U U U A U A G A U A A A A G G C U A C A - G G C U C U -

E9-1 E9-1' E9-2

16 M.c. U C C U G G A -
17 H.h. U C C U G G A -
18 N.v. U G C U G G A A U
19 H.m. U G U G G A A U
20 N.h. G A C U G G A A U
21 N.v. G U C U G G A A C
22 N.f. G C C U G G A A U
23 S.s. U C C U G G A A U
24 T.t. C G C U G G A A

25 A.t. -
26 P.t. -
27 E.c. -
28 P.v. -
29 D.s. - A A C C G C
30 M.c. - A A G G A C
31 M.c. U U A U C G C
32 M.s. U U A U C G C
33 B.f. -
34 F.h. -
35 D.d. -
36 M.x. U U U C U U -
37 A.n. -

38 E.s. A U C U A U G U G A -
39 C.r. -
40 Z.m. -
41 N.t. -

42 H.s. -
43 H.m. -
44 R.n. -
45 B.t. -
46 X.l. -
47 D.v. -
48 A.n. -
49 S.c. -
50 P.p. -
51 P.t. -
52 C.f. -
53 T.b. -
54 L.t. -
55 Z.m. -
56 T.a. -
57 O.s. -

290 300 310 320 330 340 350

r88

500	510	520	530	540	550	560	
CUACCAUGGUGACCAACGGUGA	CGGGGAUUCAGGGUUCG-AUUC	CGGAGAGAGGGAAGCUGAGAAAC	GGCUA) H.s.	1			
CUACCAUGGUGACCAACGGUGA	CGGGGAUUCAGGGUUCG-AUUC	CGGAGAGAGGGAAGCUGAGAAAC	GGCUA) M.m.	2			
CUACCAUGGUGACCAACGGUGA	CGGGGAUUCAGGGUUCG-AUUC	CGGAGAGAGGGAAGCUGAGAAAC	GGCUA) R.n.	3			
CUACCAUGGUGACCAACGGUGA	CGGGGAUUCAGGGUUCG-AUUC	CGGAGAGAGGGAAGCUGAGAAAC	GGCUA) R.n.	4			
CUACCAUGGUGACCAACGGUGA	CGGGGAUUCAGGGUUCG-AUUC	CGGAGAGAGGGAAGCUGAGAAAC	GGCUA) O.c.	5			
CUACCAUGGUGACCAACGGUGA	CGGGGAUUCAGGGUUCG-AUUC	CGGAGAGAGGGAAGCUGAGAAAC	GGCUA) X.l.	6			
CUACCAUGGUGACCAACGGUGA	CGGGGAUUCAGGGUUCG-AUUC	CGGAGAGAGGGAAGCUGAGAAAC	GGCUA) A.s.	7			
CUACCAUGGUGACCAACGGUGA	CGGGGAUUCAGGGUUCG-AUUC	CGGAGAGAGGGAAGCUGAGAAAC	GGCUA) S.c.	8			
CUACCAUGGUGACCAACGGUGA	CGGGGAUUCAGGGUUCG-AUUC	CGGAGAGAGGGAAGCUGAGAAAC	GGCUA) D.d.	9			
CUACCAUGGUGACCAACGGUGA	CGGGGAUUCAGGGUUCG-AUUC	CGGAGAGAGGGAAGCUGAGAAAC	GGCUA) D.n.	10			
CUACCAUGGUGACCAACGGUGA	CGGGGAUUCAGGGUUCG-AUUC	CGGAGAGAGGGAAGCUGAGAAAC	GGCUA) S.p.	11			
CUACCAUGGUGACCAACGGUGA	CGGGGAUUCAGGGUUCG-AUUC	CGGAGAGAGGGAAGCUGAGAAAC	GGCUA) T.t.	12			
CUACCAUGGUGACCAACGGUGA	CGGGGAUUCAGGGUUCG-AUUC	CGGAGAGAGGGAAGCUGAGAAAC	GGCUA) O.s.	13			
CUACCAUGGUGACCAACGGUGA	CGGGGAUUCAGGGUUCG-AUUC	CGGAGAGAGGGAAGCUGAGAAAC	GGCUA) Z.n.	14			
CUACCAUGGUGACCAACGGUGA	CGGGGAUUCAGGGUUCG-AUUC	CGGAGAGAGGGAAGCUGAGAAAC	GGCUA) G.n.	15			
11'	12	12'	6'	13	13'		
CCACCGUGCCCAUAUACGGU	ACGGGUGUGSAGAGCAAG-AGCCG	AGAGAGAGGGAUUCUGAGACA	AGAUUC	H.c.	16		
CCACCGUGCCCAUAUACGGU	ACGGGUGUGSAGAGCAAG-AGCCG	AGAGAGAGGGAUUCUGAGACA	AGAUUC	H.h.	17		
CCACCGUGCCCAUAUACGGU	ACGGGUGUGSAGAGCAAG-AGCCG	AGAGAGAGGGAUUCUGAGACA	AGAUUC	H.v.	18		
CCACCGUGCCCAUAUACGGU	ACGGGUGUGSAGAGCAAG-AGCCG	AGAGAGAGGGAUUCUGAGACA	AGAUUC	H.n.	19		
CCACCAAGCCUUGUACGGU	ACGGGUGUGSAGAGCAAG-AGCCG	AGAGAGAGGGAUUCUGAGACA	AGAUUC	H.h.	20		
CCACCAAGCCUUGUACGGU	ACGGGUGUGSAGAGCAAG-AGCCG	AGAGAGAGGGAUUCUGAGACA	AGAUUC	H.v.	21		
CCACCAAGCCUUGUACGGU	ACGGGUGUGSAGAGCAAG-AGCCG	AGAGAGAGGGAUUCUGAGACA	AGAUUC	H.f.	22		
CCACCAAGCCUUGUACGGU	ACGGGUGUGSAGAGCAAG-AGCCG	AGAGAGAGGGAUUCUGAGACA	AGAUUC	S.s.	23		
CCACCAAGCCUUGUACGGU	ACGGGUGUGSAGAGCAAG-AGCCG	AGAGAGAGGGAUUCUGAGACA	AGAUUC	T.t.	24		
11'	12	12'	6'	13	13'		
CUACCAAGGCGAUGAUCAGU	AGCUGGUCUGSAGAGGAUG-AGCAGC	CAUUGGGACUGAGACA	CGGCC	A.t.	25		
CUACCAAGGCGAUGAUCAGU	AGCUGGUCUGSAGAGGAUG-AGCAGC	CAUUGGGACUGAGACA	CGGCC	P.t.	26		
CUACCAAGGCGAUGAUCAGU	AGCUGGUCUGSAGAGGAUG-AGCAGC	CAUUGGGACUGAGACA	CGGCC	E.c.	27		
CUACCAAGGCGAUGAUCAGU	AGCUGGUCUGSAGAGGAUG-AGCAGC	CAUUGGGACUGAGACA	CGGCC	P.v.	28		
CUACCAAGGCGAUGAUCAGU	AGCUGGUCUGSAGAGGAUG-AGCAGC	CAUUGGGACUGAGACA	CGGCC	B.s.	29		
CUACCAAGGCGAUGAUCAGU	AGCUGGUCUGSAGAGGAUG-AGCAGC	CAUUGGGACUGAGACA	CGGCC	H.c.	30		
CUACCAAGGCGAUGAUCAGU	AGCUGGUCUGSAGAGGAUG-AGCAGC	CAUUGGGACUGAGACA	CGGCC	H.s.	31		
CUACCAAGGCGAUGAUCAGU	AGCUGGUCUGSAGAGGAUG-AGCAGC	CAUUGGGACUGAGACA	CGGCC	B.f.	32		
CUACCAAGGCGAUGAUCAGU	AGCUGGUCUGSAGAGGAUG-AGCAGC	CAUUGGGACUGAGACA	CGGCC	F.h.	33		
CUACCAAGGCGAUGAUCAGU	AGCUGGUCUGSAGAGGAUG-AGCAGC	CAUUGGGACUGAGACA	CGGCC	D.d.	34		
CUACCAAGGCGAUGAUCAGU	AGCUGGUCUGSAGAGGAUG-AGCAGC	CAUUGGGACUGAGACA	CGGCC	M.x.	35		
CUACCAAGGCGAUGAUCAGU	AGCUGGUCUGSAGAGGAUG-AGCAGC	CAUUGGGACUGAGACA	CGGCC	A.n.	36		
11'	12	12'	6'	13	13'		
CUACCAAGGCGAUGAUCAGU	AGCUGGUCUGSAGAGGAUG-AGCAGC	CAUUGGGACUGAGACA	CGGCC	E.s.	38		
CUACCAAGGCGAUGAUCAGU	AGCUGGUCUGSAGAGGAUG-AGCAGC	CAUUGGGACUGAGACA	CGGCC	C.r.	39		
CUACCAAGGCGAUGAUCAGU	AGCUGGUCUGSAGAGGAUG-AGCAGC	CAUUGGGACUGAGACA	CGGCC	Z.n.	40		
CUACCAAGGCGAUGAUCAGU	AGCUGGUCUGSAGAGGAUG-AGCAGC	CAUUGGGACUGAGACA	CGGCC	N.t.	41		
11'	12	12'	6'	13	13'		
AGGAUCAAGCAGCAGCAAAUGCA	GCUCGAAACGCUUAG-CCUAGC	CAUA		H.s.	42		
AGGAUCAAGCAGCAGCAAAUGCA	GCUCGAAACGCUUAG-CCUAGC	CAUA		M.m.	43		
AGGAUCAAGCAGCAGCAAAUGCA	GCUCGAAACGCUUAG-CCUAGC	CAUA		R.n.	44		
AGGAUCAAGCAGCAGCAAAUGCA	GCUCGAAACGCUUAG-CCUAGC	CAUA		B.t.	45		
AGGAUCAAGCAGCAGCAAAUGCA	GCUCGAAACGCUUAG-CCUAGC	CAUA		X.l.	46		
AGGAUCAAGCAGCAGCAAAUGCA	GCUCGAAACGCUUAG-CCUAGC	CAUA		D.v.	47		
AGGAUCAAGCAGCAGCAAAUGCA	GCUCGAAACGCUUAG-CCUAGC	CAUA		A.n.	48		
AGGAUCAAGCAGCAGCAAAUGCA	GCUCGAAACGCUUAG-CCUAGC	CAUA		S.c.	49		
AGGAUCAAGCAGCAGCAAAUGCA	GCUCGAAACGCUUAG-CCUAGC	CAUA		P.p.	50		
AGGAUCAAGCAGCAGCAAAUGCA	GCUCGAAACGCUUAG-CCUAGC	CAUA		P.t.	51		
AGGAUCAAGCAGCAGCAAAUGCA	GCUCGAAACGCUUAG-CCUAGC	CAUA		C.f.	52		
AGGAUCAAGCAGCAGCAAAUGCA	GCUCGAAACGCUUAG-CCUAGC	CAUA		T.b.	53		
AGGAUCAAGCAGCAGCAAAUGCA	GCUCGAAACGCUUAG-CCUAGC	CAUA		L.t.	54		
AGGAUCAAGCAGCAGCAAAUGCA	GCUCGAAACGCUUAG-CCUAGC	CAUA		Z.n.	55		
AGGAUCAAGCAGCAGCAAAUGCA	GCUCGAAACGCUUAG-CCUAGC	CAUA		T.a.	56		
AGGAUCAAGCAGCAGCAAAUGCA	GCUCGAAACGCUUAG-CCUAGC	CAUA		O.s.	57		
500	510	520	530	540	550	560

790	790	800	810	820	830	840		
U G G A G U G G C A A G U C	- U G G U G C C	A G C A G C C G C	G G U A A U U C C A G C U C C A A U	A G C G U A U (A) U A A	A A U G U G C U G C A	H.s.	1	
U G G A G U G G C A A G U C	- U G G U G C C	A G C A G C C G C	G G U A A U U C C A G C U C C A A U	A G C G U A U (A) U A A	A A U G U G C U G C A	M.m.	2	
U G G A G U G G C A A G U C	- U G G U G C C	A G C A G C C G C	G G U A A U U C C A G C U C C A A U	A G C G U A U (A) U A A	A A U G U G C U G C A	R.n.	3	
U G G A G U G G C A A G U C	- U G G U G C C	A G C A G C C G C	G G U A A U U C C A G C U C C A A U	A G C G U A U (A) U A A	A A U G U G C U G C A	R.n.	4	
U G G A G U G G C A A G U C	- U G G U G C C	A G C A G C C G C	G G U A A U U C C A G C U C C A A U	A G C G C A C (A) U C A A	A A U G U G C U G C A	D.c.	5	
U G G A G U G G C A A G U C	- U G G U G C C	A G C A G C C G C	G G U A A U U C C A G C U C C A A U	A G C G U A U (A) U A A	A A U G U G C U G C A	X.l.	6	
U G G A G U G G C A A G U C	- U G G U G C C	A G C A G C C G C	G G U A A U U C C A G C U C C A A U	A G C G U A U (A) U A A	A A U G U G C U G C A	A.s.	7	
U G G A G U G G C A A G U C	- U G G U G C C	A G C A G C C G C	G G U A A U U C C A G C U C C A A U	A G C G U A U A U A A	A A U G U G U G C A	S.c.	8	
U G X A G U G G C A A G U C	- U G G U G C C	A G C X G C C C G	G G U A A U U C C A G C U C C A A U	A G C A U A U A C A A	A A U G U G U G C A	D.d.	9	
U G G A G U G G C A A G U C	- U G G U G C C	A G C A G C C G C	G G U A A U U C C A G C U C C A A U	A G C G U A U A U A A	A A U G U G U G C A	D.n.	10	
U G G A G U G G C A A G U C	- U G G U G C C	A G C A G C C G C	G G U A A U U C C A G C U C C A A U	A G C A U A U A C A A	A A U G U G U G C A	S.p.	11	
U G G A G U G G C A A G U C	- U G G U G C C	A G C A G C C G C	G G U A A U U C C A G C U C C A A U	A G C G U A U (A) U A A	A A U G U G U G C A	T.t.	12	
U G G A G U G G C A A G U C	- U G G U G C C	A G C A G C C G C	G G U A A U U C C A G C U C C A A U	A G C G U A U A U A A	A A U G U G U G C A	O.c.	13	
U G G A G U G G C A A G U C	- U G G U G C C	A G C A G C C G C	G G U A A U U C C A G C U C C A A U	A G C G U A U A U A A	A A U G U G U G C A	Z.n.	14	
U G (A) A G U G G C A A G U C	- U G G U G C C	A G C A G C C G C	G G U A A U U C C A G C U C C A A U	A G C G U A U A U A A	A A U G U G U G C A	G.n.	15	
18						18'	2'	E18

G A C U G U G G C A A G A C	- C G G U G C C	A G C C G C C G C	G G U A A U A C C G G C A G U C C	A G U G A U G G C C	G A U C U U A U U G G G	H.c.	16
G A C U G U G G C A A G A C	- C G G U G C C	A G C C G C C G C	G G U A A U A C C G G C A G U C C	A G U G A U G G C C	G A U C U U A U U G G G	H.h.	17
A C U G U G G C A A G A C	- C G G U G C C	A G C C G C C G C	G G U A A U A C C G G C A G U C C	A G U G A U G G C C	G A U C U U A U U G G G	H.v.	18
A C U G U G G C A A G A C	- C G G U G C C	A G C C G C C G C	G G U A A U A C C G G C A G U C C	A G U G A U G G C C	G A U C U U A U U G G G	H.h.	19
G G C X G U G G C A A G X X	- X X X X X X X	X X X X X X X	X G U A A U A C C G X X X X X X	A G U G G U G G C C	G C U A U A U C U G G G	N.h.	20
G G C U G U G G C A A G U C	- C G G U G C C	A G C A G C C G C	G G U A A U A C C G A C G G C C C	A G U G G U G G C C	G A C U U U A U U G G G	M.v.	21
A C U G U G G C A A G U C	- C G G U G C C	A G C C G C C G C	G G U A A U A C C G G C A G U C C	A G U G G U G G C C	G U U U U A U U G G G	N.f.	22
C G G G U G G C A A G U C	- U G G U G C C	A G C C G C C G C	G G U A A U A C C A G C U C C G C	A G U G G U C G G G	G U G A U A C U G G G	S.s.	23
C G G G U G G U A A G U C	- U G G U G C C	A G C C G C C G C	G G U A A U A C C A G C C C G C	A G U G G U C A G G	G U G A U A C U G G G	T.t.	24

C C C C - G G C U A A C U	- U G G U G C C	A G C A G C C G C	G G U A A U A C G A A G G G G G C	A G C G U G U U C	G G A A U A C U G G G	A.t.	25	
C A C C - G G C U A A C U	- A C G U G C C	A G C A G C C G C	G G U A A U A C G A A G G G G G C	A A C G U U A A U C	G G A A U A C U G G G	P.t.	26	
C A C C - G G C U A A C U	- C G G U G C C	A G C C G C C G C	G G U A A U A C G A A G G G G G C	A A C G U U A A U C	G G A A U A C U G G G	E.c.	27	
C C A C - G G C U A A C U	- C G G U G C C	A G C A G C C G C	G G U A A U A C G A A G G G G G C	A A C G U U A A U C	G G A A U A C U G G G	P.v.	28	
C C A C - G G C U A A C U	- A C G U G C C	A G C A G C C G C	G G U A A U A C G A A G G G G G C	A A C G U U A A U C	G G A A U A C U G G G	B.s.	29	
C C C C - G G C U A A C U	- A C G U G C C	A G C A G C C G C	G G U A A U A C G A A G G G G G C	A A C G U G U U C C	G G A A U A C U G G G	H.c.	30	
C C A C - G G C U A A C U	- A U G U G C C	A G C A G C C G C	G G U A A U A C A U A G G U G G C	A A C G U U A A U C	G G A A U A C U G G G	N.h.	31	
C C A C - G G C U A A C U	- A U G U G C C	A G C A G C C G C	G G U A A U A C A U A G G U G G C	A A C G U U A A U C	G G A A U A C U G G G	N.h.	32	
G A U C - G G C U A A C U	- C G G U G C C	A G C A G C C G C	G G U A A U A C G A A G A U C C	A A C G U U A A U C	G G A A U A C U G G G	B.f.	33	
G A U C - G G C U A A C U	- C G G U G C C	A G C A G C C G C	G G U A A U A C G A A G A U C C	A A C G U U A A U	G G A A U A C U G G G	F.d.	34	
C A C C - G G C U A A C U	- C G G U G C C	A G C A G C C G C	G G U A A U A C G A A G G G G G C	A A C G U U A A U C	G G A A U A C U G G G	D.b.	35	
C A C C - G G C U A A C U	- C G G U G C C	A G C A G C C G C	G G U A A U A C G A A G G G G G C	A A C G U G U G U C	G G A A U A C U G G G	N.x.	36	
C C U C - G G C U A A C U	- C G G U G C C	A G C A G C C G C	G G U A A U A C G G G A G A G G C	A A C G U U A A U C	G G A A U A C U G G G	A.n.	37	
18						18'	2'	P18

C A U C - G G C U A A C U	- C G G U G C C	A G C A G C C G C	G G U A A U A C G G G A G A U G C	A G C G U U A U C C	G G A A U A C U U G G G	E.s.	38
C A C C - G G C U A A C U	- C U G U G C C	A G C A G C C G C	G G U A A U A C A G A G G G G G C	A A C G U U G U C C	G G A A U A C U U G G G	C.r.	39
C A U C - G G C U A A C U	- C U G U G C C	A G C A G C C G C	G G U A A U A C A G A G A G U G C	A A C G U U A U C C	G G A A U A C U U G G G	Z.n.	40
C A U C - G G C U A A C U	- C U G U G C C	A G C A G C C G C	G G U A A U A C A G A G A G U G C	A A C G U U A U C C	G G A A U A C U U G G G	N.h.	41

G G U U - G G C U A A U U	- U C G U G C C	A G C C A C C G C	G G U C A C A C G A U U A A C C C	A G U C (A) A U A G A A	U C C - - - - - G G	H.s.	42
G G U U - G G U A A A U U	- U C G U G C C	A G C C A C C G C	G G U C A U A C G A U U A A C C C	A A C (U) A A A U A U A	C U C - - - - - G G	N.h.	43
G G U U - G G U A A A U U	- U C G U G C C	A G C C A C C G C	G G U C A U A C G A U U A A C C C	A A C (U) A A A U A U A	- - - - - U U U C (G) G	R.n.	44
G G U U - G G U A A A U U	- U C G U G C C	A G C C A C C G C	G G U C A U A C G A U U A A C C C	A A G (C) U A A A C G G	- - - - - G U A C (G) G	B.t.	45
A G U U - G A C C A A A U C	- U G G U G C C	A G C C A C C G C	G G U U A U A C G A A A A A C U C	A A G U U U A C A C A	- - - - - A U U C (G) G G	X.l.	46
A U A U - G A C C A A A U C	- U G G U G C C	A G C C A C C G C	G G U U A U A C C A A U A A U A C	A A G U A A A A U U U	- - - - - U U A G U A U C	D.v.	47
U C U U A C A A A A A U C	- A C G U G C C	A G C A G C C G C	G G U A A C A C A A A G A A G G C	A A G G U U A A U C	A C U C U U A U U A G	A.n.	48
U C C U A C A A A A A U C	- U G G U G C C	A G C A G C C G C	G G U A A C A C A A A G A A G G C	A A G G U U A A U C	A U A A U - - - - - G S	C.c.	49
C U A U G G C C A A A A C	- A U G U G C C	A G C A G C C G C	G G U A A A C A A G A A U A G C	A A G G U U A A U C	G U C U C U A U U G A G	P.p.	50
C U A U G G C C A A A A C	- A U G U G C C	A G C A G C C G C	G G U A A A C A A G A A U A G C	A A G G U U A A U C	G U C C U C U A U U G A G	P.t.	51
U - - - - C A G A A A U -	- G G A U G C C	A G C C G U U G C	G G U A A U A U C U A U G C U U U	A A A U A A U A U A U	C - - - - - T.c.	C.f.	52
U - - - - C A G A A A U -	- G G A U G C C	A G C C G U U G C	G G U A A U A U C U A U G C U U U	A A A U A A U A U A U	- - - - - T.b.	S.d.	53
U - - - - C A G A A A U -	- G G A U G C C	A G C C G U U G C	G G U A A A G A C G G G G G G G C	A A G U G U C U U C	- - - - - L.t.	S.s.	54
C C C C - G G C U A A C U	- C G G U G C C	A G C A G C C G C	G G U A A A G A C G G G G G G G C	A A G U G U C U U C	G G A A U A C U U G G G	Z.n.	55
C C C C - G G C U A A C U	- C G G U G C C	A G C A G C C G C	G G U A A A G A C G G G G G G G C	A A G U G U C U U C	G G A A U A C U U G G G	T.s.	56
C C C C - G G C U A A C U	- C G G U G C C	A G C A G C C G C	G G U A A A G A C G G G G G G G C	A A G U G U C U U C	G G A A U A C U U G G G	D.a.	57

E19-3

E19-3'

P19

P19

1060 1070 1080 1090 1100 1110 1120

1060 1070 1080 1090 1100 1110 1120

H.s. 1
M.m. 2
R.n. 3
K.r.n. 4
O.c. 5
X.l. 6
A.s. 7
S.c. 8
D.r. 9
U.s. 10
D.p. 11
T.t. 12
G.m. 13
A.t. 14
F.t. 15
E.c. 16
B.v. 17
H.c. 18
M.c. 19
H.s. 20
B.f. 21
D.d. 22
M.x. 23
A.n. 24

A.t. 25
F.t. 26
E.c. 27
B.v. 28
H.c. 29
M.c. 30
H.s. 31
B.f. 32
D.d. 33
M.x. 34
A.n. 35

E.s. 36
C.r. 37
Z.m. 38
M.t. 39

H.s. 40
M.m. 41
R.n. 42
B.t. 43
X.l. 44
D.v. 45
A.n. 46
S.c. 47
P.t. 48
F.s. 49
C.b. 50
L.t. 51
S. 52
Z. 53
T.s. 54
D.s. 55

[illegible]

1480 1490 1500 1510 1520 1530 1540

```

C G G C C C C G G A - - - C A - C G G A C A - G G A U U G A C A G A U U G A U - A G C - U C U U U C U C G A U - - U C C G U G G - G U - G G H.s. 1
C G G C C C C G G A - - - C A - C G G A C A - G G A U U G A C A G A U U G A U - A G C - U C U U U C U C G A U - - U C C G U G G - G U - G G M.h. 2
C G G C C C C G G A - - - C A - C G G A C A - G G A U U G A C A G A U U G A U - A G C - U C U U U C U C G A U - - U C C G U G G - G U - G G R.n. 3
C G G C C C C G G A - - - C A - C G G A C A - G G A U U G A C A G A U U G A U - A G C - U C U U U C U C G A U - - U C C G U G G - G U - G G R.n. 4
C G G C C C C G G A - - - C A - C G G A C A - G G A U U G A C A G A U U G A U - A B C - U C U U U C U C G A U - - U C G U G G - G U - G G D.c. 5
C G G C C C C G G A - - - C A - C G G A A A - G G A U U G A C A G A U U G A U - A B C - U C U U U C U C G A U - - U C G U G G - G U - G G X.l. 6
A G G C C C C G G A - - - C A - C U G G A A - G G A U U G A C A G A U U G A G - A G C - U C U U U C U U G A U - - U C A G U G G - G U - G G A.s. 7
A G G C C C C A G A - - - C A - C A A U A A - G G A U U G A C A G A U U G A G - A G C - U C U U U C U U G A U - - U U G U G G - G U - G G S.c. 8
A A G C U U A G A - - - U A - U A G U A A - G G A U U G A C A G A C U A A A - A G A - U C U U U C A U G A U - - U C U A U A G - G U - G G D.d. 9
A G G U U C C A G A - - - C A - U A G U G A - G G A U U G A C A G A U U G A U - A G C - U C U U U C U U G A U - - U C U A U G G - G U - G G O.n. 10
A G G U U C C A G A - - - C A - U A G U G A - G G A U U G A C A G A U U G A U - A B C - U C U U U C U U G A U - - U C U A U G G - G U - G G S.p. 11
A G C G C A G A - - - C A - G A G A A G - G G A U U G A C A G A U U G A G - A B C - U C U U U C U U G A U - - U C U U U G G - G U - G G T.t. 12
A G G U U C C A G A - - - C A - U A G C A A - G G A U U G A C A G A C U G A G - A B C - U C U U U C U U G A U - - U C U A U G G - G U - G G O.s. 13
A G G U U C C A G A - - - C A - U A G C A A - G G A U U G A C A G A C U G A G - A B C - U C U U U C U U G A U - - U C U A U G G - G U - G G Z.h. 14
A G G U U C C A G A - - - C A - U A G U A A - G G - U U G A C A G A C U G A G - A B C - U C U U U C U U G A U - - U C U A U G G - G U - G G G.h. 15

```

27

28

28'

```

A G C C C C - G A - - - C A G - U A G U A A - U S A C G G U C A G G U U G A U - G A C - C U U A C C C G G A G - - - G C U A C U G - A G A G G H.c. 16
A G C C C C - G A - - - C A G - U A G U A A - U S A C G G U C A G G U U G A U - G A C - C U U A C C C G G A G - - - G C U A C U G - A G A G G H.h. 17
A G C U U C - G A - - - C U A - C A G U G - U S A C G G U C A G G U U G A U - G A C - C U U A C C A C A G A C - - - G C U G A G - A G A G G H.v. 18
G G C A C C - G A - - - C A G U G U G C A G - U S A C A G U C A G U C C G A U - G G G - C U U A C U - - U G A - - - G C C A C U G - A G A G G H.h. 19
G G A U A A - G A - - - C A G - U G A A - A - U A G A U A G U C G G G A U U G A A - G A C - U C U A C U - - U G A - - - U A G G C U G - A G A G G H.h. 20
A G G A G C - G A - - - C A G - C A G A - A - U S A C G G C C A G G U U G A C - G A C - C U U U G C - - U G A - - - A G C G C U G - A G A G G H.v. 21
A G G G C C - G A - - - C A G - C A G A - A - U S A U A G C C A G G U U G A U - G A C - C U U U G C - - U G A - - - C A A G C U G - A G A G G H.f. 22
G G G G G A - G A - - - C G G - C A G U - A - U S A C G G C C A G G C U A A C - G A C - C U U U G C - - U G A - - - G A - - - C U C G C G G - A G A G G S.s. 23
C G G G C C - G A - - - C A G - C A G G - A - U S A A G G C C A G G C U A A C - G A C - C U U U G C - - G G A - - - C G A G C U G - A G A G G T.t. 24

```

```

A G C U C U U G A C A U U C G G G U U U G G C A G U G A G A C A U U G U C C U U - C A G U A G G C U G - - - G C C C A G A A G G A.t. 25
C A C C U C U U G A C A U G C C - A G G A A C - U U A C C A G A G A U G G U U U - G G U - G C U C G A A A G A G A A C C U G A C - A C A G G P.t. 26
U G G U C U U G A C A U C C A - C G G A A U - U U U U C A G A G A U G A G A A - U G U - G C C U U C G G G A A - - - C C G U A G - A C A G G E.c. 27
U A C U C U U G A C A U C C A - C G G A A U - C C U U C A G A G A U A G A G G - A G U - G C C U U C G G G A A - - - C C G U A G - A C A G G P.v. 28
A G G U C U U G A C A U C C U - C G G A C A - A U C C U A G A G A U A G G A C - G U C - C C C U U C G G G G - - - C A G A G U G - A C A G G D.s. 29
A A G G C U U G A C A U C C U - C U G A A U - C C G A U A G A G A U A G C G G - A G U - G C C C U C G G G G A G - A C A G A G - A C A G G H.c. 30
A G G G C U - G A C A U C C A G U C A A A - G C U A U A G A G A U A U A G U - A G A - G G - U U - - - A A - - - C A U U A G - A C A G G H.c. 31
A G G G C U U G A C A U C C A G U C A A A - G C U A U A G A G A U A U A G U - A G A - G G - U U - - - A A - - - C A U U A G - A C A G G H.s. 32
C G G G C U A A A A U U G C A G U G A A U - G A U G U G S A A C A U U G U C - A G U - G - - A G C A A U C A - - - C C G C U G - G A G G D.f. 33
C G G G C U A A A A U U G A - G U G A - G U G A A A U U U A - A G U - G - - A G C A A U C A - - - C A G A A C - A C A G G F.h. 34
U A G G U U U G A C A U C C A - C G G A A C - C C U C C G A A A G G A G G - G G U - G C C C U C G G G G A C C G U A G - A C A G G D.d. 35
U G G U C U U G A C A U C C U - C A G A A U - C C U U C A G A U G A G G G - A G U - G C C C G C A A G G G A C C U G A G - A C A G G H.x. 36
A G G U U U U G A C A U C C C - C G G A A U - C U C U U G A A A C G A G G - A G U - G C C U U C G G G A G - - - C G G G G A G - A C A G G A.n. 37

```

27

28

28'

```

A G G A U U U G A C A G A U - - - C U A G G - A A G U U U G A A A G A A C G C - A G U - A C C U U C G G G U A - - - U C U A G A C - A C A G G E.s. 38
A G G G C U U G A C A U G U C - A A G A A C - C U C U C A G A A A U G G G A G - G G U - G C C C U A C G G A - - - C U U G A A C - A C A G G C.r. 39
A G G G C U U G A C A U G C C - G C G A A U - C U C U U G A A A G A G A G G - G G U - G C C C U A C G G A A - - - C G G G A C - A C A G G Z.h. 40
A U U G G C U U G A C A U G C C - G C G A A U - C C U C U U G A A A G A G A G G - G G U - G C C U U C G G G A A - - - C G G G A C - A C A G G H.t. 41

```

```

A C C U C U U G C U - - - C - A G C C U A - U A U A C C G C C A U C U U C A - G C A - A A C C C U G A U G - - - - - - - - H.s. 42
A U C U C U U G C U A A U U C - A G C C U A - U A U A C C G C C A U C U U C A - G C A - A A C C C U A A A A - - - - - - - - M.h. 43
C C U U C U U G C C - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - R.n. 44
A A U U C U U G C - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - B.t. 45
A C U U C U U G C C A A A C C C G C - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - X.l. 46
U A A A U U U G A A U A U C A - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - G U D.v. 47
A C A G U U U G A A U U U U - U A C - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - A.G S.c. 48
A U A U U U G A A U U U A - U A A A A - U A - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - U A U A - A U U - A U U A U A - U U A C - - - - - - - - A.G S.c. 49
A G C G - U A A A A U U U U U - U C U U U - A A U U A A G C U U A A U - G G - C U G - A A A U U U U U A G A - - - A A A A A A G - A A A U G P.p. 50
A G C G - U A A A A U U U U U - U C U U U - A B U U A A G C U C A U - G G - C U A - A A A U U U U U A G A - - - A A A A A A G - A A A U G P.t. 51
- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - C.f. 52
- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - T.b. 53
- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - L.t. 54
A G C C C U U G A C A U A U G - A A C A A C - A A A A C C U G U C U U A A C - G G G - A U G G U A C U A C - - - U U C A U A - - - C A G G Z.h. 55
A G C C C U U G A C A U A U G - A A C A A C - A A A A C C U G U C C U U A C - A G G - A U G G U A C U G A C - - - U U C A U A - - - C A G G T.h. 56
A G C C C U U G A C A U A U G - A A C A A C - A A A A C C U G U C C U U A C - G G G - A U G G U A C U G A C - - - U U C A U A - - - C A G G O.s. 57

```

1480 1490 1500 1510 1520 1530 1540

-	CA	C	U	C	C	U	A	A	A	U	G	C	A	S	C	A	G	U	A	C	C	U	U	U	-	M.c.	16		
-	CA	C	U	C	C	U	A	A	A	U	G	C	A	S	C	G	G	U	A	C	C	U	U	U	-	M.h.	17		
-	CA	C	U	C	C	U	A	A	A	U	G	C	A	S	C	A	S	C	A	G	U	U	U	C	U	C	M.v.	18	
-	CA	C	U	C	C	U	A	A	A	U	G	C	A	S	C	A	S	C	A	G	U	U	U	C	U	C	M.a.	19	
-	CG	U	C	C	C	A	A	A	A	U	G	C	A	S	C	A	S	C	A	G	U	U	U	C	U	C	-	-	
-	CG	C	G	A	S	C	A	G	U	G	C	A	S	C	A	G	U	U	G	A	C	C	U	U	C	U	C	M.h.	20
-	UG	C	C	C	C	U	A	A	S	U	G	C	A	S	C	A	G	U	U	U	U	U	U	C	U	C	-	M.v.	21
-	CG	C	C	C	C	U	A	A	S	U	G	C	A	S	C	A	G	U	U	U	U	U	U	C	U	C	-	M.f.	22
-	CA	C	C	C	C	C	U	A	A	S	U	G	C	A	S	C	A	G	U	U	U	U	U	C	U	C	-	S.s.	23
-	CA	C	C	C	C	U	A	A	S	U	G	C	A	S	C	A	G	U	U	U	U	U	U	C	U	C	-	T.t.	24

- C G	- C C C U U A G U U G C A G C A U U	- A t t.	25
- U G	- C C A U U A G U U G C U A - C A U U C	- P r t.	26
- U A	- U C C U U U G U U G C C A G C G G U C C	- E x c.	27
- U A	- U C C U U U G U U G C C A G C G C G U A	- P r v.	28
- U G	- A U C U U A G U U G C C A G C A U U C	- B s s.	29
- U A	- U C C U A G U U G C C A G C A G A G	- H e c.	30
- U G	- U C G U U A G U U A C U A A C A U U A A	- H e c.	31
- U G	- U C G U U A G U U A C U A A C A U U A A	- H e s.	32
- U A	- U C U U U A G U U A C A A C A G U U	- B r f.	33
- U A	- U G U U U A G U U G C C A G C A C G U U	- F r h.	34
- U A	- U G G A U A G U U G C C A G C A A G U A	- D r d.	35
- C G	- C C U U U A G U U G C C A G C C -	- M x.	36
- C G	- U U U U U A G U U G C C A U C A U U C	- A n t.	37

-U U-	U U U U A A U U(A)A C G		E.d.	38
-C G-	U C U U U A G U U B C C A		C.r.	39
-C G-	U G U U U A G U U B C C A C U A U S		Z.m.	40
-C G-	U G U U U A G U U B C C A U C G U U		N.v.	41

r105

	2320	
.....		
-----	H.s.	1
-----	H.m.	2
-----	R.n.	3
-----	R.n.	4
-----	O.c.	5
-----	X.l.	6
-----	A.s.	7
-----	S.c.	8
-----	D.d.	9
-----	O.n.	10
-----	S.p.	11
-----	T.t.	12
-----	O.s.	13
-----	Z.m.	14
-----	G.m.	15

-----	H.c.	16
-----	H.h.	17
-----	H.v.	18
-----	H.m.	19
-----	H.h.	20
-----	H.v.	21
-----	H.f.	22
-----	S.s.	23
-----	T.t.	24

-----	A.t.	25
-----	P.t.	26
-----	E.c.	27
-----	P.v.	28
-----	B.s.	29
-----	H.c.	30
-----	H.c.	31
-----	H.s.	32
-----	B.f.	33
-----	F.h.	34
-----	D.d.	35
-----	M.x.	36
-----	A.n.	37

-----	E.s.	38
-----	C.r.	39
-----	Z.m.	40
-----	N.t.	41

-----	H.s.	42
-----	H.m.	43
-----	R.n.	44
-----	B.t.	45
-----	X.l.	46
-----	D.v.	47
-----	A.n.	48
C U U A C A - -	S.c.	49
U A A A U G C C	P.p.	50
C A U G C A C C	P.t.	51
-----	C.f.	52
-----	T.b.	53
-----	L.t.	54
-----	Z.m.	55
-----	T.a.	56
-----	O.s.	57
	2320	

ACKNOWLEDGEMENTS

The present study was made possible by an F.K.F.O. assistantship granted to E.H.. We thank Erna Dams, Raymond De Baere, Lydia Hendriks, Hilde Van den Eynde and Peter Willekens for their cooperation in the correction of the srRNA computer files.

*To whom correspondence should be addressed

REFERENCES

1. Torczynski, R.M., Fuke, M., Bollon, A.P. (1985) *DNA* 4, 283-291.
2. Raynal, F., Michot, B., Bachellerie, J.-P. (1984) *FEBS Lett.* 167, 263-268.
3. Torczynski, R., Bollon, A.P., Fuke, M. (1983) *Nucl. Acids Res.* 11, 4879-4890.
4. Chan, Y.-L., Gutell, R., Noller, H.F., Wool, I.G. (1984) *J. Biol. Chem.* 259, 224-230 (1984).
5. Connaughton, J.F., Rairkar, A., Lockard, R.E., Kumar, A. (1984) *Nucl. Acids Res.* 12, 4731-4745.
6. Atmadja, J., Brimacombe, R., Maden, B.E.H. (1984) *Nucl. Acids Res.* 12, 2649-2667.
7. Nelles, L., Fang, B.L., Volckaert, G., Vandenberghe, A., De Wachter, R. (1984) *Nucl. Acids Res.* 12, 8749-8768.
8. Hogan, J.J., Gutell, R.R., Noller, H.F. (1984) *Biochemistry* 23, 3322-3330.
9. McCarroll, R., Olsen, G.J., Stahl, Y.D., Woese, C.R., Sogin, M.L. (1983) *Biochemistry* 22, 5858-5868.
10. Elwood, H.J., Olsen, G.J., Sogin, M.L. (1985) *Mol. Biol. Evol.* 2, 399-410.
11. Same reference as 10.
12. Spangler, E.A., Blackburn, E.H. (1983) *J. Biol. Chem.* 260, 6334-6340.
13. Takaiwa, F., Oono, K., Sugiura, M. (1984) *Nucl. Acids Res.* 12, 5441-5448.
14. Messing, J., Carlson, J., Hagen, G., Rubenstein, I., Olesen, A. (1984) *DNA* 3, 31-40.
15. Eckenrode, V.K., Arnold, J., Meagher, R.B. (1985) *J. Mol. Evol.* 21, 259-269.
16. Hui, I., Dennis, P.P. (1985) *J. Biol. Chem.* 260, 899-906.
17. Mankin, A.S., Kagramanova, V.K., Teterina, N.L., Rubtsov, R.M., Belova, E.M., Kopylov, A.M., Baratova, L.A., Bogdanov, A.A. (1985) *Gene* 37, 181-189.
18. Gupta, R., Lanter, J.M., Woese, C.R. (1983) *Science* 221, 656-659.
19. Leffers, H., Garrett, R.A. (1984) *EMBO J.* 3, 1613-1619.
20. Yang, D., Kaine, B.P., Woese, C.R. (1985) *System. Appl. Microbiol.* 6, 251-256.
21. Jarsch, M., Böck, A. (1985) *System. Appl. Microbiol.* 6, 54-59.
22. Lechner, K., Wich, G., Bock, A. (1985) *System. Appl. Microbiol.* 6, 157-163.
23. Olsen, G.J., Pace, N.R., Nuell, M., Kaine, B.P., Gupta, R., Woese, C.R. (1985) *J. Mol. Evol.* 22, 301-307.
24. Leinfelder, W., Jarsch, M., Böck, A. (1985) *System. Appl. Microbiol.* 6, 164-170.
25. Yang, D., Oyaizu, Y., Oyaizu, H., Olsen, G.J., Woese, C.R. (1985) *Proc. Natl. Acad. Sci. USA* 82, 4443-4447.
26. Same reference as 25.
27. Brosius, J., Dull, T., Sleeter, D.D., Noller, H. (1981) *J. Mol. Biol.* 148, 107-127.
28. Carbon, P., Ebel, J.P., Ehresmann, C. (1981) *Nucl. Acids Res.* 9, 2325-2333.

29. Green, C.S., Stewart, G.C., Hollis, M.D., Vold, B.S., Bott, K.F. (1985) *Gene* 37, 261-266.
30. Woese, C.R., Debrunner-Vosbrink, B.A., Oyaizu, H., Stackebrandt, E., Ludwig, W. (1985) *Science* 229, 762-765.
31. Iwami, M., Muto, A., Yamao, F., Osawa, S. (1984) *Mol. Gen. Genet.* 196, 317-322.
32. Frydenberg, J., Christiansen, C. (1985) *DNA* 4, 127-137.
33. Weisburg, W.G., Oyaizu, Y., Oyaizu, H., Woese, C.R. (1985) *J. Bacteriol.* 146, 230-236.
34. Same reference as 33.
35. Oyaizu, H., Woese, C.R. (1985) *System. Appl. Microbiol.* 6, 257-263.
36. Same reference as 35.
37. Tomioka, N., Sugiura, M. (1983) *Mol. Gen. Genet.* 191, 46-50.
38. Graf, L., Roux, E., Stutz, E., Kossel, H. (1982) *Nucl. Acids Res.* 10, 6369-6381.
39. Dron, M., Rahire, M., Rochaix, J.-D. (1982) *Nucl. Acids Res.* 10, 7609-7620.
40. Schwartz, Z., Kossel, H. (1980) *Nature* 238, 739-742.
41. Tohdoh, N., Sugiura, M. (1982) *Gene* 17, 213-218.
42. Eperon, I.C., Anderson, S., Nierlich, D.P. (1980) *Nature* 286, 460-467.
43. van Etten, R.A., Walberg, M.W., Clayton, D.A. (1980) *Cell* 22, 157-170.
44. Kobayashi, M., Seki, T., Yaginumi, K., Koike, K. (1981) *Gene* 16, 297-307.
45. Anderson, S., de Bruijn, M.H.L., Coulson, A.R., Eperon, I.C., Sanger, F., Young, I.G. (1982) *J. Mol. Biol.* 156, 683-717.
46. Roe, B., Ma, D.P., Wilson, R.K., Wong, J.F.H. (1985) *J. Biol. Chem.* 260, 9759-9774.
47. Clary, D.O., Wolstenholme, D.R. (1985) *Nucl. Acids Res.* 13, 4029-4045.
48. Köchel, H.G., Kuntzel, H. (1981) *Nucl. Acids Res.* 9, 5689-5696.
49. Li, M., Tzagoloff, A., Underbrink-Lyon, F., Martin, N.C. (1982) *J. Biol. Chem.* 257, 5921-5928.
50. Seilhamer, J.J., Olsen, G.J., Cummings, D.J. (1984) *J. Biol. Chem.* 259, 5167-5172.
51. Same reference as 50.
52. Sloof, P., Van den Burg, J., Voogd, A., Benne, R., Agostinelli, M., Borst, P., Gutell, R. (1985) *Nucl. Acids Res.* 13, 4171-4190.
53. Eperon, I.C., Janssen, J.W.G., Hoeijmakers, J.H.J., Borst, P. (1983) *Nucl. Acids Res.* 11, 105-125.
54. de la Cruz, V.F., Lake, J.A., Simpson, A.M., Simpson, L. (1985) *Proc. Natl. Acad. Sci. USA* 82, 1401-1405.
55. Chao, S., Sederoff, R., Levings, C.S. (1984) *Nucl. Acids Res.* 12, 6629-6644.
56. Spencer, D.F., Schnare, M.N., Gray, M.W. (1984) *Proc. Natl. Acad. Sci. USA* 81, 493-497.
57. Brennicke, A., Moller, S., Blanz, P. (1985) *Mol. Gen. Genet.* 198, 404-410.
58. Salim, M., Maden, B.E.H. (1981) *Nature* 291, 205-208.
59. Rubtsov, P.M., Musakhanov, M.M., Zakhariev, V.M., Krayev, A.S., Skryabin, K.G., Bayev, A.A. (1980) *Nucl. Acids Res.* 8, 5779-5794.
60. Carbon, P., Ehresmann, C., Ehresmann, B., Ebel, J.P. (1978) *FEBS Lett.* 94, 152-156.
61. Brosius, J., Palmer, M.L., Kennedy, P.J., Noller, H.F. (1978) *Proc. Natl. Acad. Sci. USA* 75, 4801-4805.
62. Sor, F., Fukuhara, H. (1980) *C.R. Acad. Sci. Paris Ser. D* 291, 933-936.
63. Nomenclature committee of the International Union of Biochemistry (1985) *Eur. J. Biochem.* 150, 1-5.
64. Cornish-Bowden, A. (1985) *Nucl. Acids Res.* 13, 3021-3030.
65. Woese, C.R., Gutell, R., Gupta, R., Noller, H.F. (1983) *Microbiol. Rev.* 47, 621-669.

66. Gutell, R.R., Weiser, B., Woese, C.R., Noller, H.F. (1985) *Progr. Nucl. Acid Res. Mol. Biol.* 32, 155-214.
67. Olsen, G.J., McCarroll, R., Sogin, M.L. (1983) *Nucl. Acids Res.* 11, 8037-8049.
68. Kaestner, A. (1969) *Lehrbuch der Speziellen Zoologie*, Gustaf Fisher Verlag, Stuttgart.
69. Ainsworth, G.C., Sparrow, F.K., Sussman, A.S. (1973) *The Fungi, an Advanced Treatise*, Academic Press, New York.
70. Fritsch, F.E. (1971) *The Structure and Reproduction of the Algae*, University Press, Cambridge.
71. Engler, A. (1954) *Syllabus der Pflanzenfamilien*, Gebrüder Borntraeger, Berlin.
72. Balch, W.E., Fox, G.E., Magrum, L.J., Woese, C.R., Wolfe, R.S. (1979) *Microbiol. Rev.* 43, 260-296.
73. Schnabel, R., Huet, J., Thomur, M., Zillig, W., Sentenac, A., Stetter, K.O. (1983) in *Endocytobiology*, vol. 2, pp. 895-912, E.A. Schenk, W. Schwemmler eds., De Gruyter, Berlin, New York.
74. Woese, C.R., Stackebrandt, E., Macke, T.J., Fox, G.E. (1985) *System. Appl. Microbiol.* 6, 143-151.
75. Erdmann, V.A., Wolters, J., Huysmans, E., De Wachter, R. (1985) *Nucl. Acids Res.* 13, r105-r153.