Proc. Natl. Acad. Sci. USA
Vol. 82, pp. 6955–6959, October 1985
Evolution

Rapid determination of 16S ribosomal RNA sequences for phylogenetic analyses

(reverse transcriptase/dideoxynucleotide)

DAVID J. LANE*, BERNADETTE PACE*, GARY J. OLSEN*, DAVID A. STAHL†‡, MITCHELL L. SOGIN†, AND NORMAN R. PACE*§

*Department of Biology and Institute for Molecular and Cellular Biology, Indiana University, Bloomington, IN 47405; and †Department of Molecular and Cellular Biology, National Jewish Hospital and Research Center, Denver, CO 80206

Communicated by Ralph S. Wolfe, June 26, 1985

ABSTRACT **Although the applicability of small subunit ribosomal RNA (16S rRNA) sequences for bacterial classification is now well accepted, the general use of these molecules has been hindered by the technical difficulty of obtaining their sequences. A protocol is described for rapidly generating large blocks of 16S rRNA sequence data without isolation of the 16S rRNA or cloning of its gene. The 16S rRNA in bulk cellular RNA preparations is selectively targeted for dideoxynucleotide-terminated sequencing by using reverse transcriptase and synthetic oligodeoxynucleotide primers complementary to universally conserved 16S rRNA sequences. Three particularly useful priming sites, which provide access to the three major 16S rRNA structural domains, routinely yield 800–1000 nucleotides of 16S rRNA sequence. The method is evaluated with respect to accuracy, sensitivity to modified nucleotides in the template RNA, and phylogenetic usefulness, by examination of several 16S rRNAs whose gene sequences are known. The relative simplicity of this approach should facilitate a rapid expansion of the 16S rRNA sequence collection available for phylogenetic analyses.**

The classification of organisms traditionally has been based on similarities in their morphological, developmental, and nutritional characteristics. It is now clear, however, that with microorganisms, classification based on these criteria does not necessarily correlate well with natural (i.e., evolutionary) relationships, as defined by macromolecular sequence comparisons (1). A much broader application of molecular phylogenetic analysis to the description of microbes, both eukaryotic and prokaryotic, seems desirable.

All of the available molecular methods for evaluating phylogenetic relationships (e.g., DNA·DNA and DNA·rRNA hybridization, 5S rRNA and protein sequencing, 16S rRNA oligonucleotide cataloging, enzymological patterning, etc.) have advantages and limitations. In general, macromolecular sequences seem preferred because they permit quantitative inference of relationships (2, 3). Moreover, because they accumulate, sequences are most useful in the long term.

Of the macromolecules used for phylogenetic analysis, the ribosomal RNAs, particularly 16S rRNA,[¶] have proven the most useful for establishing distant relationships because of their high information content, conservative nature, and universal distribution. Using RNase T1 oligonucleotide catalogs of 16S rRNA, Woese and his colleagues were able to establish a comprehensive outline of prokaryotic phylogeny (4). The principle of using rRNA sequences to characterize microorganisms has now gained wide acceptance (5), and its general application can be anticipated if methods for determining rRNA sequences can be simplified. The approach described here rapidly provides partial sequences of 16S rRNA that are useful for phylogenetic analysis.

MATERIALS AND METHODS

Purification of RNA Templates. Bulk, cellular RNA was purified by phenol extraction of French pressure cell lysates as detailed by Pace *et al.* (6), except that ribosomes were not pelleted before extraction. High molecular weight RNA was then prepared by precipitation with 2 M NaCl (6). Although not essential, NaCl precipitation of the RNA generally increased the amount of legible sequence data and reduced backgrounds on gels, presumably by eliminating fragmented DNA from the reactions. RNA was stored at 2 mg/ml in 10 mM Tris·HCl (pH 7.4) at −20°C.

Oligodeoxynucleotide Primers. Oligodeoxynucleotide primers were synthesized manually by using the appropriate blocked and protected nucleoside diisopropylphosphoramidites and established coupling protocols (7). Deblocked products were purified by polyacrylamide gel electrophoresis, eluted from gels in 1 mM EDTA/50 mM NH_4OAc, and adsorbed onto C_8 Bond Elut columns (Analytichem International, Harbor City, CA). The columns were washed with the loading buffer and then washed with 5 mM NH_4OAc, and finally the primers were eluted with acetonitrile/H_2O, 1:1 (vol/vol). After drying *in vacuo*, primers were dissolved in H_2O to 0.1 mg/ml. (Requests for these primers should be addressed to Susan Andrews, Institute for Molecular and Cellular Biology, Indiana University, Bloomington, IN 47405.)

Reverse Transcriptase. Reverse transcriptase from avian myeloblastosis virus (10,000 units/ml) was purchased from Seikagaku America, Inc. The enzyme and dilutions (1:10) of it in 50 mM Tris·HCl, pH 8.3/2 mM dithiothreitol/50% (wt/vol) glycerol were stored at −20°C.

Nucleotides. Deoxyadenosine 5′-[α-thio]triphosphate labeled with ^{35}S in the α-thio position (dATP[α-^{35}S]) was purchased from New England Nuclear. The 2′-deoxy- (d-) and 2′,3′-dideoxy- (dd-) nucleoside triphosphates were from P-L Biochemicals. Stock solutions of unlabeled nucleotides [10 mM in 10 mM Tris·HCl (pH 8.3) as determined spectrophotometrically] were stored frozen at −70°C. Nucleotide mixtures for the reverse transcription reactions contained 10 mM Tris·HCl (pH 8.3), 250 μM dCTP, 250 μM dGTP, 250 μM dTTP, 125 μM [α-thio]dATP, and either one dideoxynucleotide (30 μM ddCTP, 19 μM ddGTP, 30 μM ddTTP, or 1.25 μM ddATP) or no dideoxynucleotide. Aliquots sufficient for 10 reactions were stored at −20°C.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

‡Present address: Department of Veterinary Pathobiology, University of Illinois, Urbana, IL 61801.
§To whom reprint requests should be addressed.
¶16S rRNA is used here to refer to all small subunit rRNAs; actual sedimentation coefficients vary slightly from organism to organism.

Reverse Transcription Reactions. The sequencing protocol described here is a base-specific, dideoxynucleotide-terminated chain elongation method (8, 9), modified for the use of reverse transcriptase and RNA templates (10). Since sequencing reactions were only 5.0 μl, common components were mixed and then aliquots were taken to reduce pipetting errors. Untreated 0.5-ml Microfuge tubes were used throughout. DNA primers were hybridized to RNA templates in 7.5-μl reaction mixtures containing 1.5 μl of 5× hybridization buffer (500 mM KCl/250 mM Tris·HCl, pH 8.5), 3.5 μl of high molecular weight RNA (2.0 mg/ml), 1.5 μl of the desired primer (0.1 mg/ml), and 1.0 μl of H_2O. This annealing mixture was heated at 90°C for 1 min and then allowed to cool slowly over 10 min to 25°C. Of the hybridized template–primer mixture, 6.5 μl was then transferred to a tube containing 30 μCi (1 Ci = 37 GBq) of dried dATP[α-^{35}S], 6.5 μl of 5× reverse transcription buffer (250 mM Tris·HCl, pH 8.3/250 mM KCl/50 mM dithiothreitol/50 mM $MgCl_2$), and 6.5 μl of reverse transcriptase (1,000 units/ml). Three microliters of this final mixture was then added to each of six tubes that contained, respectively, 2.0 μl of ddCTP-, ddATP-, ddTTP-, ddGTP-, ddCTP-, or no-dideoxynucleotide-containing nucleotide mixture (above). The reactions were incubated for 5 min at room temperature and then 30 min at 37°C. One microliter of chase mix [1.0 mM each of dATP, dCTP, dTTP, and dGTP in 10 mM Tris·HCl (pH 8.3) containing 1,000 units/ml of reverse transcriptase] was then added to each reaction, and incubation at 37°C was continued for another 15 min. The reactions were chilled on ice and stopped by adding 6.0 μl of stop mix (86% formamide/10 mM EDTA/0.08% xylene cyanol/0.08% bromophenol blue). Reaction mixtures were heated for 2 min at 90°C immediately prior to loading onto sequencing gels. Two gels were usually run on each set of sequencing reactions: one was a 0.3-mm-thick, 40-cm-long, 8% polyacrylamide buffer gradient gel (9) (bromophenol blue run to 40 cm), and the other was an equivalent, 8% polyacrylamide nongradient gel (xylene cyanol run off the bottom to a calculated "distance" of 70 cm).

RESULTS AND DISCUSSION

In assessing the relationships of organisms to one another by the comparison of their 16S rRNA sequences, it is not important that the complete sequences of the molecules be determined. What is important is that the number of nucleotides compared is statistically meaningful and that compared sequences are appropriately aligned, such that only truly homologous sequence positions are considered.

The 16S rRNAs vary in their nucleotide sequences, but they contain regions that are conserved perfectly, or nearly so, among all organisms so far inspected. Certain of these conserved sequences, adjacent to less-conserved regions that are useful for phylogenetic evaluations, provide broadly applicable initiation sites for primer elongation sequencing techniques. Oligodeoxynucleotides, 15–20 residues in length, that are complementary to certain of the conserved sequences were synthesized and tested as primers for dideoxynucleotide-terminated sequencing reactions with reverse transcriptase and 16S rRNA templates.

Fig. 1 shows the location of three of the most useful priming sites in representative eubacterial, archaebacterial, and eukaryotic small-subunit rRNAs and indicates the extent of phylogenetically useful sequence routinely obtained from

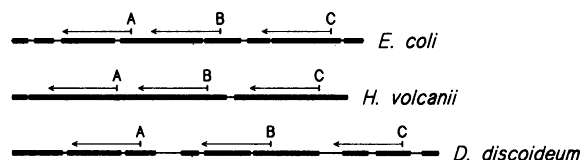

FIG. 1. Hybridizing sites of "universal" small-subunit rRNA primers. The locations of three particularly useful priming sites (A, B, and C) are shown on linear representations of the small-subunit rRNAs from *Escherichia coli* (a eubacterium) (11), *Halobacterium volcanii* (an archaebacterium) (12), and *Dictyostelium discoideum* (a eukaryote) (13). The primer sequences and their hybridizing locations in the *E. coli* 16S rRNA are G-W-A-T-T-A-C-C-G-C-G-G-C-K-G-C-T-G, positions 519–536 (A); C-C-G-T-C-A-A-T-T-C-M-T-T-T-R-A-G-T-T-T, 907–926 (B); and A-C-G-G-G-C-G-G-T-G-T-G-T-R-C, 1392–1406 (C). In these sequences K = G or T, M = A or C, R = A or G, and W = A or T. The solid boxes along the sequence lines are regions that display sufficient intrakingdom structural conservation to be generally useful in the inference of phylogenies. The arrow pointing to the left from each priming site indicates the approximate extent of the sequence data (300 nucleotides) accessible from each primer.

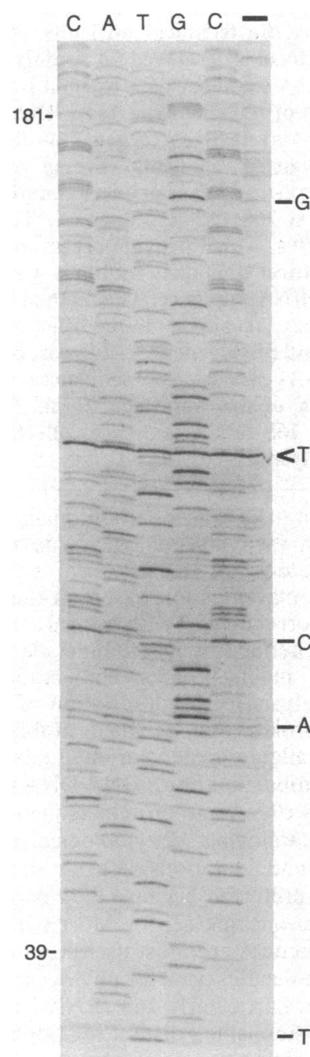

FIG. 2. Reverse transcriptase sequencing "anomalies." An autoradiogram of a typical 8% acrylamide/8 M urea buffer-gradient sequencing gel analysis (9) of reverse transcriptase reactions using bulk, high molecular weight *H. volcanii* RNA as template and the 1392–1406 primer is shown. Distances from the 3′ terminus of the primer are indicated on the left. Terminations not mediated by dideoxynucleotide incorporations in which the correct nucleotide is evident (–) and those where it is not (<) are indicated on the right. Sequencing reactions containing dideoxycytosine (C), dideoxyadenosine (A), dideoxythymidine (T), dideoxyguanosine (G), or no dideoxynucleotides (–) are indicated at the top. The two cytidine reactions facilitated alignment of bands across the gel.

each primer. The details of the primer structures are given in the figure legend. The utility of these and other primers has been tested with 16S rRNA templates from over 50 organisms, representing all three primary kingdoms—eubacteria, archaebacteria, and eukaryotes. No 16S rRNA tested has failed to serve as a template with the three primers shown in Fig. 1; therefore, these primers are operationally termed "universal."

Reverse transcriptase has been used previously to copy rRNA templates (14–17). Initial attempts to produce cDNA reverse transcripts from 16S rRNA templates using appended polyadenylate tails to create priming sites met with only partial success (14); reverse transcription was found to terminate abruptly about 30 nucleotides from the 3′-end, corresponding to the location of a widely conserved N^6-dimethyladenosine dinucleotide, m_2^6A-m_2^6A. However, it was subsequently demonstrated with *E. coli* 16S rRNA as template that if reverse transcription were initiated 5′ proximal to that modified dinucleotide, then the enzyme would proceed to the 5′ end of the template (15). Other nucleoside modifications that terminate reverse transcription from some 16S rRNAs have been identified (16), but these fall outside the sequences routinely accessible from the universal rRNA primers shown in Fig. 1.

Inaccuracies are inherent in any sequencing protocol. In order to assess inaccuracies in data gathered by using reverse transcriptase, the results of elongation from the described primers were inspected by using rRNA templates for which the gene sequences were known. These included 16S rRNAs from *E. coli* (11), mouse (18), *D. discoideum* (13), *H. volcanii* (12), and *Sulfolobus solfataricus* (unpublished data). The frequency of misidentifying or omitting a residue was approximately 1%, typical for chain elongation sequencing data gathered from only one complementary strand. The most common sources of errors were band compressions and rearrangements caused by strong secondary structure interactions in some rRNA sequences, even in the 8 M urea-containing polyacrylamide sequencing gels (19). Such artifacts generally are evident as discontinuities in band spacing in the sequencing gels or upon alignment of the derived sequence on the known 16S rRNA structure (below). Substitution of dITP for dGTP alleviates band compression phenomena to a considerable extent (20), although its routine use is unnecessary.

More common than errors with this method are ambiguities resulting from chain terminations that are not due to dideoxynucleotide incorporations. Examples of these are shown in Fig. 2. Anomalous bands are sometimes seen in all lanes, including that of the reaction not containing any dideoxynucleotide (Fig. 2). These bands appear not to result from spurious priming events (17) because (*i*) reactions performed without added primer show no dATP[α-^{35}S] incorporation and (*ii*) labeling of reverse transcripts with radioactively labeled primer rather than dATP[α-^{35}S] incorpo-

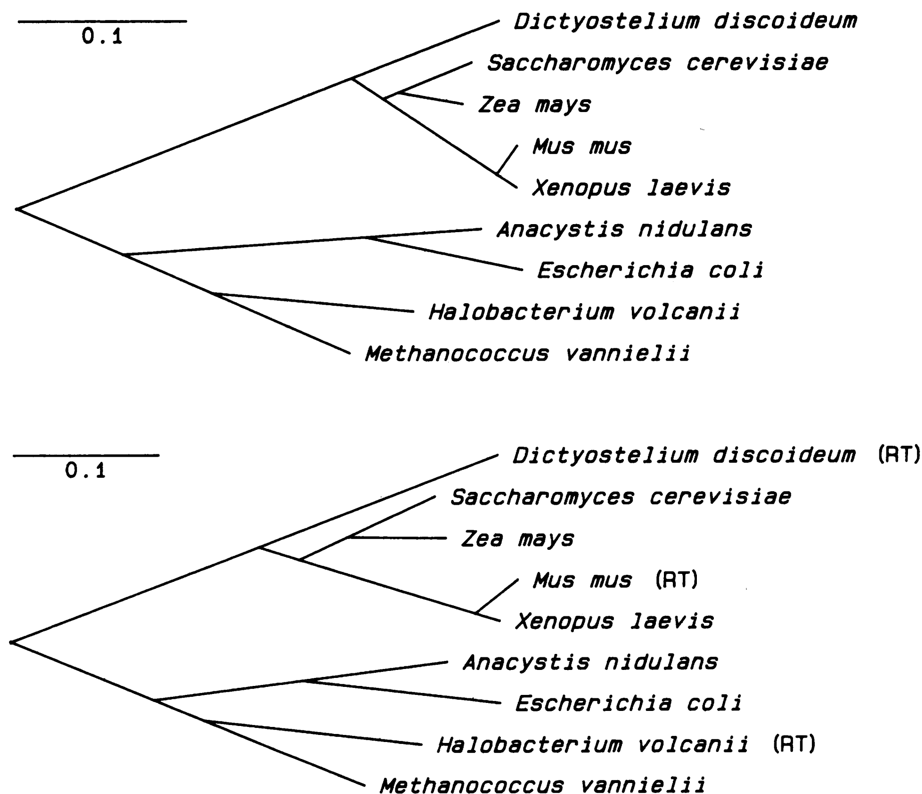

FIG. 3. Comparison of phylogenetic relationships inferred by using complete and partial 16S rRNA sequences. The illustrated networks ("trees") were deduced by the matrix method (24), omitting regions of ambiguous sequence alignment from homology calculations. Alignment gaps were assigned half the weight assigned to a nucleotide in that position. While determining the optimal tree, the mean square difference between the estimated evolutionary distance separating each pair of sequences and the corresponding tree distance was weighted by the statistical uncertainty of the distance estimate (26). An iterative program that follows the path of "steepest descent" in the optimization parameter (25, 27) was used to determine the tree topology which best fit, by the weighted least-mean-square difference criterion, the homology data. The scale bar represents an evolutionary distance of 0.1 K_{nuc} (K_{nuc} = average number of nucleotide changes per sequence position). The networks are based on complete 16S rRNA sequences (11–13, 18, 28–32) using about 950 nucleotide positions after elimination of regions of ambiguous alignment (*Upper*) or are based on blocks of sequence determined by reverse transcription from the 1392–1406 and 907–926 primers of *H. volcanii*, *D. discoideum*, and mouse rRNA and homologous blocks from the other sequences (*Lower*). Only about 350 nucleotide positions are compared. However, it should be noted that because of the lesser amount of interkingdom sequence conservation, fewer unambiguously alignable blocks of sequence are available for these comparisons than are available (see Fig. 1) for intrakingdom comparisons.

ration does not appreciably affect the distribution or relative intensities of the anomalous bands. They are most straightforwardly interpreted as points where the reverse transcriptase is, with some frequency, released from the template (21). The anomalous bands superficially appear to result from pauses (15, 22), in that the enzyme generally continues into more primer-distal sequences. However, addition of nonradioactive nucleotides and reverse transcriptase to complete reactions does not diminish their intensities.

There seems to be no single explanation for the anomalous bands generated by reverse transcriptase. Some result from fragmented template in rRNA preparations and are alleviated by further purification of the 16S rRNA after denaturation. Others likely result from modified residues (14, 16, 22), although no ambiguities corresponding to pseudouridine, N^2-methylguanosine (m^2G), 2′-O-methyladenosine (mA), or the dinucleotide N^5-methylcytidine-N^2-methylguanosine (m^5C-m^2G) were detectable. An unidentified adenosine modification in *D. discoideum* rRNA (position 1745; ref. 13) did correspond to a gel band anomaly. No particularly striking effects on the avian myeloblastosis virus reverse transcriptase attributable to the secondary structure of the RNA template have been observed. Tertiary structure interactions in the rRNA, while so far unknown, conceivably influence the enzyme.

Since most reverse transcriptase molecules read through the anomalous positions, the template nucleotide often is revealed as the most prominent band (Fig. 2). Positions where the residue assignment is not clear are scored as "N," for sequence alignment and comparison purposes. Ambiguities in the derived nucleotide sequence merely reduce the extent of sequence available for comparison with other sequences. For phylogenetic purposes, even minor inaccuracies in nucleotide identification incur little penalty. Random errors impart the appearance of a fast "evolutionary clock." On the other hand, assignment of nucleotide identity on the basis of homology with a sequence from another organism (which might be assumed related) results in systematic errors, perhaps forcing incorrect affilations.

After the collection of partial sequence data from a novel organism, they must be aligned with other available sequences. Variabilities in the primary and secondary structures of small subunit rRNAs are now sufficiently well understood that alignment of a novel sequence on the known skeleton is straightforward, keying on regions of conserved sequence and secondary structure (23). By alignment of novel sequences on the existing rRNA framework, many sequence errors due to band compressions or rearrangements on sequencing gels may also be detected. For establishing phylogenetic relationships, it is imperative to employ only sequences that are unambiguously homologous (24). Regions of ambiguous homology must be eliminated from analyses.

There are two approaches to establishing the phylogenetic relationships of an aligned sequence to others in the data collection. In the first approach, phylogenetic "trees" are generated using the available 16S rRNA sequences. Nucleic acid and protein sequences have been extensively used in phylogenetic tree construction, although use of the method with 16S rRNA is limited at this time by the small number of sequences available; only about 25 complete 16S rRNA sequences have been reported. As the reference collection of complete and partial 16S rRNA sequences expands, however, quantitative phylogenetic trees may provide the best means for relating organisms.

It has been shown previously that phylogenetic trees constructed by using certain limited regions of the 16S rRNA had topologies identical to that obtained by using complete sequences (13, 25). Fig. 3 illustrates that this is also true for the sequence blocks accessed by reverse transcription from the universal rRNA primers. The upper tree was constructed by using complete 16S rRNA sequences. The lower tree was constructed by using reverse transcription data from mouse, *D. discoideum*, and *H. volcanii* templates and the homologous blocks of sequence from the other 16S rRNAs. As is evident in the figure, the reverse transcription data generate the same branching pattern as the complete sequence data, with approximately the same branch point depths. This is an impressive correspondence for interkingdom comparisons.

The second approach to relating organisms by using sequence information gathered as described here uses the collection of RNase T1 oligonucleotide catalogs established by Woese and his colleagues over the past decade (4). RNase T1 oligonucleotide sequences are implicit in the continuous sequences, so novel rRNAs are readily related to those in the Woese data collection by the occurrence of group-specific oligonucleotides or by binary association coefficients. Because of their extensive nature (*ca.* 400 organisms; ref. 33), the oligonucleotide data currently are the best resource for relating organisms by their 16S rRNA sequences. The methodology described here provides rapid and convenient access to the appropriate sequence information.

The general sequencing procedures employed here are in wide use. The novel aspect is the use of evidently universally applicable, oligodeoxynucleotide primers to access 16S rRNA sequences for phylogenetic characterizations. Because of its rapidity (from cell pellet to 800–1,000 nucleotides of sequence in *ca.* 3 days) and technical simplicity, the approach is amenable to the screening of large numbers of organisms. Therefore, phylogenetic analysis by 16S rRNA sequences could become a standard tool of laboratories concerned with characterizing organisms of uncertain affiliation.

We thank Drs. Carl Woese, Robin Gutell, and Olin Kew for useful discussions and Dr. Richard Gumport for providing facilities for the synthesis of the first primer. This paper is dedicated to Carl Woese. This work was supported by National Institutes of Health Grant GM34527 to N.R.P.

1. Stackebrandt, E. & Woese, C. R. (1981) in *Molecular and Cellular Aspects of Microbial Evolution*, eds. Carlisle, M. J., Collins, J. R. & Moseley, B. E. B. (University Press, Cambridge), pp. 1–31.
2. Zuckerkandl, E. & Pauling, L. (1965) *J. Theor. Biol.* **8**, 357–366.
3. Fitch, W. M. & Margoliash, E. (1967) *Science* **155**, 279–284.
4. Fox, G. E., Stackebrandt, E., Hespell, R. B., Gibson, J., Maniloff, J., Dyer, T. A., Wolfe, R. S., Gupta, R., Bonen, L., Lewis, B. J., Stahl, D. A., Luehrsen, K. R., Chen, K. N. & Woese, C. R. (1980) *Science* **209**, 457–463.
5. Murray, R. G. E. (1984) in *Bergey's Manual of Systematic Bacteriology*, eds. Kreig, N. R. & Holt, J. G. (Williams & Wilkins, Baltimore), Vol. 1, pp. 31–34.
6. Pace, B., Matthews, E. A., Johnson, K. D., Cantor, C. R. & Pace, N. R. (1982) *Proc. Natl. Acad. Sci. USA* **79**, 36–40.
7. Caruthers, M. H., Beaucage, S. L., Becker, C., Efcavitch, J. W., Fisher, E. F., Galluppi, G., Goldman, R., de Haseth, P., Matteucci, M., McBride, L. & Stabinsky, Y. (1983) in *Gene Amplification and Analysis*, eds. Papas, T. S., Rosenberg, M. & Chirikjian, J. G. (Elsevier, New York), Vol. 3, pp. 1–26.
8. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5463–5467.
9. Biggin, M. D., Gibson, T. J. & Hong, G. F. (1983) *Proc. Natl. Acad. Sci. USA* **80**, 3963–3965.
10. Hamlyn, P. H., Brownlee, G. G., Cheng, C.-C., Gait, M. J. & Milstein, C. (1978) *Cell* **15**, 1067–1075.
11. Brosius, J., Palmer, M. L., Kennedy, R. J. & Noller, H. F. (1978) *Proc. Natl. Acad. Sci. USA* **75**, 4801–4805.
12. Gupta, R., Lanter, J. M. & Woese, C. R. (1983) *Science* **221**, 656–659.
13. McCarroll, R., Olsen, G. J., Stahl, Y. D., Woese, C. R. & Sogin, M. L. (1983) *Biochemistry* **22**, 5858–5868.

14. Hagenbüchle, O., Santer, M., Steitz, J. A. & Mans, R. J. (1978) *Cell* **13**, 551–563.
15. Youvan, D. C. & Hearst, J. E. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 3751–3754.
16. Youvan, D. C. & Hearst, J. E. (1981) *Nucleic Acids Res.* **9**, 1723–1741.
17. Qu, L. H., Michot, B. & Bachellerie, J.-P. (1983) *Nucleic Acids Res.* **11**, 5903–5920.
18. Raynal, F., Michot, B. & Bachellerie, J.-P. (1984) *FEBS Lett.* **167**, 263–268.
19. Brownlee, G. G. & Cartwright, E. M. (1977) *J. Mol. Biol.* **114**, 93–117.
20. Mills, D. R. & Kramer, F. R. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 2232–2235.
21. Myers, J. C. & Spiegelman, S. (1978) *Proc. Natl. Acad. Sci. USA* **75**, 5329–5333.
22. Inoue, T. & Cech, T. I. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 648–652.
23. Woese, C. R., Gutell, R. R., Gupta, R. & Noller, H. F. (1983) *Microbiol. Rev.* **47**, 621–669.
24. Fitch, W. M. & Smith, T. F. (1983) *Proc. Natl. Acad. Sci. USA* **80**, 1382–1386.
25. Olsen, G. J. (1983) Dissertation (University of Colorado Denver, CO), p. 128.
26. Kimura, M. & Ohta, T. (1972) *J. Mol. Evol.* **2**, 87–90.
27. Pace, N. R., Stahl, D. A., Lane, D. J. & Olsen, G. J. (1985) *Adv. Microb. Ecol.*, in press.
28. Tomioka, N. & Sugiura, M. (1983) *Mol. Gen. Genet.* **191**, 46–50.
29. Messing, J., Carlson, J., Hagen, G., Rubenstein, I. & Oleson, A. (1984) *DNA* **3**, 31–40.
30. Jarsch, M. & Böck, A. (1985) *Syst. Appl. Microbiol.* **6**, 54–59.
31. Rubtsov, P. M., Musakhanov, M. M., Zakharyev, V. M., Krayev, A. S., Skryabin, K. G. & Bayev, A. A. (1980) *Nucleic Acids Res.* **8**, 5779–5794.
32. Salim, M. & Maden, B. E. H. (1981) *Nature (London)* **291**, 205–208.
33. Woese, C. R., Weisburg, W. G., Hahn, C. M., Paster, B. J., Zablen, L. B., Lewis, B. J., Macke, T. J., Ludwig, W. & Stackebrandt, E. (1985) *Syst. Appl. Microbiol.* **6**, 25–33.