

22619-87-0; *N*⁶-DNP-lysine, 1094-76-4; selenocystine, 29621-88-3; cystine, 56-89-3; selenocysteine, 10236-58-5.

References

- Chiu, D., Fletcher, B., Stults, F., Zakowski, J., & Tappel, A. (1975) *Fed. Proc., Fed. Am. Soc. Exp. Biol.* 34, 925, Abstr.
- Condell, R. A., & Tappel, A. L. (1982) *Biochim. Biophys. Acta* 709, 304-309.
- Cone, J. E., DelRio, R. M., Davis, J. N., & Stadtman, T. C. (1976) *Proc. Natl. Acad. Sci. U.S.A.* 73, 2659-2663.
- Epp, O., Ladenstein, R., & Wendel, A. (1983) *Eur. J. Biochem.* 133, 51-69.
- Flohe, L., Gunzler, W. A., & Schock, H. H. (1973) *FEBS Lett.* 32, 132-134.
- Flohe, L., Gunzler, W. A., & Ladenstein, R. (1976) in *Glutathione: Metabolism and Function* (Arias, I. M., & Jakoby, W. B., Eds.) pp 115-138, Raven Press, New York.
- Forstrom, J. W., Zakowski, J. J., & Tappel, A. L. (1978) *Biochemistry* 17, 2639-2644.
- Kesner, L., Muntwyler, E., Griffin, G., & Quaranta, P. (1967) *Methods Enzymol.* 11, 94-108.
- Kondo, H., Moriuchi, F., & Sunamoto, J. (1981) *J. Org. Chem.* 46, 1333-1336.
- Kraus, R. J., & Ganther, H. E. (1980) *Biochem. Biophys. Res. Commun.* 96, 1116-1122.
- Kraus, R. J., Prohaska, J. R., & Ganther, H. E. (1980) *Biochim. Biophys. Acta* 615, 19-26.
- Ladenstein, R., Epp, O., Bartels, K., Jones, A., Huber, R., & Wendel, A. (1979) *J. Mol. Biol.* 134, 199-218.
- Ladenstein, R., Epp, O., Huber, R., & Wendel, A. (1981) in *Selenium in Biology and Medicine* (Spallholz, J. E., Martin, J. L., & Ganther, H. E., Eds.) pp 33-43, Avi, Westport, CT.
- Oh, S.-H., Ganther, H. E., & Hoekstra, W. G. (1974) *Biochemistry* 13, 1825-1829.
- Reed, D. J., Babson, J. R., Beatty, P. W., Brodie, A. E., Ellis, W. W., & Potter, D. W. (1980) *Anal. Biochem.* 106, 55-62.
- Reich, H. J., Hoeger, C. A., & Willis, W. W., Jr. (1982) *J. Am. Chem. Soc.* 104, 2936-2937.
- Rotruck, J. T., Pope, A. L., Ganther, H. E., Swanson, A. B., Hafeman, D. G., & Hoekstra, W. G. (1973) *Science (Washington, D.C.)* 179, 588-590.
- Sanger, F. (1945) *Biochem. J.* 39, 507-515.
- Swenson, R. P., Williams, C. H., Jr., & Massey, V. (1982) *J. Biol. Chem.* 257, 1937-1944.
- Twiss, D. F. (1914) *J. Chem. Soc.* 105, 1672-1678.

Nucleotide Sequence of the *Dictyostelium discoideum* Small-Subunit Ribosomal Ribonucleic Acid Inferred from the Gene Sequence: Evolutionary Implications[†]

Robert McCarroll,[‡] Gary J. Olsen, Yvonne D. Stahl, Carl R. Woese,[§] and Mitchell L. Sogin*

ABSTRACT: We have determined the T₁ oligonucleotide catalog and the complete nucleotide sequence of the gene for the *Dictyostelium discoideum* small-subunit ribosomal RNA. The gene does not appear to contain an intron and has relatively

low homology with other eukaryotic small-subunit rRNAs. The homology data indicate that *D. discoideum* diverged from the mainstream of eukaryotic descent at the earliest branch yet characterized by molecular phylogeny.

The deepest branchings within the eukaryotic line of descent are represented by members of the "kingdom" Protoctista.¹ It has not been possible to infer consistent phylogenies for these simple eukaryotes by using classical taxonomic approaches, i.e., comparative studies of phenotypes. This, in part, is a consequence of the enormous physiological, cytological, and biochemical diversity within the group.

Genealogical relationships between such diverse organisms can be inferred from comparisons of the amino acid or nucleotide sequences of functionally equivalent biopolymers (Zuckerkanl & Pauling, 1965). Analyses based upon amino acid sequence changes in cytochrome *c* (Fitch, 1976; McLaughlin & Dayhoff, 1973), ferredoxin (Ysunobu & Ta-

naka, 1973; Rao & Hall, 1977; Rao & Cammack, 1981), superoxide dismutase (Asada et al., 1980; Fridovich, 1976; Lumsden & Hall, 1975), etc. have proven extremely useful in confirming, refining, and extending classical eukaryotic phylogeny, but sometimes they produce unconvincing junctions at the deepest phylogenetic levels where their limits as molecular "chronometers" are approached. By comparison, the ribosomal RNAs seem to be better suited for revealing phylogenetic relationships. They are universally distributed and are functionally equivalent in all cells. In addition, the size and the conservative nature of large portions of the 16-18S and 23-28S rRNAs permit the measurement of the largest phylogenetic distances (Stackebrandt & Woese, 1981; Otsuka et al., 1983). Yet, other regions of the molecule change more rapidly, and this allows the measurement of close phylogenetic relationships. Finally, the sequence homology that exists between the rRNAs of the eukaryotes and both kingdoms of the prokaryotes permits the evaluation of relative rates of

[†] From the Department of Molecular and Cellular Biology, National Jewish Hospital and Research Center, Denver, Colorado 80206, and the Department of Biochemistry, Biophysics and Genetics, University of Colorado Health Sciences Center, Denver, Colorado 80262. Received May 25, 1983. This investigation was supported by National Institutes of Health Research Grant GM23464 to M.L.S. and by NASA Grant NSG 7044 to C.R.W.

[‡] Present address: Department of Genetics, University of Washington, Seattle, WA 98195.

[§] Present address: Department of Genetics and Development, University of Illinois, Urbana, IL 61801.

¹ Membership in this division is "defined by exclusion: its members are neither animals (which develop from blastula), plants (which develop from an embryo), fungi (which lack undulipodia and develop from spores), nor prokaryotes" (Margulis & Schwartz, 1982).

evolution among all eukaryotic rRNA sequences.

Evolutionary relationships inferred from 5S and 5.8S rRNA sequences (Olsen & Sogin, 1982), as well as comparisons of the small-subunit rRNA T₁ oligonucleotide catalogs (see below), indicate that the earliest divergence from the mainstream of eukaryotic descent yet characterized by molecular phylogeny is represented by *Dictyostelium discoideum*. This finding is further corroborated by comparisons of the primary structure of putative tryptophan tRNA (Peffley & Sogin, 1981) and valine tRNA genes from *D. discoideum* with homologous tRNAs from other eukaryotes (M. L. Sogin, unpublished data). Because of the considerable statistical uncertainty of comparisons between 5S and 5.8S rRNA sequences (which is due to their small number of independently varying positions and paucity of "functional domains"), it is necessary to confirm independently the inferred genealogical relationships. Pairwise comparisons of the small-subunit ribosomal RNA (srRNA) can be used to determine more reliable homologies because these molecules are large (ca. 1470–2000 nucleotides) and contain numerous functional domains.

We have determined the T₁ oligonucleotide catalog of the *D. discoideum* 17S ribosomal RNA and the complete nucleotide sequence of its gene. The gene does not appear to contain an intron, and as with the other RNAs that are integral components of the cytoplasmic translation apparatus, the most divergent of the characterized eukaryotic small-subunit ribosomal RNA sequences is that of *D. discoideum*. The divergence represented by *D. discoideum* reflects an early branching in the eukaryotic line of descent rather than a "fast" evolutionary clock.

Experimental Procedures

Materials. Restriction enzymes were purchased from New England Biolabs. Terminal transferase and cordycepin [α -³²P]triphosphate were purchased from New England Nuclear. *Eco*RI endonuclease was isolated as described by Greene et al. (1974), omitting the DEAE-cellulose step. Polynucleotide kinase was isolated from bacteriophage T4 am E4314 infected *Escherichia coli* as described by Uhlenbeck & Cameron (1977). Carrier-free [γ -³²P]ATP was prepared by using the procedures of Reeve and Huang (Johnson & Walseth, 1978). Low gel temperature agarose and Bio-Gel agarose beads were purchased from Bio-Rad Laboratories.

Growth and Labeling of Cells. Unlabeled RNA was isolated from whole cells of *D. discoideum* strain Ax3 grown in HL5 + 1% glucose (Cocucci & Sussman, 1970). RNA labeled in vivo with ³²P was isolated from cells grown in phosphate-depleted Mes-HL5 + 1% glucose containing 0.1 mCi/mL H₃³²PO₄ (ICN) (Olsen & Sogin, 1982).

Cloning. *D. discoideum* nuclear DNA was prepared as previously described and inserted into the *Eco*RI site of plasmid pMB9 (Sogin & Olsen, 1980). Recombinant plasmids carrying the tetracycline-resistance gene were used to transform the tetracycline-sensitive host *E. coli* HB101. Clones were grown on tetracycline-containing L-agar and screened for ribosomal RNA genes by using a slight modification of the colony-hybridization technique (Grunstein & Hogness, 1975). Ribosomal RNA coding sequences were located by incubating the colony-hybridization filters with in vitro ³²P-labeled 17S or 25S rRNA probes in 5X SSC,² 50% deionized

formamide, and 0.5% SDS at 42 °C for 16 h. The filters were washed 30 min in hybridization buffer at 42 °C, which was followed by two 30-min washes in 2X SSC at room temperature. Colonies capable of hybridizing with the ³²P-labeled rRNA probe were identified by autoradiography with intensifying screens at -70 °C. All work was carried out under P2-EK1 containment conditions.

Preparation of Ribosomal RNA Hybridization Probes. Total cytoplasmic RNAs were isolated as previously described (Sogin & Olsen, 1980) and were fractionated by electrophoresis through 2.8% polyacrylamide-SDS gels (Bis to acrylamide ratio 1:20) in E buffer. The regions of the gel containing 17S- and 25S-sized RNA were located by the "UV shadowing" technique (Hassur & Whitlock, 1974). The regions of the gel containing the 17S and 25S rRNAs were excised, and the RNA was subjected in situ to partial alkaline hydrolysis, in 50 mM NaHCO₃, pH 9.0, at 90 °C for 80 min. Subsequent to elution from the gel and concentration by ethanol precipitation, the fragments were labeled in vitro with polynucleotide kinase and [γ -³²P]ATP (Maxam & Gilbert, 1980).

Plasmid DNA Isolations and Preparation of DNA Fragments. Plasmid DNA was isolated on cesium chloride-ethidium bromide gradients as described by Maniatis et al. (1982). Whenever possible, uniquely sized, labeled DNA restriction fragments were fractionated in agarose bead columns equilibrated in 50 mM Tris (pH 7.2), 0.5 mM NaOAc, and 1 mM EDTA. Fractions containing the DNA were pooled and concentrated by ethanol precipitation. When labeled restriction fragments could not be resolved on the agarose bead columns, fragments greater than 300 nucleotides were isolated by fractionation in 0.75 or 1.5% low gel temperature agarose slab gels built in E buffer. Subsequent to electrophoresis, the region of the gel containing the DNA fragment, as identified by UV shadowing or autoradiography, was excised and placed in a vial containing 5 volumes of 20 mM Tris-acetate (pH 8.0) and 1 mM EDTA. The gel was melted by heating to 65 °C and extracted with H₂O-saturated phenol equilibrated with 20 mM Tris-acetate (pH 8.0) and 1 mM EDTA. The aqueous fractions containing the DNA fragment were extracted with ether and concentrated by ethanol precipitation. DNA fragments smaller than 300 nucleotides were isolated on 4–8% polyacrylamide gels.

DNA Sequencing. Restriction fragments labeled at their 5'-termini with [γ -³²P]ATP and polynucleotide kinase (Maxam & Gilbert, 1980) or at their 3'-termini with cordycepin [α -³²P]triphosphate and terminal transferase (Tu & Cohen, 1980) were restricted with a second enzyme and refractionated on agarose bead columns, agarose gels, or polyacrylamide gels. The resulting DNA fragments labeled at one end were sequenced according to the chemical procedures of Maxam & Gilbert (1980). The hydrazine reaction in the presence or absence of 1 M NaCl was used for site-specific cleavage of dCyd and dCyd plus dThd. The dimethyl sulfate reaction was used for dGuo-specific chemistries. Alkali treatment at 90 °C was used for the dAdo and dCyd cleavage. KMnO₄ was used for dThd- plus dGuo-specific cleavage (Rubin & Schmid, 1980). The products of the sequencing reactions were resolved on 25, 20, 8, or 5% gels (Air et al., 1976). The sequencing strategies are displayed in Figure 1.

RNase T₁ Oligonucleotide Cataloging. Ten micrograms of in vivo labeled 17S rRNA was digested with 1 μ g of RNase T₁. The products were separated by two-dimensional fingerprinting (Sanger et al., 1965). The compositions and sequences of the oligomers were determined by electrophoretic analysis of the RNase U₂, RNase A, and RNase T₁ over-

² Abbreviations: SSC, 0.15 M NaCl and 0.015 M sodium citrate; SDS, sodium dodecyl sulfate; Bis, N,N'-methylenebis(acrylamide); E buffer, 40 mM Tris-acetate, pH 7.0, and 1 mM EDTA; Tris, tris(hydroxymethyl)aminomethane; EDTA, ethylenediaminetetraacetic acid; Mes, 4-morpholineethanesulfonic acid.

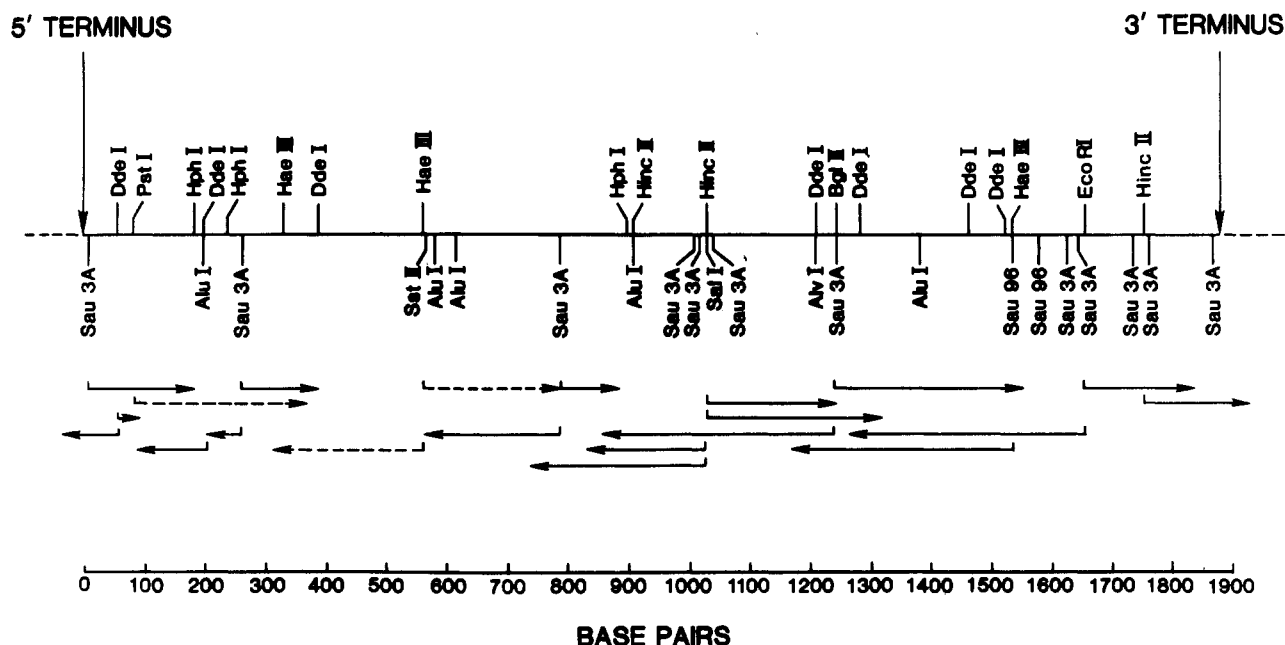


FIGURE 1: Restriction map and sequencing strategies used to determine DNA sequence of *D. discoideum* small-subunit ribosomal RNA gene. The region defining the *D. discoideum* small-subunit rRNA resides within contiguous 2.7- and 4.75-kb *EcoRI* restriction fragments. The cleavage sites for various restriction endonucleases were determined by a combination of single and double digests or by partial digestion of asymmetrically labeled restriction fragments. Solid arrows indicate the extent of sequence data read from a particular restriction site with 5'-end-labeled fragments. Dashed arrows indicate the extent of sequence data read from a particular restriction site with 3'-end-labeled fragments.

cutting products (Uchida et al., 1974).

Results

The *D. discoideum* ribosomal DNA genes are reiterated approximately 200-fold. The rDNA repeat units are 44 kilobases in length and are arranged as nonchromosomal palindromic dimers (Cockburn et al., 1978). The entire small-subunit rRNA coding region resides within contiguous 2.7- and 4.75-kb *EcoRI* restriction fragments (Cockburn et al., 1976; Maizels, 1976; Sogin & Olsen, 1980). A detailed restriction map of the 17S rRNA coding region (Figure 1) was derived via agarose and acrylamide gel analyses of single and double restriction endonuclease digestions, as well as single enzyme digests of purified restriction fragments labeled at their termini with [γ - 32]ATP.

Sequence Analysis of rDNA. The protocols of Maxam & Gilbert (1980) were used to sequence asymmetrically labeled restriction fragments. The strategies employed are summarized in Figure 1. Of the 1875 nucleotides, 1873 were unambiguously identified. The two unidentified residues lie within regions that are conserved in all eukaryotic small-subunit rRNA sequences. If the numbering system for the aligned sequence displayed in Figure 2 is used, position 610 is a purine and, by homology with other eukaryotic sequences, is probably a G residue. Similarly, position 632 is likely to be an A residue. The completed primary structure, deduced from the sequences of overlapping restriction fragments, is shown in Figure 2 aligned with the small-subunit ribosomal RNA sequences from *Saccharomyces cerevisiae* (Rubstov et al., 1980), *Xenopus laevis* (Salim & Maden, 1981), *E. coli* (Brosius et al., 1978), *Zea mays* chloroplast (Schwartz & Kössel, 1980), *Halobacterium volcanii* (Gupta et al., 1983), and the *Aspergillus nidulans* mitochondrion (Küntzel & Köchel, 1981).

RNase T₁ Oligonucleotide Analysis. The sequence of the *D. discoideum* small-subunit rRNA coding region was verified by comparing it with the 17S rRNA T₁ oligonucleotide catalog (previously determined). (The oligonucleotide sequences, molar yields, and locations within the aligned *D. discoideum*

sequence displayed in Figure 2 are provided in the supplementary material. See paragraph at end of paper regarding supplementary material.) The compositions and molar yields for the oligonucleotides are in substantial agreement with the DNA sequence. Therefore, it appears that there are no introns within the coding region and that the DNA sequence does not represent a "pseudogene".

Correlation of the oligonucleotide catalog with the complete coding sequence permits the placement of most of the modified residues in the *D. discoideum* small-subunit rRNA. We have listed the T₁ oligonucleotide sequences that contain modified residues and their locations within the aligned *D. discoideum* sequence displayed in Figure 2. We have positioned nine of the thirteen 2'-O-methylated residues (Table I). Five of the 2'-O-methyl group locations (positions 29, 470, 640, 681, and 1544) are identical with methylation sites in *X. laevis* 18S rRNA. (There are 34 2'-O-methyl residues in the *X. laevis* srRNA.) Additional modifications that were found at analogous positions in other species include (1) two N⁶-dimethyl-A residues at positions 1956 and 1957 (these modifications have been found in all srRNAs exclusive of the organellar rRNAs), (2) an unidentified U modification (position 1267) in a highly conserved region of the srRNA [*Rhodospiridium toruloides*, *Blastocladiella emersonii* (M. L. Sogin and C. R. Woese, unpublished data), and *Lemna* (C. R. Woese and T. Dyer, unpublished data)] also have an unidentified U modification at the analogous site. In *X. laevis*, the modification has been determined to be an N-acetylpsouridine (Salim & Maden, 1981).], (3) an unidentified A modification at position 1744 (*Lemna* has an A modification at position 1744 or 1745), and (4) a modified A residue at either position 506 or 507 (The altered mobility of the predicted RNase A digestion product, AAU, is similar to AAC and, hence, must reflect a modification of one of the A residues. *R. toruloides* has a modified residue at position 506. The analogous site in *X. laevis* is adjacent to a 2'-O-methylation.) Other modifications include a presumptive N-acetylcytosine at position 1949, a pseudouridine at position 1073, pseudouridine residues in the oligonucleotides Ψ G and (Ψ ,U)G (positions not determined), and

Table I: Terminal and Modified RNase T₁ Oligonucleotides of the *D. discoideum* Small-Subunit rRNA

oligomer ^a	location ^b
termini	
pUAACUG	6
AUCAUUUN _{OH}	1969
modified oligomers ^c	
C-2'-O-mAG	nd ^d
A-2'-O-mCG	nd
2'-O-mGUG	nd
2'-O-mUG	nd
ψG	nd
(U,ψ)G	nd
N*CCG	nd
AψCAG	1072
m ₂ ² Am ₂ ² ACCUG	1956
CAAUAAC-2'-O-mAG	1585
ACU*CAACUG	1265
UAUC-N-acCG	1944
UU-2'-O-mAAAAAG	679
UCUCA-2'-O-mAAG	40
AAACUUA-2'-O-mAAG	1207
UCAU-2'-O-mAUG	29
2'-O-mUAAUCCAG	640
AUUC-2'-O-mUAUAAG	1335
A*AUUCCUUG	1744
C-2'-O-mAAAUUACUCAUCCCAAUACG	470
AAUUAAC(U,2'-O-mUC)UUAG	1533
AC,(A,A*)U,AAAU,AU,C,AAU,	504
AC,C,U,AU,C,C,U,U,U,U,U,G	

^a Oligonucleotide sequences were determined as described under Experimental Procedures. Each oligomer occurs once per molecule. ^b Locations are in the aligned numbering system of Figure 2. ^c ψ, pseudouridine; *, unidentified modification. ^d nd, not determined.

an uncharacterized modification in the oligonucleotide NCCG (position not determined).

Homology Calculations. Because there are large variations in the lengths of the small-subunit rRNAs, it is necessary to align their sequences before meaningful homologies can be computed. Rather than relying only upon primary structures

to quantitatively construct best fit alignments, we have used an algorithm that considers both primary and "phylogenetically proven" secondary structures (a helix is considered to be phylogenetically proven if it is found in a significant number of srRNAs and if its formation is independent of absolute primary structure conservation; several compensating base changes must be found that maintain helical structure) (Noller & Woese, 1981). Initially, the sequences were aligned according to primary structure. The locations of evolutionarily conserved secondary structures were then used to align sequences where length variation occurred. The differences in sequence lengths were compensated by appropriate nucleotide gaps. This method has the advantage of taking into account the biological relevance of helical structures that are phylogenetically conserved.

The pairwise homologies between representative small-subunit rRNAs are presented in Table II. We have tested a variety of techniques for calculating homologies between aligned sequences from various organisms. All methods generated similar results. The homology values presented in the upper right portion of Table II represent pairwise comparisons of all homologous nucleotide positions. A weighting factor of 0.5 was used where insertions or deletions were required for aligning the compared molecules (a maximum of five adjacent insertions or deletions was considered). This treatment prevents large gaps (which could represent a single event in evolution) from dominating the homology estimate. Thus, the homologies were calculated as

$$\text{homology} = \frac{\text{match positions}}{[\text{match positions} + \text{mismatch positions} + (0.5)\text{gaps}]}$$

The pairwise homology values were converted into "evolutionary distances" (average number of base changes per sequence position) according to Hori & Osawa (1979). These distances were converted into phylogenetic trees where the segment lengths were assigned to minimize the mean square standard deviation (using the Hori and Osawa measure of variance) of the tree path lengths, relative to the pairwise

Table II: Homologies between Small-Subunit rRNA Sequences^a

organism ^b	fractional homology with												
	chloroplasts								mitochondria				
	<i>X.l.</i>	<i>S.c.</i>	<i>D.d.</i>	<i>H.v.</i>	<i>E.c.</i>	<i>E.g.</i>	<i>C.r.</i>	<i>Z.m.</i>	human	mouse	<i>P.a.</i>	<i>S.c.</i>	<i>A.n.</i>
<i>X. laevis</i>		0.77	0.68	0.50	0.48	0.46	0.46	0.47	0.39	0.40	0.38	0.38	0.37
<i>S. cerevisiae</i>	0.89		0.71	0.51	0.49	0.47	0.48	0.47	0.40	0.40	0.40	0.40	0.38
<i>D. discoideum</i>	0.86	0.87		0.51	0.49	0.48	0.48	0.47	0.40	0.41	0.40	0.42	0.41
<i>H. volcanii</i>	0.63	0.64	0.63		0.59	0.59	0.60	0.61	0.45	0.45	0.45	0.43	0.44
<i>E. coli</i>	0.60	0.59	0.60	0.68		0.72	0.75	0.75	0.45	0.44	0.50	0.48	0.49
<i>E. gracilis</i> chl	0.60	0.60	0.60	0.68	0.84		0.77	0.79	0.45	0.46	0.52	0.48	0.50
<i>C. reinhardtii</i> chl	0.58	0.60	0.60	0.68	0.84	0.88		0.82	0.43	0.43	0.49	0.47	0.49
<i>Z. mays</i> chl	0.58	0.59	0.60	0.69	0.85	0.90	0.90		0.45	0.46	0.49	0.47	0.49
human mito	0.43	0.45	0.44	0.50	0.50	0.51	0.49	0.51		0.76	0.44	0.40	0.42
mouse mito	0.44	0.44	0.44	0.50	0.50	0.52	0.49	0.51	0.79		0.43	0.43	0.42
<i>Paramecium aurelia</i> mito	0.50	0.51	0.52	0.55	0.64	0.64	0.63	0.62	0.47	0.46		0.46	0.47
<i>S. cerevisiae</i> mito	0.49	0.49	0.50	0.53	0.60	0.59	0.59	0.59	0.44	0.46	0.55		0.54
<i>A. nidulans</i> mito	0.46	0.45	0.49	0.49	0.60	0.58	0.58	0.58	0.46	0.45	0.57	0.62	

^a The upper right half of the table gives the fraction of homologous positions containing identical nucleotides for all pairs of aligned small-subunit rRNA sequences. The internal length variation has been treated by scoring locations with a deletion in one sequence opposite a nucleotide in the second sequence as half a mismatch. When both sequences in the pair have deletions at the same location, the position was not counted. A maximum of five adjacent insertions or deletions was considered. The lower left half of the table contains the fractional homologies between pairs of sequences when the analysis is limited to the summation of conserved and semiconserved sequences in the small-subunit rRNAs. With the numbering system in Figure 2, these regions correspond to nucleotides 9-65, 355-573, 606-694, 938-1107, 1184-1299, 1348-1416, 1582-1641, 1701-1816, and 1929-1980. ^b Sequence data from *X. laevis* (Salim & Maden, 1981), *S. cerevisiae* (Rubtsov et al., 1980), *H. volcanii* (Gupta et al., 1983), *E. coli* (Brosius et al., 1978), *E. gracilis* chloroplast (Graf et al., 1982), *C. reinhardtii* chloroplast (Dron et al., 1982), *Z. mays* chloroplast (Schwarz & Kössel, 1980), human mitochondrion (Eperon et al., 1980), mouse mitochondrion (Van Etten et al., 1980), *P. aurelia* mitochondrion (J. J. Seilhamer, unpublished data), *S. cerevisiae* mitochondrion (Sor & Fukuhara, 1980), and *A. nidulans* mitochondrion (Küntzel & Köchel, 1981).

FIGURE 2: Sequence of *D. discoideum* small-subunit ribosomal RNA coding region aligned with other small-subunit rRNAs. The sequence of the *D. discoideum* small-subunit rRNA is shown aligned with those from *X. laevis* (Salim & Maden, 1981), *S. cerevisiae* (Rubtsov et al., 1980), *H. volcanii* (Gupta et al., 1983), *E. coli* (Brosius et al., 1978), *Z. mays* chloroplast (Schwarz & Kössel, 1980), and *A. nidulans* mitochondrion (Küntzel & Köchel, 1981). The *D. discoideum* sequence was determined as described under Experimental Procedures by using the sequencing

[illegible]

strategies depicted in Figure 1. Initially, the small-subunit sequences were aligned according to primary structure. The locations of evolutionarily conserved secondary structures were then used to refine the alignment where length variation occurred. The differences in sequence length were compensated by introducing appropriate gaps (-) in the sequences. A numbering system for the aligned sequences, as well as numbering systems for each sequence, is provided in the figure.

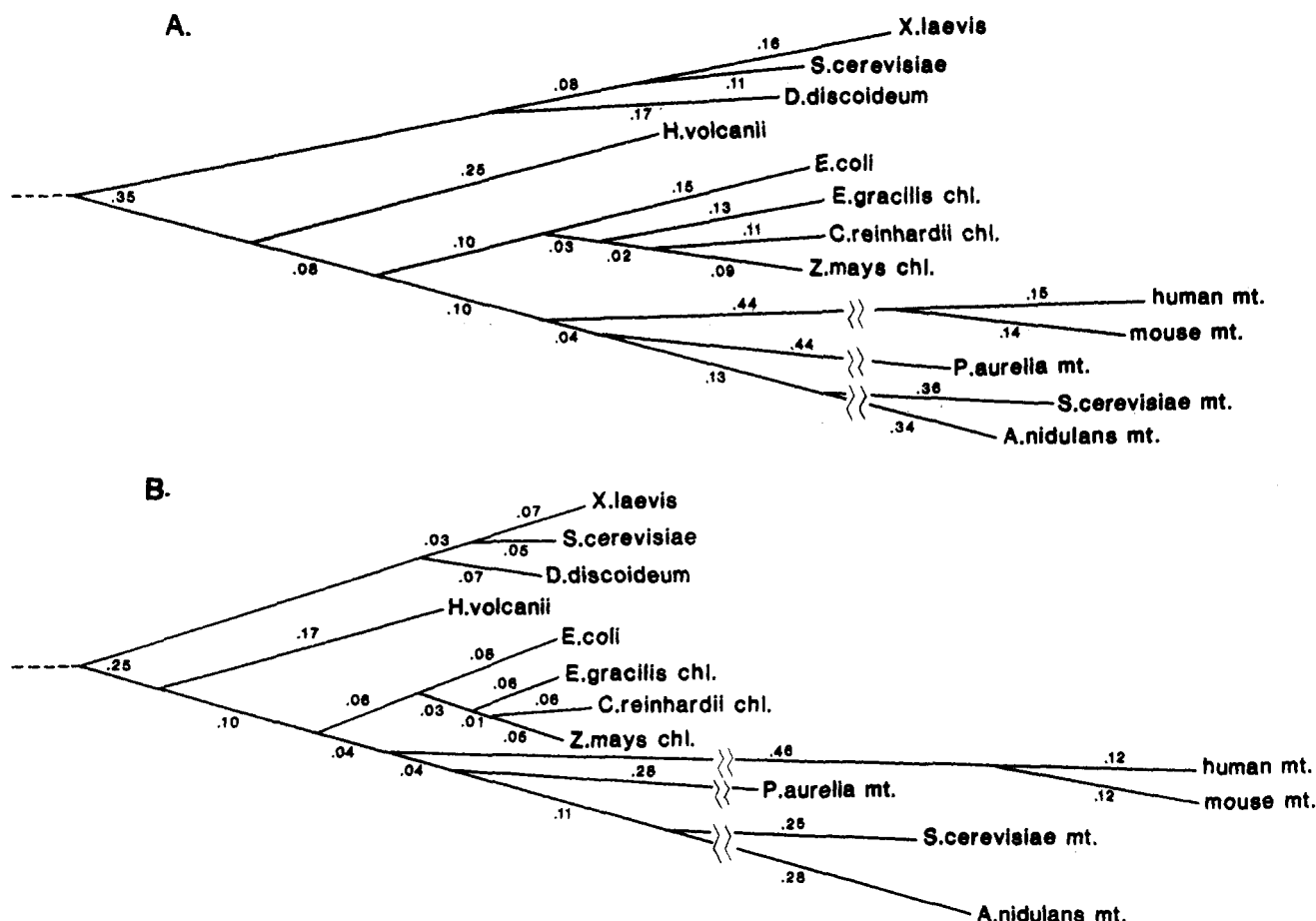


FIGURE 3: Phylogeny of small-subunit ribosomal RNA sequences. Phylogenetic trees were inferred by using the homology data in Table II for the entire small-subunit rRNA (panel A) or the summation of conserved and semiconserved regions (panel B). The trees were constructed as described under Results. The evolutionary distance between nodes of the tree is given alongside the segment connecting them and is represented in the horizontal component of their separation in the figure.

distance data. (We arrive at very similar results by minimizing other measures of error, in particular, the mean square fractional difference between the tree paths and the distance data.) This method was used to infer the phylogenetic tree in Figure 3A from comparisons of complete small-subunit ribosomal RNA sequences. This phylogeny confirms the genealogical relationships derived from the 5.8S rRNA homologies; *D. discoideum* is seen to diverge from the mainstream of eukaryotic descent at the earliest branch yet characterized by molecular phylogeny.

Only a minimal phylogenetic framework for the eukaryotes can be constructed from the limited data base of complete small-subunit rRNA sequences and T_1 oligonucleotide catalogs. Quantitative phylogenies among the eukaryotes could be constructed more rapidly if the sequence analyses undertaken were restricted to selected portions of the small-subunit rRNA. Comparisons of aligned small-subunit ribosomal RNA sequences, as well as the T_1 oligonucleotide catalogs representing over 200 prokaryotic organisms (Fox et al., 1980), reveal the existence of "universal" and "eukaryote-specific" sequences (regions that are conserved among all organisms or all eukaryotes, respectively), "semiconserved" sequences (regions of intermediate conservation), and "nonconserved" sequences (regions that display very high rates of genetic drift). In Figure 4, the sequence homologies between three eukaryotic ribosomal RNAs are plotted as a function of nucleotide position. This representation of the homology data reveals the interspersed nature of several conserved and semiconserved sequences among the nonconserved regions within the srRNAs. Interspersed nature of conserved and semiconserved among nonconserved

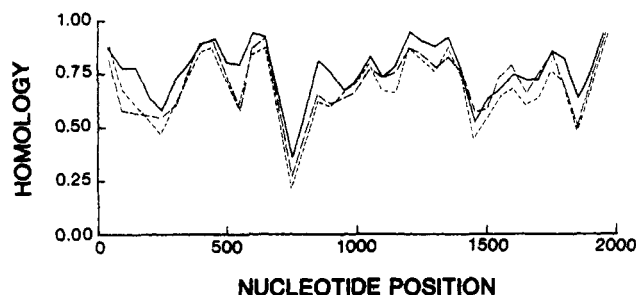


FIGURE 4: Pattern of sequence conservation in eukaryotic small-subunit ribosomal RNAs. Pairwise homologies for 100 nucleotide segments along the length of aligned eukaryotic srRNAs in Figure 2 were calculated as described under Results. Homologies for overlapping segments were plotted at increments of 50 nucleotides for *X. laevis* vs. *S. cerevisiae* (solid line), *D. discoideum* vs. *X. laevis* (dotted line), and *D. discoideum* vs. *S. cerevisiae* (dashed line).

regions has also been reported for 23S-28S rRNAs (Otsuka et al., 1983). The conserved sequences are essential for establishing meaningful alignments between small-subunit rRNAs but are not sufficient for constructing molecular phylogenies; because of a lack of sequence variation, the conserved regions do not contribute information about sequence divergence. The high rates of nucleotide substitutions in the nonconserved regions diminish their value in the construction of molecular phylogenies involving distantly related organisms, but they are valuable for quantitating very close genealogical relationships. Semiconserved regions may be used in pairwise sequence comparisons when inferring phylogenies if (1) they display statistically significant variation, (2) their

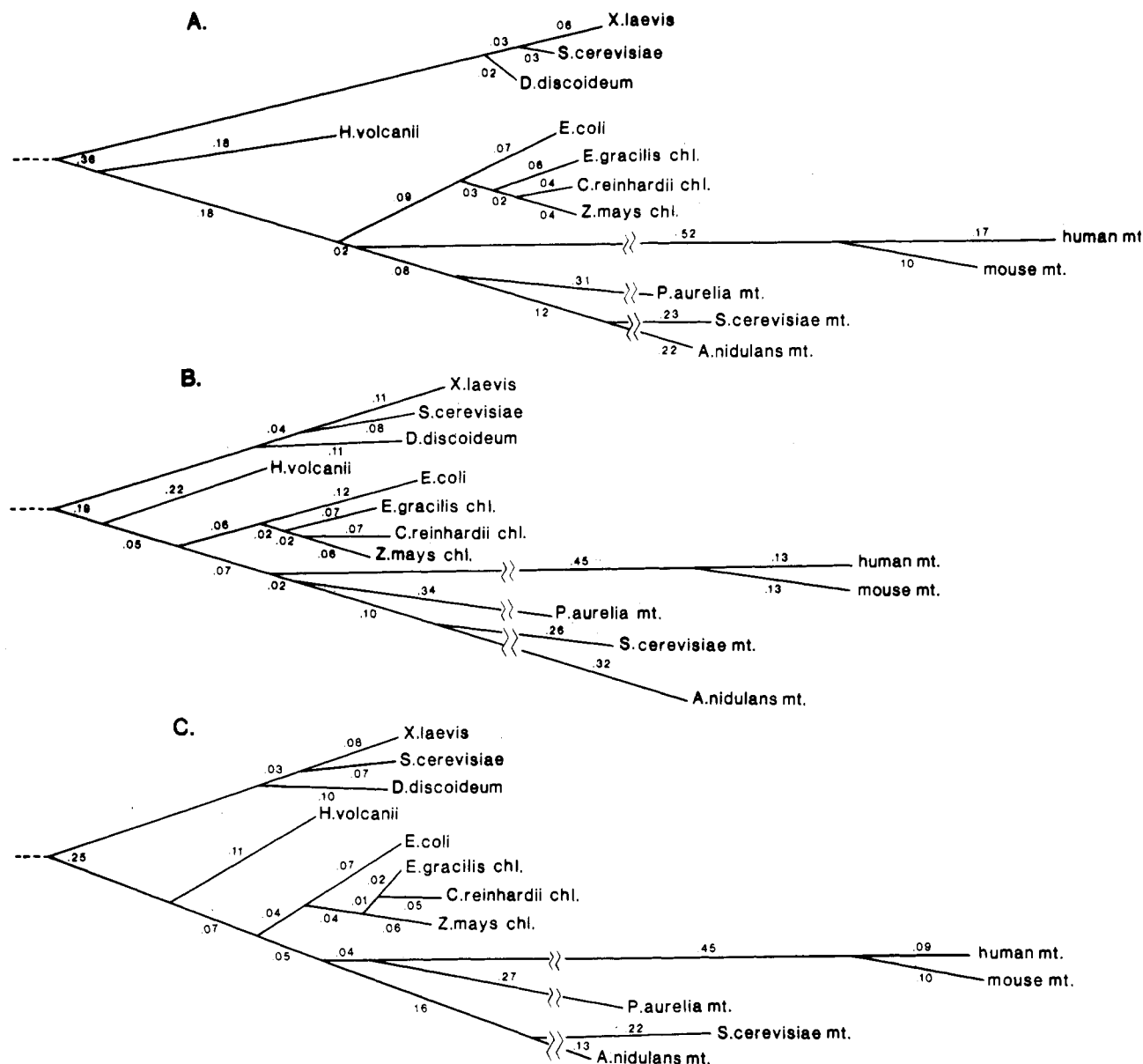


FIGURE 5: Phylogenetic relationships derived from comparisons of partial srRNA sequences. Homologies for conserved and semiconserved sequences in three regions of the srRNAs were determined and used to infer phylogenetic trees as described under Results. By use of the aligned numbering system in Figure 2, the trees were constructed from regions 355-573 plus 606-694 (panel A), 938-1107 plus 1184-1299 (panel B), and 1582-1641 plus 1701-1816 (panel C). The evolutionary distance between nodes of the tree is given alongside the segment connecting them and is represented in the horizontal component of their separation in the figure.

average rate of nucleotide substitution is consistent with that of the entire molecule, (3) they can be aligned with analogous eukaryotic and eubacterial sequences, and (4) they span a statistically significant number of independently variable nucleotide positions (200-400). To test the validity of using partial rRNA sequences for phylogenetic analyses, we have compared the phylogenetic tree based upon complete sequences with one derived from the summation of conserved and semiconserved regions (ca. 850 nucleotides). The homology values presented in the lower left portion of Table II represent pairwise comparisons of the homologous conserved and semiconserved nucleotide positions. These comparisons were used to construct the phylogenetic tree in Figure 3B as described above. Although the two trees in Figure 3 display identical branching patterns, the mean square standard deviation of the tree path lengths relative to the pairwise distance data (and, hence, the statistical accuracy of the tree) is improved when the regions of extensive structural variation are excluded.

We have also examined the possibility of inferring phylogenetic relationships from only a few of the conserved and

semiconserved regions in the small-subunit rRNAs. Figure 5 displays trees inferred from three regions of semiconserved sequence (referred to as regions I, II, and III; see Figure 5 legend). The phylogenies deduced from sequence regions I and II agree with those inferred from the complete sequences. The rearrangement of the chloroplast branching order in the tree inferred from region III reflects statistical fluctuation in the local sequence homologies (two changes in 170 nucleotides defined by this region are sufficient to cause this rearrangement). Similarly, the rearrangement of mitochondrial branching order is minor. In general, however, this analysis supports the notion that reliable phylogenies can be inferred from small-subunit rRNA homology computations that are restricted to conserved and semiconserved sequences.

Discussion

Homology Analyses. A homology matrix for the complete small-subunit rRNAs as well as the summation of conserved and semiconserved srRNA regions is presented in Table II. It can be seen that the homologies of the *D. discoideum* se-

quence with those of *X. laevis* and *S. cerevisiae*, are low, relative to the *X. laevis*/yeast values. This observation of a low homology of the *D. discoideum* sequence with other eukaryotes is reinforced by the T₁ oligonucleotide analyses of the small-subunit rRNAs from *B. emersonii* (M. L. Sogin and C. R. Woese, unpublished data) and *Lemna* (C. R. Woese and T. Dyer, unpublished data). The simplest interpretation of these data is that *D. discoideum* diverged from the mainstream of eukaryotic descent well before the branching of the Animalia, Fungi, and Plantae. (Comparisons of the *Lemna* T₁ oligonucleotide catalogs with other eukaryotic small-subunit rRNA sequences indicate that the radiation of plants occurred approximately at the same time as the animal-fungal arborization.) This interpretation is consistent with the generally low homologies between the 5S rRNA, 5.8S rRNA, and tRNA sequences of *D. discoideum* and those of other eukaryotes.

An examination of the available ribosomal RNA sequence data failed to identify other protoctists that might represent an earlier branching. Comparisons of the small-subunit rRNA T₁ oligonucleotide catalogs places *B. emersonii* near the yeast. Molecular phylogenies inferred from the 5.8S rRNA sequences also group *Tetrahymena pyriformis* (Fujiwara & Ishikawa, 1982) and *Acanthamoeba castellanii* (MacKay & Doolittle, 1981) with the Fungi. The 5.8S rRNA sequences of *Cryptothecodinium cohnii* (Hinnebusch et al., 1981), *Physarum polycephalum* (Otsuka et al., 1982, 1983) and *Crithidia fasciculata* (Schnare & Gray, 1982) indicate that those organisms represent branchings that are slightly deeper than that represented by *T. pyriformis*; this level of branching is comparable to the radiation of the Animalia, Fungi, and the Plantae (Olsen, 1983).

The low homology of the *D. discoideum* small-subunit rRNA sequence with those of other eukaryotes could represent: (1) a rapid rate of evolutionary drift (a fast evolutionary clock), (2) convergent evolution in the rRNAs of other eukaryotes, or (3) the earliest branching known in the eukaryotes. The divergence of the *D. discoideum* sequence does not represent a fast evolutionary clock. By comparison with a very distantly related organism, e.g., *E. coli*, it is possible to identify altered "clock speeds". If a sequence is subject to an aberrant clock speed, its divergence from a distantly related sequence will differ from the divergence of close relatives and the distantly related control sequence; the higher the mutation rate for a given sequence, the greater the divergence from a distantly related sequence. As can be seen from Table II, the homology between *E. coli* or *H. volcanii* and *D. discoideum* is similar to the homology between the other eukaryotic sequences and *E. coli* or *H. volcanii*. This theme is echoed by the phylogenetic trees in Figures 3 and 5. Phylogenetic trees are more revealing than raw homology data because they integrate the data for all pairs of sequences. The trees, whether inferred from complete sequences or various regions of conserved and semiconserved sequence, consistently represent the *D. discoideum* srRNA as slightly less divergent from those of the prokaryotes than are the *S. cerevisiae* and *X. laevis* sequences. Consequently, the data are most consistent with *D. discoideum* being the deepest branching of these eukaryotes. On the other hand, the divergence of the mitochondrial sequences is consistent with relatively fast clock speeds; they have diverged more from the eukaryotes than have the other eubacterial sequences.

The potential effects of sequence convergence upon the reliability of molecular phylogenies is a recurring question. Simple metabolic functions may have arisen independently

more than once, but there is no evidence that selection of a biochemical capability results in the selection of a particular nucleotide or amino acid sequence. Indeed, there are numerous examples where functionally equivalent macromolecules show no statistically significant sequence homology. For example, the three-dimensional structure and subunit composition of the vertebrate photoreceptor protein rhodopsin is very similar to that of bacterial rhodopsin (Dratz & Hargrave, 1983). Both proteins contain seven transmembrane helices and form similar types of photoproducts. Sequence comparisons of these proteins fail to reveal meaningful structural homologies. Another example can be drawn from the serine proteases, which have been extensively studied in both prokaryotes and eukaryotes. In many cases, they have similar three-dimensional structures and are functionally equivalent, but their primary structures are completely unrelated (Kraut, 1977). Whether this sequence diversity represents extreme divergence from a common ancestral sequence or functional convergence of independent sequences is irrelevant. In either case, functional selection does not result in sequence convergence.

Convergence within a given molecular phylogeny can be detected by comparison with phylogenetic trees that have been derived from functionally independent markers. If the trees are nearly identical, they must reflect divergent evolution from a common ancestor; it is improbable that multiple, functionally independent traits would converge at the same rate during evolution. Similarly, comparisons of phylogenies derived from several, independently variable functional domains of a single macromolecule can be used to detect convergence. The small-subunit rRNA contains numerous functional domains, distributed throughout the length of the molecule. As pointed out above, phylogenies inferred from different portions of the molecule (which almost certainly represent different functional domains), as well as phylogenies inferred from other components of the translation apparatus, are consistent with each other. Therefore, the deep divergence of the *D. discoideum* rRNA sequences is not a consequence of convergent evolution.

The methods used for constructing the phylogenetic relationships displayed in Figures 3 and 5 produce "unrooted" trees. The total trees can be viewed as collections of subtrees whose geometries are defined by sequence homologies within groups of "closely related" organisms. The root of each subtree is then determined by sequence comparisons of organisms within that subtree to one or more organisms that are known (by other criteria) to lie outside of the subtree.³ We have used several prokaryotic small-subunit rRNA sequences in the present analysis as "outside references" for establishing the root of the eukaryotic subtree. This analysis leads to the conclusion that the earliest branching within the eukaryotic subtree is that represented by *D. discoideum*. Similarly, we conclude that the mitochondria represent the deepest characterized branching within the eubacterial line of descent; the node corresponding to this branching is the eubacterial node closest to the eukaryotic subtree. The *H. volcanii* sequence is a representative of the archaeobacterial line of descent and, as previously shown (Gupta et al., 1983), lies outside of both the eukaryotic and the eubacterial subtrees.

Since there is no known line of descent to serve as an outside reference for the branching of the three fundamental lines of descent, the archaeobacteria, the eubacteria, and the eukaryotes (a condition that always holds for the deepest branching of

³ This description is intended to qualitatively explain the use of outside "reference" sequences. In practice, all tree segment lengths are evaluated simultaneously, and the relationships between sequences in different subtrees do influence the structure within the subtrees.

a phylogeny), there is no objective method for placing the root of the total tree. In the absence of a rationale for doing otherwise, the common ancestor to these has been arbitrarily placed so that the distance from the ancestor to the contemporary organisms is similar for the eukaryotic and eubacterial lines of descent. The resulting diagrams show the eukaryotes diverging prior to the separation of the two prokaryotic kingdoms from each other. This affiliation of the eubacteria and the archaeobacteria was not detected in comparisons of T₁ oligonucleotides (Woese & Fox, 1977) due to the lower resolution of the catalogs relative to complete sequence information.

The endosymbiotic theory for the origins of eukaryotic cells postulates that the chloroplasts and mitochondria arose from symbiotic relationships involving ancestral eukaryotes and early prokaryotes (Margulis, 1981).⁴ Although there also is evidence for polyphyletic origins of nuclear genes, it appears that the nuclear-defined ribosomal RNAs are derived from a common ancestor, termed the ancestral "urkaryote" (Woese & Fox, 1977). This view is consistent with the trees in Figure 3, which include three distinct "eukaryotic" lines of descent, the chloroplast, the mitochondrion, and the urkaryote. Several features of the tree are significant in this respect. The radiation of the chloroplasts is well represented in the tree, and there are no substantial evolutionary gaps since the divergence of this group from the *E. coli* line of descent. Although the data are consistent with a monophyletic origin, the chloroplast radiation cannot be compared with that of the corresponding eukaryotes until their nuclear-defined, small-subunit rRNA sequences are determined.

The available mitochondrial srRNA sequences also appear to be monophyletic. We arrive at this same conclusion using a variety of sequence-alignment and homology-computation techniques. Although different sequence alignments were employed and our analyses included a greater number of positions, the differences between our conclusions and those of Kuntzel & Köchel (1981) that mitochondria are polyphyletic probably reflect differences in the effect of fast evolutionary clock speeds upon the methods of tree construction.

If mitochondria are of monophyletic origin, then there should be a direct correspondence of the branching points in the mitochondrial and urkaryotic subtrees. In particular, the node at which yeast and *X. laevis* diverge (the animal/fungal division) in the urkaryotic tree corresponds to the node at which yeast and *A. nidulans* mitochondria separate from the human and mouse mitochondria. In Figure 3B this latter point lies only 0.043 unit from the point at which the mitochondrial line diverged from the rest of the eubacteria. Since the *D. discoideum* branch lies 0.026 unit deeper within the urkaryotic line, its mitochondrion is expected to branch from a point quite close to the *E. coli*/chloroplast/mitochondrial common ancestor. The determination of the phylogenetic status of the *D. discoideum* mitochondrial rRNA is of substantial interest since, if the above is true, it is to be expected that eukaryotes that represent substantially deeper branches than *D. discoideum* will not have mitochondria as we know them.

It is probable that the features common to organisms representing the major divisions of the eukaryotic kingdom will be reflections of features possessed by their common ancestors. The identification of these divisions requires the inference of reliable phylogenies on the basis of comparisons of multiple, functionally independent characteristics that can be found in

all organisms. The small-subunit ribosomal RNAs appear to satisfy this need.

Acknowledgments

We thank Dr. J. J. Seilhamer and D. J. Cummings for providing us with the *P. aurelia* mitochondrial small-subunit rRNA sequence prior to publication and to Dr. R. Gupta for providing us with the *H. volcanii* srRNA sequence prior to its publication. We also thank Dr. Norman Pace for his critical reading of the manuscript.

Supplementary Material Available

Documentation of T₁ oligonucleotide catalogs (Table III) (1 page). Ordering information is given on any current masthead page.

References

- Asada, K., Kanematsu, S., Okada, S., & Hayakawa, T. (1980) in *Chemical and Biochemical Aspects of Superoxide and Superoxide Dismutase* (Bannister, J. V., & Hill, H. A. O., Eds.) pp 136-153, Elsevier/North-Holland, Amsterdam.
- Brosius, J., Palmer, M. L., Kennedy, P. J., & Noller, H. F. (1978) *Proc. Natl. Acad. Sci. U.S.A.* 75, 4801-4805.
- Cockburn, A. F., Newkirk, M. J., & Firtel, R. A. (1976) *Cell (Cambridge, Mass.)* 9, 605-613.
- Cockburn, A. F., Taylor, W. C., & Firtel, R. A. (1978) *Chromosoma* 70, 19-29.
- Cocucci, S. M., & Sussman, M. (1970) *J. Cell Biol.* 45, 399-407.
- Dratz, E. A., & Hargrave, P. A. (1983) *Trends Biochem. Sci. (Pers. Ed.)* 8, 128-131.
- Dron, M., Rahire, M., & Rochemaix, J. D. (1982) *Nucleic Acids Res.* 10, 7609-7619.
- Eperon, I. C., Anderson, S., & Nierlich, D. P. (1980) *Nature (London)* 286, 460-467.
- Fitch, W. M. (1976) *J. Mol. Evol.* 8, 13-40.
- Fox, G. E., Stackebrandt, E., Hespell, R. B., Gibson, J., Maniloff, J., Dyer, T., Wolfe, R. S., Balch, W., Tanner, R., Magrum, L. J., Zablen, L. B., Blackmore, R., Gupta, R., Luehrs, K. R., Bonen, L., Lewis, B. J., Chen, K. N., & Woese, C. R. (1980) *Science (Washington, D.C.)* 209, 457-463.
- Fridovich, I. (1976) in *Free Radicals in Biology* (Pryor, W. A., Ed.) p 239, Academic Press, New York.
- Fujiwara, H., & Ishikawa, H. (1982) *Nucleic Acids Res.* 10, 5173-5182.
- Graf, L., Roux, E., Stutz, E., & Kössel, H. (1982) *Nucleic Acids Res.* 10, 6369-6381.
- Greene, P. J., Betlach, M. C., & Boyer, H. W. (1974) in *Methods in Molecular Biology* (Wickner, R. B., Ed.) pp 87-111, Marcel Dekker, New York.
- Grunstein, M., & Hogness, D. S. (1975) *Proc. Natl. Acad. Sci. U.S.A.* 72, 3961-3965.
- Gupta, R., Lanter, J., & Woese, C. R. (1983) *Science (Washington, D.C.)* 221, 656-665.
- Hassur, S. M., & Whitlock, H. W. (1974) *Anal. Biochem.* 69, 162-164.
- Hinnebusch, A. G., Klotz, L. C., Blanken, R. L., & Loeblich, A. R. (1981) *J. Mol. Evol.* 17, 334-347.
- Hori, H., & Osawa, S. (1979) *Proc. Natl. Acad. Sci. U.S.A.* 76, 381-385.
- Johnson, R. A., & Walseth, T. F. (1978) in *Advances in Cyclic Nucleotide Research* (Brooker, G., Greengard, P., & Robinson, G. A., Eds.) pp 135-167, Raven Press, New York.
- Kraut, J. (1977) *Annu. Rev. Biochem.* 46, 331-358.

⁴ It has also been suggested that similar events gave rise to other eukaryotic structures, e.g., 9 + 2 flagella (Margulis, 1981).

- Küntzel, H., & Köchel, H. G. (1981) *Nature (London)* 293, 751-755.
- Lumsden, J., & Hall, D. O. (1975) *Nature (London)* 257, 670-672.
- MacKay, R. M., & Doolittle, W. F. (1981) *Nucleic Acids Res.* 9, 3321-3334.
- Maizels, N. (1976) *Cell (Cambridge, Mass.)* 9, 431-438.
- Maniatis, T., Fritsch, E. F., & Sambrook, J. (1982) in *Molecular Cloning: A Laboratory Manual*, pp 93-94, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
- Margulis, L. (1981) in *Symbiosis in Cell Evolution* (Brewer, P., Ed.) W. H. Freeman, San Francisco.
- Margulis, L., & Schwartz, K. V. (1982) in *Five Kingdoms* (Brewer, P., Ed.) W. H. Freeman, San Francisco.
- Maxam, A. M., & Gilbert, W. (1980) *Methods Enzymol.* 65, 499-560.
- McLaughlin, P. J., & Dayhoff, M. O. (1973) *J. Mol. Evol.* 2, 99-116.
- Noller, H. F., & Woese, C. R. (1981) *Science (Washington, D.C.)* 212, 403-411.
- Olsen, G. J. (1983) Ph.D. Thesis, University of Colorado Health Sciences Center, Denver, CO.
- Olsen, G. J., & Sogin, M. L. (1982) *Biochemistry* 21, 2335-2343.
- Otsuka, T., Nomiyama, H., Sakaki, Y., & Takagi, Y. (1982) *Nucleic Acids Res.* 10, 2379-2385.
- Otsuka, T., Nomiyama, H., Yoshida, H., Kukita, T., Kuhara, S., & Sakaki, Y. (1983) *Proc. Natl. Acad. Sci. U.S.A.* 80, 3163-3167.
- Peffley, D. M., & Sogin, M. L. (1981) *Biochemistry* 20, 4015-4021.
- Rao, K. K., & Hall, D. O. (1977) in *The Evolution of Metalloenzymes, Metalloproteins and Related Materials* (Leigh, G. J., Ed.) p 39, Symposium Press, London.
- Rao, K. K., & Cammack, R. (1981) in *Molecular and Cellular Aspects of Microbial Evolution* (Carlile, M. J., Collins, J. F., & Moseley, B. E. B., Eds.) Cambridge University Press, Cambridge.
- Rubin, C. M., & Schmid, C. W. (1980) *Nucleic Acids Res.* 8, 4613-4619.
- Rubstov, P. M., Musakhanov, M. M., Zakharyev, V. M., Krayev, A. S., Skryabin, K. G., & Bayev, A. A. (1980) *Nucleic Acids Res.* 8, 5779-5794.
- Salim, M., & Maden, B. E. H. (1981) *Nature (London)* 291, 205-208.
- Sanger, F., Brownlee, G. G., & Barrell, B. G. (1965) *J. Mol. Biol.* 13, 373-398.
- Schnare, M. N., & Gray, M. W. (1982) *Nucleic Acids Res.* 10, 2085-2092.
- Schwarz, Zs., & Kössel, H. (1980) *Nature (London)* 283, 739-742.
- Sogin, M. L., & Olsen, G. J. (1980) *Gene* 8, 231-238.
- Sor, F., & Fukuhara, H. (1980) *C. R. Hebd. Seances Acad. Sci., Ser. D* 291, 933-936.
- Stackebrandt, E., & Woese, C. R. (1981) in *Molecular and Cellular Aspects of Microbial Evolution* (Carlile, M. J., Collins, J. F., & Moseley, B. E. B., Eds.) Cambridge University Press, Cambridge.
- Tu, C. P., & Cohen, S. N. (1980) *Gene* 10, 177.
- Uchida, T., Bonen, L., Schaup, H. W., Lewis, B. J., Zablen, L., & Woese, C. R. (1974) *J. Mol. Evol.* 3, 63-77.
- Uhlenbeck, O. C., & Cameron, V. (1977) *Nucleic Acids Res.* 4, 85-98.
- Van Etten, R. A., Walberg, M. W., & Clayton, D. A. (1980) *Cell (Cambridge, Mass.)* 22, 157-170.
- Woese, C. R., & Fox, G. E. (1977) *Proc. Natl. Acad. Sci. U.S.A.* 74, 5088-5090.
- Ysunobu, K. T., & Tanaka, M. (1973) in *Iron-Sulfur Proteins* (Lovenberg, W., Ed.) Vol. 2, p 27, Academic Press, New York.
- Zukerkandl, E., & Pauling, L. (1965) *J. Theor. Biol.* 8, 357-366.