Eukaryotes versus Prokaryotes: An Estimate of Evolutionary Distance

Abstract. *The divergence of nucleated organisms and bacteria was 2.6 times more remote in evolution than the divergences of the nucleated organisms into separate kingdoms, as evidenced by genetic changes in cytochrome c and transfer RNA. The development of the genetic code through the differentiation of transfer RNA's for different amino acids was still more remote in evolution. The overall rates of transfer RNA evolution in bacteria and nucleated organisms were comparable.*

Recently, biologists have realized that the most basic division between organisms is not that between animals and plants. Rather, the difference between the eukaryotes (nucleated organisms) and the prokaryotes (bacteria and blue-green algae), which is marked by presence or absence of true nuclei, mitosis, meiosis, and certain cell organelles, is the most important distinction in the course of evolution (*1*). We now add further evidence derived from protein and nucleic acid sequences and tentatively quantify this major divergence.

We estimated evolutionary distance by the quantity of difference in unit characters between kingdoms; the quantity of difference between prokaryotes and eukaryotes was compared to the quantity of difference between separate eukaryote kingdoms. The unit characters which we counted are defined by three criteria. (i) The unit characters are determined by coding in the chromosmal DNA. (ii) Each unit character can be independently altered by a point mutation (single nucleotide change) or other single genetic event, such as the insertion of a small section of DNA into the gene. (iii) A change in one unit character can alter the efficiency of the organism and can influence the effect of natural selection on the organism. A well-known example of such a change is the substitution of valine for glutamic acid in position six of the β chain of hemoglobin, which results in sickle-cell anemia (*2*).

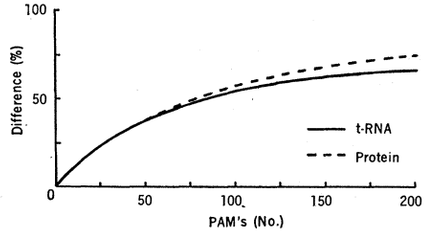

Fig. 1. The relation of the observed percentage of difference between sequences to PAM's (accepted point mutations per 100 positions). These are correction curves for superimposed mutations. The solid line represents comparisons of tRNA sequences; the dashed line represents comparisons of protein sequences (*5*).

The unit characters which we used are the individual amino acids making up proteins and the nucleotides composing the metabolically functional ribonucleic acids (RNA). The information for our estimate of evolutionary distance was elicited from the sequences of two types of molecules. The first type is cytochrome c, a part of the electron transport system. Although this protein functions in the mitochondria of eukaryotes, the code for its structure is contained in the nuclear DNA (*3*). The related protein cytochrome c_2, from the photoanaerobe *Rhodospirillum rubrum*, appears to function in the photosynthetic electron transfer chain of this bacterium (*4*). The second type of molecule is composed of the transfer RNA's (tRNA's), which transport specific amino acids into a protein chain as it is being synthesized.

Twenty-two sequences were used for comparisons in the cytochrome group and eleven in the tRNA's (*5*). We have used as many sequences of cytochrome c and tRNA as possible in order to average the variations in individual species. The enzymatic modifications of the tRNA nucleotides were ignored in our comparisons.

The cytochrome sequences contain an average of 108 unit characters and the tRNA's have 58 unit characters. The number of unit characters in each sequence was determined by removing from the total sequence those positions which would not have changed independently in the course of evolution. For both the cytochromes and tRNA's, it is very probable that additions or deletions of several adjacent residues were due to a single event and thus should be compared as a single change; therefore, all but one of the residues of such an addition were removed before comparisons. For the tRNA's, the anticodon was removed because it is part of the definition of a tRNA as to type; generally, any change in the anticodon would alter the basic function of the molecule. Also, one side of each base-paired region was removed from the tRNA sequences. These base-paired regions maintain the conformation of the molecule (*6*), and in all of the known sequences only 5 percent of the pairs are mismatched. Whenever one member of a pair mutates, selection pressure evidently favors a second mutation in the pair to reestablish the bonding; thus the evolution of the two sides is not independent.

In comparing two sequences from diverse organisms which have evolved separately for some time, we expect that any one position might have undergone superimposed or parallel mutations. Obviously these multiple changes became more numerous over longer periods of evolution. A direct count of changed positions would underestimate the actual number of separate changes and thus reduce the apparent evolutionary distance. Therefore, using the curves of Fig. 1, we have estimated the true number of changes. These curves were based on a model of random mutations (*5*). The model reflected the un-

Cytochrome c comparisons

	Plant	Fungi	Prokaryotes (c_2)
Animals	40.5	44.9	66.1
Plant		49.3	69.0
Fungi			74.3

Comparisons within tRNA types

Types of tRNA	Eukaryote vs. Eukaryote	Eukaryote vs. Prokaryote
tRNA^{Phe}	6	16.5
tRNA^{Tyr}		23.5
tRNA^{Val}		22.5
tRNA^{Ser}	9	

Comparisons between tRNA types

Types of tRNA	tRNA^{Tyr}	tRNA^{Val}	tRNA^{Ser}
tRNA^{Phe}	22.4	20.9	24.1
tRNA^{Tyr}		25.1	21.7
tRNA^{Val}			26.1

Fig. 2. The average numbers of amino acid differences in comparisons between cytochrome sequences and of nucleotide differences in comparisons between tRNA's. The cytochrome figures are averages of from 1 to 51 comparisons of sequences of about 108 units; those for tRNA are averages of from 1 to 9 comparisons of sequences of 58 units (*5*).

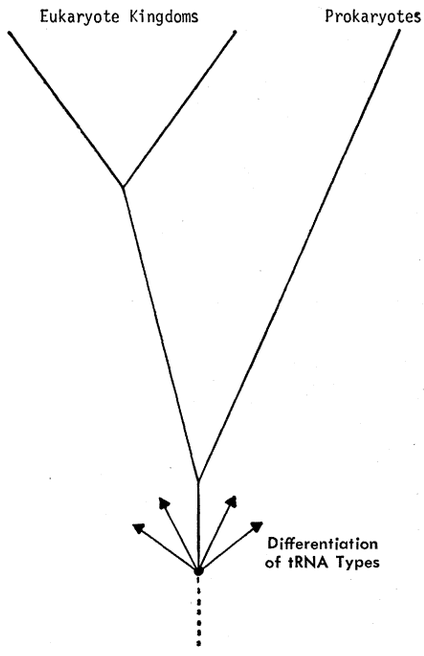

Fig. 3. The evolutionary distances between eukaryote kingdoms, between eukaryotes and prokaryotes, and between tRNA types. Not shown are the further differentiations of each type of tRNA with the divergences of groups of organisms. Those events giving rise to greater amounts of differences are shown as happening proportionately earlier in evolutionary history.

equal frequency of occurrence of the amino acids and nucleotides and the unequal rates of change of one amino acid or nucleotide to another, based on the information actually observed in the total available data on sequences. We express the adjusted number of accepted point mutations per 100 positions in units called PAM's (*5*).

Sequences of each eukaryote and each prokaryote were compared, and the number of differences was averaged within each group (Fig. 2). Then sequences of members of different kingdoms of eukaryotes were compared and averaged. Animals, green plants, and fungi were considered to be three separate kingdoms which diverged from each other within a relatively short time (*7*). In each comparison, the percentage of difference in unit characters was converted to PAM's. The average distance between the different eukaryote kingdoms is 58 PAM's for cytochrome and 15 PAM's for tRNA's. Between the eukaryotes and prokaryotes there are 137 PAM's for cytochrome and 53 for tRNA's. Although four different types of tRNA are represented in the data, complete series of sequences are not known for all of them. At least two types contribute to each estimate of distance; thus, we considered the tRNA's as equivalent to two sets of comparisons comprising 116 unit characters. To combine the data from the different kinds of molecules in proportion to the numbers of unit characters counted, the cytochrome average was weighted as 48 percent and added to the tRNA average, which was weighted as 52 percent. The eukaryote kingdoms are separated by an average of 36 PAM's, whereas the prokaryotes are 93 PAM's distant from the eukaryotes (Fig. 3). The quantity of difference, in unit character changes, from the eukaryotes to the prokaryotes is 2.6 times as great as the difference between eukaryote kingdoms. Using an alternate method, we calculated the evolutionary distance ratio for the combined data by weighting the ratio from each type of molecule in inverse proportion to its variance. This procedure yields a similar value of 2.7 for the combined ratio. The standard error in this figure is 0.4, or 14.7 percent (*8*). Our results are consistent with those of Jukes (*9*) who, in an unadjusted comparison of four tRNA's, estimated that the divergence of bacteria and plants was about twice as remote as the divergence of yeasts and higher plants.

Thus, it can be seen that there is greater similarity among the eukaryote groups than between any of these and the prokaryotes, and that the animals, green plants, and fungi developed from a common ancestral stock which differs greatly from the prokaryotes.

The divergences of these major groups of living organisms can be related to an even earlier group of events in a very primitive organism—the development of the genetic code through the proliferation and distinction of genes for different tRNA molecules capable of transporting different amino acids. Using the same technique and data on tRNA as given above, we made comparisons between the tRNA's which carry different amino acids (Fig. 2). The differences between these tRNA types average 62 PAM's. The ratio of the evolutionary distance between tRNA types to the evolutionary distance between prokaryotes and eukaryotes is 1.2. In Fig. 2 this differentiation is placed at a distance proportional to the scale of the other divergences.

The relationship of time to the amount of change in unit characters is of great interest. One might expect the rate of change in sequences to correspond to the rate of change in morphology and thus that the bacteria would show less change than the eukaryotes. On the other hand, one might expect the bacterial sequences to change more rapidly, as indicated by the ease of producing mutant strains in the laboratory.

We can compare the rate of change in tRNA in the two major groups of organisms. Within the prokaryote line, and separately within the eukaryote line, we compared tRNA's of different types, for example, valine tRNA and phenylalanine tRNA from *Escherichia coli*. The time interval is the same for these comparisons because the divergence of the tRNA's took place in a common ancestor before the divergence of the eukaryote and prokaryote lines. A number of separate counts of this type are possible for the tRNA's of phenylalanine, valine, and tyrosine. There is an average of 22.2 changes for the bacterial line and an average of 22.7 changes for the eukaryote lines, with a standard deviation of ± 2.8 changes. From these data, it is evident that the overall rates of tRNA evolution in the bacterial and eukaryote lines are comparable. We have therefore drawn Fig. 3 as though the rates were the same.

P. J. McLaughlin
M. O. Dayhoff
National Biomedical Research Foundation, 11200 Lockwood Drive, Silver Spring, Maryland 20901

References and Notes

1. L. Margulis, *Science* **161**, 1020 (1968); L. Sagan, *J. Theor. Biol.* **14**, 225 (1967); A. Allsopp, *New Phytol.* **68**, 591 (1969).
2. M. Murayama, *Clin. Chem.* **14**, 578 (1967).
3. T. S. Work, J. L. Coote, M. Ashwell, *Fed. Proc.* **27**, 1174 (1968); S. Yamagata and R. Sato, *J. Biochem. (Tokyo)* **64**, 549 (1968); F. Sherman, J. W. Stewart, E. Margoliash, J. Parker, W. Campbell, *Proc. Nat. Acad. Sci. U.S.* **55**, 1498 (1966).
4. For cytochrome c_2 sequence determination, see K. Sletten, K. Dus, H. deKlerk, M. D. Kamen, *J. Biol. Chem.* **243**, 5492 (1968); K. Dus, K. Sletten, M. D. Kamen, *ibid.*, p. 5507. Cytochrome c_2 is the only sequenced protein from a prokaryote which can be quantitatively compared with eukaryote sequences, for it is clearly related to the eukaryote cytochromes along the whole length of the molecule. A comparable amount of change has occurred in all regions. The sequence of cytochrome c_{551} from *Pseudomonas fluorescens* has only a few points of resemblance with eukaryote cytochromes, mainly confined to the first portion of the sequence. We cannot unambiguously determine the correspondence of the last portion of this sequence to the other cytochromes. Thus there is no valid basis for a quantitative comparison of unit characters using the cytochrome c_{551}.
5. M. O. Dayhoff, Ed., *Atlas of Protein Sequence and Structure 1969* (National Biomedical Research Foundation, Silver Spring, Md., 1969),

vol. 4. The experimental determinations of sequences by numerous workers are referenced, and the sequences are listed in the Data Section, edited by L. T. Hunt and P. J. McLaughlin. The sequences used as data for this paper are those for cytochrome c from human, horse, pig, gray whale, rabbit, gray kangaroo, chicken, king penguin, Pekin duck, pigeon, snapping turtle, Puget Sound dogfish, Pacific lamprey, fruit fly, screwworm fly, silkworm moth, tobacco hornworm moth, wheat, *Neurospora crassa*, baker's yeast, and *Candida krusei* and for cytochrome c_2 from *Rhodospirillum rubrum*. The tRNA sequences used are phenylalanine tRNA from baker's yeast, wheat, and *Escherichia coli*; tyrosine tRNA from baker's yeast, torula yeast, and *E. coli*; valine tRNA from baker's yeast, torula yeast, and *E. coli*; serine tRNA from baker's yeast and rat. Two sequences not listed in the *Atlas* are valine tRNA from *E. coli* [M. Yaniv and B. G. Barrell, *Nature* **222**, 278 (1969)] and phenylalanine tRNA from *E. coli* [B. G. Barrell and F. Sanger, *Fed. Eur. Biochem. Soc. Lett.* **3**, 275 (1969)]. Two recently determined sequences of cytochrome c from green plants support our calculations [D. Boulter, M. V. Laycock, J. A. M. Ramshaw, E. W. Thompson, in *Phytochemical Phylogeny*, J. B. Harborne, Ed. (Academic Press, London and New York, in press)]. Because cytochrome c changes more slowly than most proteins for which we can calculate mutation rates, a larger and more representative sample of protein types might give higher numbers for the evolutionary distances. However, the ratio of evolutionary distances would not be affected. For a table of mutation rates of several proteins, see P. S. McLaughlin and M. O. Dayhoff, in *Atlas of Protein Sequence and Structure 1969*, M. O. Dayhoff, Ed. (National Biomedical Research Foundation, Silver Spring, Md., 1969), vol. 4, chap. 6. For information on PAM's for proteins, see M. O. Dayhoff, R. V. Eck, C. M. Park, in *ibid.*, chap. 9.

6. R. W. Holley, J. Apgar, G. A. Everett, J. T. Madison, M. Marquisee, S. H. Merrill, J. R. Penswick, A. Zamir, *Science* **147**, 1462 (1965); F. Cramer, H. Doepner, F. v.d. Haar, E. Schlimme, H. Seidel, *Proc. Nat. Acad. Sci. U.S.* **61**, 1384 (1968); D. H. Metz and G. L. Brown, *Biochemistry* **8**, 2329 (1969).
7. This is supported by our sequence comparisons and by biological evidence [R. H. Whittaker, *Science* **163**, 150 (1969)].
8. The variance of mutation data conforms closely to the theoretical formula for a binomial distribution, npq, where n is the length of the sequence, p is the fraction of positions which have changed, and q is the fraction of positions which have not changed. We used this formula in estimating the error.
9. T. H. Jukes, *Space Life Sci.* **1**, 469 (1969).
10. We thank Drs. L. T. Hunt, W. C. Barker, and C. M. Park for helpful discussions and assistance. Supported by NIH grant GM-08710 and NASA contract 21-003-002.

10 April 1970 ■

Scanning Electron Microscopy of Developing Plant Organs

Abstract. *Shoot apices and young meristematic leaves can be examined directly with the scanning electron microscope without prior fixation or metal coating. The form of the shoot apex, cellular organization, and leaf arrangement (phyllotaxis) can be observed, perhaps as they have never been visualized before.*

A large volume of literature is available on phyllotaxis and on the form, structure, and development of shoot apices and leaves. Ordinarily, the results of these investigations are based on observations of sectioned material with the light microscope or, in some instances, of fresh material with a dissecting microscope. Permanent photographic records of entire living shoot apices are often difficult to obtain. In some instances, material is chemically fixed and stained, and the entire object is photographed in the conventional way (*1*). These preparations are often excellent, but they frequently lack a three-dimensional quality.

Einert *et al.* (*2*) have examined *Lilium* apices with the scanning electron microscope (SEM); however, their preparative procedure involved freeze-drying and exposure to acetone, which resulted in some distortion of the tissues.

It was recently shown (*3*) that surfaces of fragments of mature leaves may be examined in the wet and fresh condition with the SEM without recourse to metal coating to prevent an accumulation of surface charge. We found the same to be true for meristematic regions of the plant.

We wish to report that the SEM is an excellent instrument for examining delicate meristematic regions such as the shoot apex itself and young leaf primordia or floral appendages, at least in many species. Moreover, the SEM offers resolution and depth of field not always attainable by light microscopy. Figure 1 is representative of our research with the shoot tips of *Tropaeolum*. Older leaves were removed until the shoot apex and young leaf primordia were visible. Then the shoot tip was attached to a specimen stub with an animal-hide glue. No fixatives, de-

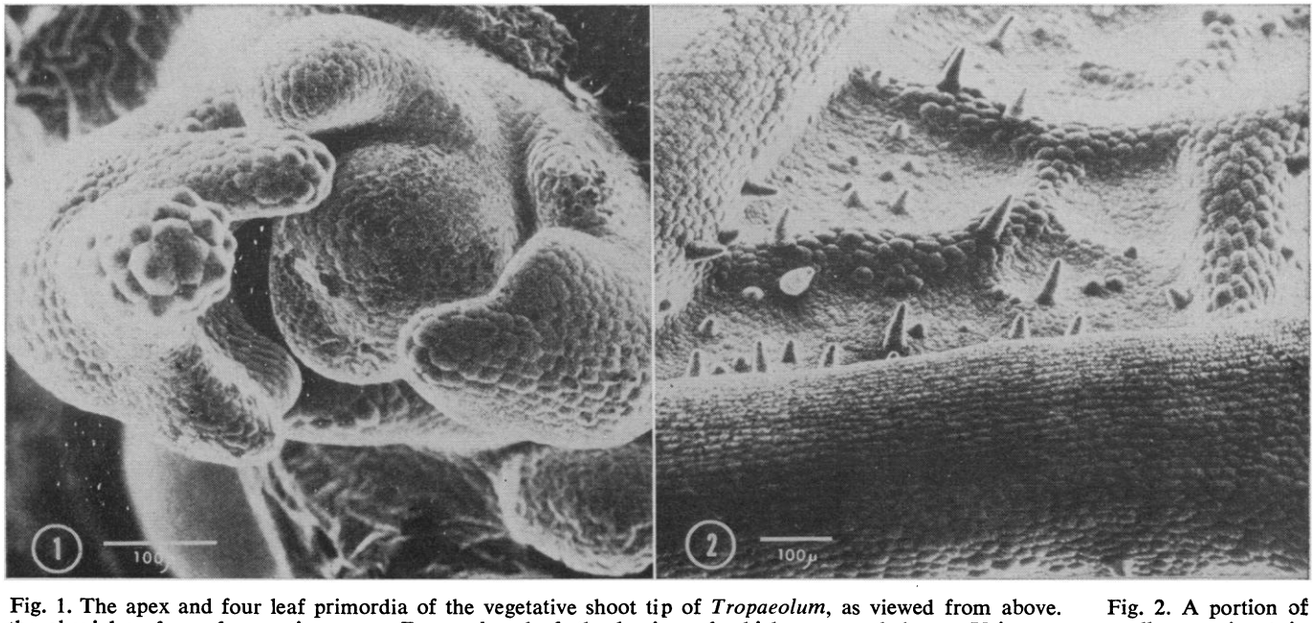

Fig. 1. The apex and four leaf primordia of the vegetative shoot tip of *Tropaeolum*, as viewed from above. Fig. 2. A portion of the abaxial surface of an entire young *Tropaeolum* leaf, the lamina of which measured 1 cm. Veins are normally prominent in leaves of *Tropaeolum*.