

On the evolutionary origin of the plant mitochondrion and its genome

(ribosomal RNA genes/chlorophytes/metaphytes/phylogeny)

MICHAEL W. GRAY*, ROBERT CEDERGREN†, YVON ABEL‡, AND DAVID SANKOFF‡

*Department of Biochemistry, Dalhousie University, Halifax, Nova Scotia, B3H 4H7 Canada; and †Département de Biochimie and ‡Centre de Recherches Mathématiques, Université de Montréal, Montréal, Québec, H3C 3J7 Canada

Communicated by Carl R. Woese, December 1, 1988

ABSTRACT Higher plants occupy very different positions in the mitochondrial and nuclear lineages of global phylogenetic trees based on conserved regions of small subunit (SSU) and large subunit (LSU) rRNA sequences. In the nuclear subtree, plants branch off late, at a position reflecting a massive radiation of the major multicellular (and some unicellular) groups; in the mitochondrial subtree, in contrast, plants branch off early, near the point of connection between the mitochondrial and eubacterial lineages. Moreover, in the nuclear lineage, plants branch together with the unicellular green alga *Chlamydomonas reinhardtii*, whereas in the mitochondrial lineage (in both SSU and LSU trees), metaphytes and chlorophyte branch separately. Statistical evaluation indicates that the anomalous branching position of higher plants in the mitochondrial lineage is not a treeing artifact attributable to the relatively rapid rate of sequence divergence of non-plant mitochondrial rRNA sequences. In considering alternative biological explanations for these results, we are led to propose that the rRNA genes in plant mitochondria may be of more recent evolutionary origin than the rRNA genes in other mitochondria. This proposal has implications for monophyletic vs. polyphyletic scenarios of mitochondrial origin and is consistent with other evidence indicating that plant mtDNA is an evolutionary mosaic.

An origin of chloroplasts and mitochondria from bacterial endosymbionts (1–3) is strongly supported by comparisons of small subunit (SSU) and large subunit (LSU) rRNA sequences (4). Phylogenetic trees based on such analyses not only confirm the eubacterial ancestry of the rRNA genes encoded by the genomes of chloroplasts (5) and mitochondria (5–8) but have further traced the specific origin of these genes to the cyanobacteria (chloroplasts) (8) and the α -subdivision of the purple bacteria (mitochondria) (9). Interest is now shifting to the question of whether the postulated endosymbioses occurred only once (10) or more than once (11, 12) (i.e., whether mitochondria and chloroplasts are monophyletic or polyphyletic in origin).

A monophyletic origin of mitochondria most readily reconciles the fact that the basic function of mtDNA is the same in all eukaryotes examined (13). However, extraordinary differences in size, arrangement, and mode of expression are exhibited by the range of mitochondrial genomes (14), and this structural diversity has hampered efforts to elucidate the evolutionary history of mitochondria (4). To circumvent this difficulty, we developed a method, based on analysis of a highly conserved structural core in rRNA, that generates global phylogenetic trees interrelating very distant organisms and organelles (8). We have now analyzed and compared phylogenies determined in parallel from separate SSU and LSU rRNA data bases, using an improved version of the

treeing algorithm, which incorporates the “bootstrap” statistical technique (15, 16) to evaluate the significance of the results. This has provided a better perspective on mitochondrial phylogeny than was previously possible. Here, we discuss further data bearing on the question of a polyphyletic origin of mitochondria, data that raise the possibility that the rRNA genes in plant mitochondria originated separately, at a later time in evolution, than the mitochondrial rRNA genes of other eukaryotes.

METHODOLOGY

A complete description of the phylogenetic methodology, including a detailed treatment of the molecular cladistics problem, local optimization using temporary constraints, and bootstrap analyses, is presented elsewhere (17, 18).

To test the possible effect of unequal rates of sequence divergence on tree topology, particularly in the mitochondrial lineage, we have used methods designed specifically to avoid the “long branches attract” artifact (see below). Cavender and Felsenstein (19) and Lake (20) each present criteria for choosing the best of three possible binary trees derived from a set of four sequences. To apply these criteria, we first constructed separate consensus rRNA sequences for animal mitochondria, plant mitochondria, and eubacteria. For the SSU analysis, we used only those eubacteria closest to the mitochondrial subtree (i.e., the purple bacteria, α , β , and γ , Fig. 1), whereas in the LSU analysis we used the *Escherichia coli* sequence alone as the eubacterial representative. In aligning these consensus sequences with each other and with the corresponding SSU or LSU rRNA sequence of *Chlamydomonas reinhardtii* mitochondria, we discarded all positions involving a deletion or an insertion in any of the sequences as well as positions at which more than two of the four possible nucleotides appeared in more than one of the three consensus sequences. (Since the Cavender–Felsenstein criteria have been worked out for two-state characters only, all sequences were rewritten in terms of purines vs. pyrimidines prior to that part of the analysis.) This left 458 (Lake) or 466 (Cavender–Felsenstein) of 489 positions for the SSU data base and 620 (Lake) or 625 (Cavender–Felsenstein) of 642 positions for the LSU data base. In applying these methods to positions containing alternatives (in one consensus sequence), each possibility was analyzed separately, and the results were weighted fractionally, as appropriate.

RESULTS AND DISCUSSION

Congruency of Phylogenetic Trees. Statistically validated and highly concordant SSU and LSU rRNA-based phylogenies have been presented in ref. 18. Both trees confirm the eubacterial character of the chloroplast and mitochondrial lineages, with the same affiliations between organellar and

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.

Abbreviations: SSU, small subunit; LSU, large subunit.

eubacterial subtrees in each case. All mitochondrial sequences form a single grouping that in the SSU tree branches with the α -subdivision of the purple bacteria (Fig. 1), a result confirming that of Yang *et al.* (9). Moreover, the relative branching order (Fig. 1) for four mitochondrial SSU rRNA sequences (wheat, mouse, *Aspergillus nidulans*, *Paramecium primaurelia*) is the same as that determined in ref. 9 using a distance matrix method applied to a larger number of positions. The mitochondrial/eubacterial topology is not an artifact of the especially pronounced structural similarities between plant mitochondrial and eubacterial rRNA sequences (cf. ref. 21), because when the plant mitochondrial data are omitted, there is no change in the point at which the mitochondrial subtree branches from the eubacterial one.

In the nuclear and mitochondrial subtrees of both the SSU and LSU trees, plant sequences occupy very different positions. In the nuclear subtree, plants branch off at about the same point as animals and fungi (Fig. 2 and ref. 22), a position reflecting a massive and relatively late radiation of these groups (23). In the mitochondrial subtree (Fig. 1), in contrast, higher plants cluster very near the root. Also, in the mito-

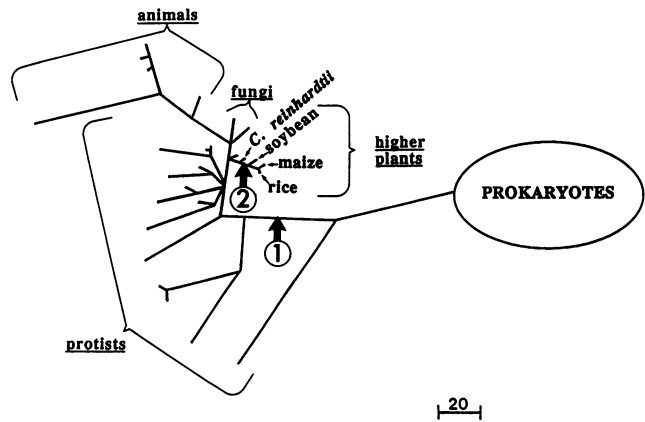


FIG. 2. Nuclear portion of the unrooted SSU phylogenetic tree described in Fig. 1 (see ref. 18), showing the putative proto-mitochondrial endosymbioses discussed in the text. Step 1, early endosymbiosis postulated to have given rise to the mitochondria/mitochondrial genomes of most eukaryotes; step 2, late endosymbiosis postulated to have contributed the rRNA genes (and perhaps other components) of higher plant mitochondria.

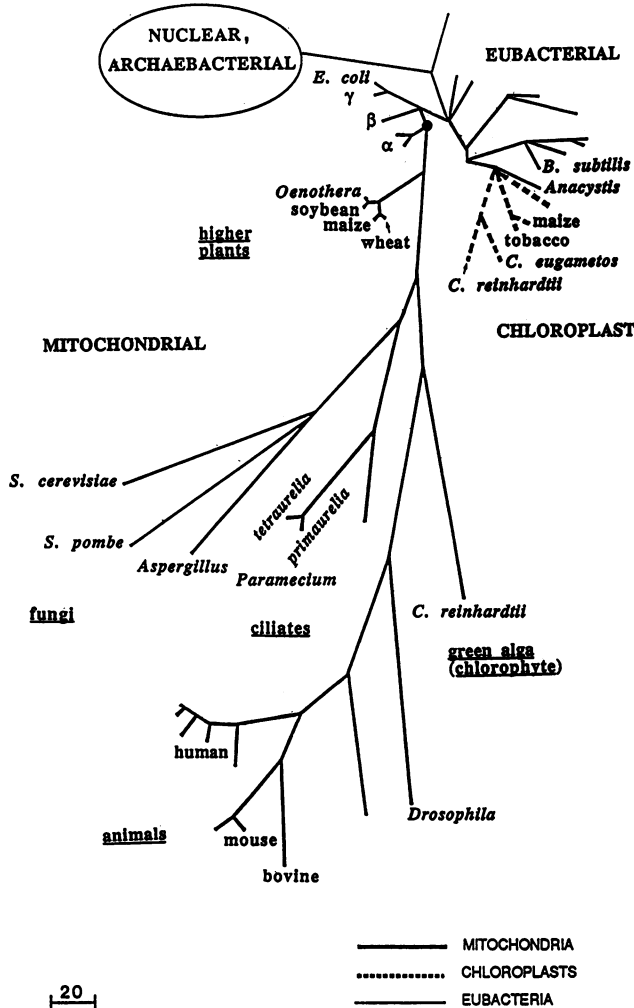


FIG. 1. Eubacterial-chloroplast-mitochondrial portion of an unrooted phylogenetic tree inferred from a SSU rRNA sequence data base (see ref. 18), comprising a total of 76 sequences (9 archaeobacterial, 15 eubacterial, 26 eukaryotic nuclear, 5 chloroplast, and 21 mitochondrial). The length of each branch is proportional to the number of fixed mutational events represented by its two endpoints, as indicated by the scale. This tree is redrawn from data presented in ref. 18, where the identity of all organisms at terminal nodes is given and bootstrap values indicating the robustness of each internal branch are provided.

chondrial subtree, plant and green algal (*C. reinhardtii*) sequences branch separately (Fig. 1), whereas in the nuclear lineage, these sequences form a clade (Fig. 2 and ref. 22), as they do also in the chloroplast lineage (Fig. 1). The clustering of the nuclear SSU rRNA sequences of green algae and higher plants is consistent with traditional phylogenies that group chlorophytes and metaphytes together (24). The green alga/higher plant dichotomy is, in fact, the major incongruity between the nuclear and mitochondrial topologies. Evidently, this anomaly lies not in the branching position of the chlorophyte mitochondrial sequence but in the position of the plant mitochondrial sequences (these branch in the same place relative to other mitochondrial sequences whether or not the chlorophyte sequence is included in the analysis).

Is the Anomalous Position of Plant Mitochondria a "Slow Clock vs. Fast Clock" Artifact? A notable feature of the mitochondrial subtrees (both SSU and LSU) is the long branch lengths leading to ciliates, fungi, animals, and chlorophyte, reflecting the relatively high rate of primary sequence divergence in the mitochondrial genes of most eukaryotes (e.g., refs. 25, 26). In contrast, plant mitochondrial genes diverge in sequence at a much lower rate (27). Under probabilistic models, there is a risk that the parsimony criterion will group rapidly evolving lines vs. slowing evolving lines, even when this is not the correct branching configuration (28). It could be argued, therefore, that a treeing artifact has incorrectly grouped the conservative plant mitochondrial sequences together with the eubacteria, to the exclusion of the chlorophyte and other more rapidly evolving mitochondrial sequences.

To address this objection, we have made use of the Cavender and Felsenstein (19) and Lake (20) methods to evaluate possible topologies (Fig. 3) relating the chlorophyte (*C. reinhardtii*) mitochondrial rRNA sequences and consensus sequences for the corresponding rRNA species of animal mitochondria, plant mitochondria, and eubacteria. For the SSU rRNA data, the Cavender-Felsenstein criterion strongly supports the parsimony solution (tree 1: plant mitochondria + eubacteria vs. chlorophyte mitochondria + animal mitochondria), whereas the Lake criterion strongly supports this configuration for the LSU rRNA data (Table 1). In applying the Cavender-Felsenstein criterion to the LSU data, both tree 1 and tree 3 (chlorophyte mitochondria + plant mitochondria vs. animal mitochondria + eubacteria) are strongly supported against tree 2 (plant mitochondria +

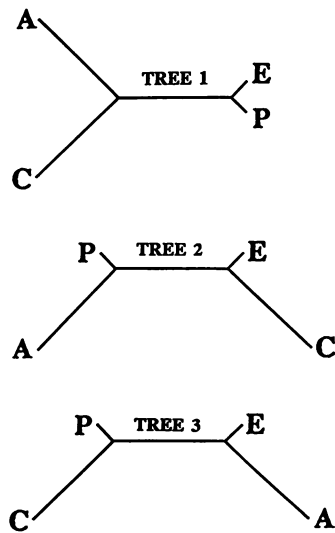


FIG. 3. Possible alternative tree topologies resulting from a four-way pairwise comparison of homologous chlorophyte (*C. reinhardtii*) mitochondrial (C), plant mitochondrial (P), animal mitochondrial (A), and eubacterial (E) rRNA sequences. See Table 1.

animal mitochondria vs. chlorophyte mitochondria + eubacteria), with little to choose between the former two. In applying the Lake criterion to the SSU data, none of the three possibilities is preferred. In summary, the parsimony solution (tree 1, Fig. 3) is clearly preferred in two of four tests, with no clear preference in the other two. Thus, to the extent that we are able to test for a long branches attract artifact, we conclude that differential rate of sequence divergence is not responsible for the early branching position of plant mitochondria.

Position of Higher Plants in the Mitochondrial Lineage.

Leaving aside for the moment the issue of tree topology, the fact remains that plant mitochondrial rRNAs are substantially more similar to their eubacterial/chloroplast counterparts than they are to their homologs in other mitochondria (4). In addition, there is evidence (4) that wheat mitochondrial SSU rRNA contains posttranscriptionally modified nucleosides that are characteristic of eubacterial SSU rRNA,

Table 1. Tests of the long branches attract explanation of the position of plant mitochondria in the rRNA phylogenies (see Fig. 3)

Trees compared	Cavender-Felsenstein criterion*			
	SSU		LSU	
	Score	Supports	Score	Supports
2 vs. 3	0.0024	3 (+)	-0.0166	3 (+++)
3 vs. 1	-0.0213	1 (+++)	0.0036	3 (+)
1 vs. 2	0.0189	1 (+++)	0.0130	1 (+++)
Tree	Lake criterion†			
	SSU		LSU	
	1.5‡		4.4§	
	1.1‡		1.1‡	
Tree 3	2.3‡		0.1‡	

*In the Cavender-Felsenstein test, a large positive value of the criterion favors the first of the alternative trees being compared, whereas a negative value favors the second alternative. (+), Tree weakly supported; (+++), tree strongly supported.

†In the Lake test, a statistic is computed for each tree, and this value is χ^2 -distributed if that tree does not account for the data any better than other configurations (χ^2 , $df = 1$).

‡Not significant.

§Significant, $P < 0.05$.

including a 3-methyluridine (m^3U) residue at a site homologous to position 1498 (m^3U) in *E. coli* 16S rRNA (29). This latter modification is not present in yeast (30), hamster (31), or *Tetrahymena* (32) mitochondrial SSU rRNA. Finally, plant mitochondria contain a unique species of 5S rRNA (33), encoded by plant mtDNA (34); neither this RNA nor the corresponding gene has been found in the mitochondria of any other eukaryote. These observations must be accommodated in any reconstruction of the evolutionary history of the mitochondrial lineage.

Another consideration is the very different way in which genes are organized and expressed in the small [16 kilobase pairs (kbp)] *C. reinhardtii* mitochondrial genome and the large (200–2400 kbp) plant mitochondrial genomes. *C. reinhardtii* mtDNA approximates the size of animal mtDNA, and the two share many of the same features of gene organization and expression (35), including extensive physical and transcriptional linkage of genes, processing of long cotranscripts by discrete endonucleolytic scissions, and a virtual absence of any 5'-untranslated region in mature mRNAs. In all of these characteristics of mitochondrial genome organization and expression, *C. reinhardtii* differs from higher plants (36), with no indication that the two shared a common mitochondrial ancestor as recently as they shared a common nuclear (or chloroplast) ancestor.

How Can We Rationalize the Anomalous Position of Plant Mitochondrial rRNA Sequences? There seems to us to be two ways to explain the quite different branching position of plants within the nuclear and mitochondrial lineages and to account for the especially strong eubacterial characteristics of plant mitochondrial rRNAs. Either (i) the mitochondrial rRNA genes of plants have diverged relatively little from the rRNA genes of an ancient eubacterial ancestor of all mitochondria or (ii) the mitochondrial rRNA genes of higher plants have been acquired more recently than the mitochondrial rRNA genes of other eukaryotes.

Alternative 1 (monophyletic origin). This scenario assumes that plant mitochondrial SSU and LSU rRNA sequences, which are demonstrably slowly evolving (Fig. 1 and ref. 36), are representative of the ancestral types present in a eubacteria-like protomitochondrion that was the ancestor of all contemporary mitochondria. Characteristics common to eubacteria and plant mitochondria (posttranscriptional rRNA modifications, a 5S rRNA gene) are considered to be ancestral traits that must have been present in the last common ancestor of the two.

If mitochondria are monophyletic, and if the protomitochondrion was established at an early stage in eukaryotic evolution (step 1 in Fig. 2), then we must postulate that mitochondrial rRNAs and their genes evolved away from this ancestral pattern in all eukaryotic groups *except* higher plants. This would have had to involve not only primary sequence divergence but also (i) loss of the capacity to carry out at least some posttranscriptional modifications in rRNA and (ii) loss of the 5S rRNA gene. Such divergence and loss would have had to occur a number of times independently in the course of evolution (in all those eukaryotic groups that according to the nuclear tree, branched off both before and after higher plants) to account for the selective retention of an ancestral pattern in plant mitochondria. Because this scenario implies that the common ancestor of green algae and higher plants also possessed the primitive (i.e., plant mitochondria-like) rRNA sequence pattern, we must further postulate a selective sequence divergence away from the ancestral pattern in the mitochondrial rRNA genes of *C. reinhardtii* but not those of higher plants. At the same time, however, the nuclear rRNA genes of *C. reinhardtii* and higher plants obviously diverged relatively little (Fig. 2). Finally, we must also assume that the mitochondrial translation system could have accommodated, readily and repeat-

edly, to the loss of 5S rRNA, a species known to be essential in other protein synthesizing systems (37).

Alternative 2 (biphyletic origin). An alternative interpretation of the data presented here is that, insofar as rRNA genes are concerned, plant mitochondria shared a more recent common ancestor with eubacteria than did the mitochondria of other eukaryotes. This implies that plant mitochondrial rRNA genes were acquired at a relatively late stage in the evolution of the eukaryotic cell—i.e., *after* the separation of *C. reinhardtii* and higher plants from their last common ancestor (at step 2 in Fig. 2). This would not only account for the unusually high structural similarity between plant mitochondrial and eubacterial rRNA genes but could also explain the selective presence of a 5S rRNA gene in plant mtDNA, assuming this gene was acquired at the same time as the SSU and LSU rRNA genes.

Implications of this "Secondary Acquisition" Hypothesis. The most likely route by which rRNA genes could have been introduced selectively into the lineage leading to higher plants would have been through a secondary symbiotic (presumably endosymbiotic) event. We suppose that the host cell taking up such a(n) (endo)symbiont was a primitive green alga already containing a mitochondrion. Thus, acquisition of novel mitochondrial rRNA genes would have had to involve (i) formation of a(n) (endo)symbiosis between host cell (a eukaryote) and a donor cell (a eubacterium), (ii) lateral transfer of rRNA genes from (endo)symbiont to the mitochondrial genome of the host, and (iii) integration of transferred rRNA genes into the mitochondrial genome, presumably with displacement of rRNA genes already present. A number of observations are relevant to each of these premises.

(a) The pivotal role of symbiosis in cell evolution is now well established (38), and there are numerous examples of endosymbioses involving eukaryotic host cells and bacterial symbionts (38, 39). Thus, there is ample precedence for the coexistence of mitochondria and symbionts in the same cell. Although some endosymbioses of this type may be transient and unstable, others are obligate and involve an absolute interdependence of host and endosymbiont (39). In the latter case, mutual adaptation of the partners can occur very rapidly (40). Such adaptation must involve accommodation at the level of nuclear and symbiont genomes but could conceivably involve the mitochondrial genome of the host, as well.

(b) Members of the α -subdivision of the purple bacteria maintain an intimate, usually intracellular, relationship with contemporary eukaryotic cells and include such plant symbionts as the agrobacteria and rhizobacteria. These eubacteria (the evolutionary source of plant mitochondrial rRNA genes) have recently been implicated in DNA exchanges with their eukaryotic hosts (41, 42).

(c) Transfer of genetic information from symbiont to mitochondrial genome might be expected to be favored in a proto-algal host cell deficient in either mtDNA or mitochondrial function, as exemplified by *C. reinhardtii* mutants devoid of mtDNA (43) or lacking cytochrome oxidase and cyanide-sensitive respiration (44). Although the latter are nuclear gene mutants, mutations in mitochondrial genes might be expected to produce a similar (respiratory-deficient) phenotype. Such algal mutants, if they were able to survive as obligate photoautotrophs, would be ideal candidate hosts for the type of lateral gene transfer postulated here.

(d) A distinctive feature of the plant mitochondrial genome is its ability to incorporate foreign genetic information, both chloroplast (45) and nuclear (46). Prominent among such "promiscuous DNA" are rRNA and tRNA genes. Although there is no evidence that the former are functional in plant mitochondria, some of the latter give rise to mature tRNAs, and in at least one case there is suggestive evidence that a

promiscuous chloroplast tRNA may actually function in translation in plant mitochondria (47). Thus, the plant mitochondrial genome may be considered a mosaic, with evidence for the acquisition of genetic information (and possibly even active genes) from several distinct sources in the course of evolution (48).

If the rRNA genes of plant mtDNA were indeed a secondary acquisition, might other plant mitochondrial genes have been acquired at the same time from the postulated endosymbiont? Like rRNA genes, tRNA genes in plant mitochondria (those not obviously derived from chloroplast DNA) are, on average, substantially more similar to their eubacterial homologs than they are to the same tRNA genes in other mitochondria (49). Lateral transfer of a block of closely linked rRNA and tRNA genes [as found, e.g., in the *Bacillus subtilis* genome (50)] might account for this fact. Other late acquisitions might be the genes for those proteins (e.g., the α -subunit of mitochondrial F₁ ATPase and a homolog of *E. coli* ribosomal protein S13) (see ref. 4) uniquely encoded by plant mtDNA.

Are Mitochondria Polyphyletic? Could the entire plant mitochondrion have been derived by a second endosymbiosis? Such an event would have had to involve complete displacement or subsumption of the original mitochondrion of the host cell by the new endosymbiont that would become the contemporary mitochondrion of plants. This in turn would presumably have necessitated a transitional state in which the original mitochondrion coexisted with the new symbiont. Assuming that the genome of the new endosymbiont contained many more genes than presently exist in plant mtDNA, there would then have had to be massive loss of endosymbiont genetic information and/or its transfer to the nuclear genome of the host—a prime assumption of any earlier endosymbiosis that produced the original mitochondrion (2). A process of "mitochondrial succession" (displacement of a preexisting mitochondrial system by a new one) would necessarily require extensive accommodation at the level of the nuclear genome to reestablish the coordinated nuclear-cytoplasmic interactions that are essential for mitochondrial biogenesis and function (13). Such accommodation might not be all that implausible, given the relative rapidity with which it has been observed to occur between symbiont and host in some contemporary situations (40).

CONCLUSIONS

Validation or rejection of the secondary acquisition hypothesis we present here will obviously require further comparative data; in particular, additional mitochondrial genomes within the chlorophyte-metaphyte grouping must be examined. If the thesis developed here is correct, then we should expect to see an abrupt transition from a prototypical chlorophyte mitochondrial genome (if such can be defined) to a prototypical metaphyte mitochondrial genome somewhere within the green algal lineage leading directly to higher plants. Further studies of the α -purple bacteria could also provide valuable information bearing on our hypothesis. The idea that the rRNA genes of plant mitochondria originated in a separate event at a later stage in evolution would be further strengthened if it could be demonstrated that these genes derive from a *different subgroup* of the α -purple bacteria than the rRNA genes of other mitochondria.

That the modern eukaryotic cell is an evolutionary mosaic (1, 38) can no longer be disputed. It is also clear, however, that the genomes within a eukaryotic cell are not absolutely isolated from one another but that genetic information has moved from one subcellular compartment to another in the course of evolution. Thus, we must also recognize that genomes themselves are mosaics to different extents and that lateral gene flow constitutes an important part of their

evolutionary history. The mosaic nature of the plant mitochondrial genome (48) is now well established. The data and arguments presented here add a new dimension and perspective to this problem.

We thank Suzanne Drolet for carrying out the Cavender–Felsenstein and Lake procedures, colleagues (cited in ref. 18) who provided us with unpublished sequences, and Lisa Laskey for her assistance in the preparation of this manuscript. This work was supported by individual operating grants from the Natural Sciences and Engineering Research Council of Canada (NSERC) (to M.W.G., R.C., and D.S.), an NSERC infrastructure grant and a CRAY computer time allotment (to R.C. and D.S.), and a Medical Research Council of Canada operating grant (to M.W.G.). M.W.G., R.C., and D.S. are Fellows in the Evolutionary Biology Program of the Canadian Institute for Advanced Research.

1. Margulis, L. (1970) *Origin of Eukaryotic Cells* (Yale Univ. Press, New Haven, CT).
2. Gray, M. W. & Doolittle, W. F. (1982) *Microbiol. Rev.* **46**, 1–42.
3. Gray, M. W. (1983) *BioScience* **33**, 693–699.
4. Gray, M. W. (1988) *Biochem. Cell Biol.* **66**, 325–348.
5. McCarroll, R., Olsen, G. J., Stahl, Y. D., Woese, C. R. & Sogin, M. L. (1983) *Biochemistry* **22**, 5858–5868.
6. Küntzel, H. & Köchel, H. G. (1981) *Nature (London)* **293**, 751–755.
7. Köchel, H. G. & Küntzel, H. (1982) *Nucleic Acids Res.* **10**, 4795–4801.
8. Gray, M. W., Sankoff, D. & Cedergren, R. J. (1984) *Nucleic Acids Res.* **12**, 5837–5852.
9. Yang, D., Oyaizu, Y., Oyaizu, H., Olsen, G. J. & Woese, C. R. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 4443–4447.
10. Cavalier-Smith, T. (1987) *Ann. N.Y. Acad. Sci.* **503**, 55–71.
11. Raven, P. J. (1970) *Science* **169**, 641–646.
12. Stewart, K. D. & Mattox, K. R. (1984) *J. Mol. Evol.* **21**, 54–57.
13. Attardi, G. & Schatz, G. (1988) *Annu. Rev. Cell Biol.* **4**, 289–333.
14. Gray, M. W. (1982) *Can. J. Biochem.* **60**, 157–171.
15. Diaconis, P. & Efron, B. (1983) *Sci. Am.* **248**, 116–130.
16. Felsenstein, J. (1985) *Evolution* **39**, 783–791.
17. Sankoff, D., Abel, Y., Cedergren, R. J. & Gray, M. W. (1988) in *Classification and Related Methods of Data Analysis*, ed. Bock, H. H. (North-Holland, Amsterdam), pp. 385–394.
18. Cedergren, R., Gray, M. W., Abel, Y. & Sankoff, D. (1988) *J. Mol. Evol.* **28**, 98–112.
19. Cavender, J. A. & Felsenstein, J. (1987) *J. Classif.* **4**, 57–71.
20. Lake, J. A. (1987) *Mol. Biol. Evol.* **4**, 167–191.
21. Spencer, D. F., Schnare, M. N. & Gray, M. W. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 493–497.
22. Gunderson, J. H., Elwood, H., Ingold, A., Kindle, K. & Sogin, M. L. (1987) *Proc. Natl. Acad. Sci. USA* **84**, 5823–5827.
23. Sogin, M. L., Elwood, H. J. & Gunderson, J. H. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 1383–1387.
24. Ragan, M. A. & Chapman, D. J. (1978) *A Biochemical Phylogeny of the Protists* (Academic, New York).
25. Brown, W. M., Prager, E. M., Wang, A. & Wilson, A. C. (1982) *J. Mol. Evol.* **18**, 225–239.
26. Morin, G. B. & Cech, T. R. (1988) *Nucleic Acids Res.* **16**, 327–346.
27. Wolfe, K. H., Li, W.-H. & Sharp, P. M. (1987) *Proc. Natl. Acad. Sci. USA* **84**, 9054–9058.
28. Felsenstein, J. (1978) *Syst. Zool.* **27**, 401–410.
29. Schnare, M. N. & Gray, M. W. (1982) *Nucleic Acids Res.* **10**, 3921–3932.
30. Klootwijk, J., Klein, I. & Grivell, L. A. (1975) *J. Mol. Biol.* **97**, 337–350.
31. Dubin, D. T., Taylor, R. H. & Davenport, L. W. (1978) *Nucleic Acids Res.* **5**, 4385–4397.
32. Schnare, M. N., Heinonen, T. Y. K., Young, P. G. & Gray, M. W. (1986) *J. Biol. Chem.* **261**, 5187–5193.
33. Spencer, D. F., Bonen, L. & Gray, M. W. (1981) *Biochemistry* **20**, 4022–4029.
34. Bonen, L. & Gray, M. W. (1980) *Nucleic Acids Res.* **8**, 319–335.
35. Gray, M. W. & Boer, P. H. (1988) *Phil. Trans. R. Soc. London Ser. B*, **319**, 135–147.
36. Sederoff, R. R. (1987) *Am. Nat.* **130**, S30–S45.
37. Erdmann, V. A. (1976) *Prog. Nucleic Acids Res. Mol. Biol.* **18**, 45–90.
38. Margulis, L. (1981) *Symbiosis in Cell Evolution* (Freeman, San Francisco).
39. Sitte, P. & Hansmann, P. (1986) *Prog. Bot.* **48**, 30–55.
40. Jeon, K. W. (1987) *Ann. N.Y. Acad. Sci.* **503**, 359–371.
41. Furner, I. J., Huffman, G. A., Amasino, R. M., Garfinkel, D. J., Gordon, M. P. & Nester, E. W. (1986) *Nature (London)* **319**, 422–427.
42. Carlson, T. A. & Chelm, B. K. (1986) *Nature (London)* **322**, 568–570.
43. Gillham, N. W., Boynton, J. E. & Harris, E. J. (1987) *Curr. Genet.* **12**, 41–47.
44. Wiseman, A., Gillham, N. W. & Boynton, J. E. (1977) *J. Cell Biol.* **73**, 56–77.
45. Stern, D. B. & Palmer, J. D. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 1946–1950.
46. Schuster, W. & Brennicke, A. (1987) *EMBO J.* **6**, 2857–2863.
47. Marechal, L., Runeberg-Roos, P., Grienerberger, J. M., Colin, J., Weil, J. H., Lejeune, B., Quetier, F. & Lonsdale, D. M. (1987) *Curr. Genet.* **12**, 91–98.
48. Schuster, W. & Brennicke, A. (1988) *Plant Sci.* **54**, 1–10.
49. Joyce, P. B. M., Spencer, D. F., Bonen, L. & Gray, M. W. (1988) *Plant Mol. Biol.* **10**, 251–262.
50. Green, C. J., Stewart, G. C., Hollis, M. A., Vold, B. S. & Bott, K. F. (1985) *Gene* **37**, 261–266.