Proc. Natl. Acad. Sci. USA
Vol. 87, pp. 5317–5321, July 1990
Evolution

Different fates of the chloroplast *tufA* gene following its transfer to the nucleus in green algae

(elongation factor Tu/gene transfer/organelle genomes/gene duplication/multigene family)

SANDRA L. BALDAUF*, JAMES R. MANHART†, AND JEFFREY D. PALMER*

*Department of Biology, Indiana University, Bloomington, IN 47405; and †Department of Biology, Texas A&M University, College Station, TX 77843

Communicated by David L. Dilcher, May 8, 1990 (received for review March 23, 1990)

ABSTRACT **Previous work suggested that the *tufA* gene, encoding protein synthesis elongation factor Tu, was transferred from the chloroplast to the nucleus within the green algal lineage giving rise to land plants. In this report we investigate the timing and mode of transfer by examining chloroplast and nuclear DNA from the three major classes of green algae, with emphasis on the class Charophyceae, the proposed sister group to land plants. Filter hybridizations reveal a chloroplast *tufA* gene in all Ulvophyceae and Chlorophyceae and in some but not all Charophyceae. One charophycean alga, *Coleochaete orbicularis*, is shown to contain an intact but highly divergent chloroplast *tufA* gene, whose product is predicted to be nonfunctional in protein synthesis. We propose that a copy of the *tufA* gene was functionally transferred from the chloroplast to the nucleus early in the evolution of the Charophyceae, with chloroplast copies of varying function being retained in some but not all of the subsequently diverging lineages. This proposal is supported by the demonstration of multiple *tufA*-like sequences in *Coleochaete* nuclear DNA and in nuclear DNA from all other Charophyceae examined.**

Chloroplasts and mitochondria encode only a small subset of the proteins necessary for their function, the rest being encoded in the nucleus and posttranslationally imported into the organelles. Characterization of organelle proteins encoded by nuclear genes shows that many are eubacterial in nature (1, 2). According to endosymbiotic theory these genes arose by direct transfer from the organelles, which once were free-living eubacteria. The conservation of gene content among organelles of distantly related taxa suggests that most gene transfer occurred early in organelle evolution (3, 4). However, evidence of modern gene transfer has been accumulating (5, 6), suggesting that the process continues, albeit at a greatly reduced rate.

The plant *tufA* gene encodes the chloroplast protein synthesis elongation factor Tu (EF-Tu). A chloroplast-localized *tufA* has been sequenced from *Euglena gracilis* (7) and from chlorophycean (*Chlamydomonas reinhardtii*, ref. 5) and ulvophycean (*Codium fragile*, M. Kuhsel and J.D.P., unpublished data) green algae. However, *tufA* is missing from the chloroplast DNA (cpDNA) of all examined land plants, including a bryophyte (*Marchantia polymorpha*, ref. 8), and has been found in the nuclear DNA (ncDNA) of the land plant *Arabidopsis thaliana* (5). Phylogenetic analysis suggests that *tufA* was transferred from the chloroplast to the nucleus within the green algal lineage giving rise to land plants (5).

Five classes of green algae are recognized, with most taxa being assigned to the classes Charophyceae, Chlorophyceae, and Ulvophyceae (9). To further characterize the transfer of the *tufA* gene, we have investigated its structure and subcellular location in members of these three classes of green algae by a combination of filter hybridization and gene sequencing. Among the Charophyceae, an unusual chloroplast *tufA*‡ was found in the genus *Coleochaete*, the proposed sister group to land plants (10, 11).

MATERIALS AND METHODS

Filaments of *Spirogyra maxima* and *Sirogonium melanosporum* were obtained from unialgal cultures grown in soil/water medium (12) on a 16:8 hr light:dark cycle at 20°C ± 2°C under fluorescent light at an illumination of 50 $\mu mol \cdot m^{-2} \cdot s^{-1}$. *Nitella translucens* and *Chara connivens* were grown in aquaria in a soil/water solution in a greenhouse. *Coleochaete orbicularis* was grown at 20°C in D11 solution (13) under 24-hr fluorescent light. Plants of *Cladophora* sp. were collected from a stream at the Matthaei Botanical Gardens of the University of Michigan.

cpDNAs and ncDNAs of the above algae were extracted from total DNA preparations (14) by centrifugation in cesium chloride and bisbenzimide H33258 (15). Other algal DNAs were generously provided by J. E. Boynton (Duke University), A. W. Coleman (Brown University), M. Li-Weber (Max-Planck-Institute, F.R.G.), C. Lemieux (Universite Laval, Quebec), and R. Meints (Oregon State University). Nonflowering land plant cpDNAs were prepared as described (16). Crucifer ncDNAs were extracted from Percoll-gradient isolated nuclei (17). *Brassica campestris* mitochondrial DNA was purified using DNase I (18) and cpDNA by density gradient centrifugation (19).

The single *Coleochaete orbicularis* chloroplast *tufA* was mapped to two adjacent cpDNA *Pst* I fragments of 2.1 and 8.6 kilobases (kb), and a 5.8-kb *Hin*dIII fragment was found to overlap the junction between the two *Pst* I fragments. Complete sequencing of the 2.1-kb *Pst* I fragment showed that it contained the bulk of the *tufA* gene. The remaining 5′ end of the gene was then determined from the 5.8-kb *Hin*dIII fragment using synthetic primers. Sequences were determined for both strands by dideoxy chain-termination, and all restriction sites were sequenced across.

Restriction enzyme digestion, agarose gel electrophoresis, Southern transfer, preparation of ^{32}P-labeled probes, and hybridization were performed as described (20) with minor modifications (5). Hybridizations were at 60°C, and filters were washed after hybridization, once at room temperature and three times at 60°C in 0.3 M NaCl/30 mM sodium citrate/0.5% SDS.

Deduced protein sequences were aligned by eye, maximizing the alignment of identical and conserved amino acids (21), while introducing a minimum of gaps. Amino acids were scored for all informative sites as equally weighted, unordered, multistate characters. Gaps were scored as missing data. Phylogenetic trees were calculated by parsimony criteria

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: cpDNA, chloroplast DNA; ncDNA, nuclear DNA; EF-Tu, elongation factor Tu.

‡The sequence reported in this paper has been deposited in the GenBank data base (accession no. M34286).

using PAUP version 3.0a (D. Swofford, Illinois Natural History Survey) on a Macintosh II computer. Tree topologies were searched with the general heuristic algorithm of PAUP. Separate searches were done using four different addition regimes (i.e., simple, closest, asis, and random) holding four trees at each step (except for the random addition). All searches yielded the same single shortest tree.

RESULTS

Distribution of *tufA* in cpDNAs of Green Algae and Land Plants. Within the three classes of green algae analyzed, a *tufA*-specific probe shows strong hybridization to cpDNAs from all examined members of the Ulvophyceae and Chlorophyceae and some but not all members of the Charophyceae (Fig. 1). Of the three orders of Charophyceae examined, members of the Charales (*Chara* and *Nitella*) show strong *tufA* signals, a member of the Coleochaetales (*Coleochaete*) gives only weak signals, and no signal at all can be detected in the Zygnematales (*Spirogyra* and *Sirogonium*). No *tufA* hybridization is seen to cpDNA from several basally derived branches of tracheophytes (Fig. 1), further supporting the notion that *tufA* is absent from all land plant cpDNAs (5, 8, 22). Note that all hybridizations were conducted at conditions under which a probe for the chloroplast *rbcL* gene hybridized strongly to all cpDNAs tested (Fig. 1 *Bottom*).

The apparent presence of *tufA* in cpDNAs from three classes of green algae suggests that the gene was ancestrally present in green algal cpDNA (5). This is also consistent with the proposed endosymbiotic origin of chloroplasts. The apparent lack of *tufA* in some charophycean and all land plant cpDNAs supports the previous designation of the Charophyceae as the algal lineage giving rise to land plants (11) and suggests that transfer of the gene occurred in the charophycean lineage.

An Unusually Divergent Chloroplast *tufA* from *Coleochaete*. Of the weakly hybridizing *tufA* signals of *Coleochaete* (Fig. 1) only the 2.1-kb *Pst* I fragment was found to correspond to a cpDNA fragment. This and an overlapping *Hin*dIII fragment were further examined by sequencing, and a single open reading frame of 1245 bp encoding an EF-Tu-like sequence was found.‡ This deduced EF-Tu sequence of 415 amino acids aligns throughout its length with all other chloroplast and cyanobacterial EF-Tus except for extensions of three amino acids each at the amino and carboxyl termini. None of the internal insertions/deletions characteristic of other eubacterial, eukaryotic, or mitochondrial EF-Tus are found.

The *Coleochaete* sequence is, however, unusually divergent. Sequence similarity is <55% overall with all other known EF-Tus (Table 1). In contrast, the lowest level of similarity found among all other chloroplast EF-Tus is 71% (*Codium fragile* versus *Cryptomonas* ϕ). The EF-Tus of *Arabidopsis thaliana*, a flowering plant, and *Thermotoga maritima*, a member of probably the earliest diverging group of eubacteria (27), are still 64% identical (Table 1).

The *Coleochaete* sequence also differs considerably at what are otherwise conserved amino acid positions. Of the 297 positions that are identical in sequence in nearly all cyanobacterial and chloroplast EF-Tus, 105 are altered in *Coleochaete* (Fig. 2). Of these changes, ≈25% involve nonconservative amino acid substitutions (21). The *Coleochaete* EF-Tu also differs at 22 positions that are nearly universally conserved in all eubacteria, archaebacteria, and eukaryotes (Fig. 2).

Mutations at amino acid positions 24, 236, and 394 (Fig. 2), corresponding to positions 20, 222, and 375 of *Escherichia coli* EF-Tu (24), have been characterized in *E. coli*. Position 24 lies in a phosphate-binding loop (28), and a glycine substitution at this position in *E. coli* results in a 10-fold reduction in GDP-binding and a 3-fold reduction in the overall rate of protein synthesis (29). *Coleochaete* EF-Tu contains the potentially much more disruptive substitution of a phenylalanine at this position (21, 30). Substitutions of either an aspartate at position 236 or a valine or threonine at position 394 in *E. coli* all result in the production of frameshifting errors *in vivo* and a decrease in cell growth rate (31). As shown in Fig. 2, both positions are altered in *Coleochaete* EF-Tu.

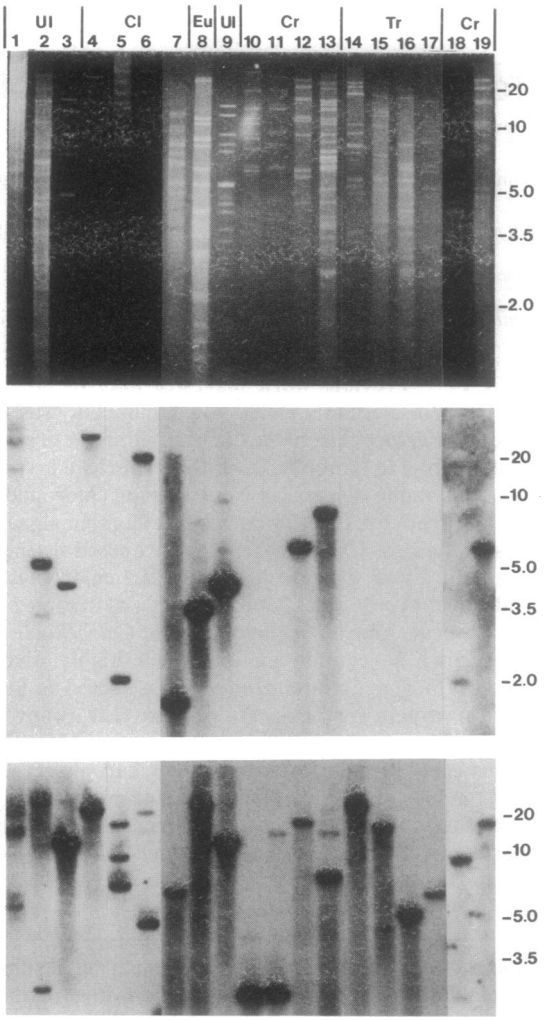

FIG. 1. *tufA* hybridizes to cpDNAs of all green plants except some charophycean algae and all land plants. (*Top*) Portions of three 0.9% agarose gels containing restriction digests of ulvophycean (Ul), chlorophycean (Cl), charophycean (Cr), euglenophycean (Eu), and tracheophyte (Tr) cpDNAs. Lanes: 1, *Acetabularia mediterranea*, *Eco*RI; 2, *Cladophora* sp., *Eco*RI; 3 and 9, *Codium fragile*, *Cla* I; 4, *Chlorella* sp., *Kpn* I; 5, *Chlamydomonas eugametos*, *Eco*RI; 6, *Pandorina morum*, *Bgl* II; 7, *Chlamydomonas reinhardtii*, *Eco*RI; 8, *Euglena gracilis*, *Eco*RI; 10, *Sirogonium melanosporum*, *Nco* I; 11, *Spirogyra maxima*, *Nco* I; 12 and 19, *Chara connivens*, *Pst* I; 13, *Nitella translucens*, *Eco*RI; 14, *Equisetum arvense*, *Xho* I; 15, *Lycopodium digitatum*, *Eco*RV; 16, *Psilotum nudum*, *Eco*RV; 17, *Selaginella* sp., *Hin*dIII; 18, *Coleochaete orbicularis*, *Pst* I. Nylon filter replicas of the gels shown were hybridized sequentially with a 940-base-pair (bp) *Nco* I–*Stu* I fragment internal to the nuclear *tufA* gene of *Arabidopsis* (*Middle*) and a 1167-bp *Pst* I–*Hin*dIII fragment internal to the pea chloroplast *rbcL* gene (*Bottom*). Approximate sizes in kb are indicated on the right.

The *Coleochaete* EF-Tu is also unusual in having a net charge of +21, whereas all other EF-Tus are close to neutrality, ranging from +4 (*Chlamydomonas*, *Codium*) to −7 (*Micrococcus luteus*, ref. 32). The predicted extensions of the *Coleochaete* sequence at the amino and carboxyl termini are also unique among EF-Tus. Thus, in a number of respects, the *Coleochaete* cpDNA sequence encodes by far the most divergent EF-Tu known.

Table 1. Percent amino acid identity* among EF-Tu sequences

		1	2	3	4	5	6	7	8	9	10	11
Thermotoga	1	—										
Saccharomyces	2	63.3	—									
Escherichia	3	70.8	64.9	—								
Anacystis	4	64.8	61.7	75.3	—							
Spirulina	5	64.3	61.4	73.2	81.8	—						
Cyanophora	6	64.1	61.4	72.5	82.3	81.1	—					
Cryptomonas	7	64.2	60.8	70.0	79.3	77.1	80.3	—				
Euglena	8	64.3	62.4	71.4	78.8	76.7	75.5	76.2	—			
Chlamydomonas	9	64.6	61.7	68.4	75.2	74.5	75.5	75.4	78.4	—		
Codium	10	62.2	58.3	67.0	72.5	72.3	73.2	71.2	77.1	72.8	—	
Arabidopsis	11	63.6	61.4	68.9	76.5	74.8	75.5	73.2	77.4	76.7	73.0	—
Coleochaete	12	48.2	48.4	51.3	51.2	51.9	53.2	50.4	54.4	53.4	52.1	51.0

Sources of additional *tufA* sequences: *Anacystis nidulans* (23), *Arabidopsis thaliana* and *Chlamydomonas reinhardtii* (5), *Codium fragile* (M. Kuhsel and J.D.P., unpublished data), *Cyanophora paradoxa* (M. Kraus and W. Loeffelhardt, personal communication), *Cryptomonas* ϕ (S. Douglas, personal communication), *Escherichia coli* (24), *Euglena gracilis* (7), *Saccharomyces cerevisiae* (mitochondrion, ref. 25), *Spirulina platensis* (26), and *Thermotoga maritima* (27).
*Adjusted for missing data.

The *Coleochaete tufA* Is Derived from a Chloroplast *tufA*. The exceptional divergence of the *Coleochaete tufA* sequence raises the possibility that the gene was not originally present in the chloroplast but was acquired by it from a foreign source via lateral gene transfer. Although the lack of internal insertions/deletions characteristic of other EF-Tus suggests that the *Coleochaete* sequence is of chloroplast/cyanobacterial origin, cladistic analysis was used to further investigate the evolutionary origin of the gene.

A single shortest tree constructed from parsimony analysis of 15 EF-Tu sequences places the *Coleochaete* EF-Tu well within a clade of chloroplast-encoded proteins (Fig. 3). The exceptionally long terminal branch leading to *Coleochaete* further emphasizes the extensive divergence unique to this lineage. Thus, it seems that the *Coleochaete* sequence was derived from a green algal chloroplast *tufA* gene, which has evolved at an accelerated rate in the lineage leading to *Coleochaete*.

Distribution of *tufA* in Green Algal ncDNA. The sporadic distribution of *tufA* in the cpDNA of charophycean algae suggests that the gene may have been established in the nucleus early in the evolution of the lineage. Examination of ncDNAs from three of the six recognized orders of Charophyceae (9) shows strong *tufA* hybridization in all ncDNAs tested (Fig. 4). The weaker signals seen in the Zygnematales are roughly in proportion to the smaller amounts of DNA loaded in these lanes (Fig. 4). In contrast, no signal was found in the ncDNA of the chlorophycean alga *Chlamydomonas* (ref. 35; S.L.B., unpublished data).

The strength of the signals and their similarity in intensity among all of the charophycean ncDNAs (when normalized relative to the amount of DNA loaded in each lane) suggest that these signals are not due to bacterial DNA contamination, although this possibility cannot be ruled out. However, since the algae used were obtained from separate sources, and grown at different times in various media and under different conditions, the chances of their being contaminated to similar extents seems unlikely.

The presence of multiple *tufA*-hybridizing bands in the Charophyceae ncDNAs is in contrast to the single *tufA* found

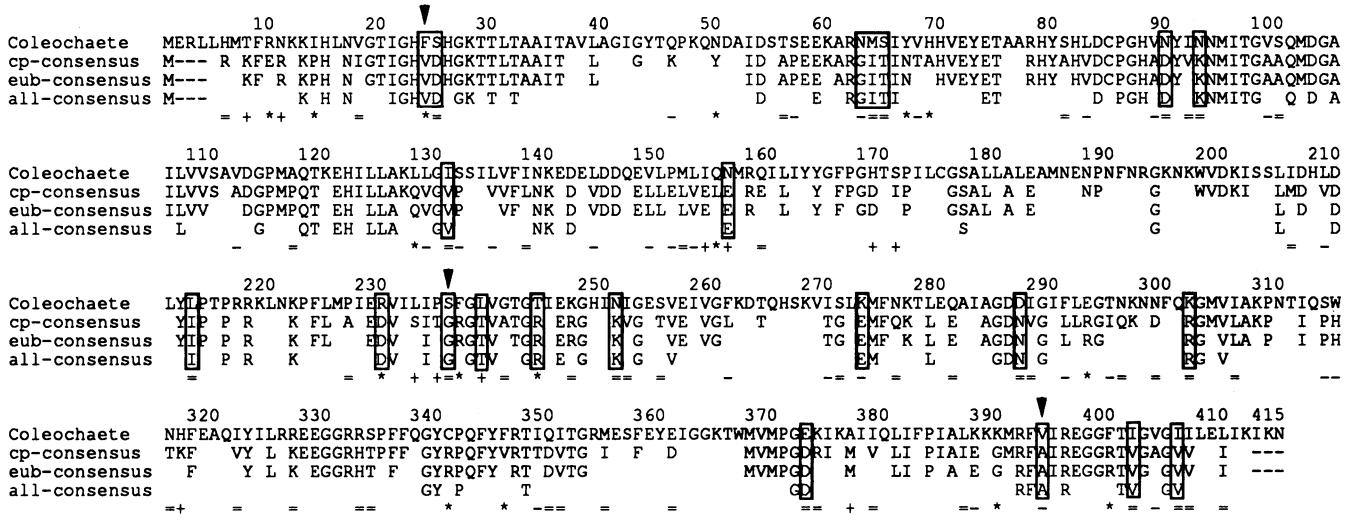

FIG. 2. Comparison of *Coleochaete* EF-Tu with EF-Tu consensus sequences. The *Coleochaete* deduced amino acid sequence is presented in its entirety together with consensus sites for cyanobacteria and chloroplasts (cp-consensus); eubacteria, cyanobacteria, and chloroplasts (eub-consensus); and eubacteria, eukaryotes, and archaebacteria (all-consensus). The cp-consensus includes sites identical for 6 or more of 8 chloroplast and cyanobacterial sequences, the eub-consensus includes sites identical for 11 or more of 13 eubacterial and organellar sequences, and the all-consensus includes sites identical for 23 or more of 27 total sequences. Changes in the *Coleochaete* sequence relative to the consensus sequences are indicated below as very conservative (=), conservative (−), nonconservative (+), and very nonconservative (*) changes, as defined by Dayhoff *et al.* (21). Amino acid positions at which *Coleochaete* differs from all consensus sequences are enclosed in boxes. Sites specifically discussed in the text are denoted with arrowheads.

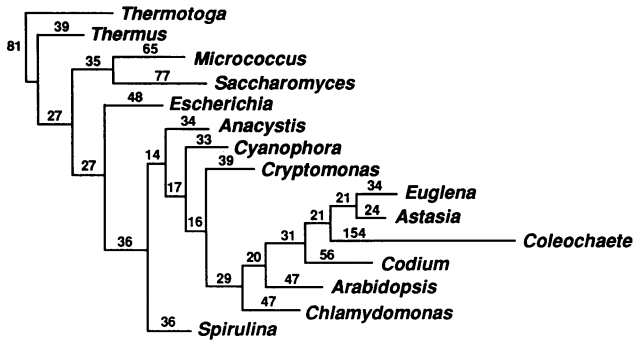

FIG. 3. The *Coleochaete tufA* is derived from a green algal chloroplast gene. The tree shown is the single shortest tree found by cladistic analysis of 15 eubacterial EF-Tu amino acid sequences. The tree has a total length of 1108 steps and a consistency index of 0.66, excluding autapomorphies. *Thermotoga maritima*, representing the earliest known branch of eubacteria (27), is used to root the tree. Horizontal branches are drawn to scale with lengths indicated numerically above the branches. *tufA* sequences in addition to those reported in Table 1 are *Astasia longa* (33), *Micrococcus luteus* (32), and *Thermus thermophilus* (34).

in *Arabidopsis thaliana* (5). However, multiple *tufA* signals are also found in the ncDNAs of all crucifers examined other than *Arabidopsis* (Fig. 5), suggesting that *tufA* exists as a multigene family in these nuclear genomes. Reduction in copy number of multigene families in *Arabidopsis*, whose nuclear genome is unusually small in size, has been noted previously (36).

DISCUSSION

The *tufA* gene appears to be present in the cpDNA of all ulvophycean and chlorophycean green algae based on filter hybridization (Fig. 1) and the sequencing of an apparently "normal" chloroplast *tufA* from a member of each class (Fig. 3: *Codium*, Ulvophyceae; *Chlamydomonas*, Chlorophyceae). However, *tufA* is missing from the cpDNAs of all land plants based on its absence from a bryophyte (8), probably the earliest diverging group of land plants (11), and from all other lineages of vascular plants examined (Fig. 1; refs. 5 and 22). This suggests that *tufA* was probably transferred to the nucleus after separation of the major green algal classes, since all appear to have at least a vestige of the gene in their

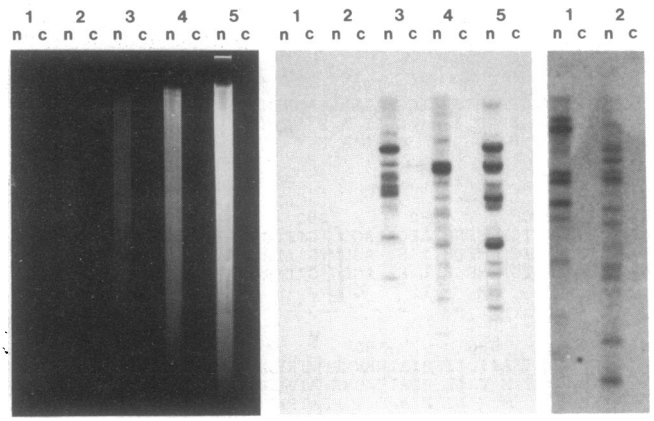

FIG. 4. *tufA* hybridizes to ncDNA of charophycean green algae. (*Left*) A 0.9% agarose gel containing restriction digests of ncDNA (n) and cpDNA (c). Lanes: 1, *Spirogyra maxima*, *Nco* I; 2, *Sirogonium melanosporum*, *Nco* I; 3, *Chara connivens*, *Pst* I; 4, *Nitella translucens*; 5, *Coleochaete orbicularis*, *Eco*RI. A nylon filter replica of the gel shown was hybridized with a 940-bp *Nco* I–*Stu* I fragment internal to the nuclear *tufA* gene of *Arabidopsis* (*Center* and *Right*). (*Right*) A 5-fold longer exposure of the first four lanes shown in *Middle*.

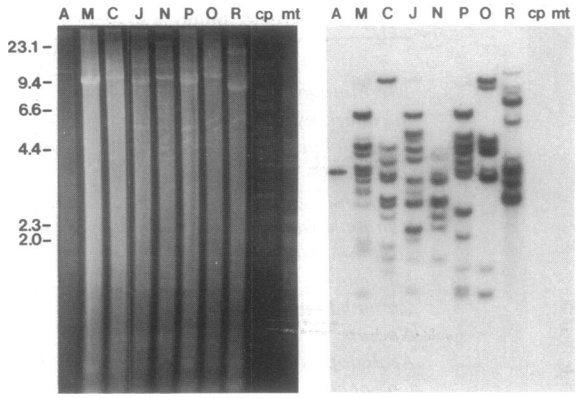

FIG. 5. *tufA* is a multicopy gene in all Brassicaceae except *Arabidopsis*. One microgram of ncDNA from *Arabidopsis thaliana* (A) and 10 μg of ncDNAs from *Brassica campestris* (M), *B. carinata* (C), *B. juncea* (J), *B. nigra* (N), *B. napus* (P), *B. oleracea* (O), and *Raphanus sativus* (R) was digested with *Hin*dIII, whereas 100 ng each of *B. nigra* cpDNA (cp) and mitochondrial DNA (mt) were digested with *Eco*RI. DNAs were fractionated on a 0.9% agarose gel (*Left*), blotted onto a nylon filter, and hybridized (*Right*) with a gel-isolated gene-internal fragment of the *Arabidopsis tufA* (940-bp *Nco* I–*Stu* I). Sizes in kb of *Hin*dIII fragments are indicated on the left.

cpDNA, but before the emergence of land plants. If this is the case, then the transfer must have occurred in the lineage leading to land plants, the Charophyceae.

The actual distribution of *tufA* within cpDNA of the Charophyceae—i.e., its apparent presence in Charales but absence from Zygnematales and land plants—is not congruent with the phylogeny of green plants as proposed by Bremer *et al.* (11). Their analysis, based on a broad range of phenotypic characters, places the Coleochaetales and Charales as closer to land plants than the Zygnematales and the Coleochaetales as the sister group to land plants. This branching pattern is also supported by the apparent shared gain of a cpDNA tRNA intron by *Coleochaete*, the Charales, and land plants (37) and by the results of 5S ribosomal RNA sequence analysis (38). Possible explanations of the distribution of *tufA* are (*i*) transfer of the gene to the nucleus early in the charophycean lineage, with subsequent loss of the cpDNA copy in at least two descendent lineages (land plants and Zygnematales); (*ii*) two independent transfers of the gene, one in the Zygnematales and one in the Coleochaetales/land plant lineage; or (*iii*) a single transfer in the common ancestor of Zygnematales and land plants, with these two being sister groups. We favor the first of these explanations for the following reasons.

The hybridization of *tufA* to ncDNA of Zygnematales, Charales, and Coleochaetales (Fig. 5) supports transfer of *tufA* early within the Charophyceae, in a common ancestor to all three orders. This common ancestor would then have retained copies of the gene in the chloroplast and the nucleus through the divergence of each lineage, after which the cpDNA copy followed distinct evolutionary paths in each. In the Zygnematales, the cpDNA copy would have been lost entirely—in parallel with its loss from the common ancestor of land plants—whereas in the Charales it appears to have been retained. In *Coleochaete*, the cpDNA copy also seems to have been retained but may no longer encode a functional elongation factor (see below).

One possibility that cannot be ruled out is that the *tufA* gene transfer occurred even earlier in a common ancestor of the Charophyceae and other classes of green algae. This explanation is most consistent with the phylogenetic analysis (Fig. 3), in which *Coleochaete* and *Arabidopsis* do not come out as sister taxa. However, the extreme divergence of the

Coleochaete EF-Tu suggests that it may not be an accurate representative of the charophycean lineage.

Parallel, independent transfers of the gene in the Zygnematales and Coleochaetales would seem to be a striking coincidence considering that the gene has been stably maintained in the chloroplast genome of all other algal groups examined (Figs. 1 and 3; M. Kuhsel and J.D.P., unpublished data). One possible explanation for such an apparent coincidence would be the existence of a predisposition within the charophycean lineage favoring transfer of this specific gene.

From the standpoint of *tufA* alone, the most parsimonious explanation for its distribution among charophycean and land plant cpDNA is that the Zygnematales, instead of the Charales and Coleochaetales, are the closest sister group to land plants. However, as stated above, this explanation is in conflict with phenotypic (11) and molecular (37, 38) data.

Although the *Coleochaete* sequence was apparently derived from a long lineage of normal EF-Tus, its level of divergence seems incompatible with known EF-Tu function. EF-Tu is central to protein synthesis catalytically and in controlling the rate and fidelity of protein synthesis (39). This probably accounts, at least in part, for the slow rate of EF-Tu sequence evolution (Table 1, Fig. 3). Thus, it seems unlikely that the vast changes unique to the *Coleochaete* EF-Tu could be compatible with the retention of normal function. For instance, changes predicted to disrupt nucleotide binding, such as the phenylalanine substitution at position 24 in *Coleochaete* EF-Tu, may result in a reduced level of protein since EF-Tu is unstable in the absence of bound nucleotide (40). Such changes could affect the rate and accuracy of protein synthesis as both require high levels of EF-Tu.

The long-term maintenance of the *Coleochaete tufA* as an open reading frame despite the accumulation of numerous mutations (Table 1, Fig. 4) suggests that selection is acting to maintain the gene. Because EF-Tu is a multifunctional protein, one possibility is that *Coleochaete* chloroplast *tufA* encodes a protein retaining some less constrained subset of its original functions. For example, in *E. coli* EF-Tu is known to regulate gene expression (41, 42) and adenylate cyclase activity (43) and to prime RNA-dependent RNA synthesis (44). Alternatively, the *Coleochaete* chloroplast-encoded EF-Tu may have evolved an entirely new function, as apparently has occurred in other cases of gene duplication (45).

Coleochaete tufA is somewhat reminiscent of the *atp9* gene of *Neurospora* and *Aspergillus*, which occurs in the mitochondrion and nucleus (46, 47). Only the nuclear copies of *atp9* appear to be expressed in these fungi, although developmentally restricted expression of the mitochondrial gene has been suggested. However, the sequences of the mitochondrial and nuclear genes are very different and their evolutionary relationship is unclear (46).

Acquisition of function by a relocated gene requires the gain of compartment-specific regulatory sequences, upstream and downstream, and an in-frame, amino-terminal transit peptide sequence. Since this process probably requires several independent events, a successfully transferred gene probably exists for a period of time in the nucleus before becoming functional. This coexistence of two subcellular copies of a gene could then be followed either by loss of one subcellular copy or the evolution by one copy of a new or altered function. In the Zygnematales and in land plants, as in the case of most transferred organelle genes, the cpDNA copy appears to have been lost. However, this is apparently not the case in *Coleochaete* nor probably in the Charales. The Charales are particularly intriguing since their chloroplast-encoded *tufA* appears to be well conserved and it is possible that, in this lineage, the cpDNA copy may retain its central function as an elongation factor in protein synthesis.

We thank J. E. Boynton, A. W. Coleman, B. Dudock, N. W. Gillham, M. Li-Weber, L. J. Moore, R. Meintz, and C. Lemieux for providing various algal chloroplast and nuclear DNAs and S. Douglas, M. Kraus, M. Kuhsel, and W. Loeffelhardt for providing unpublished sequences. We also thank M. Kuhsel, C. Morden, J. Nugent, and K. Wolfe for critical reading of the manuscript. This work was supported by National Institutes of Health Grant GM35087 to J.D.P. and National Science Foundation Grants BSR8600167 and BSR8906126 to J.R.M.

1. Martin, W. & Cerff, R. (1986) *Eur. J. Biochem.* **159,** 323–331.
2. Froman, B. E., Tait, R. C. & Gottlieb, L. D. (1989) *Mol. Gen. Genet.* **217,** 126–131.
3. Attardi, G. & Schatz, G. (1988) *Annu. Rev. Cell Biol.* **4,** 289–333.
4. Palmer, J. D. (1985) *Annu. Rev. Genet.* **19,** 325–354.
5. Baldauf, S. L. & Palmer, J. D. (1990) *Nature (London)* **344,** 262–265.
6. Ohyama, K., Kohchi, T., Sano, T. & Yamada, Y. (1988) *Trends Biochem. Sci.* **13,** 19–22.
7. Montandon, P.-E. & Stutz, E. (1983) *Nucleic Acids Res.* **11,** 5877–5891.
8. Ohyama, K., Fukuzawa, H., Kohchi, T., Shirai, H., Sano, T., Sano, S., Umesono, K., Shiki, Y., Takeuchi, M., Chang, Z., Aota, S.-i., Inokuchi, H. & Ozeki, H. (1986) *Nature (London)* **322,** 572–574.
9. Mattox, K. R. & Stewart, K. D. (1984) in *Systematics of the Green Algae*, eds. Irvine, D. E. G. & John, D. M. (Academic, London), pp. 29–72.
10. Graham, L. E. (1984) *Am. J. Bot.* **71,** 604–608.
11. Bremer, K., Humphries, C. J., Mishler, B. D. & Churchill, S. P. (1987) *Taxon* **36,** 339–349.
12. Hoshaw, R. W., Wells, C. V. & McCourt, R. M. (1987) *J. Phycol.* **23,** 267–273.
13. Graham, J. M., Auer, M. T., Canale, R. P. & Hoffman, J. P. (1982) *J. Great Lakes Res.* **8,** 100–111.
14. Doyle, J. J. & Doyle, J. L. (1987) *Phytochem. Bull.* **19,** 11–15.
15. Goff, L. J. & Coleman, A. W. (1988) *J. Phycol.* **24,** 357–368.
16. Milligan, B. G. (1989) *Plant Mol. Biol. Rep.* **7,** 144–149.
17. Watson, J. C. & Thompson, W. F. (1986) *Methods Enzymol.* **118,** 57–74.
18. Kolodner, R. & Tewari, K. K. (1972) *Proc. Natl. Acad. Sci. USA* **69,** 1830–1834.
19. Palmer, J. D. (1986) *Methods Enzymol.* **118,** 167–185.
20. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab., Cold Spring Harbor, NY).
21. Dayhoff, M. O., Eck, R. V. & Park, C. M. (1972) in *Atlas of Protein Sequence and Structure*, ed. Dayhoff, M. O. (Natl. Biochem. Res. Found., Silver Spring, MD), pp. 89–99.
22. Sugiura, M. (1989) *Annu. Rev. Cell Biol.* **5,** 51–70.
23. Meng, B. Y., Shinozaki, K. & Sugiura, M. (1989) *Mol. Gen. Genet.* **216,** 25–30.
24. Yokota, T., Sugisaki, H., Takanami, M. & Kaziro, Y. (1980) *Gene* **12,** 25–31.
25. Nagata, S., Tsunetsugu-Yokota, Y., Naito, A. & Kaziro, Y. (1983) *Proc. Natl. Acad. Sci. USA* **80,** 6192–6196.
26. Buttarelli, F. R., Calogero, R. A., Tiboni, O., Gualerzi, C. O. & Pon, C. L. (1989) *Mol. Gen. Genet.* **217,** 97–104.
27. Bachleitner, M., Ludwig, W., Stetter, K. O. & Schleifer, K. H. (1989) *FEMS Microbiol. Lett.* **57,** 115–120.
28. Jurnak, F. (1985) *Science* **230,** 32–36.
29. Jacquet, E. & Parmeggiani, A. (1988) *EMBO J.* **7,** 2861–2867.
30. Grantham, R. (1974) *Science* **185,** 862–864.
31. Vijgenboom, E. & Bosch, L. (1989) *J. Biol. Chem.* **264,** 13012–13017.
32. Ohama, T., Yamao, F., Muto, A. & Osawa, S. (1987) *J. Bacteriol.* **169,** 4770–4777.
33. Siemeister, G., Buchholz, C. & Hachtel, W. (1990) *Mol. Gen. Genet.* **220,** 425–432.
34. Seidler, L., Peter, M., Meissner, F. & Sprinzl, M. (1987) *Nucleic Acids Res.* **15,** 9263–9277.
35. Watson, J. C. & Surzycki, S. J. (1983) *Curr. Genet.* **7,** 201–210.
36. Meyerowitz, E. M. (1987) *Annu. Rev. Genet.* **21,** 93–111.
37. Manhart, J. R. & Palmer, J. D. (1990) *Nature (London)* **345,** 268–270.
38. Hori, H., Lim, B.-L. & Osawa, S. (1985) *Proc. Natl. Acad. Sci. USA* **82,** 820–823.
39. Thompson, R. C. (1988) *Trends Biochem. Sci.* **13,** 91–93.
40. Lapadat, M. A. & Spremulli, L. L. (1989) *J. Biol. Chem.* **264,** 5510–5514.
41. Vijgenboom, E., Nilsson, L. & Bosch, L. (1988) *Nucleic Acids Res.* **16,** 10183–10212.
42. Travers, A. (1973) *Nature (London)* **244,** 15–17.
43. Reddy, P., Miller, D. & Peterkofsky, A. (1986) *J. Biol. Chem.* **261,** 11448–11451.
44. Brown, S. & Blumenthal, T. (1976) *J. Biol. Chem.* **251,** 2749–2753.
45. Prager, E. M. & Wilson, A. C. (1988) *J. Mol. Evol.* **27,** 326–335.
46. van den Boogaart, P., Samallo, J. & Agsteribbe, E. (1982) *Nature (London)* **298,** 187–189.
47. Brown, T. A., Ray, J. A., Waring, R. B., Scazzocchio, C. & Davies, R. W. (1984) *Curr. Genet.* **8,** 489–492.