

# Molecular phylogeny of the *atpB* and *atpE* genes of the brown alga *Pylaiella littoralis*

WILLIAM F. MARTIN<sup>1</sup>, STEFAN JOUANNIC<sup>2</sup> AND SUSAN LOISEAUX-DE GOËR<sup>2\*</sup>

<sup>1</sup>Institut für Genetik, Biozentrum, Technische Universität Braunschweig, Postfach 3329, W-3300 Braunschweig, Germany

<sup>2</sup>Centre d'Etudes Océanologiques et de Biologie Marine, CNRS-UPR 4601, BP 74, F 29682 Roscoff Cedex, France

Recent phylogenetic studies suggest that plastid ribosomal RNA genes from *Pylaiella littoralis* have a cyanobacterial origin, whereas their Rubisco genes are related to the homologous  $\alpha$ - and  $\beta$ -purple eubacterial genes. We have constructed a phylogenetic tree based upon the *atpB* and *atpE* sequences, including the same range of taxa (chlorophytes, chromophytes, cyanobacteria,  $\alpha$ - and  $\gamma$ -purple eubacteria) and using the same methods as previously described for *rbcL* genes. This phylogenetic tree clearly shows that the *atpB* and *atpE* genes of this brown alga are more closely related to their cyanobacterial homologues than to those of  $\alpha$ - or  $\gamma$ -purple eubacteria. Different hypotheses that could explain the apparently composite origin of red-chromophyte plastomes are discussed.

**Key words:** Chromophyte algae, plastidial genome, phylogeny.

## Introduction

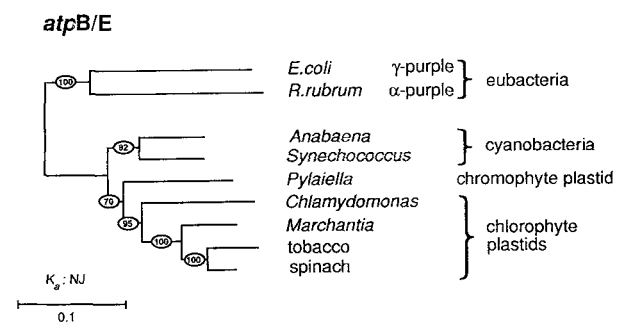
The *atpB* and the *atpE* genes encode, respectively, for the  $\beta$  and  $\epsilon$  subunits of the CF<sub>1</sub>/ATPase complex. In the plastid genome of the brown alga *Pylaiella littoralis* (L.) Kjellm., these two genes are organised in a cluster and separated by a 14-nucleotide spacer region (Jouannic *et al.*, 1992).

Comparison of the *atpB* gene sequences in organisms from eubacteria to land plants shows that this gene is highly conserved. Given its presence in all organisms and its large size, we predicted that it could be a good model for phylogenetic studies. In trees inferred from ribosomal RNA gene sequences all plastid lineages appear as descendants of the cyanobacterial lineage (Markowicz & Loiseaux-de Goër, 1991; Douglas & Turner, 1991; Somerville *et al.*, 1993). However, in trees inferred from Rubisco gene sequences the red-chromophyte plastid lineage emerges from the  $\alpha$ - (type 1) and  $\beta$ -purple eubacterial lineage while, in contrast, the green plastid lineage emerges from the cyanobacterial lineage (Assali *et al.*, 1990, 1991; Martin *et al.*, 1992). Thus, we constructed a phylogenetic tree based upon the *atpB/E* sequences to determine whether these protein-coding genes in red-chromophyte plastids have a different origin from that of ribosomal RNA genes.

## Materials and methods

Aligned amino acid sequences for the *atpB* and *atpE* genes were taken from Jouannic *et al.* (1992). Corresponding codon insertions were introduced into the nucleotide sequences. Amino- and carboxyterminal regions of length heterogeneity were deleted from the alignment, yielding

for comparison 1914 nucleotide positions in each set of sequences (*atpB* + *atpE*). Divergence at non-synonymous sites between aligned nucleotide sequences (alignment available upon request from the authors) was measured as  $K_a$  with the weighted pathway method (Li *et al.*, 1985). The matrix of distance values was analysed with the neighbour-joining method (Saitou & Nei, 1987). Parsimony bootstrap analysis (Felsenstein, 1985) was performed for first and second codon positions (100 replicates); this



**Fig. 1.** Neighbour-joining tree (Saitou & Nei, 1987) for values of divergence at non-synonymous sites between concatenated *atpB* and *atpE* coding regions as measured with the weighted pathway method (Li *et al.*, 1985), showing bootstrap confidence intervals at nodes. The scale bar indicates a length of  $K_a = 0.1$ ; branch lengths (horizontal) are drawn to scale. The leftmost branch of the tree was arbitrarily 'bent' to permit display of species names in a convenient manner. Species are: *Escherichia coli*, a  $\gamma$ -purple eubacterium (Kanazawa *et al.*, 1985); *Rhodospirillum rubrum*, an  $\alpha$ -purple eubacterium (Falk *et al.*, 1985); *Synechococcus* 6301 and *Anabaena* PCC 7120, cyanobacteria (Curtis, 1988); *Pylaiella littoralis*, a chromophyte (Jouannic *et al.*, 1992); and the chlorophytes *Chlamydomonas reinhardtii* (Woessner *et al.*, 1986, 1987), *Marchantia polymorpha* (Umesono *et al.*, 1988), *Nicotiana tabacum* (Kazuo *et al.*, 1983) and *Spinacia oleracea* (Zurawski *et al.*, 1982).

\* To whom correspondence should be addressed.

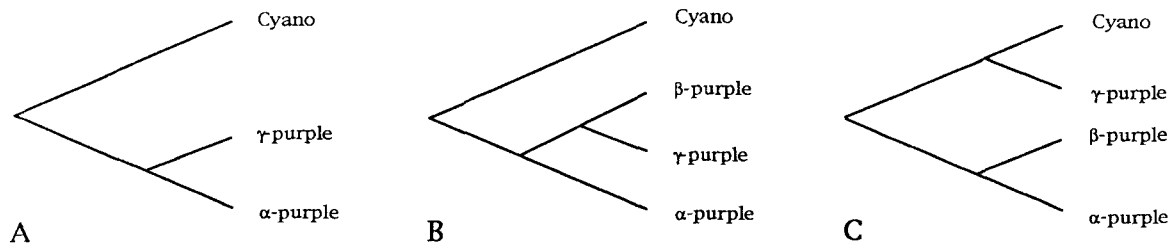


Fig. 2. Schematic representation of topologies obtained for phylogenetic trees based upon eubacterial sequences of the *atpB/E* genes (A), the ribosomal RNA genes (B) and the Rubisco genes (C) from cyanobacteria (cyano),  $\alpha$ -purple eubacteria ( $\alpha$ -purple),  $\beta$ -purple eubacteria ( $\beta$ -purple) and  $\gamma$ -purple eubacteria ( $\gamma$ -purple).

procedure approximates the removal of synonymous sites from the nucleotide data set (Martin *et al.*, 1992).

## Results and Discussion

The phylogenetic tree based on the *atpB* and *atpE* genes is shown in Fig. 1. The topologies of trees based upon either gene alone are similar to that of *atpB* and *atpE* taken together (not shown). As in trees based upon ribosomal RNA sequences, and in contrast to those based upon Rubisco genes, the  $\alpha$ - and  $\gamma$ -purple eubacteria group together on a different lineage from that of cyanobacteria and all plastids. In the same way, two plastid lines emerge from the cyanobacterial lineage: the green plastid lineage comprising plastids of green plants and a green alga, and the chromophyte lineage. Unfortunately, no *atpB/E* sequences from rhodophyte plastids are available. Consequently, the *atpB* and *atpE* genes from *P. littoralis* appear to have a cyanobacterial origin, as do the 16S and 23S ribosomal RNA genes. However, the topology of the tree, as of those inferred from ribosomal RNA sequences, does not allow us to argue in favour of monophyly or polyphyly of plastomes in the absence of any information on the timing of endosymbiosis.

A similar result has recently been obtained by Leitsch & Kowallik (1992) for a phylogenetic tree based upon amino acid sequences of *atpB* from a different brown alga, *Dictyota dichotoma*. These authors obtained bootstrap confidence intervals with equivalent or slightly lower values than ours, perhaps because of the different methods used. The *D. dichotoma* sequences have not been added to our data, but would certainly group with those of *P. littoralis* given the high homology between amino acid sequences of the two *atpB* subunits (90.6%).

In contrast to the rRNA and *atpB/E* genes, the *rbcL* and *rbcS* genes of *P. littoralis*, other chromophytes and rhodophytes are clearly more closely related to  $\alpha$ - and  $\beta$ -purple eubacterial homologues than to those of cyanobacteria and chlorophytes. In order to reconcile these data, one could postulate a chimaeric nature for rhodophyte and chromophyte plastomes. This composite plastidial origin could be the result of a mixture of entire genomes or could come from lateral transfers of genes. Different possibilities can be considered. The chimaeric nature could have

existed before the endosymbiosis, resulting from exchanges between prokaryotes such as cyanobacteria and purple eubacteria. It could also have appeared in the host cell, after endosymbiosis, between different endosymbionts or between an endosymbiont and mitochondria which derived from  $\alpha$ -purple eubacteria (Yang *et al.*, 1985). In this latter case, however, one must explain the close relationship of the Rubisco genes of cyanobacteria and  $\gamma$ -purple eubacteria and of  $\alpha$ - and  $\beta$ -purple eubacteria (Fig. 2C) in contrast to the close relationship between  $\gamma$  and  $\beta$ -purple eubacterial rRNA genes well separated from those of cyanobacteria (Fig. 2B). Unfortunately, no *atpB/E* sequences were available from  $\beta$ -purple eubacteria for inclusion in the present study, but the topology of the tree (Fig. 2A) clearly shows that  $\gamma$ -purple *atpB* genes are more closely related to the  $\alpha$ -purple eubacterial than to the cyanobacterial homologues.

Alternatively, the progenitors of purple eubacteria and cyanobacteria may have possessed two or more *rbcL* genes, one of which was differentially lost several times along the lineages giving rise to rhodophyte plastids, chromophyte plastids and  $\alpha$ - and  $\beta$ -purple eubacteria on one hand and to cyanobacteria, green plastids and  $\gamma$ -purple eubacteria on the other. The latter scenario cannot be excluded because extant purple and cyanobacteria often possess multiple copies of photosynthesis genes (Vrba & Curtiss, 1989; Gibson *et al.*, 1991; Lill & Nelson, 1991; W. Martin unpublished data), and their antecedents might also have possessed these multiple genes.

## Acknowledgements

We wish to thank the Gesellschaft für Biotechnologische Forschung, Braunschweig, Germany, for the use of their computer facilities.

## References

- ASSALI, N.E., MACHE, R. & LOISEAUX-DE GOËR, S. (1990). Evidence for a composite phylogenetic origin of the plastid genome of the brown alga *Pylaiella littoralis* (L) Kjellm. *Plant Mol. Biol.*, **15**: 307–315.
- ASSALI, N.E., MARTIN, W.F., SOMERVILLE, C.C. & LOISEAUX-DE GOËR, S. (1991). Evolution of the Rubisco operon from prokaryotes to algae: structure and analysis of the *rbcS* gene of the brown alga *Pylaiella littoralis*. *Plant Mol. Biol.*, **17**: 853–863.

- CURTIS, E.C. (1988). Structure, organization and expression of cyanobacterial ATP synthase genes. *Photosynthesis Res.*, **18**: 223–244.
- DOUGLAS, S.E. & TURNER, S. (1991). Molecular evidence for the origin of plastids from a cyanobacterium-like ancestor. *J. Mol. Evol.*, **33**: 267–273.
- FALK, G., HAMPE, A. & WALKER, J.E. (1985). Nucleotide sequence of the *Rhodospirillum rubrum* *atp* operon. *Biochem. J.*, **228**: 391–407.
- FELSENSTEIN, J. (1985). Confidence limits on phylogenies: an approach using the bootstrap. *Evolution*, **39**: 783–791.
- GIBSON, J.L., FALCONE, D.L. & TABITA, F.R. (1991). Nucleotide sequence, transcriptional analysis and expression of genes encoded within the form I CO<sub>2</sub> fixation operon of *Rhodobacter sphaeroides*. *J. Biol. Chem.*, **266**: 14646–14653.
- JOUANNIC, S., KERBOURCH, C., KLOAREG, B. & LOISEAUX-DE GOËR, S. (1992). Nucleotide sequences of the *atpB* and the *atpE* genes of the brown alga *Pylaiella littoralis* (L.) Kjellm. *Plant. Mol. Biol.*, **18**: 819–822.
- KANAZAWA, H., KAYANO, T., KIYASU, T. & FUTAI, M. (1982). Nucleotide sequence of the genes for beta and epsilon subunits of proton-translocating ATPase from *Escherichia coli*. *Biochem. Biophys. Res. Commun.* **105**: 1257–1264.
- KAZUO, S., HIROSHI, D., AKIRA, K. & MASAHIRO, S. (1983). Overlap and cotranscription of the genes for the beta and epsilon subunits of tobacco chloroplast ATPase. *Gene*, **24**: 147–155.
- LEITSCH, C. & KOWALLIK, K.V. (1992). Nucleotide sequence and phylogenetic implication of the ATPase subunits  $\beta$  and  $\epsilon$  encoded in the chloroplastic genome of the brown alga *Dictyota dichotoma*. *Plant Mol. Biol.*, **19**: 289–298.
- LI, W.H., WU, C.I. & LUO, C.C. (1985). A new method for estimating synonymous and non-synonymous rates of nucleotide substitution considering the relative likelihood of nucleotide and codon changes. *Plant Mol. Biol.*, **2**: 150–174.
- LILL, H. & NELSON, N. (1991). The *atp1* and *atp2* operons of the cyanobacterium *Synechocystis* sp PCC 6803. *Plant Mol. Biol.*, **17**: 641–652.
- MARKOWICZ, Y. & LOISEAUX-DE GOËR, S. (1991). Plastid genomes of Rhodophyta and Chromophyta constitute a distinct lineage which differs from that of Chlorophyta and have a composite phylogenetic origin, perhaps like that of Euglenophyta. *Curr. Genet.*, **20**: 427–430.
- MARTIN, W.F., SOMERVILLE, C.C. & LOISEAUX-DE GOËR, S. (1992). Molecular phylogenies of plastid origins and algal evolution. *J. Mol. Evol.* **35**: 385–403.
- SAITOU, N. & NEI, M. (1987). The neighbor joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.*, **4**: 406–425.
- SOMERVILLE, C.C., JOUANNIC, S., MARTIN, W.F., KLOAREG, B. & LOISEAUX-DE GOËR, S. (1993). Sequence, secondary structure and phylogeny of the chloroplastic 23S rRNA gene from the brown alga *Pylaiella littoralis*. *Plant Mol. Biol.* (in press).
- UMESONO, K., INOKUCHI, H., SHIKI, Y., TAKEUCHI, M., CHANG, Z., FUKUZAWA, H., KOHCHI, T., SHIRAI, H., OHYAMA, K. & OZEKI, H. (1988). Structure and organization of *Marchantia polymorpha* chloroplast genome. II. Gene organization of the large single copy region from *rps12* to *atpB*. *J. Mol. Biol.*, **203**: 299–331.
- VRBA, J.M. & CURTIS, S.E. (1989). Characterization of a four-member *psaA* gene family from the cyanobacterium *Anabaena* PCC 7120. *Plant Mol. Biol.*, **14**: 81–92.
- WOESSNER, J.P., GILLHAM, N.W. & BOYNTON, J.E. (1986). The sequence of the chloroplast *atpB* gene and its flanking regions in *Chlamydomonas reinhardtii*. *Gene*, **44**: 17–28.
- WOESSNER, J.P., GILLHAM, N.W. & BOYNTON, J.E. (1987). Chloroplast genes encoding subunits of the H<sup>+</sup>-ATPase complex of *Chlamydomonas reinhardtii* are rearranged compared to higher plants: sequence of the *atpE* gene and location of the *atpF* and *atpI* genes. *Plant Mol. Biol.*, **8**: 151–158.
- YANG, D., OYAIZU, Y., OZAIZU, H., OLSON, G.J. & WOESE, C.R. (1985). Mitochondrial origins. *Proc. Natl. Acad. Sci. USA*, **82**: 4443–4447.
- ZURAWSKI, G., BOTTOMLEY, W. & WHITFIELD, P.R. (1982). Structures of the genes for the beta and epsilon subunits of spinach chloroplast ATPase indicate a dicistronic mRNA and an overlapping translation stop/start signal. *Proc. Natl. Acad. Sci. U.S.A.* **79**: 6260–6264.

(Received 13 November 1992; revised 27 January 1993; accepted 29 January 1993)