



HAL
open science

The arabidopsis chloroplast ribosomal protein L21 is encoded by a nuclear gene of mitochondrial origin

Jean-Luc Gallois, P. Achard, G. Green, R. Mache

► **To cite this version:**

Jean-Luc Gallois, P. Achard, G. Green, R. Mache. The arabidopsis chloroplast ribosomal protein L21 is encoded by a nuclear gene of mitochondrial origin. *Gene*, 2001, 274 (1-2), pp.179-185. 10.1016/S0378-1119(01)00613-8 . hal-01594066

HAL Id: hal-01594066

<https://hal.science/hal-01594066>

Submitted on 26 Sep 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

The *Arabidopsis* chloroplast ribosomal protein L21 is encoded by a nuclear gene of mitochondrial origin[☆]

Jean-Luc Gallois¹, Patrick Achard, Gaynor Green, Régis Mache^{*}

Laboratoire de Génétique Moléculaire des Plantes, Université J. Fourier and Centre National de la Recherche Scientifique,
BP 53, F-38041 Grenoble Cedex 09, France

Received 20 December 2000; received in revised form 6 June 2001; accepted 11 July 2001
Received by W. Martin

Abstract

Many chloroplast genes of cyanobacterial origin have been transferred to the nucleus during evolution and their products are re-addressed to chloroplasts. The *RPL21* gene encoding the plastid r-protein L21 has been lost in higher plant chloroplast genomes after the divergence from bryophytes. Based on phylogenetic analysis and intron conservation, we now provide evidence that in *Arabidopsis* a nuclear *RPL21c* gene of mitochondrial origin has replaced the chloroplast gene. The control of expression of this gene has been adapted to the needs of chloroplast development by the acquisition of plastid-specific regulatory promoter *cis*-elements. © 2001 Elsevier Science B.V. All rights reserved.

Keywords: *RPL21* genes; Higher plants; Phylogeny; Gene expression; Promoter *cis*-elements

1. Introduction

Molecular data and phylogenetic results concur to support the hypothesis of a cyanobacterial origin of chloroplasts. After the engulfment of a cyanobacterium in a primitive eukaryotic cell cyanobacterial genes were either lost or transferred into the host genome, resulting in the structure of the present chloroplast chromosome. The organization of this chromosome is remarkably conserved in all evolutionary distant lineages, from algae to angiosperms, giving evidence that the transfer of cyanobacterial genes to the nucleus occurred rapidly after endosymbiosis. However, a few genes were transferred much later, including components of the translational apparatus. For instance, the *tufA* gene, encoding plastid elongation factor Tu, is present in chloroplasts of algae and has been transferred to the nucleus in an algal lineage that gave rise to land plants (Baldauf and Palmer, 1990). The *rpl22* gene coding for the

ribosomal protein (r-protein) L22 represents another interesting example of late transfer. In legumes, the *rpl22* gene has been transferred into the nucleus, possibly 100 million years before the loss from the chloroplast genome (Gantt et al., 1991). During this time the nuclear *rpl22* gene has been preserved suggesting that it was active. In a recent study, Martin et al. (1998) identified 44 functional nuclear genes out of 210 genes examined that originated from a cyanobacterial ancestral genome and whose products are addressed to chloroplasts. This means that present-day chloroplast proteins are mainly of cyanobacterial origin, regardless of where the corresponding genes are localized, in plastids or in the nucleus. Because the mitochondrial genome originated from an endosymbiotic α -proteobacterium (Anderson et al., 1998; Gray et al., 1999), the same principle should be true for mitochondrial proteins. Taken together, the question arises whether present-day nuclear genes of cyanobacterial origin could also provide proteins that are relocated to mitochondria and/or whether nuclear genes of α -proteobacterial origin could provide proteins that are relocated to plastids. In the present paper we addressed the second part of this question. The *RPL21* gene is still present in the chloroplast chromosome of many algae and the lower land plant *Marchantia polymorpha*, a bryophyte, but is lost from the chloroplast genomes of higher plants. Based on the relative low conservation of chloroplast L21 protein sequences between *Marchantia* and spinach, it was proposed several

Abbreviations: r-protein, ribosomal protein; RT-PCR, reverse transcriptase-polymerase chain reaction

[☆] The nucleotide sequence data reported here appeared in the EMBL databank under the Accession numbers Y15964 and AJ278909.

^{*} Corresponding author. Tel.: +33-4-76-51-46-67; fax: +33-4-76-51-43-36.

E-mail address: regis.mache@ujf-grenoble.fr (R. Mache).

¹ Present address: Department of Cellular and Developmental Biology, John Innes Centre, Norwich, NR4 7UH, UK.

years ago that the spinach nuclear gene might be of mitochondrial origin (Martin et al., 1990). But phylogenetic analyses were lacking to support the hypothesis. Detailed analysis of prokaryotic-type *RPL21* genes in *Arabidopsis thaliana*, a higher plant, shows that the nuclear gene replacing the chloroplast gene is not related to cyanobacterial species and is grouped with the *Caenorhabditis elegans* mitochondrial L21-like gene. The expression of this gene has been adapted to chloroplast needs during evolution as demonstrated by the existence of chloroplast r-protein-specific regulatory elements in the promoter of this gene.

2. Materials and methods

2.1. Cloning and sequence analyses

An *Arabidopsis* genomic library (ecotype Columbia) has been screened with the α -³²P-labeled EST 146E8T7 (Accession number: T75800) that covers two-thirds of the *RPL21c* cDNA. Standard procedures (Sambrook et al., 1989) were used for genomic DNA isolation, subcloning into pBlue-script II (Stratagene), RNA isolation, hybridization and sequencing, using an Applied Biosystems 373 XL automated DNA sequencer.

2.2. RT-PCR

Arabidopsis RNA (1–3 μ g) was used for RT-PCR. Primers were as follows: adenosine phosphoribosyl transferase, APT, CCTTTCCTTAAGCTCTG, TCCCAGAATCGCTAAGATTGCC; *RPL21c*, AGAGGATCCCAAAGTCTCGAGCTTAGAG, GAATGGCAAAGGTACA; *RP-L21m*, TCGTGCTCTCCGATGGCG, AGGCCTTTTTCATCATTTC. The 1 kb ladder (Gibco-BRL) was used as a size marker.

2.3. Gel shift assays

Gel retardation analyses were performed as described by Lagrange et al. (1997) using 30–60 μ g *Arabidopsis* whole protein extracts and the following annealed oligonucleotides specific to elements of the *RPL21c* promoter. Monomers of the S2 site, dimers of GT1 and S1 sites or mutants (GT1M and S1M) were used: S2, 5'-TAAGTAATTTTGTATAA-TACAACTGTGA-3' and 3'-TTAAAACAATATTATG-TTTGACACTTTGACA-5'; GT1, 5'-AAACCAATTTGA-A-3' and 3'-TTTGGTTAAACTT-5'; GT1M, 5'-AAAG-GAATTTGAA-3' and 3'-TTTCCTTAAACTT-5'; S1, 5'-GGGGTACAATGGTAATTTT-3' and 3'-CCCCATGTT-ACCATTAAAA-5'; and S1M, 5'-GGGGTAGAGCTGT-AATTTT-3' and 3'-CCCCATCTCGACATTAATAA-5'. The fragments were [α -³²P]CTP-labeled and gel-purified. Electrophoresis was performed on 6% non-denaturing polyacrylamide gels.

2.4. Computer analyses

Sequences were analyzed using the GCG software (University of Wisconsin) and internet-available bioinformatic programs. Parsimony, neighbor-joining and Fitch–Margoliash tree building were performed through the Pasteur (<http://bioweb.pasteur.fr>) and Infobiogen (<http://www.infobiogen.fr>) servers. Trees were drawn with the Phylodendron tree printer program at <http://iubio.bio.indiana.edu/>. The L21 r-protein sequences from prokaryotes and the prokaryotic-like L21 r-protein sequences from plants were retrieved from available databanks. The different sequences were aligned using the GCG pileup program and the conserved core sequences were used for phylogenetic analysis. Because the L21 protein deduced from a *Physcomitrella patens* EST (Accession number: AW145090) is too short it was not included in the data used for phylogenetic tree building. The sequences corresponding to the cytosolic 60S r-protein L21 are not included because they present no similarities with L21 r-proteins from either chloroplast or from bacteria. Prediction of organelle localization was made using the following web sites: <http://www.cbs.dtu.dk/services/targetP/>; <http://www.cbs.dtu.dk/services/chloroP/>; <http://www.psорт.nibb.ac.jp:8800/>.

3. Results

3.1. The *Arabidopsis* genome contains two *RPL21* genes of prokaryotic origin

An *Arabidopsis* genomic clone was isolated using *Arabidopsis* EST encoding the 50S chloroplast r-protein L21 as a probe. After sequencing, the corresponding gene, termed *RPL21c*, and the adjacent genes were identified (Accession number: Y15964). The adjacent genes code for a Ca-dependent protein kinase (Accession number: D21806) and a phosphoenolpyruvate carboxykinase (Accession number: CAB38935). The *RPL21c* gene is located on chromosome 1 (Accession number: *gi6289106*). This gene is unique in *Arabidopsis* as verified by Southern blot hybridization (data not shown) and it is functional as shown by the isolation of the corresponding cDNA clone (Accession number: Z49787). It is orthologous to the spinach chloroplast *RPL21* gene (76% identity and 86% similarity). A second putative gene located on chromosome 4 was sequenced and annotated in the framework of the *Arabidopsis* Genome Initiative (Accession number: *gi7269995*). This gene encodes SP8-like protein that includes a L21 r-protein signature. We re-examined the sequence data and found that the putative gene can be separated into two different genes, one encoding a L21 r-protein, the other one encoding the SP8-like protein. We termed the L21-encoding gene *RPL21m* (Accession number: AJ278909) to distinguish it from *RPL21c*. One EST (Accession number: AI995545) corresponding exclusively to exons of the *RPL21m* gene

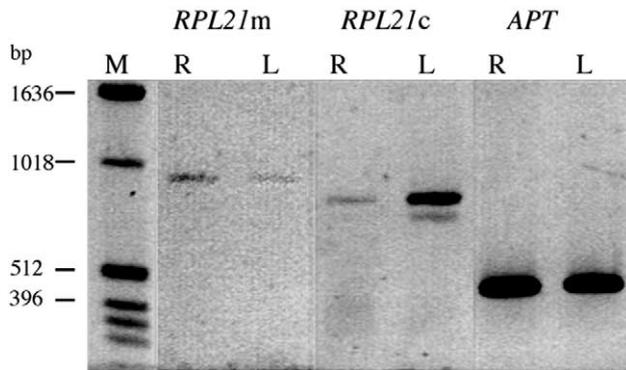


Fig. 1. Expression of *RPL21* genes in *Arabidopsis*. Semi-quantitative measurement of mRNAs is made by RT-PCR (27 cycles). Primers corresponding to *RPL21c*, *RPL21m* and constitutively expressed adenosine phosphoribosyl transferase gene, *APT*, taken as a control, were used. cDNAs were analyzed by electrophoresis using RNAs isolated from roots (R) or leaves (L). M, size markers.

is present in the dbEST databank, supporting our hypothesis that the *RPL21m* gene is expressed independently of the SP8-like gene. In order to confirm the existence of *RPL21m* mRNA, we performed Northern and RT-PCR analyses. Northern analysis revealed that *RPL21m* encodes a transcript of the expected size (0.9 kb; data not shown).

The RT-PCR analyses show that both the *RPL21m* and the *RPL21c* genes are constitutively expressed, but that only the *RPL21c* gene is highly expressed in leaves (Fig. 1). From these data, we conclude that two independent *RPL21* genes exist in *A. thaliana*.

Computer analysis of the *Arabidopsis RPL21* genes using different algorithms (see Section 2) clearly predicts a chloroplast-targeting N-terminal sequence for the *RPL21c* gene product and a mitochondrial-targeting N-terminal sequence for the *RPL21m* gene product. It was previously shown by Sanchez et al. (1996) that antibodies directed against the orthologous spinach chloroplast L21 protein recognized an *Arabidopsis* chloroplast protein of the same size. These antibodies do not cross-react with mitochondrial proteins, thus confirming the chloroplast localization of the *RPL21c* gene product. We assume that *RPL21m* encodes a mitochondrial protein because only one L21 protein exists per organelle and L21m protein is closer than L21c protein to the mitochondrial protein of *C. elegans* (32% identity versus 28%).

3.2. The chloroplast *RPL21c* gene is of mitochondrial origin

The comparison of the *RPL21c* and *RPL21m* proteins with the spinach plastid L21 r-protein and the L21 proteins

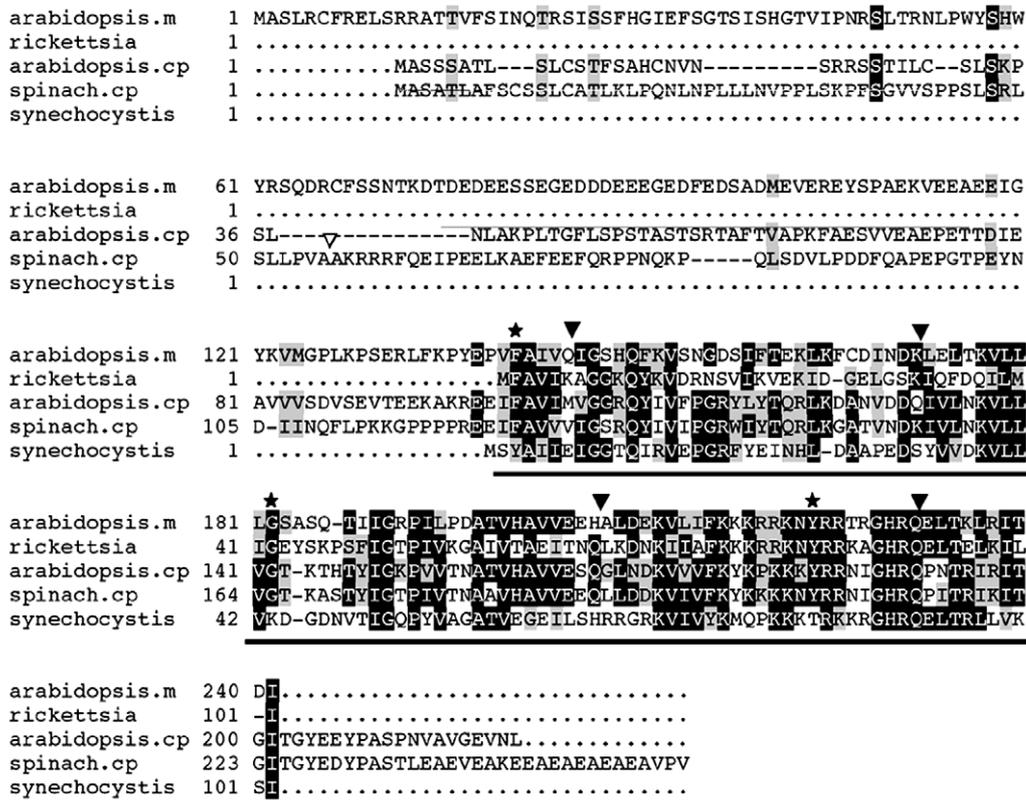


Fig. 2. Amino acid sequence alignment of chloroplast or mitochondrial L21 r-proteins from spinach and *Arabidopsis* with their respective eubacterial ancestors, *Synechocystis* sp. and *Rickettsia prowazekii*, respectively. Spinach-cp, *Spinacia oleracea* chloroplast; Arabidopsis-cp, *A. thaliana* chloroplast; Arabidopsis-m, *A. thaliana* mitochondria. Corresponding positions of introns in the nucleotide sequence of higher plant *RPL21* genes are indicated (black triangles). The cleavage site leaving the mature L21 in spinach chloroplast is reported (open triangle). The conserved prokaryotic core of L21 proteins is underlined. The two amino acids specific to the α -proteobacterial lineage are indicated by a star. Identical amino acids are boxed in black, and similar amino acids are boxed in gray.

from ancestors of chloroplast (a cyanobacterium, *Synechocystis* sp.) or of mitochondria (an α -proteobacterium, *Rickettsia prowazekii*) is shown in Fig. 2. The higher plant L21 r-proteins have conserved a central core similar to the bacterial L21 amino acid sequences. In the N-terminal region, the plant L21 proteins have acquired a targeting sequence followed by a long protein extension not significantly

related to each other, probably because the selection pressure on sequences encoding these elements was not very high. In contrast to higher plants, the L21 r-protein of the lower plant *M. polymorpha* is encoded in the chloroplast genome and is much more similar to the cyanobacterial L21 protein, lacking the N-terminal extension and resulting in a similar size (data not shown).

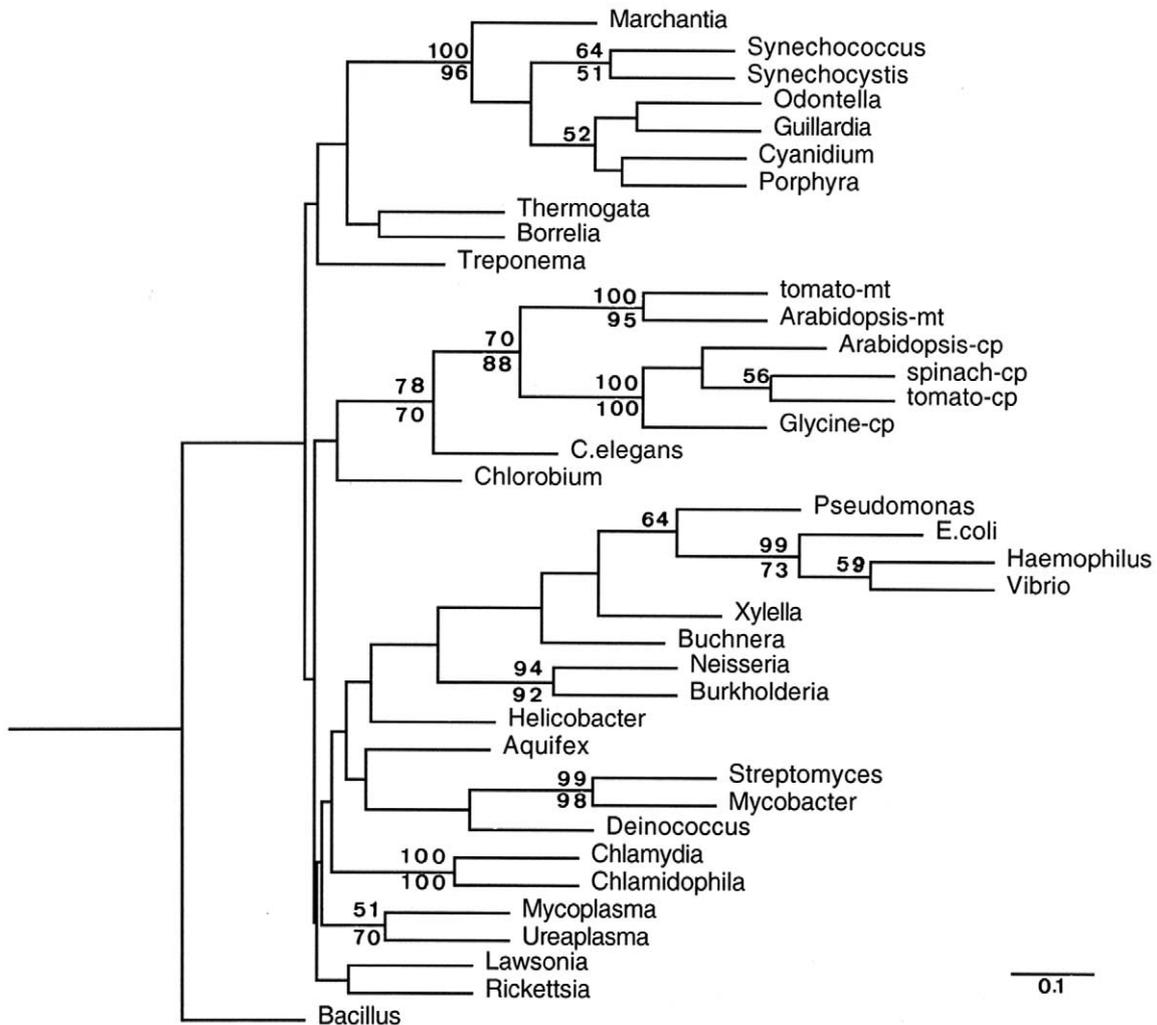


Fig. 3. Protein phylogeny of chloroplast and mitochondrial L21 r-proteins. The unrooted tree was constructed by the Fitch–Margoliash least-square distance method using 50S L21 r-proteins from plants, from bacteria and from *C. elegans*. The core regions of amino acid sequences (Fig. 2) have been used. The L21 protein sequences and their Accession numbers are as follows. (1) Lower plants, chloroplast-encoded: *Cyanidium caldarium*, Cyanidium, a rhodophyte, AF022186; *Porphyra purpurea*, Porphyra, a rhodophyte, P51209; *Guillardia theta*, Guillardia, a cryptophyte, O78429; *Marchantia polymorpha*, Marchantia, a bryophyte, P06387. (2) Higher plants, nuclear-encoded: *Spinacia oleracea*, Spinacia-cp, P24613; *Arabidopsis thaliana*, Arabidopsis-cp, P51412; Arabidopsis-mt, deduced from AL161578; *Lycopersicon esculentum*, Tomato-cp, EST AW031100, and Tomato-mt, EST AI485205; *Glycine maxima*, Glycine-cp, AW831287; Glycine-mt, EST AW234792. (3) Animals: *Caenorhabditis elegans*, *C. elegans*, Z46937. (4) Bacteria: *Aquifex aeolicus*, Aquifex, AAC07544; *Bacillus subtilis*, Bacillus, CAA42108; *Borrelia burgdorferi*, Borrelia, AAV67129; *Buchnera* sp., Buchnera, BAB13091; *Burkholderia pseudomallei*, Burkholderia, AAG01347; *Chlamydia trachomatis*, Chlamydia, AAC68017; *Chlamydomonas reinhardtii*, Chlamidophila, AAF38077; *Chlorobium vibrioforme*, Chlorobium, CAA55804; *Deinococcus radiodurans*, Deinococcus, AAF09679; *Escherichia coli*, *E. coli*, AE000399; *Haemophilus influenzae*, Haemophilus, AAC22536; *Helicobacter pylori*, Helicobacter, AAD07365; *Lawsonia intracellularis*, Lawsonia, AAC67129; *Mycobacterium tuberculosis*, Mycobacter, CAB03777; *Mycoplasma pneumoniae*, Mycoplasma, AAB961591; *Neisseria meningitidis*, Neisseria, AL162758; *Pseudomonas aeruginosa*, Pseudomonas, AAG01347; *Rickettsia prowazekii*, Rickettsia, CAA15179; *Streptomyces coelicolor*, Streptomyces, CAB75378; *Synechococcus* sp., *Synechococcus*, BAA3700; *Synechocystis* sp., *Synechocystis*, P74266; *Thermogata maritima*, Thermogata, AAD36526; *Treponema pallidum*, Treponema, AAC65713; *Ureaplasma urealyticum*, Ureaplasma, AAF306201; *Vibrio cholerae*, Vibrio, AAF93608; *Xylella fastidiosa*, Xylella, AAF85223. Bootstrap values above 50, using 100 replications, using Fitch–Margoliash (upper) or parsimony (lower) methods are indicated. The scale bar indicates 0.1 amino acid substitutions per site.

To examine the origin of chloroplast L21 protein in different plant species, we compared the amino acid sequences of the conserved core regions. RPL21c is closer to RPL21m (46% identities) than to the chloroplast-encoded L21 protein of *M. polymorpha* (31%). Using parsimony, neighbor-joining and Fitch–Margoliash least-square distance methods, phylogenetic trees that include prokaryotic and prokaryotic-like L21 sequences were constructed using 103 positions common to 38 different species. Protein sequences of cytoplasmic 60S L21 r-proteins that possess a different signature than prokaryotic L21 r-proteins were not included in the trees. Only the Fitch–Margoliash tree is shown (Fig. 3). Results indicate that plant RPL21 r-proteins belong to two different clades. Similar results were obtained using parsimony or neighbor-joining methods of analysis. High bootstrap values were obtained by using the different methods of phylogenetic analyses showing that the chloroplast *RPL21* genes of lower plants (algae and the bryophyte *M. polymorpha*) have a different origin than the nuclear *RPL21* genes encoding the chloroplast L21 r-protein in higher plants. The clade of lower plants is clearly of cyanobacterial origin as expected. The clade of higher plants includes mitochondrial L21 r-proteins suggesting a mitochondrial origin. Interestingly, the recently sequenced genome of the worm *C. elegans* encodes a L21-like protein (Accession number: CAA87060) that is included in the mitochondrial group and that shares a common origin with higher plant L21 r-proteins. These results strongly suggest a mitochondrial origin for the higher plant clade and consequently for genes encoding chloroplast L21. Independently of these phylogenetic aspects, we analyzed higher plant *RPL21* gene organization. Genes encoding either mitochondrial or chloroplast L21 gene products possess common features. The five exons and four introns are present at exactly the same position in the *RPL21m* and *RPL21c* genes (Fig. 2), indicating a common origin. We also searched for specific residues, unique to the products of nuclear genes of mitochondrial origin. Analysis of the sequence data of the 38 species used for phylogenetic tree construction (Fig. 3) showed that the F, G and Y residues that are indicated in the sequences reported in Fig. 2 are constantly and specifically present in the products of *RPL21* genes of mitochondrial origin. These three residues might represent the signature of these genes. Altogether, the

results show a mitochondrial origin for higher plant nuclear genes encoding chloroplast RPL21 proteins.

3.3. Promoter elements controlling the expression of the *RPL21c* gene

The replacement of the chloroplast gene product by a product of a nuclear gene of mitochondrial origin raises the question of the control of gene expression in order to adapt the synthesis of the *RPL21c* gene to chloroplast development in mesophyll cells, in parallel with leaf development. As shown in Fig. 1, the mitochondrial *RPL21m* gene shows a different expression pattern compared to the chloroplast *RPL21c* gene. The *RPL21c* is preferentially expressed in photosynthetic tissues. Several *cis*-acting elements have been characterized in promoters of nuclear genes encoding plastid proteins. Among these elements are the GT1 box (for review see Zhou, 1999), and the S1 and S2 sites (Villain et al., 1994; Lagrange et al., 1997). Sequences corresponding to these three sites are also present in the promoter region of the *RPL21c* gene as indicated in Table 1. These elements bind *trans*-acting factors present in *Arabidopsis* as shown by gel shift assays using appropriate probes (Fig. 4). The protein/DNA interactions are specific as shown by probing with mutated S1 site and mutated GT1 box or by competition with cold S2 site probe. Altogether, these experiments show that the *RPL21c* gene has acquired regulatory elements found in other nuclear genes encoding chloroplast r-proteins.

4. Discussion

Results presented in this paper indicate a mitochondrial origin for the nuclear *RPL21c* gene encoding a chloroplast r-protein in *Arabidopsis*. Phylogenetic results and intron positions support this conclusion. *RPL21c* is distant from the *M. polymorpha* chloroplast-encoded *rpl21* gene and bootstrap analysis places *RPL21c* outside the entire lineage of chloroplast-encoded genes and concomitantly from their cyanobacterial ancestor. In contrast, *RPL21c* is placed in a sister clade with a group of mitochondrial origin. Mitochondria descend from an eubacteria different from cyanobacteria and related to the α -proteobacterial group. *Rickettsia prowazekii* represents one of the closest eubacterial

Table 1

Common promoter *cis*-elements of the spinach *RPL21*, *RPS1*, and of *Arabidopsis RPL21c* genes^a

Origin	S1 site	S2 site	GT1 box
Spinach <i>RPS1</i> (1)	227–TTC ATGGTA ACA–248	254–CAC ATACAC CT–264	268–GTTAG TTAA AAGA–280
Spinach <i>RPL21</i> (2)	163–CAA ATGGTA TTA–152 137–CAA ATGGTA TTA–126	302–CCC ATACA TTA–292	234–GTT GGTTAA TAAT–246
<i>A. thaliana RPL21c</i> (3)	78–ACA ATGGTA T–68	429–ATA ATACA AC–419 243–TAA ATACA CTA–233	102–TTT GGTTAA ACTT–114

^a Sequences are numbered from the transcription start site, using promoter P2 (Lagrange et al., 1997). (1) Villain et al. (1994); (2) Lagrange et al. (1997); (3) Accession number: gi6289106. The S1 and GT1 boxes in the spinach *RPS1* promoter have similar binding properties (Villain et al., 1994).

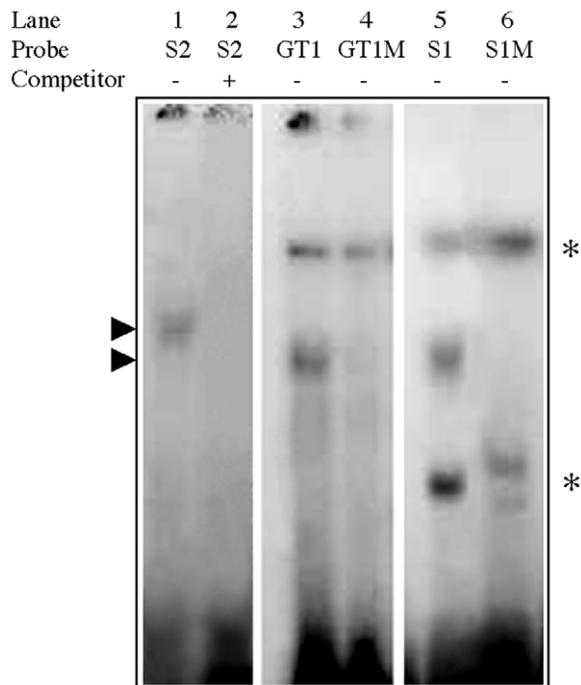


Fig. 4. Gel shift assays for *Arabidopsis* S2, GT1 and S1 sites. The labeled S2 site fragments (lanes 1 and 2), GT1 box and S1 site dimers (lanes 3 and 5) or mutated GT1M and S1M dimers (lanes 4 and 6) were incubated with protein extracts. S2 site fragments were incubated in the absence (–) or presence (+) of a 100-fold molar excess of unlabeled wild-type fragments (lanes 1 and 2). Protein/DNA complexes are indicated by arrowheads. The asterisks indicate non-specific bands.

genomes to mitochondria (Andersson et al., 1998). In our phylogenetic analyses, the L21 of *R. prowazekii* is not closely related to the higher plant L21 clade. This is probably related to the lack of accuracy in the branching of different bacterial groups as verified by the very low bootstrap values obtained after introducing 100 replicates in the different tree constructions. This is indeed related to the fact that bacterial phylogeny can not be determined by using a short protein as L21 only.

The *RPL21c* gene is present in one copy, as for other genes encoding plastid r-proteins. Its localization on chromosome 1 suggests that the *RPL21c* gene originated from a duplication of the *RPL21m* gene located on chromosome 4 or from a polyploidization event as recently suggested (*Arabidopsis* Genome Initiative, 2000). The neighbor genes surrounding *RPL21m* (a SP8-like gene and an unknown gene) have not been conserved during this duplication or polyploidization.

The replacement of a chloroplast-encoded gene product by a nuclear-encoded gene product requires a number of steps including the acquisition of a sequence for chloroplast targeting. In addition, in the case of *RPL21c*, the original mitochondrial exons had to be adapted to code for a protein that fits the chloroplast ribosome architecture.

We have considered the question of expression of a gene formerly adapted to mitochondria and whose product is

now destined to chloroplasts. As shown in Fig. 1, *RPL21m* is constitutively expressed in different tissues. The promoter regions of the *RPL21c* and *RPL21m* genes are not conserved (data not shown). To adapt the expression to chloroplasts the promoter region of the *RPL21c* gene must have acquired several specific *cis*-elements. Interestingly, these elements are present in several nuclear genes encoding chloroplast r-proteins of spinach and *Arabidopsis* (Zhou, 1999; Lagrange et al., 1997) suggesting that coordinated synthesis of different plastid r-proteins is made via similar *cis*-elements.

The recruitment of a nuclear gene of mitochondrial origin to replace the function of a formerly chloroplast-located gene is unusual. It probably represents an exception to the rule. The products of many genes of cyanobacterial origin that are encoded in the nucleus are re-located to plastids (Martin et al., 1998; Martin and Herrmann, 1998). The reason for the exceptional channeling of nuclear gene of mitochondrial origin to chloroplasts remains speculative. However, regarding the quantity of genes of cyanobacterial origin that have been successfully transferred to the nucleus during evolution, one can easily envisage that some transfer events were unsuccessful, i.e. the integration of the gene into the nuclear DNA did not occur in a region favorable for the linkage of a sequence encoding a plastid transit peptide necessary for the transport of the protein to the plastid. In such a case, replacement by a gene of mitochondrial origin becomes possible, if gene duplication, integration of the duplicated gene in a favorable context and compatibility of function are assured.

Acknowledgements

We thank Gérard Clabault for advice and help, and Professor N.-H. Chua for the gift of the GT1 box (BoxII) and mutated GT1M box oligonucleotides. We acknowledge the financial support of Région Rhône-Alpes (Programme “La Semence”) and of the EEC (ESSA 3 Programme).

References

- Andersson, S.G., Zomorodipour, A., Andersson, J.O., Sicheritz-Ponten, T., Alsmark, U.C., Podowski, R.M., Naslund, A.K., Eriksson, A.S., Winkler, H.H., Kurland, C.G., 1998. The genome sequence of *Rickettsia prowazekii* and the origin of mitochondria. *Nature* 396, 133–140.
- Arabidopsis* Genome Initiative, 2000. Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* 408, 796–815.
- Baldauf, S.L., Palmer, J.D., 1990. Evolutionary transfer of the chloroplast *tufA* gene to the nucleus. *Nature* 344, 262–265.
- Gantt, J.S., Baldauf, S.L., Calie, P.J., Weeden, N.F., Palmer, J.D., 1991. Transfer of *rpl22* to the nucleus greatly preceded its loss from the chloroplast and involved the gain of an intron. *EMBO J.* 10, 3073–3078.
- Gray, M.W., Burger, G., Lang, B.F., 1999. Mitochondrial evolution. *Science* 283, 1476–1481.
- Lagrange, T., Gauvin, S., Yeo, H.J., Mache, R., 1997. S2F, a leaf-specific

- trans-acting factor, binds to a novel cis-acting element and differentially activates the RPL21 gene. *Plant Cell* 9, 1469–1479.
- Martin, W., Herrmann, R.G., 1998. Gene transfer from organelles to the nucleus: how much, what happens, and why? *Plant Physiol.* 118, 9–17.
- Martin, W., Stoebe, B., Goremykin, V., Hansmann, S., Hasegawa, M., Kowallik, K., 1998. Gene transfer to the nucleus and the evolution of chloroplasts. *Nature* 393, 162–165.
- Martin, W., Lagrange, T., Li, Y.F., Bisanz-Seyer, C., Mache, R., 1990. Hypothesis for the evolutionary origin of the chloroplast ribosomal protein L21 of spinach. *Curr. Genet.* 18, 553–556.
- Sambrook, J., Maniatis, T., Fritsch, E.F., 1989. *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- Sanchez, H., Fester, T., Kloska, S., Schroder, W., Schuster, W., 1996. Transfer of rps19 to the nucleus involves the gain of an RNP-binding motif which may functionally replace RPS13 in Arabidopsis mitochondria. *EMBO J.* 15, 2138–2149.
- Villain, P., Clabault, G., Mache, R., Zhou, D.X., 1994. S1F binding site is related to but different from the light-responsive GT-1 binding site and differentially represses the spinach rps1 promoter in transgenic tobacco. *J. Biol. Chem.* 269, 16626–16630.
- Zhou, D.X., 1999. Regulatory mechanism of plant gene transcription by GT-elements and GT-factors. *Trends Plant Sci.* 4, 210–214.