RESEARCH ARTICLE

INTERNATIONAL MICROBIOLOGY

Presence of multiple group I introns closely related to bacteria and fungi in plastid 23S rRNAs of lichen-forming *Trebouxia*

Eva M. del Campo,[1]* Leonardo M. Casano,[1] Francisco Gasulla,[2] Eva Barreno[2]

[1]Department of Plant Biology, University of Alcalá, Alcalá de Henares, Spain. [2]Cavanilles Institute of Biodiversity and Evolutionary Biology (CBIBE), Botany Department, Faculty of Biology, University of Valencia, Valencia, Spain

Summary. The chloroplast-encoded large subunit ribosomal RNA gene of several free-living green algae contains group I introns at *Escherichia coli* genic positions 1917, 1931, 1951, and 2449. Herein we report the presence of group I introns at these positions within the chloroplast-encoded large subunit ribosomal RNA gene of several lichen-forming green algae belonging to the *Trebouxia* genus. In contrast to the introns inserted at position 2449, all introns inserted at positions 1917, 1931, and 1951 contained LAGLIDADG homing endonuclease genes. Phylogenetic analyses show that: (i) introns inserted at positions 1917, 1931, and 1951 are closely related to introns located at homologous insertion sites in bacterial rDNA genes; and (ii) introns inserted at position 2449 are closely related to fungal introns located at homologous insertion sites in mitochondrial rDNA genes. The symbiogenetic thalli of some lichens are proposed as the likely setting of horizontal transfer of genetic material among distantly related organisms such as bacteria, fungi, and green algae. [**Int Microbiol** 2009; 12(1):59-67]

Key words: *Trebouxia* ssp. · group I introns · plastid 23S rRNA · lichens · horizontal transfer

Introduction

Group I introns are widespread elements of DNA, found more frequently in eukaryotes than in prokaryotes [23]. More than half (~58%) of identified group I introns are located within the nuclear genome, with the rest divided between chloroplasts and mitochondria. A remarkable feature of group I introns is their ability to colonize new insertion sites, resulting in their spread within the genome [23]. Generally, intron insertion is catalyzed by endonucleases that recognize and cleave the target gene [37]. These proteins, called homing endonucleases (HEs), are encoded by open reading frames within introns called homing endonuclease genes (HEG). Some HEs act via flexible homing-site recognition, without significant loss of binding affinity or cleavage efficiency. This flexible but specific recognition strategy could explain the sporadic distribution of group I introns, and suggests that many of these transposable elements have undergone horizontal transfer into different species and genes. Along this line of thinking, Goddard and Burt [19] proposed that frequent horizontal transmission is necessary for the long-term persistence of HEGs.

Lichens are complex organisms with a specific biological organization [3]. Their thalli are the result of cyclical symbiotic associations [9] between at least two very different organisms, a heterotrophic fungus (mycobiont) and a photosynthetic partner or photobiont (cyanobacteria or unicellular green algae). The symbiotic interaction between the mycobiont and the photobiont(s) occurs through a process of symbiogenesis and results in the lichen thallus, which is a unique entity, or holobiont [28]. This context of close contact among

*Corresponding author: Eva M. del Campo
Department of Plant Biology
University of Alcalá
28871 Alcalá de Henares, Spain
Tel. 34-918856432 Fax 34-918855066
E-mail: eva.campo@uah.es

mycobiont and photobiont(s) may favor horizontal transfer events, making lichens a useful tool for studying the evolutionary history of genetic elements, especially in the case of potentially mobile elements like group I introns. However, very few studies have focused on group I introns in the nuclear rDNAs of lichen-forming algae [4,18], and there is even less information about group I introns within their organelles. To investigate these genetic elements and evaluate their evolutionary importance, we examined the nature and distribution of group I introns in chloroplasts of several *Trebouxia* species, which are the most common lichen photobionts.

Accordingly, portions of the plastid LSU rDNA that exhibited a high frequency of group I introns among *Chlorophyta* were sequenced in a number of different *Trebouxia* species: *Trebouxia angustilobata* Beck (Tan); *T. jamesii* (Hildr. & Ahmadjian) Gärtner (Tja); *T. erici* Ahmadjian; *T. gelatinosa* Ahmadjian; *T . glomerata* (Waren) Ahmadjian; *T. irregularis* Hildr. & Ahmadjian; *T. suecica* Beck (Tsu); *T. showmanii* (Hildr. & Ahmadjian) Gärtner (Tsh); and *Trebouxia* sp. (Tsp) isolated from a Valencian population (Spain) of the lichen *Ramalina farinacea* (L.) Ach. After sequencing, 31 different group I introns were studied by analyzing their structure and their possible evolutionary history in the context of lichen symbiosis.

Materials and methods

Biological material. Axenically cultured photobionts were used. The following photobionts were obtained from the Culture Collection of Algae (SAG) at the University of Göttingen (Germany): *Trebouxia angustilobata* (strain 2204), *Trebouxia erici* (strain 32.85), *Trebouxia glomerata* (strain 100.80), *Trebouxia irregularis* (strain 33.85), *Trebouxia showmanii* (strain 2009), and *Trebouxia suecica* (strain 2207). Other cultured photobionts were obtained from the Culture Collection of Algae at the University of Texas (Austin, TX, USA): *Trebouxia gelatinosa* (strain 906) and Trebouxia *jamesii* (strain 2233). *Trebouxia* sp. was isolated in our laboratories from a population of the lichen *Ramalina farinacea* (L.) Ach., according to Gasulla (PhD unpublished data). The thalli were collected in the air-dried state on *Quercus rotundifolia* Lam. at Sierra del Toro (Castellón, Spain; 39°54′16″ N, 0°48′22″ W). Samples were frozen at –20ºC until the isolation experiment, 3 months after collection. Organisms were cultured in Bold 3N medium [5] in a growth chamber in white light (25 μmol m^{-2}s^{-1}) at 15ºC, under 14 h:10 h light:dark cycles.

Nucleic acids extraction and purification, cDNA synthesis, and PCR amplification. Total DNA was isolated from each strain of *Trebouxia* algae following the procedure of Cenis (1992) [8]. Isolated DNAs were amplified by PCR with specific primers covering positions 759–2536 in the *Escherichia coli* 23S rRNA [31]: 23ST1: 5′-CCG-AAAGGCGTAGTCGATGG-3′; 23ST2 5′-CAGTAATGCTTCATAGGG-TC-3′; 23ST3 5′-GAAGGGAGCTTGACTGCAAG-3′; 23ST4: 5′-CCGTC-GATGTGAACTCTTGG-3′. Oligonucleotides were designed based on the nucleotide sequence of the LSU rDNA from *Trebouxia aggregata* (DDBJ/EMBL/GenBank accession no. L43542, [32]) and on partial sequences obtained in our laboratory. All reactions were done under the following cycling conditions: one cycle at 94ºC for 2 min; 30 cycles of 94ºC

for 30 s, 55ºC for 30 s, and extension at 72ºC for 2 min. These cycles were followed by a final extension step at 72ºC for 7 min. Total RNA was extracted from each *Trebouxia* photobiont with PureYield RNA Midiprep System (Promega, Madison, WI, USA), according to the manufacturer's instructions. Isolated RNAs were used as templates for reverse transcription and further PCR amplification with the Access RT-PCR System (Promega). Reverse transcription reactions were conducted at 48ºC for 60 min. The obtained cDNAs were amplified by PCR with appropriate primer pairs under the following cycling conditions: 94ºC for 2 min; 30 cycles of: 94ºC for 30 s, 55ºC for 30 s, and 68ºC for 1 min. Amplification products were subjected to electrophoresis through 1.5% agarose gels. The corresponding bands were excised from the gel, purified with a DNA gel extraction kit (Qiaex II; Qiagen GmbH, Hiden Germany), and directly sequenced using the Big DyeTM Terminator Cycle Sequence Ready Reaction Kit II (Applied Biosystems, Foster City, CA, USA). The obtained reaction products were separated by automated Multicapillary Electrophoresis and further analyzed on an ABI Prism 3730 Genetic Analyzer (Applied Biosystems).

Phylogenetic analyses. All of the sequences determined in this work were compared with group I introns downloaded from GenBank [http://www.ncbi.nlm.nih.gov]. Sequences were aligned using the default parameters implemented in Muscle 3.6 [13] and then adjusted with BioEdit 7.0 [20]. For constructing "host" cell phylogenies of *Trebouxia* based on LSU-rDNA sequences, all 651 nucleotide sites were included, resulting in a total of 275/141 variable/parsimony-informative sequence positions. The data set for constructing phylogenies of introns inserted at positions 1917, 1931, and 1951 comprised 28 taxa and 1600 characters after excluding 466 sequence positions (1–47 and 1648–2066), resulting in 1514/778 variable/parsimony-informative sequence positions. The data set for constructing phylogenies of introns inserted at position 2449 consisted of 14 taxa and 1081 characters, resulting in a total of 366/175 variable/parsimony-informative sequence positions. The data set for constructing phylogenies of HE encoded within the studied intron was made up of 27 taxa and 122 characters after excluding 105 sequence positions (1–42 and 163–227). All of the sequences were included in the phylogenetic analyses but gaps were treated as missing characters. All data sets were subjected to maximum-parsimony (MP), maximum-likelihood (ML), and distance analyses using PAUP 4.0b10 [38] and also to a Bayesian approach using MrBayes 3.1.2 [24]. In the MP and ML analyses, a heuristic search was performed using random stepwise addition, the tree bisection–reconnection (TBR), and the branch-swapping algorithm to find the best tree. For the ML analyses, the transition/transversion ratio was estimated from the minimum evolutionary tree as implemented in PAUP. Likelihood settings corresponded to the HKY85 model. MP analyses were also carried out via quartet puzzling analysis, yielding values greater than 0.7 for the consistency index (CI) and below 0.3 for the homoplasy index (HI). Distance analyses were done by using the neighbor-joining method [35]. Trees were built from a general time-reversible distance matrix calculated using the minimum evolution criterion [34] in conjunction with the HKY85 distance correction [21]. For the minimum evolution method, a heuristic search procedure was used and site-to-site rate heterogeneity was modeled as a gamma distribution [39]. Bootstrap probabilities [16] were calculated to estimate the robustness of the clades from 1000 replicates of the data. Phylogenetic Bayesian analyses (MrBayes 3.1.2 [24]) were performed using the GTR model [33]. Markov chain Monte Carlo method (MCMC) from a random starting tree was initiated in the Bayesian inference and run for 10^5 generations. Posterior probabilities for each node were inferred from the resulting consensus tree using the "sumt" option of MrBayes.

Results and Discussion

Group I introns in the plastid LSU rDNA of *Trebouxia* phycobionts. A relatively high number of group I introns have been reported for free-living green algae within organellar LSU rDNAs, including *Trebouxiophyceae* [27]. Nevertheless, very little information is currently avail-

Table 1. Group I introns in the plastid rDNA of *Trebouxia* species[a]

Organism	Intron name	rDNA insertion site	Intron size (nt)[b]	Classification	HEG size (aa)[b]
Trebouxia sp.	Tsp.cL1917	1917–1918	1018	IB4	171
	Tsp.cL1951	1951–1952	1010	IA3	99
	Tsp.cL2449	2449–2450	627	IA1	–
T. angustilobata	Tan.cL1917	1917–1918	922	IB4	138
	Tan.cL1931	1931–1932	733	IB4	165
T. suecica	Tsu.cL1917	1917–1918	1207	IB4	171
	Tsu.cL1931	1931–1932	676	IB4	165
T. showmanii	Tsh.cL1951	1951–1952	814	IA3	100
T. jamesii	Tja.cL1917	1917–1918	934	IB4	151
	Tja.cL2449	2449–2450	350	IA1	–
T. gelatinosa	Tge.cL2449	2449–2450	322	IA1	–
T. irregularis	Tir.cL2449	2449–2450	511	IA1	–
T. erici	Ter.cL2449	2449–2450	405	IA1	–
T. glomerata	Tgl.cL2449	2449–2450	356	IA1	–

[a]Nomenclature of group I introns is according to [26]. Intron classification was carried out according to [7] and [29].
[b]nt: nucleotides; aa: amino acids.

able for symbiotic lichen-forming green algae. For this reason, we searched for introns within the plastid-encoded rRNA large subunit gene from a number of common phycobionts of the *Trebouxia* genus. This study is part of a global investigation focused on the phylogenetic relationships among lichen phycobionts based on comparisons of the plastid-encoded rRNA genes and the evolutionary history of their introns within the context of lichen symbiosis. Here, we amplified two regions of the gene in which a relatively high frequency of group I introns has been observed in other members of the *Chlorophyta* [32]. These two regions (from position 1363 to position 2081, and from position 2335 to position 2486 in *E. coli* LSU rDNA) were amplified with specific primer pairs (23ST1/23ST2 and 23ST3/23ST4, respectively) and sequenced in nine *Trebouxia* species. The obtained sequences are accessible in GenBank: *Trebouxia angustilobata* (FJ694755), *T. erici* (EU352795), *T. gelatinosa* (FJ687229), *T. glomerata* (FJ687228), *T. irregularis* (EU725861), *T. jamesii* (EU352 794), *T. showmanii* (FJ694756), *T. suecica* (FJ694757) and *Trebouxia* sp. (EU600236) from the lichen *Ramalina farinacea* (L.) Ach. All of the sequences of the chloroplast LSU rDNA of *Trebouxia* phycobionts available to date, except that of *Trebouxia aggregata*, which has already been published [27], were determined by our laboratories. We also determined experimentally the precise 5′ and 3′ ends of these introns and their splicing in vivo by amplifying and sequencing the corresponding cDNAs (data not shown).

Fourteen different group I introns were identified among the *Trebouxia* species investigated, as summarized in Table 1. This relatively high number of introns contrasts with the fact that, generally, for the free-living *Trebouxiophyceae* described to date no introns within the plastid LSU rDNA have been reported [10].

All introns analyzed in this study (Table 1) have characteristics indicating that they are of the group I type: (i) the base-paired stems P3, P4, P5, P6, P7, and P8; and (ii) canonical conserved sequences P, Q, R, and S (Table 2), characteristic of subgroup IA and IB intron cores [29]. The phylograms depicted in Fig. 1B, C indicate that the analyzed introns are grouped with other introns of the same subgroup, and inserted at the same position within the gene encoding the plastid LSU rDNA.

LAGLI-DADG HEs within group I introns in the chloroplast LSU rDNA of *Trebouxia* photobionts. To obtain relevant information about intron transfer processes, we searched for open reading frames (ORFs) within the group I introns under study. The results, summarized in Table 1, indicated that eight of the 14 studied introns potentially encoded HEs. Introns inserted at positions 1917, 1931, and 1951 contained ORFs encoding putative proteins of 99–171 amino acids (Tan.cL1917, Tja.cL1917, Tsp.cL1917, Tsu.cL1917, Tan.cL1931, Tsu.cL1931, Tsh.cL1951, and Tsp.cL1951). Interestingly, none of the introns inserted at position L2449 encoded a protein. The ORF localization within the corresponding introns followed a characteristic pattern depending on the intron subgroup: the IA1

Table 2. Conserved core elements of *Trebouxia* plastid LSU rDNA introns

Subgoup	Intron name	P4 (P)	P4′ (Q)	P7 (R)	P7′ (S)
IA1	Ter.cL2449	UGCAAA	UUUGCAGCG	CGUUCAACGACUA	UAGUCU
	Tge.cL2449	UGCAAG	UUUGCAGCG	CGUUCAACGACUA	UAGUCU
	Tgl.cL2449	UGCAAA	UUUGCAGCG	CGUUCAACGACUA	AGGUCU
	Tir.cL2449	UGCAAA	UUUGCAGCG	CGUUCAACGACUA	UAGUCU
	Tja.cL2449	UGCAAA	UUUGCAGCG	CGUUCAACGACUA	UAGUCU
	Tsp.cL2449	UGCAAA	UUUGCAGCG	CGUUCAACGACUA	UAGUCU
IA3	Tsh.cL1951	ACGGUG	UACCGCGGG	CCUGUAACGACUG	UAGUCU
	Tsp.cL1951	ACGGUG	UACCGCGGG	CCUGUAACGACUG	UAGUCU
IB4	Tan.cL1917	UGCUGG	UCAGCAGGA	UCCUCAGAGACUA	UAGUCC
	Tja.cL1917	UGCUGG	UCAGCAGGA	UCCUCAGAGACUA	UAGUCC
	Tsp.cL1917	UGCUGG	UCAGCAGGA	UCCUCAGAGACUA	UAGUCC
	Tsu.cL1917	UGCUGG	UCAGCAGGA	UCCUCAGAGACUA	UAGUCC
	Tan.cL1931	UGCUGG	UCAGCAGGG	CCCUCAGAGACUA	UUGUCC
	Tsu.cL1931	CGCUAG	UCAGCAGGG	CCCUCAGAGACUA	UUGUCA

Nomenclature of group I introns is according to [26]. Conserved sequences of the subgroup IA and IB intron cores are according to [29].

subgroup introns featured the loop of the P8 helix; the IA3 and IB4 subgroups featured an extension from the peripheral loops towards the intron core structure.

More importantly, all of these putative proteins contained one LAGLIDADG motif, identified by using Pfam v21.10 [17], characteristic of the HE involved in intron mobility. The LAGLIDADG family of HE comprises more than 200 proteins and is the most diverse of the HE families since its members are present in the genomes of plant and algal chloroplasts, fungal and protozoan mitochondria, bacteria, and archaea [37]. One reason for the wide distribution of these proteins appears to be their ability to invade unrelated types of intervening sequences, including group I introns. Although LAGLIDADG proteins display almost no similarity in terms of primary sequence, they share a conserved three-dimensional structure and use a flexible homing-site recognition strategy without significant loss of binding affinity or cleavage efficiency [37]. The biochemical basis of this mechanism is to make undersaturated DNA–protein contacts across long DNA target sites. Taking into account the role of the group I ribozyme in splicing and homing, it is probable that HE ORFs are under strong selective pressure to maintain both the catalytic RNA function in splicing and the production of a functional HE protein.

Relationship between introns at positions L1917, L1931 and L1951 with bacterial group I introns. With respect to host evolution, an examination of the LSU rDNA tree (Fig. 1B) shows that the members of the *Trebouxia* genus analyzed in this study form two distinct main lineages: lineage 1, defined by *T. angustilobata, T.gelatinosa, T. jamesii, Trebouxia* sp.*, T. showmanii*, and *T. suecica*; and lineage 2, defined by *T. erici, T. glomerata*, and *T. irregularis*. The monophyletic origin of each of these two lineages is supported by high bootstrap values. When the different introns were mapped onto the host phylogeny (Fig. 1B), it is noteworthy that introns inserted at positions 1917, 1931, and 1951 were restricted to lineage 1 and seemed to be inherited by a common ancestor. This fact is consistent with a stable inheritance (or loss) of the 1917, 1931, and 1951 group I introns throughout the evolutionary history of these *Trebouxia* species of lineage 1 (however, this hypothesis is based on limited data). For instance, the distribution of the 1917 introns indicates their retention in most of *Trebouxia* species and their possible loss in *T. gelatinosa* (Fig. 1B,C). To test the hypothesis of inheritance of the 1917 introns from a common ancestor, the partition homogeneity test (PHT) was implemented [15]. This test indicates whether the phylogenetic history of the L1917 introns differs from that of the
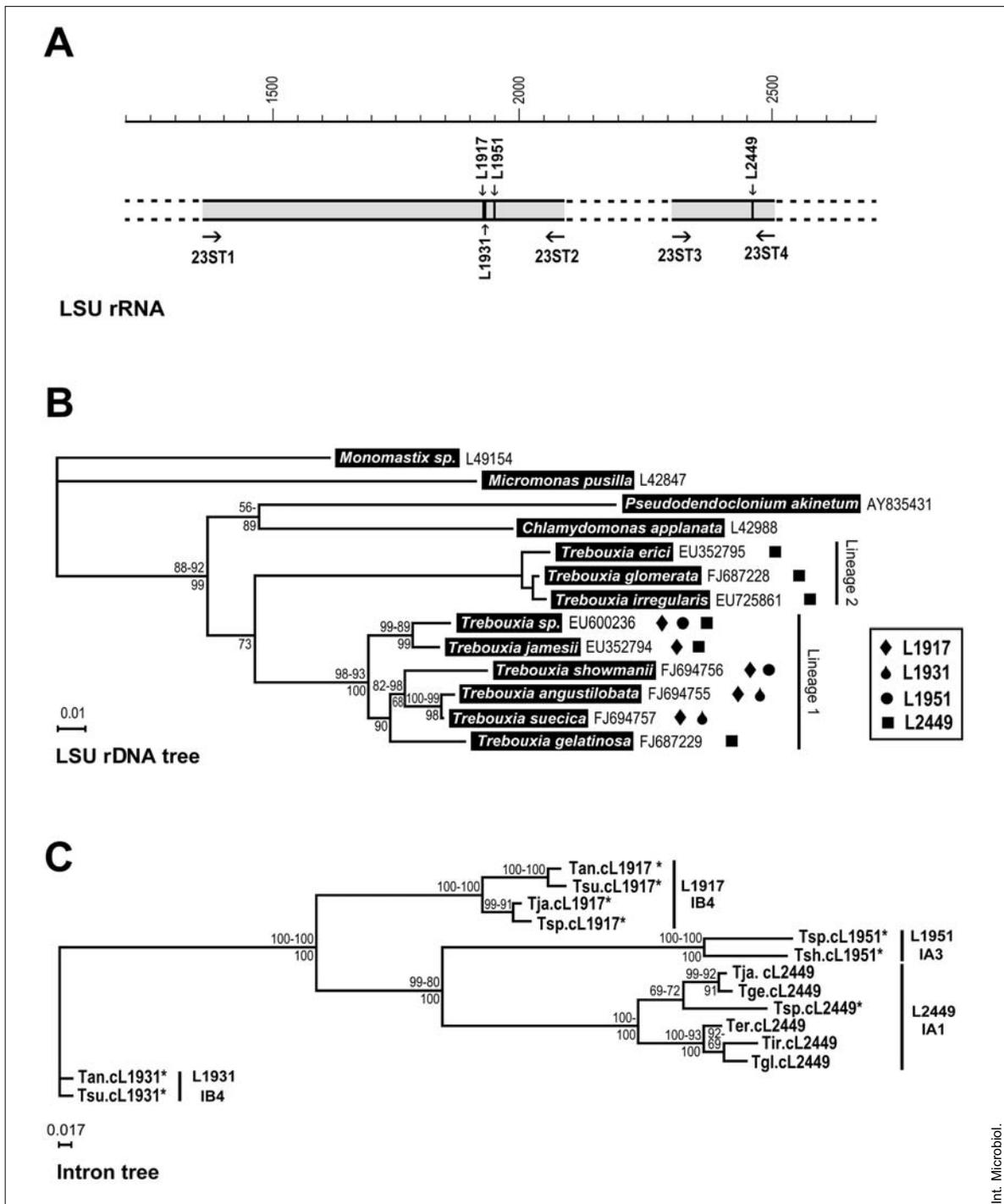
Fig. 1. Genetic map and phylogenies of plastid LSU rDNA and their group I introns from *Trebouxia* spp. (**A**) Genetic map of the sequenced portions of plastid LSU rDNA from *Trebouxia* spp. (1363–2081 and 2335–2486 positions in *Escherichia coli*). Light gray boxes represent the amplified and further sequenced portions of the gene. Discontinuous lines indicate non-amplified portions. Primer pairs: 23ST1/23ST2 and 23ST3/23ST4, which amplified these two portions of the gene, respectively, are depicted at the bottom as horizontal arrows. Intron positions (1917, 1931, and 2449) are indicated with vertical arrows. The scale bar at the top indicates nucleotide positions in *E. coli*. (**B**) Phylogeny of chloroplast LSU rDNA. The analysis of chloroplast LSU rDNA was based on 647 aligned sequence positions (200 parsimony-informative sites). Different introns inserted at fixed positions and labeled with specific symbols (see legend at the right) are mapped onto the host phylogeny. (**C**) Phylogeny of chloroplast LSU rDNA group I introns. The analysis was based on 1511 sequence positions (569 parsimony-informative sites). Asterisks indicate the presence of putative HEs. Vertical bars on the right denote introns that share insertion positions as well as intron subgroups (indicated at right of the bar). Phylograms were inferred with the maximum-likelihood method. Values at the left of the slash mark were inferred with a Bayesian analysis; those at the right of the slash mark, with a maximum-parsimony method; those below the branches indicate bootstrap values. Bootstrap values were recorded only for those lineages supported by 50% of all (1000) replicates. The scale bar indicates the inferred number of substitutions per site.
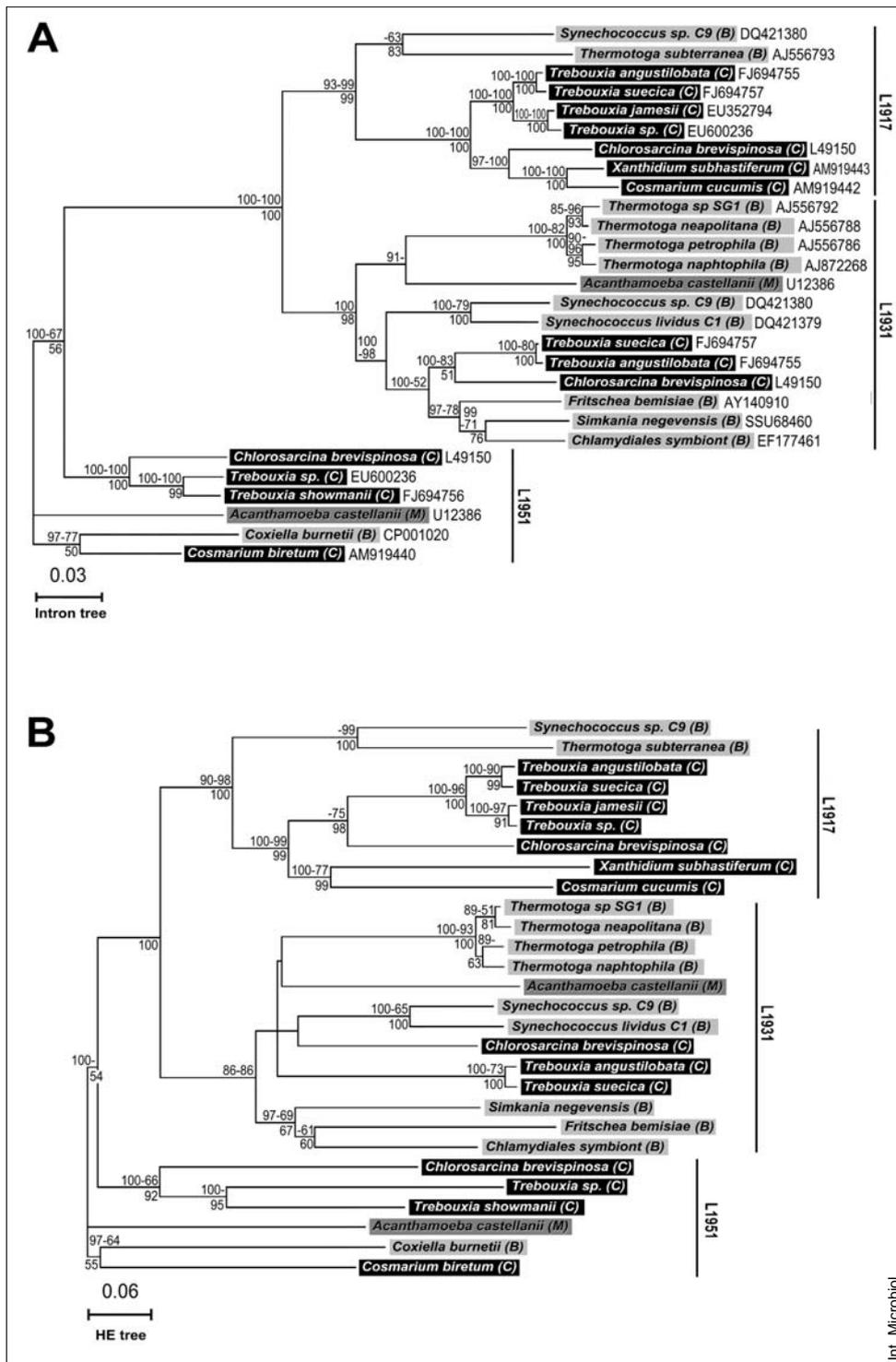
Fig. 2. Unrooted phylogenies of plastid group I introns inserted at positions L1917, L1931, and L1951 and those of other similar group I introns. (**A**) Phylogram based on intron sequences. The analysis of introns inserted at positions 1917, 1931, and 1951 was based on 1981 aligned sequence positions (778 parsimony-informative sites). The phylogram was inferred with the maximum-likelihood method. (**B**) Phylogram based on the corresponding HE proteins. The analysis of HEs encoded within the introns inserted at positions 1917, 1931, and 1951 was based on 100 aligned sequence positions. Data sets were subjected to distance analyses using the neighbor-joining method and the minimum evolution criterion. In both phylograms, bacterial introns (*B*) are shown on a light-gray background, chloroplast introns (*C*, all from green algae) on a black background, and introns from mitochondria of *Acanthamoeba castellanii* (*M*, all from Ascomycota) on a dark-gray background. Vertical bars on the right of the tree denote groups shareing insertion positions in the LSU rDNA. Values at the left of the slash mark were inferred with a Bayesian analysis, those at the right of the slash mark, with a maximum-parsimony method; those below the branches indicate bootstrap values (1000 replicates). Values were recorded only for lineages supported by 50% of all replicates. The scale bar indicates the inferred number of substitutions per site.

host. A data partition was established between exonic (positions 1363–1917 of plastid LSU rDNA) and intronic (corresponding to all 1917 introns) sequences. The results indicated that these two datasets did not differ according to the PHT (P = 1), supporting the initial hypothesis of vertical transmission of L1917 introns from a common ancestor.

A search of the NCBI databases with BLAST using the obtained *Trebouxia* group I introns L1917, L1931, and L1951 showed significant homologies with introns in the LSU rDNA of distantly related organisms such as prokaryotes in addition to chloroplasts from green algae and mitochondria of the ameboid protist *Acanthamoeba castellanii*. The phylograms in Fig. 2
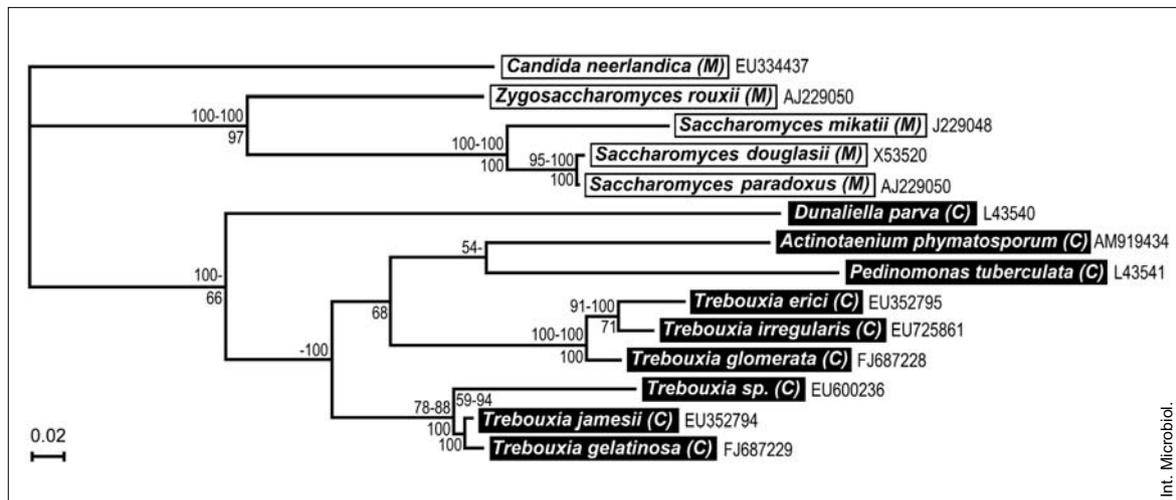
Fig. 3. Unrooted phylogeny of plastid and mitochondrial L2449 group I introns. The analysis of introns inserted at position 2449 was based on 859 aligned sequence positions (151 parsimony-informative sites). Phylogram was inferred with the maximum-likelihood method. Values at the left of the slash mark were inferred with a Bayesian analysis, those at the right of the slash mark, with a maximum-parsimony method; those below the branches indicate bootstrap values (1000 replicates). Values were recorded only for those lineages supported by 50% of all replicates. The scale bar indicates the inferred number of substitutions per site. Fungal introns are in black letters on a white background; those from green algae are in white letters on a black background.

correspond to introns inserted at positions 1917, 1931 and 1951 (Fig. 2A) and the encoded HE (Fig. 2B). Our inferred phylogenetic trees indicated that *Trebouxia* introns and their HE form a cluster with all other known introns or HE with the same insertion site in bacteria and chloroplasts from other green algae (Fig. 2). These findings are consistent with the hypothesis of horizontal transmission of these genetic elements among green algae chloroplasts and bacteria. Accordingly, it was proposed that the acquisition of intron-HEG elements (precisely at positions 1917, 1931, and 1951) by several thermophilic bacteria and cyanobacterial strains occurred through horizontal transfer from eukaryotic organisms [22,30]. The biological connections among these organisms and the mode of group I intron transmission is unclear. In the case of the introns within the mitochondrial LSU rDNA of *Acanthamoeba castellanii*, Haugen et al. (2007) pointed out that the bacteria *Simkania negevensis* is capable of growing and persisting in acanthamoebal cells, indicating a potential association between these two organisms which harbor 1931 introns. The presence of nonphotosynthetic bacteria has also been detected in lichen thalli by electron microscopy [11,36]. More recently, in situ hybridization experiments and fluorescence-labeled nucleic acid probes complementary to rRNAs allowed the improved characterization of bacteria within lichen thalli [6].

Relationship between introns at position L2449 with yeast mitochondrial group I introns. In contrast to 1917 introns, comparisons of the chloroplast LSU rDNA and 2449 group I intron phylogenies of *Trebouxia*

species provided evidence against a stable ancestry of the intron within this genus. As shown in Fig. 1B,C, the phylogenies of 2449 introns are not congruent with those of the host gene (chloroplast LSU rDNA). To investigate incongruence between these phylogenies, the PHT [14] was applied, which established a data partition between exonic (positions 1363–1917 of the plastid LSU rDNA gene) and intronic (corresponding to all 2449 introns) sequences. The results suggested that 2449 introns have different phylogenetic histories than the host since the two datasets differed significantly according to the PHT ($P < 0.001$). Such phylogenetic incongruence is strong evidence for horizontal transmission among different *Trebouxia* species. Nonetheless, these analyses alone do not guarantee the phylogenetic relationships of donor and recipient species. For example, it is not clear whether the transfers occurred among different members of *Trebouxia* or whether the genus repeatedly exchanged elements with more distantly-related taxa, as in the case of introns 1917, 1931, and 1951. To address this issue, a BLAST search of the NCBI databases was carried out using the sequences of 2449 introns from each of the different *Trebouxia* species. Significant matches with introns within the mitochondrial LSU rDNA of several *Saccharomyces* yeasts were found. Interestingly, the 2449 position in the mitochondrial LSU rDNA in yeasts is occupied by an ω-HE-containing group I self-splicing intron [25]. In 1999, Goddard and Burt [19] provided compelling evidence of horizontal transmission of intact mitochondrial DNA sequences among yeast species by surveying 20 species of yeast for a

particular HEG class known as ω (also known as r1 and Sc LSU.1). Figure 3 shows a phylogram of introns inserted at position 2449 containing a broad mixture of sequences from green algae chloroplast and fungal (*Ascomycetes*) mitochondria. The fact that several chloroplast sequences are closely related to a number of fungal mitochondrial sequences strongly indicates a history of horizontal transmissions. Interestingly, none of the L2449 introns included in Fig. 3 had any ORFs, except Tsp.cL2449, which encoded a putative protein of 71 amino acids. This relatively short protein could result from a degeneration of the encoding ORF, rendering a non-functional ω-HE, since functional ω elements in yeasts have a half-length of more than 200 amino acids. Additionally, we observed different intron states among the studied *Trebouxia* species: (i) an intron with a residual ω-HEG, in *Trebouxia* sp.; (ii) an intron without HEG, in *T. erici, T. gelatinosa, T. glomerata, T. irregularis,* and *T. jamesii*; and (iii) no intron, in *T. angustilobata, T. showmanii*, and *T. suecica* (Fig. 1B). These data are consistent with the perpetual cycle of invasion by transmission, followed by degeneration, eventual loss, and then re-invasion proposed for the ω element in yeast [19]. Again, the observed different intron states were not congruent with the host phylogeny based on plastid LSU rRNA sequences (Fig. 1), reinforcing the notion that these introns were transferred horizontally among *Trebouxia* species, organelles (chloroplasts-mitochondria), and/or organisms (algae-fungus).

Possible horizontal transmission of group I introns among lichen symbionts. Lichen associations are one of the most successful lifestyles among the fungi, representing almost 20% of the 65,000 described fungi [12]. Almost all of the 13,500 lichen-forming species are Ascomycetes (lichen-forming fungi represent 40% of all described Ascomycetes) and only about 50 are Basidiomycetes. These data are consistent with the fact that some of our studied *Trebouxia* introns showed a significant similarity to those of distantly related organisms such as prokaryotes and fungi. The very restricted and sporadic phylogenetic distribution of group I introns (especially within bacteria, mitochondria, and chloroplasts) suggested that they have been frequently transferred horizontally. Ahmadjian (1993) [1] suggested that lichen symbionts exchange genetic material and, a decade later, Friedl et al. [18] proposed that lichenization facilitates intron lateral transfer. In addition, in the last few years, the existence of "mycosomes" in cultures of *Trebouxia erici* has been described [2]. Mycosomes are defined as wall-deficient reproductive fungal propagules associated with algal and plant plastid development. Interestingly, the isolated mycosomes from *T. erici* were identified as a variant of the original lichen fungus of *Cladonia cristatella*, which contained *T. erici* as photobiont [2]. It is noteworthy that the strain of *Trebouxia erici* sequenced in our study (SAG 32.85 identical to UTEX 911) was isolated from the same species of lichen. Mycosomes may facilitate an interchange of genetic material between photobionts and fungi. Lichenoidal fossils suggest that the early steps of lichenization occurred in a marine ecosystem where a large number of free-living cyanobacteria, algae, and fungi were in close association, a necessary step in the evolution of symbiosis [40]. It is plausible to propose the symbiogenetic thalli of some lichens as a possible scenario for the horizontal transfer of genetic material among distantly related organisms such as bacteria, fungi, and green algae.

References

1. Ahmadjian V (1993) The lichen symbiosis. John Wiley, New York
2. Atsatt PR (2003) Fungus propagules in plastids: the mycosome hypothesis. Int Microbiol.6:17-26
3. Barreno E (2004) Hongos simbiontes: Líquenes, Micoficobiosis y Micorrizas. In: Izco J (ed) Botánica. McGraw-Hill Interamericana, Madrid
4. Bhattacharya D, Friedl T, Damberger S (1996) Nuclear-encoded rDNA group I introns: origin and phylogenetic relationships of insertion site lineages in the green algae. Mol Biol Evol 13:978-989
5. Bold HC, Parker BC (1962) Some supplementary attributes in the classification of Chlorococcum species. Arch Mikrobiol 42:267-288
6. Cardinale M, Vieira de Castro J Jr, Muller H, Berg G, Grube M (2008) In situ analysis of the bacterial community associated with the reindeer lichen *Cladonia arbuscula* reveals predominance of Alphaproteobacteria. FEMS Microbiol Ecol 66:63-71
7. Cech TR, Damberger SH, Gutell RR (1994) Representation of the secondary and tertiary structure of group I introns. Nat Struct Biol 1:273-280
8. Cenis JL (1992) Rapid extraction of fungal DNA for PCR amplification. Nucleic Acids Res 20:2380
9. Chapman MJ, Margulis L (1998) Morphogenesis by symbiogenesis. Int Microbiol 1:319-326
10. de Cambiaire JC, Otis C, Turmel M, Lemieux C (2007) The chloroplast genome sequence of the green alga *Leptosira terrestris*: multiple losses of the inverted repeat and extensive genome rearrangements within the Trebouxiophyceae. BMC Genomics 8:213
11. De los Ríos A, Wierzchos J, Sancho LG, Green TGA, Ascaso C (2005) Ecology of endolithic lichens colonizing granite in continental Antarctica. Lichenologist 37:383-395
12. DePriest PT (2004) Early molecular investigations of lichen-forming symbionts: 1986-2001. Annu Rev Microbiol 58:273-301
13. Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res 32:1792-1797
14. Farris JS (1989) The retention index and homoplasy excess. Syst Zool 38:406-407
15. Farris JS, Källersjö M, Kluge AG, Bult D (1995) Testing significance of incongruence. Cladistics 10:315-319

16. Felsenstein J (1985) Confidence limits on phylogenies: An approach using the bootstrap. Evolution 39:783-791

17. Finn RD, Mistry J, Schuster-Bockler B, et al. (2006) Pfam: clans, web tools and services. Nucleic Acids Res 34:D247-251

18. Friedl T, Besendahl A, Pfeiffer P, Bhattacharya D (2000) The distribution of group I introns in lichen algae suggests that lichenization facilitates intron lateral transfer. Mol Phylogenet Evol 14:342-352

19. Goddard MR, Burt A (1999) Recurrent invasion and extinction of a selfish gene. Proc Natl Acad Sci USA 96:13880-13885

20. Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symposium Series 41:95-98

21. Hasegawa M, Kishino H, Yano T (1985) Dating the human-ape split by a molecular clock of mitochondrial DNA. J Mol Evol 22:160-174

22. Haugen P, Bhattacharya D, Palmer JD, Turner S, Lewis LA, Pryer KM (2007) Cyanobacterial ribosomal RNA genes with multiple, endonuclease-encoding group I introns. BMC Evol Biol 7:159

23. Haugen P, Simon DM, Bhattacharya D (2005) The natural history of group I introns. Trends Genet 21:111-119

24. Huelsenbeck JP, Ronquist F (2001) MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics 17:754-755

25. Jacquier A, Dujon B (1983) The intron of the mitochondrial 21S rRNA gene: distribution in different yeast species and sequence comparison between *Kluyveromyces thermotolerans* and *Saccharomyces cerevisiae*. Mol Gen Genet 192:487-499

26. Johansen S, Haugen P (2001) A new nomenclature of group I introns in ribosomal DNA. RNA 7:935-936

27. Lucas P, Otis C, Mercier JP, Turmel M, Lemieux C (2001) Rapid evolution of the DNA-binding site in LAGLIDADG homing endonucleases. Nucleic Acids Res 29:960-969

28. Margulis L, Barreno E (2003) Looking at lichens. BioScience 53:776-778

29. Michel F, Westhof E (1990) Modelling of the three-dimensional architecture of group I catalytic introns based on comparative sequence analysis. J Mol Biol 216:585-610

30. Nesbo CL, Doolittle WF (2003) Active self-splicing group I introns in 23S rRNA genes of hyperthermophilic bacteria, derived from introns in eukaryotic organelles. Proc Natl Acad Sci USA 100:10806-10811

31. Ohnishi M, Murata T, Nakayama K, et al. (2000) Comparative analysis of the whole set of rRNA operons between an enterohemorrhagic *Escherichia coli* O157:H7 Sakai strain and an *Escherichia coli* K-12 strain MG1655. Syst Appl Microbiol 23:315-324

32. Pombert JF, Lemieux C, Turmel M (2006) The complete chloroplast DNA sequence of the green alga *Oltmannsiellopsis viridis* reveals a distinctive quadripartite architecture in the chloroplast genome of early diverging ulvophytes. BMC Biol 4:3

33. Rodriguez F, Oliver JL, Marin A, Medina JR (1990) The general stochastic model of nucleotide substitution. J Theor Biol 142:485-501

34. Rzhetsky A, Nei M (1992) Statistical properties of the ordinary least-squares, generalized least-squares, and minimum-evolution methods of phylogenetic inference. J Mol Evol 35:367-375

35. Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol 4:406-425

36. Souza-Egipsy V, Wierzchos J, Garcia-Ramos JV, Ascaso C (2002) Chemical and ultrastructural features of the lichenvolcanic/sedimentary rock interface in a semiarid region (Almeria, Spain). Lichenologist 34:155-167

37. Stoddard BL (2005) Homing endonuclease structure and function. Q Rev Biophys 38:49-95

38. Swofford DL (2003) PAUP. Phylogenetic analysis using parsimony (and other methods), version 4. Sinauer Associates, Sunderland, MA

39. Yang Z (1994) Maximum likelihood phylogenetic estimation from DNA sequences with variable rates over sites: approximate methods. J Mol Evol 39:306-314

40. Yuan X, Xiao S, Taylor TN (2005) Lichen-like symbiosis 600 million years ago. Science 308:1017-1020