ELSEVIER

FEMS Microbiology Letters 233 (2004) 153–157

FEMS MICROBIOLOGY Letters

www.fems-microbiology.org

The non-photosynthetic, pathogenic green alga *Helicosporidium* sp. has retained a modified, functional plastid genome

Aurélien Tartar *, Drion G. Boucias

Department of Entomology and Nematology, University of Florida, Gainesville, FL 32611-0620, USA

Received 24 December 2003; received in revised form 4 February 2004; accepted 5 February 2004

First published online 21 February 2004

Abstract

A fragment of the *Helicosporidium* sp. (Chlorophyta: Trebouxiophyceae) plastid genome has been sequenced. The genome architecture was compared to that of both a non-photosynthetic relative (*Prototheca wickerhamii*) and a photosynthetic relative (*Chlorella vulgaris*). Comparative genomic analysis indicated that *Helicosporidium* and *Prototheca* are closely related genera. The analyses also revealed that the *Helicosporidium* sp. plastid genome has been rearranged. In particular, two ribosomal protein-encoding genes (*rpl19* and *rps23*) appeared to have been transposed, or lost from the *Helicosporidium* sp. plastid genome. RT-PCR reactions demonstrated that the retained plastid genes were transcribed, suggesting that, despite rearrangement(s), the *Helicosporidium* sp. plastid genome has remained functional. The modified plastid genome architecture is a novel apomorphy that indicates that the Helicosporidia are highly derived green algae, more so than *Prototheca* spp. As such, they represent a promising model to study organellar genome reorganizations in parasitic protists.

Keywords: *Helicosporidium*; Green algae; Plastid genome; Genome reduction; Pathogen

1. Introduction

The Helicosporidia are obscure pathogenic protists that have been reported in a wide range of invertebrate hosts [1]. They are characterized by the formation of a highly resistant cyst that encloses three ovoid cells and a diagnostic filamentous cell [2]. To date, it remains unclear whether the Helicosporidia possess a free-living stage or are obligate pathogens that exist outside their hosts only as cysts. A new *Helicosporidium* sp. was recently isolated in Florida [1]. Morphological and molecular data compiled on this organism have demonstrated that the Helicosporidia are non-photosynthetic green algae, and they are related to *Prototheca*, another non-photosynthetic, parasitic algal genus [1,3–5]. Furthermore, sequencing of a chloroplast gene [4] provided evidence that *Helicosporidium* spp., as prototechans [6,7], have retained a modified chloroplast (plastid) and chloroplast genome.

Cryptic, modified chloroplasts (and their genomes) have been reported in a variety of non-photosynthetic protists, including the green algae *Prototheca wickerhamii* [6], the euglenoid *Astasia longa* [8], the stramenopiles *Pteridomonas danica* and *Ciliophrys infusionum* [9] and the apicomplexan parasites *Plasmodium falciparum* and *Toxoplasma gondii* [10]. Sequence information on secondary non-photosynthetic plastid genomes is accumulating, showing that these genomes are much smaller than that of photosynthetic relatives, but they have remained functional. A widely accepted hypothesis is that the reduction in size can be explained by the loss of most of the genes involved in photosynthesis. The remaining genes have been selectively retained because they are involved in other essential plastid function(s). However, the number of retained plastid genes varies depending on the species, suggesting that secondary non-photosynthetic plastids have been retained for different reasons. As reviewed by Williams and Keeling [11], the

* Corresponding author. Tel.: +1-352-392-1901x147; fax: +1-352-392-0190.

E-mail address: dgb@mail.ifas.ufl.edu (A. Tartar).

0378-1097/$22.00
doi:10.1016/j.femsle.2004.02.006

plastid genomes of parasitic organisms (*P. falciparum*, *P. wickerhamii*) tend to be more reduced.

The *Helicosporidium* sp. plastid genome is expected to be similar to that of *P. wickerhamii* (estimated at 54 kb, [6]). However, its exact size has yet to be determined. Pulse field electrophoresis techniques did not allow for visualization and size estimation of this genome because, as it was previously reported for *Chlorella vulgaris* [12], organellar, circular DNA molecules failed to enter the gel [13]. In an effort to better characterize the *Helicosporidium* vestigial chloroplast, sequencing of the genome has been initiated. Here, we report the sequencing of a portion of the plastid genome, as well as comparative genomic analysis using two close relatives: the *P. wickerhamii* plastid genome [6] and the *C. vulgaris* chloroplast genome [14].

2. Materials and methods

2.1. Helicosporidium isolate and culture conditions

The *Helicosporidium* sp. was originally isolated from a black fly larvae [1]. It was maintained in vitro in Sabouraud Maltose agar supplemented with 2% Yeast extract (SMY) at 25 °C. Helicosporidial cells produced on these plates were inoculated into flasks containing SMY broth and shaken at 23 °C on a rotary shaker (250 rpm) for 3–4 days. Cells were collected by centrifugation and used for DNA extraction.

2.2. DNA extraction and PCR amplification

Cellular DNA was extracted as previously described [3,4], using the MasterPure Yeast DNA purification kit (Epicentre). The *Helicosporidium* sp. elongation factor gene *tufA* was amplified using the degenerate primers TufAf and TufAr (Table 1). The resulting amplification product was gel-extracted and sequenced. Gene-specific primers (GSPs) were designed from the *Helicosporidium* sp. *tufA* sequence and used in combination with primers designed from genes predicted to be located on a locus close to *tufA* within the chloroplast genome. The use of the fMET and rpl2R primers (Table 1) allowed for the amplification and subsequent sequencing of the 5′ and 3′ flanking regions, respectively.

Table 1
Primer names and sequences

Primer name	Sequences
TufAf	AAYATGATTACAGGTGCTGC
TufAr	ACGTAAACTTGTGCTTCAAA
fMET	GGGTAGAGCAGTCTGGTAGC
rpl2R	CCTTCACCACCACCATGCG

2.3. RNA extraction and RT-PCR

Helicosporidium sp. cells were frozen under liquid nitrogen and ground into a fine powder. Total RNA was isolated using TriReagent, according to the manufacturer's protocol. To prevent any DNA contamination, *Helicosporidium* RNA was treated with RNase free DNase before being resuspended in formamide and stored at −70 °C. Prior to storage, an aliquot of the RNA suspension was used to spectrophotometrically estimate the final concentration. Upon utilization, stored RNA was reprecipitated in 4 volumes of 100% ethanol and 0.2 M sodium acetate (pH 5.2) and suspended in distilled water. First-strand cDNA synthesis was performed using 1 μg of total RNA, a *tufA* gene specific primer and the Thermoscript RT-PCR system from Life Technologies, following the manufacturer's directions. The *tufA* primer was then combined with a *rps12* and a *rps7* gene-specific primers in two separate reactions (Fig. 2(b)) that were performed under the same conditions: 30 cycles of 94 °C for 30 s, 50 °C for 30 s, and 72 °C for 3 min.

3. Results

3.1. Analysis of the plastid genome sequence

Portions of the plastid genome were readily PCR-amplified from *Helicosporidium* sp. total genomic DNA. A similar technique, based on the PCR amplification of overlapping sequences, was recently used to sequence the entire *Eimeria tenella* apicoplast genome [15]. A 3348 bp fragment was amplified and sequenced from *Helicosporidium* sp. (GenBank Accession No. AY498714). Sequence comparison analyses demonstrated that the fragment contains four open reading frames (ORFs), corresponding to the elongation factor *tufA* and the ribosomal proteins *rps12*, *rps7* and *rpl2*. In addition, the 5′ end of the sequenced ptDNA fragment includes a portion of the proline tRNA (tRNA-P) gene. All five *Helicosporidium* sp. plastid genes are similar to homologous genes sequenced from both *P. wickerhamii* and *C. vulgaris* chloroplast genomes. Furthermore, phylogenies reconstructed from a *tufA* alignment identified *Helicosporidium* sp. as a sister taxon to *P. wickerhamii* (data not shown).

The overall organization of the sequenced *Helicosporidium* sp. ptDNA fragment is presented in Fig. 1. The *tufA*, *rps7* and *rps12* genes are known as the *str*-(streptomycin) cluster. This cluster is conserved across archeabacteria and eubacteria, including chloroplasts as intracellular descendants of the latter [16]. Not surprisingly, the *str*-cluster is also conserved in *Helicosporidium* sp. plastid genome (Fig. 1). The *Helicosporidium* sp. ptDNA has an organization that is very similar to that *P. wickerhamii*, especially in regard to the location of the *rpl2* gene. In both *Helicosporidium* sp. and *P. wickerhamii* ptDNA, this gene is

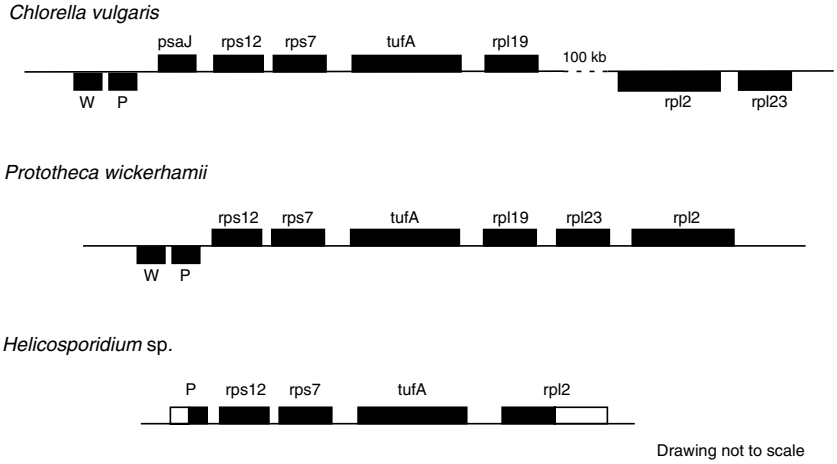

Fig. 1. Comparison of the *Helicosporidium* sp. plastid genome fragment with that of non-photosynthetic (*Prototheca wickerhamii*) and photosynthetic (*Chlorella vulgaris*) close relatives. The sequenced regions are in black. The direction of transcription is from left to right for genes depicted above the lines and from right to left for those shown below the line.

located close to the 3′ end of the *str*-cluster. This common organization differs from that of *C. vulgaris* and other photosynthetic green algae (such as the ancestral *Nephroselmis olivacea* [17]), suggesting that the common ancestor of *Helicosporidium* sp. and *P. wickerhamii* possessed a rearranged chloroplast genome. Rearrangements included the fusion of the *rpl2* cluster and *str*-cluster and may have been associated with the loss of photosynthesis.

Despite these similarities, the *Helicosporidium* sp. ptDNA fragment is also remarkably different from that of *P. wickerhamii* (Fig. 1). First, two genes, corresponding to the ribosomal proteins *rpl19* and *rps23*, have yet to be been found in *Helicosporidium* sp. These genes are not at their expected loci, suggesting that they may have been transposed within the plastid genome, or lost. As noted by Stoebe and Kowallik [16], modifications in chloroplast genomes occur mainly in form of gene losses. Therefore, even if only a portion of the ptDNA has been sequenced, a likely hypothesis is that both *rpl19* and *rps23* have been lost from the *Helicosporidium* sp. plastid genome. In addition to the absence of the *rpl19* and *rps23* genes, the orientation of the *str*-cluster in relation to the tRNA-P gene is different in *Helicosporidium* sp.: the tRNA-P gene is located on the same strand as the *str*-cluster and is transcribed in the same direction (Fig. 1). In contrast, the *Prototheca* tRNA-P orientation is similar to photosynthetic relatives such as *C. vulgaris* and *N. olivacea*, suggesting that it represents an ancestral type among green algae. Overall, the *Helicosporidium* ptDNA fragment (Fig. 1) is characterized by a unique, derived organization, which differs from that of *P. wickerhamii*.

3.2. RT-PCR reactions

As presented in Fig. 2, the *str*-cluster was successfully amplified from *Helicosporidium* sp. cDNA, demonstrating

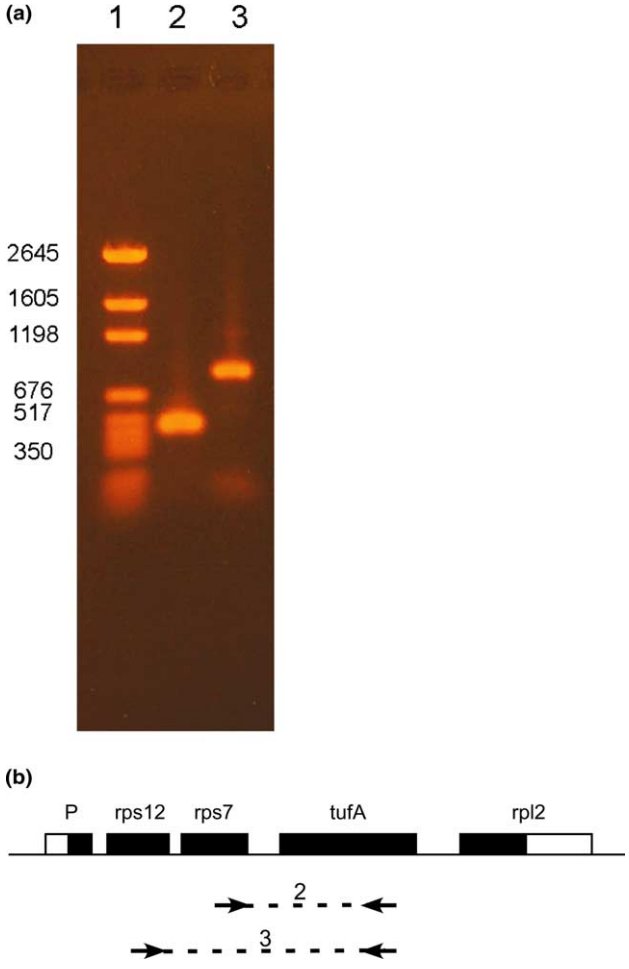

Fig. 2. RT-PCR amplification of the *Helicosporidium* sp. *str*-cluster. (a) RT-PCR products run on a 1% agarose gel, and viewed under UV light. The product in lane 2 was obtained using a combination of gene specific primers corresponding to the *rps7* (forward) and *tufA* (reverse) genes. The product in lane 3 was obtained with *rps12* (forward) and *tufA* (reverse) gene specific primers. DNA markers (pGEM) are shown in lane 1. (b) Schematic illustration of RT-PCR reactions.

that the ptDNA genes are expressed. Nuclear gene primers failed to amplify any products (not shown), indicating that the cDNA synthesis was specific and that the amplification of the *str*-cluster was not due to DNA contamination. The RT-PCR products showed that the *str*-cluster genes are transcribed on the same mRNA molecule in an operon-like manner reminiscent of the chloroplast bacterial origin [16]. Importantly, the fact that plastid genes are expressed suggests that the *Helicosporidium* sp. plastid genome has remained functional.

4. Discussion

Previous phylogenetic analyses [3–5] have demonstrated that the Helicosporidia are close relatives of the non-photosynthetic algae *Prototheca* spp. (Chlorophyta; Trebouxiophyceae). In accordance with these analyses, *Helicosporidium* spp. are believed to possess a *Prototheca*-like plastid and a plastid genome [4]. Although the *Helicosporidium* sp. plastid has yet to be observed in microscopic examination, the combined PCR and RT-PCR amplifications presented in this study showed that *Helicosporidium* sp., as *P. wickerhamii*, has retained plastid genes, including the conserved *str*-cluster, that are expressed in helicosporidial cells. The presence of a transcribed ptDNA in *P. wickerhamii* has been demonstrated by Northern Blot analysis [6]. To date, the function of these vestigial organelles remains unclear.

A fragment of the *Helicosporidium* sp. ptDNA was sequenced and its architecture was compared to that of similar chloroplast genome fragments previously sequenced from both non-photosynthetic and photosynthetic relatives. These comparative genomic analyses revealed that the *Helicosporidium* sp. ptDNA is most similar to that of *P. wickerhamii*, confirming that these two organisms arose from a common, most recent ancestor [3,4]. However, a number of dissimilarities were also identified, suggesting that the Helicosporidia possess a unique, more derived plastid genome that has experienced additional reorganization events. One hypothesis that may explain these differences in genome architecture involves gene losses, and the possibility that the *Helicosporidium* sp. plastid genome may be more reduced than the 54 kb *P. wickerhamii* ptDNA.

Concordant with the hypothesis that the helicosporidial ptDNA has been reduced in size is the fact that the nuclear genome appeared reduced as well. Karyotyping analyses showed that this genome is composed of nine chromosomes, ranging from 700 to 2000 kb [13]. Summing up the sizes of individual chromosomes gave a 10.5 Mb estimate for the *Helicosporidium* sp. nuclear genome, three times smaller than the genome of one of *Helicosporidium* sp. closest relatives, *C. vulgaris* (38.8 Mb) [12]. Genome reduction is a common pattern observed for both pathogenic prokaryotes [18] and eukaryotes [19], and it is associated with the evolution toward pathogenicity and an obligate, host-dependent, minimalist lifestyle. Interestingly, biological observations that include the existence of a very specific infectious cyst stage [1] and the ability to replicate intracellularly within insect hemocytes [20] have shown that the Helicosporidia possess characteristics that have not been reported for *Prototheca* spp. and that suggest that *Helicosporidium* spp. are more derived toward an obligate pathogenic lifestyle. Such observations concur with the hypothesis that the *Helicosporidium* sp. plastid genome may be smaller than that of *P. wickerhamii*.

The generation of the complete sequence of the *Helicosporidium* sp. plastid genome will provide information on the extent of the genome reduction and rearrangement event(s). It is interesting to note that, besides Apicomplexa, the Helicosporidia are the only known plastid-containing, intracellular pathogens. The Apicomplexa are a group of intracellular parasites that includes the malaria causative agent *P. falciparum*. Recent evidence showed that these parasites arose from a photosynthetic ancestor and have retained a modified chloroplast (named apicoplast) [10]. To date, it is unclear if the *Helicosporidium* sp. plastid genome is similar to the 35-kb apicoplast genome but, as noted by Williams and Keeling [11], the Helicosporidia represent a remarkable opportunity to compare the evolution of non-photosynthetic plastids in two unrelated groups of intracellular pathogens. They may also prove to be a better model to study the transition from a free-living, autotrophic stage to a parasitic, heterotrophic stage and the impact of this transition on both nuclear and plastid genomes (gene losses and transfers) because the phylogenetic affinity of *Helicosporidium* spp. and its relationships to both non-photosynthetic and photosynthetic relatives have been well established [3,4], in contrast to the situation for Apicomplexa.

Acknowledgements

The authors thank Verena Blaeske and Michael Botts for technical assistance during *Helicosporidium* sp. in vitro cultures; the ICBR Sequencing Facility at the University of Florida; and Byron Adams, Jim Maruniak, Dave Clark and Bill Farmerie for critical comments on an earlier draft of the manuscript. This work is supported by a grant from the National Science Foundation (NSF, MCB-0131017). Florida Agricultural Experiment Station Journal Series No. R-09949.

References

[1] Boucias, D.G., Becnel, J.J., White, S.E. and Bott, M. (2001) In vivo and in vitro development of the protist *Helicosporidium* sp. J. Eukaryot. Microbiol. 48, 460–470.

[2] Keilin, D. (1921) On the life history of *Helicosporidium parasiticum* n.g. sp., a new species of protist parasite in the larvae of *Dashelaea obscura* Winn (Diptera: Ceratopogonidae) and in some other arthropods. Parasitology 13, 97–113.
[3] Tartar, A., Boucias, D.G., Adams, B.J. and Becnel, J.J. (2002) Phylogenetic analysis identifies the invertebrate pathogen *Helicosporidium* sp. as a green alga (Chlorophyta). Int. J. Syst. Evol. Microbiol. 52, 273–279.
[4] Tartar, A., Boucias, D.G., Becnel, J.J. and Adams, B.J. (2003) Comparison of plastid 16S rDNA (*rrn16*) genes from *Helicosporidium* spp.: evidence supporting the reclassification of Helicosporidia as green algae (Chlorophyta). Int. J. Syst. Evol. Microbiol. 53, 1719–1723.
[5] Ueno, R., Urano, N. and Suzuki, M. (2003) Phylogeny of the non-photosynthetic green micro-algal genus *Prototheca* (Trebouxiophyceae, Chlorophyta) and related taxa inferred from SSU and LSU ribosomal DNA partial sequence data. FEMS Microbiol. Lett. 223, 275–280.
[6] Knauf, U. and Hachtel, W. (2002) The genes encoding subunits of ATP synthase are conserved in the reduced plastid genome of the heterotrophic alga *Prototheca wickerhamii*. Mol. Genet. Genom. 267, 492–497.
[7] Melville, P.A., Benites, N.R., Sinhorini, I.L. and Costa, E.O. (2002) Susceptibility and features of the ultrastructure of *Prototheca zopfii* following exposure to copper sulfate, silver nitrate and chlorexidine. Mycopathologia 156, 1–7.
[8] Gockel, G. and Hachtel, W. (2000) Complete gene map of the plastid genome of the non-photosynthetic euglenoid flagellate *Astasia longa*. Protist 151, 347–351.
[9] Sekiguchi, H., Moriya, M., Nakayama, T. and Inouye, I. (2002) Vestigial chloroplasts in heterotrophic stramenopiles *Pteridomonas danica* and *Ciliophrys infosium* (Dictyochophyceae). Protist 153, 157–167.
[10] Wilson, R.J.M. (2002) Progress with parasite plastids. J. Mol. Biol. 319, 257–274.
[11] Williams, B.A.P. and Keeling, P.J. (2003) Cryptic organelles in parasitic protists and fungi. Adv. Parasitol. 54, 9–67.
[12] Higashiyama, T. and Yamada, T. (1991) Electrophoretic karyotyping and chromosomal gene mapping of *Chlorella*. Nucleic Acids Res. 19, 6191–6195.
[13] Tartar, A. (2004) *Incertae sedis* no more: the phylogenetic affinity of Helicosporidia. Dissertation, University of Florida, Gainesville.
[14] Wakasugi, T., Nagai, T., Kapoor, M., Sugita, M., Ito, M., Ito, S., Tsudzuki, J., Nakashima, K., Tsudzuki, T., Suzuki, Y., Hamada, A., Ohta, T., Inamura, A., Yoshinaga, K. and Sugiura, M. (1997) Complete nucleotide sequence of the chloroplast genome from the green alga *Chlorella vulgaris*: the existence of genes possibly involved in chloroplast division. Proc. Natl. Acad. Sci. USA 94, 5967–5972.
[15] Cai, X., Lorraine Fuller, A., McDougald, L.R. and Zhu, G. (2003) Apicoplast genome of the coccidian *Eimeria tenella*. Gene 321, 39–46.
[16] Stoebe, B. and Kowallik, K.V. (1999) Gene-cluster analysis in chloroplast genomics. Trends Genet. 15, 344–347.
[17] Turmel, M., Otis, C. and Lemieux, C. (1999) The complete chloroplast DNA sequence of the green alga *Nephroselmis olivacea*: insights into the architecture of ancestral chloroplast genomes. Proc. Natl. Acad. Sci. USA 96, 10248–10253.
[18] Moran, N.A. (2002) Microbial minimalism: genome reduction in bacterial pathogens. Cell 108, 583–586.
[19] Vivares, C.P., Gouy, M., Thomarat, F. and Metenier, G. (2002) Functional and evolutionary analysis of a eukaryotic parasitic genome. Curr. Opin. Microbiol. 5, 499–505.
[20] Bläske, V.-U. and Boucias, D.G. (2004) Influence of *Helicosporidium* spp. (Chlorophyta: Trebouxiophyceae) infection of the development and survival of three noctuid species. Environ. Entomol. 33 (1), 54–61.