ISEP XIV

Comparison of plastid 16S rRNA (*rrn16*) genes from *Helicosporidium* spp.: evidence supporting the reclassification of Helicosporidia as green algae (Chlorophyta)

Aurélien Tartar,[1] Drion G. Boucias,[1] James J. Becnel[2] and Byron J. Adams[3]

Correspondence
Aurélien Tartar
aurelien@ufl.edu

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

[2]Center for Medical, Agricultural and Veterinary Entomology, USDA, ARS, Gainesville, FL 32604, USA

[3]Microbiology & Molecular Biology Department, Brigham Young University, Provo, UT 84602-5253, USA

The Helicosporidia are invertebrate pathogens that have recently been identified as non-photosynthetic green algae (Chlorophyta). In order to confirm the algal nature of the genus *Helicosporidium*, the presence of a retained chloroplast genome in Helicosporidia cells was investigated. Fragments homologous to plastid 16S rRNA (*rrn16*) genes were amplified successfully from cellular DNA extracted from two different *Helicosporidium* isolates. The fragment sequences are 1269 and 1266 bp long, are very AT-rich (60·7 %) and are similar to homologous genes sequenced from non-photosynthetic green algae. Maximum-parsimony, maximum-likelihood and neighbour-joining methods were used to infer phylogenetic trees from an *rrn16* sequence alignment. All trees depicted the Helicosporidia as sister taxa to the non-photosynthetic, pathogenic alga *Prototheca zopfii*. Moreover, the trees identified *Helicosporidium* spp. as members of a clade that included the heterotrophic species *Prototheca* spp. and the mesotrophic species *Chlorella protothecoides*. The clade is always strongly supported by bootstrap values, suggesting that all these organisms share a most recent common ancestor. Phylogenetic analyses inferred from plastid 16S rRNA genes confirmed that the Helicosporidia are non-photosynthetic green algae, close relatives of the genus *Prototheca* (Chlorophyta, Trebouxiophyceae). Such phylogenetic affinities suggest that *Helicosporidium* spp. are likely to possess *Prototheca*-like organelles and organelle genomes.

INTRODUCTION

The Helicosporidia, a unique group of invertebrate pathogens, have been detected in insects, collembolans, mites, crustaceans and trematodes and they also have been isolated from ditch water samples (Kellen & Lindegren, 1973; Sayre & Clark, 1978; Purrini, 1984; Avery & Undeen, 1987; Pekkarin, 1993). These pathogens have been found in Europe, South America, North America, Asia and Africa (Keilin, 1921; Weiser, 1970; Kellen & Lindegren, 1973; Hembree, 1979; Seif & Rifaat, 2001). Although *Helicosporidium* spp. seem to be ubiquitous, they have been studied so little that their occurrence and their importance as invertebrate pathogens are unclear. Significantly, their taxonomic position has remained *incertae sedis*. After the first description of a *Helicosporidium* sp. by Keilin (1921), this organism has been thought to be either a protozoan or a fungus, but it has never been clearly associated with any other known protist.

Recently, a *Helicosporidium* sp. was isolated from larvae of the black fly *Simulium jonesii* Stone and Snoddy collected in Florida (Boucias *et al.*, 2001). Microscopic observation of the vegetative growth of *Helicosporidium* sp. under *in vivo* and *in vitro* conditions led Boucias *et al.* (2001) to associate this protist with green algae, particularly the unicellular, non-photosynthetic and pathogenic algae belonging to the

This paper was presented at the XIVth meeting of the International Society for Evolutionary Protistology in Vancouver, Canada, 19–24 June 2002.

Published online ahead of print on 28 March 2003 as DOI 10.1099/ijs.0.02559-0.

Abbreviations: ML, maximum-likelihood; MP, maximum-parsimony; NJ, neighbour-joining.

The GenBank/EMBL/DDBJ accession numbers for the sequences of the plastid 16S rRNA genes of the black fly *Helicosporidium* and the weevil *Helicosporidium* are respectively AF538864 and AF538865.

genus *Prototheca*. Boucias *et al.* (2001) noticed that, as protothecans, the vegetative cells of *Helicosporidium* sp. undergo one or two cell divisions within a pellicle. This pellicle eventually splits open and releases either two or four daughter cells. This association between *Helicosporidium* and *Prototheca* was surprising, but was later confirmed by molecular sequence comparisons (Tartar *et al.*, 2002). Phylogenetic analyses of several *Helicosporidium* sp. genes (rDNA, actin and β-tubulin) all identified this organism as a member of the green algae clade (Chlorophyta). Moreover, a nuclear 18S rDNA phylogeny of the Chlorophyta depicted *Helicosporidium* sp. as a close relative of *Prototheca wickerhamii* and *Prototheca zopfii*, within the class Trebouxiophyceae. Based on both morphological and molecular evidence, the transfer of the genus *Helicosporidium* to Chlorophyta, Trebouxiophyceae was proposed.

Prototheca spp. have been shown to be closely related to the photoautotrophic genus *Chlorella* (Chlorophyta, Trebouxiophyceae), based on phylogenetic analyses inferred from the nuclear 18S rRNA and the plastid 16S rRNA genes (Huss *et al.*, 1999; Nedelcu, 2001). The plastid 16S rRNA gene (*rrn16*) is a chloroplast gene. Despite having lost their photosynthetic abilities, non-photosynthetic green algae such as protothecans have been found to retain vestigial, degenerate chloroplasts called leucoplasts. The presence of such plastids has been demonstrated extensively in the non-photosynthetic green algae of the genus *Polytoma* (Lang, 1963; Siu *et al.*, 1976), which are closely related to *Chlamydomonas* (Chlorophyta, Chlorophyceae). To our knowledge, there are no records of microscopic observations of a leucoplast in a *Prototheca* sp. cell. However, the plastid genome of *Prototheca wickerhamii* has recently been isolated and partially sequenced (Knauf & Hachtel, 2002). Similar to the situation described previously for plastid genomes in non-photosynthetic plants (reviewed by Hachtel, 1996), this genome is highly reduced in size but is believed to be functional.

In order to confirm *Helicosporidium* sp. as a green alga and as a close relative of the genus *Prototheca*, the presence of plastid DNA in helicosporidial cells was investigated. Herein, we report the PCR amplification and sequencing of plastid 16S rDNA homologues from two different isolates of *Helicosporidium*. This gene was targeted because it has been previously sequenced for numerous photosynthetic and non-photosynthetic algal species (Nedelcu, 2001), thereby allowing us to infer a plastid 16S rDNA phylogeny of the Chlorophyta that includes the genus *Helicosporidium*.

METHODS

***Helicosporidium* isolates.** The first *Helicosporidium* sp. was isolated from the black fly *Simulium jonesii* as described previously (Boucias *et al.*, 2001). A second *Helicosporidium* sp. has since been isolated from the weevil *Cyrtobagous salviniae* (Coleoptera: Curculionidae). This insect is a biological control agent for the aquatic weed *Salvinia molesta* (Goolsby *et al.*, 2000). Both isolates were successfully amplified in *Helicoverpa zea* larvae, as described previously (Boucias *et al.*, 2001). Cysts produced in *Helicoverpa zea* larvae were purified by gradient centrifugation on Ludox and grown in artificial medium (TNM-FH insect medium, supplemented with gentamicin and 5 % fetal bovine serum; Sigma-Aldrich) before harvest and DNA extraction. The two isolates will be referred to as weevil *Helicosporidium* and black fly *Helicosporidium*.

DNA extraction and amplification. Helicosporidial DNA was extracted according to Boucias *et al.* (2001) using the Masterpure Yeast DNA extraction kit from Epicentre Technologies. Cellular DNA was used as a template for the PCR amplification of *rrn16* using chloroplast 16S rRNA gene-specific primers ms-5′ and ms-3′ listed by Nedelcu (2001). PCR products were gel-purified with the QiaxII gel extraction kit (Qiagen) and cloned in pGEM-T vectors using the pGEM-T easy vector systems (Promega). Positive clones were sent to the Interdisciplinary Core for Biotechnology Research (ICBR) at the University of Florida for sequencing.

Phylogenetic analyses. The two plastid 16S rDNA sequences from *Helicosporidium* spp. were aligned with homologous sequences available in GenBank. The alignment was obtained using CLUSTAL X software with default parameters (Thompson *et al.*, 1997) and optimized manually. Analyses of the aligned sequences were performed in PAUP* version 4.0 beta 10 (Swofford, 2000), using maximum-parsimony (MP), maximum-likelihood (ML) and neighbour-joining (NJ) methods. The statistical model of DNA substitution used in ML analyses was determined by likelihood ratio tests as implemented in MODELTEST (Posada & Crandall, 1998). This program identified the General Time Reversible model (GTR + I + G) as being the most appropriate for our data. MP analyses were performed using the default parameters in PAUP*. NJ analyses were based on the two-parameter method of Kimura, but other models including the likelihood model used for ML analyses and HK85 were also used. Branch support for MP and NJ analyses was assessed by bootstrapping (100 replicates). Relative rate tests were performed using Tajima's test (Tajima, 1993), as implemented in MEGA version 2.1 (Kumar *et al.*, 2001). Our alignment, as well as the resulting MP and ML trees, can be obtained from TreeBase (Morell, 1996; http://www.treebase.org), with the study accession number S819.

RESULTS

Fragments homologous to plastid 16S rRNA genes were successfully amplified from both *Helicosporidium* cellular DNA preparations. The fragment lengths were 1269 bp for the weevil *Helicosporidium* 16S rDNA and 1266 bp for the black fly *Helicosporidium* 16S rDNA. The two sequences were very similar, but they were not identical. This is concordant with results obtained when comparing other genes from the two isolates: 18S rDNA, actin and β-tubulin fragments sequenced from the two isolates also showed some differences in nucleotide sequences (data not shown). The helicosporidial plastid 16S rRNA genes are very AT-rich: 60·7 % for both *rrn16* sequences. Such a deviation from homogeneity is common in non-photosynthetic alga genes; for example, the A + T content of the *Prototheca zopfii* plastid 16S rDNA is 63·1 % (Nedelcu, 2001).

The two plastid 16S rRNA gene sequences were compared with 21 homologous sequences from algal species belonging to two major classes of Chlorophyta, Trebouxiophyceae and Chlorophyceae. Both classes include some non-photosynthetic species. Phylogenetic reconstructions using NJ and MP methods produced the same tree, presented in

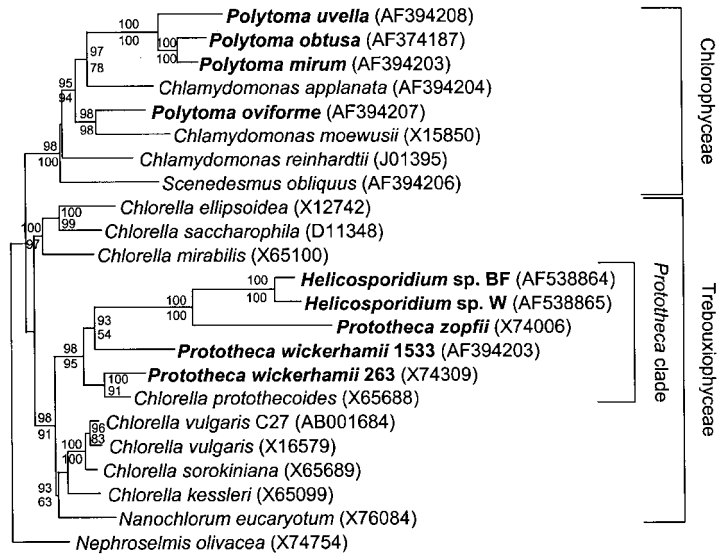

Fig. 1. Phylogenetic tree based on plastid 16S rDNA sequences. *Helicosporidium* spp. are depicted as Trebouxiophyceae, members of a strongly supported *Prototheca* clade and sister taxa to *Prototheca zopfii*. The letters BF and W respectively refer to the black fly and the weevil *Helicosporidium*. Non-photosynthetic taxa are in bold. Branch lengths correspond to evolutionary distances. Numbers above and below nodes represent the results of bootstrap analyses (100 replicates) using MP and NJ methods, respectively. Only values greater than 50 % are shown. All but the helicosporidial sequences were downloaded from GenBank. Accession numbers are indicated after each species name.

Fig. 1. Likelihood analyses resulted in a similar tree, with one topological difference (discussed below).

The MP/NJ tree (Fig. 1) was rooted with the plastid 16S rDNA sequence of *Nephroselmis olivacea*, a member of the class Prasinophyceae, which is thought to include descendants of the earliest-diverging green algae (Turmel *et al.*, 1999). The relationships among green algal taxa depicted in Fig. 1 are consistent with affiliations suggested previously by other phylogenetic studies (Bhattacharya & Medlin, 1998; Huss *et al.*, 1999; Nedelcu, 2001; Tartar *et al.*, 2002). First, both classes (Trebouxiophyceae and Chlorophyceae) appear monophyletic. Within the Chlorophyceae, two non-photosynthetic clades can be identified (Fig. 1): *Polytoma uvella*, *Polytoma obtusum* and *Polytoma mirum* are monophyletic and are sister taxa to *Chlamydomonas applanata*, whereas *Polytoma oviforme* is more closely related to *Chlamydomonas moewusii*. A paraphyletic *Polytoma* has been demonstrated previously by Nedelcu (2001) based on nuclear 18S rDNA and plastid 16S rDNA phylogenies. Only one non-photosynthetic clade exists among the Trebouxiophyceae (as identified by Nedelcu, 2001). This clade is strongly supported by bootstrap values and it includes *Helicosporidium* spp., *Prototheca* spp. and *Chlorella protothecoides*, an auxotrophic, mesotrophic, but photosynthetic species. The genus *Prototheca* appears paraphyletic, as shown previously by nuclear 18S rDNA and plastid 16S rDNA phylogenies (Huss *et al.*, 1999; Nedelcu, 2001). In our tree (Fig. 1), the two *Helicosporidium* isolates are depicted as being monophyletic, sister taxa to *Prototheca zopfii*, and this relationship is supported by maximal bootstrap values. This is consistent with a previously published nuclear 18S rDNA phylogeny (Tartar *et al.*, 2002), which associated the black fly *Helicosporidium* and *Prototheca zopfii*.

Likelihood analyses depicted the same relationships, except for *Prototheca wickerhamii* strain 1533. In the ML tree (not shown; available at http://www.treebase.org), the two strains of *Prototheca wickerhamii* appeared monophyletic, sister taxa to *Chlorella protothecoides*. Kishino–Hasegawa tests (Hasegawa *et al.*, 1985) showed that the ML tree is significantly different from the MP/NJ tree, using ML or MP parameters and optimality criteria (data not shown). However, it should be noted that the lack of stability of the *Prototheca* taxa has already been reported (Nedelcu, 2001) and that it did not have any influence on the position of the *Helicosporidium* spp. in our analyses. As shown in Fig. 1, *Helicosporidium* spp. always appear closely related to *Prototheca zopfii*, within the *Prototheca* clade (*sensu* Nedelcu, 2001).

DISCUSSION

Based on morphological comparisons, Lindegren & Hoffman (1976) introduced the hypothesis that there may be more than one species of *Helicosporidium*. However, to date, there is only one named species (*Helicosporidium parasiticum* Keilin 1921), and few comparative analyses have occurred. Here, we report the discovery of two *Helicosporidium* isolates, from different hosts. These isolates exhibit some polymorphism in all known nucleotide sequences (nuclear 18S rDNA, actin, β-tubulin and plastid 16S rDNA), suggesting that they can be differentiated at a molecular level. However, it remains unclear whether these nucleotide differences are significant and sufficient to propose that our two isolates represent different strains or species. A thorough characterization of these two isolates is currently under way.

The presence of plastid genes suggests strongly that *Helicosporidium* cells may contain a degenerate plastid (leucoplast) and plastid DNA. By itself, the existence of a degenerate chloroplast in *Helicosporidium* cells provides strong arguments in favour of Helicosporidia being non-photosynthetic green algae, but is not sufficient. Indeed, some protozoans (Apicomplexa) have also been shown to possess a degenerate, vestigial chloroplast (apicoplast) with

a functional genome (Wilson, 2002). This plastid has been proposed to derive from an endosymbiotic interaction with a red alga (secondary symbiosis). However, the algal nature of *Helicosporidium* has already been suggested by morphological observations (Boucias *et al.*, 2001) and strongly supported by phylogenetic analyses inferred from several nuclear genes (Tartar *et al.*, 2002). Therefore, helicosporidial cells are likely to possess a plastid similar to other non-photosynthetic Chlorophyta, derived from a primary endosymbiosis.

Comparative analyses of the plastid gene sequences confirm that Helicosporidia are closely related to non-photosynthetic algae in the class Trebouxiophyceae (Chlorophyta). In all our phylogenetic trees, *Helicosporidium* spp. appear as members of the *Prototheca* clade (as defined by Nedelcu, 2001), sister taxa to *Prototheca zopfii*. The positions of *Helicosporidium* spp. are identical in phylogenies based on nuclear 18S rRNA genes (Tartar *et al.*, 2002). Similar to the situation observed in the 18S rDNA phylogeny, the branch leading to the *Helicosporidium*+*Prototheca zopfii* clade is the longest of the tree, suggesting that this association could be an artifact due to long-branch attraction. However, it should be noted that *Helicosporidium* spp. are depicted in exactly the same position even if *Prototheca zopfii* is removed from the sequence alignment and that their relationship with *Prototheca wickerhamii* is still very strongly supported (data not shown). Therefore, this relationship is probably not an artifact.

Based on our phylogenetic analyses (Tartar *et al.*, 2002; this study), we propose that the Helicosporidia should be included in the *Prototheca* clade defined by Nedelcu (2001). The clade is consistently and strongly supported by resampling tests, suggesting that *Helicosporidium* spp., *Prototheca* spp. and *Chlorella protothecoides* may have arisen from a common ancestor. Within the clade, the relationships are less robust: the genus *Prototheca* has always appeared paraphyletic, and *Chlorella protothecoides*, despite being proposed to be the closest green relative of *Prototheca* spp., has never appeared in a basal position (Huss *et al.*, 1999; Nedelcu, 2001; Tartar *et al.*, 2002). In our trees (Fig. 1), these ambiguities remain. However, additional resolution may be obtained inside the *Prototheca* clade by adding more taxa and/or by using other genes, such as protein-encoding genes, which are likely to exhibit a lower rate of nucleotide substitution. For the plastid 16S rRNA genes, Tajima's tests (Tajima, 1993) showed that the substitution rates inside the *Prototheca* clade are heterogeneous and are significantly different from the photosynthetic clade rates (data not shown). These differences in evolutionary rate may explain why the relative position of *Prototheca wickerhamii* strain 1533 changes with the type of method used (Nedelcu, 2001; this study).

Phylogenetic affinities, and the presence of a plastid gene, suggest that *Helicosporidium* spp. are likely to possess a plastid genome similar to *Prototheca wickerhamii*. In this non-photosynthetic alga, the size of the chloroplast (leucoplast) genome has been estimated to be 54 100 bp, which is much smaller than the 150 kb chloroplast DNA of the photosynthetic relative *Chlorella vulgaris* (Knauf & Hachtel, 2002). This decrease in size is common in all secondarily non-photosynthetic green plants and algae (Hachtel, 1996) and has been explained by the loss of most of the plastid genes that were involved in photosynthesis. However, some plastid genes have been selectively retained, suggesting that they may encode essential protein products. In *Prototheca*, the functions of these proteins are not known (Knauf & Hachtel, 2002). In Apicomplexa, retained plastid ORFs have been associated with the apicoplast's hypothetical primary functions: fatty acid and isoprenoid biosynthesis (reviewed by Wilson, 2002).

Additionally, *Prototheca wickerhamii* is also known to possess a very characteristic mitochondrial genome. As reviewed by Nedelcu *et al.* (2000), the *Prototheca*-like mitochondrial genome represents an ancestral type among green algae that features (among other characteristics) a larger size (45–55 kb) and a more complex set of protein-coding genes than the derived, *Chlamydomonas*-like mitochondrial genome. Having shown that the Helicosporidia are non-photosynthetic green algae and close relatives of the genus *Prototheca*, our hypothesis is that *Helicosporidium* spp. possess *Prototheca wickerhamii*-like organelles and organelle genomes, i.e. a highly reduced plastid genome and an ancestral type of mitochondrial genome.

ACKNOWLEDGEMENTS

The authors would like to acknowledge the technical support of Genie White (USDA-ARS) as well as the ICBR Sequencing Facility at the University of Florida. This work is supported by a grant from the National Science Foundation (NSF, MCB-0131017). Florida Agricultural Experiment Station Journal Series no. R-09030.

REFERENCES

Avery, S. W. & Undeen, A. H. (1987). The isolation of microsporidia and other pathogens from concentrated ditch water. *J Am Mosq Control Assoc* **3**, 54–58.

Bhattacharya, D. & Medlin, L. (1998). Algal phylogeny and the origin of land plants. *Plant Physiol* **116**, 9–15.

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

Goolsby, J. A., Tipping, P. W., Center, T. D. & Driver, F. (2000). Evidence for a new *Cyrtobagous* species (Coleoptera: Curculionidae) on *Salvinia minima* Baker in Florida. *Southwest Entomol* **25**, 299–301.

Hachtel, W. (1996). DNA and gene expression in nonphotosynthetic plastids. In *Handbook of Photosynthesis*, pp. 349–355. Edited by M. Pessarakli. New York: Marcel Dekker.

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

Hembree, S. C. (1979). Preliminary reports of some mosquito pathogens from Thailand. *Mosq News* **39**, 575–582.

Huss, V. A. R., Frank, C., Hartmann, E. C., Hirmer, M., Kloboucek, A., Seidel, B. M., Wenzeler, P. & Kessler, E. (1999). Biochemical taxonomy and molecular phylogeny of the genus *Chlorella* sensu lato (Chlorophyta). *J Phycol* **35**, 587–598.

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.

Kellen, W. R. & Lindegren, J. E. (1973). New host records for *Helicosporidium parasiticum*. *J Invertebr Pathol* **22**, 296–297.

Knauf, U. & 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 Genomics* **267**, 492–497.

Kumar, S., Tamura, K., Jakobsen, I. B. & Nei, M. (2001). MEGA2: molecular evolutionary genetics analysis software. *Bioinformatics* **17**, 1244–1245.

Lang, N. J. (1963). Electron-microscopic demonstration of plastids in *Polytoma*. *J Protozool* **10**, 333–339.

Lindegren, J. E. & Hoffman, D. F. (1976). Ultrastructure of some developmental stages of *Helicosporidium* sp. in the navel orangeworm *Paramyelois transitella*. *J Invertebr Pathol* **27**, 105–113.

Morell, V. (1996). TreeBASE: the roots of phylogeny. *Science* **273**, 569.

Nedelcu, A. M. (2001). Complex pattern of plastid 16S rRNA gene evolution in nonphotosynthetic green algae. *J Mol Evol* **53**, 670–679.

Nedelcu, A. M., Lee, R. W., Lemieux, C., Gray, M. W. & Burger, G. (2000). The complete mitochondrial sequence of *Scenedesmus obliquus* reflects an intermediate stage in the evolution of the green algal mitochondrial genome. *Genome Res* **10**, 819–831.

Pekkarin, M. (1993). Bucephalid trematode sporocysts in brackish-water *Mytilus edulis*, new host of a *Helicosporidium* sp. (Protozoa: Helicosporida). *J Invertebr Pathol* **61**, 214–216.

Posada, D. & Crandall, K. A. (1998). MODELTEST: testing the model of DNA substitution. *Bioinformatics* **14**, 817–818.

Purrini, K. (1984). Light and electron microscope studies on *Helicosporidium* sp. parasitizing orbitid mites (Oribatei, Acarini) and collembola (Apterygota: Insecta) in forest soils. *J Invertebr Pathol* **44**, 18–27.

Sayre, R. M. & Clark, T. B. (1978). *Daphinia magna* (Cladocera: Chydoroidea): a new host of a *Helicosporidium* sp. (Protozoa: Helicosporidia). *J Invertebr Pathol* **31**, 260–261.

Seif, A. I. & Rifaat, M. M. (2001). Laboratory evaluation of a *Helicosporidium* sp. (Protozoa: Helicosporida) as an agent for the microbial control of mosquitoes. *J Egypt Soc Parasitol* **31**, 21–35.

Siu, C., Swift, H. & Chiang, K. (1976). Characterization of cytoplasmic and nuclear genomes in the colorless alga *Polytoma*. I. Ultrastructural analysis of organelles. *J Cell Biol* **69**, 352–370.

Swofford, D. L. (2000). *PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods)*, version 4. Sunderland, MA: Sinauer Associates.

Tajima, F. (1993). Simple methods for testing the molecular evolutionary clock hypothesis. *Genetics* **135**, 599–607.

Tartar, A., Boucias, D. G., Adams, B. J. & 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.

Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F. & Higgins, D. G. (1997). The CLUSTAL_X Windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res* **25**, 4876–4882.

Turmel, M., Otis, C. & 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 U S A* **96**, 10248–10253.

Weiser, J. (1970). *Helicosporidium parasiticum* Keilin infection in the caterpillar of a hepialid moth in Argentina. *J Protozool* **17**, 440–445.

Wilson, R. J. M. (2002). Progress with parasite plastids. *J Mol Biol* **319**, 257–274.