

Phylogenetic analysis of phagotrophic, phototrophic and osmotrophic euglenoids by using the nuclear 18S rDNA sequence

A. N. Müllner,¹ D. G. Angeler,² R. Samuel,¹ E. W. Linton³
and R. E. Triemer³

¹ Vienna University, Institute of Botany, Department of Higher Plant Systematics and Evolution, Rennweg 14, A-1030 Vienna, Austria

² Departamento de Ecología, Facultad de Ciencias, Edif. de Biología, Módulo Dpt Ecología/Despacho 201, Campus de Cantoblanco, Ctra. de Colmenar, km. 15, E-28049 Madrid, Spain.

³ Rutgers University, Department of Cell Biology and Neuroscience, Nelson Biological Laboratories, 604 Allison Road, Piscataway, NJ 08854–8082, USA

Author for correspondence: A. N. Müllner. Tel: +43 1 42 77 54163. Fax: +43 1 42 77 9541. e-mail: muellner@pflaphy.pph.univie.ac.at

Phylogenetic analyses of 35 strains including 25 previously published sequences and 10 which have been newly sequenced, representing two species of *Euglena*, five species of *Phacus* and three species of *Astasia*, were carried out using the SSU rDNA. Parsimony, distance and maximum-likelihood inferred phylogenies support (1) monophyly of the euglenoids, (2) kinetoplastids as the sister group, (3) the phagotrophic *Petalomonas cantuscyni* Cann et Pennick anchoring the base of the euglenoid lineage, (4) evolution of phototrophy within the euglenoids from a single event, (5) multiple origins of osmotrophic euglenoids and (6) polyphyly of the genera *Euglena* Ehrenberg and *Phacus* Dujardin. Analyses also indicate that *Lepocinclis* Perty, *Trachelomonas* Ehrenberg and *Astasia* Dujardin are polyphyletic. In addition, the results suggest that neither the Euglenales nor the Eutreptiales form a monophyletic lineage, thus questioning currently available classifications. Concerning the phagotrophic mode of nutrition, the data suggest that the feeding apparatus arose multiple times.

Keywords: euglenoids, Euglenozoa, molecular phylogeny, 18S rDNA

INTRODUCTION

Euglenoids are an ancient and diverse lineage of unicellular eukaryotic flagellates found predominantly in freshwater, but also in marine, soil and parasitic environments. Although they can be clearly distinguished from all other protists by a number of cytological (mitosis), morphological (pellicle structure) and physiological (paramylon) autapomorphies, the phylogenetic relationships within this group are still far from being fully resolved. For example, the genera *Euglena*, *Astasia* and *Khawkinia* are nearly identical morphologically and are distinguished by only a few diagnostic features. *Euglena* has a chloroplast, a stigma and a paraxonemal swelling associated with the eyespot. *Khawkinia* has a stigma and swelling, but no chloroplast. *Astasia* lacks a chloroplast, a stigma and a

swelling. However, there is substantial evidence that *Astasia* has lost these morphological features but still retains a major portion of the chloroplast genome (Gockel *et al.*, 1994a, b; Siemeister & Hachtel, 1990a, b). In 1928, Reichenow suggested that colourless and photosynthetic forms be united into a single group that was distinct from the phagotrophic forms (Reichenow, 1928). Unfortunately, this view was widely opposed (Hall & Jahn, 1929) at the time and never gained acceptance. How these taxa should be organized (separate genera, subgenera or species within a genus) is still in question.

Since the description of the genus *Euglena* by Ehrenberg (1838), over 150 species have been described. Various authors have grouped the species differently based on a variety of criteria. Pringsheim (1956) divided the genus into six subgenera, none of which held the name of the type. Subgenera were distinguished largely by the morphology and distribution of the chloroplasts and the pyrenoids. He created the subgenus, *Limpidae*, to accommodate what he believed were the colourless derivatives of *Euglena*, in particular the genera *Khawkinia* and *Astasia*.

This paper was presented at the XIIIth meeting of the International Society for Evolutionary Protistology in České Budějovice, Czech Republic, 31 July–4 August 2000.

The GenBank accession numbers for the sequences of the nuclear 18S rRNA genes determined in this study are AF283306, AF283307, AF283309–AF283311, AF283313–AF283316 and AY004245.

Leedale (1967) included five of Pringsheim's subgenera in his taxonomy, but maintained *Khawkinea* and *Astasia* as separate genera. Bourrelly (1970) created eight subgenera. Three of the subgenera were new and five were equivalent to those of Pringsheim. Bourrelly rejected Pringsheim's *Limpidae*, and recognized *Astasia* as a distinct genus containing the subgenus *Khawkinea*. The taxonomy of these three genera varied depending on the weight placed on the presence of the chloroplast, stigma and paraflagellar swelling. Previous studies using SSU rDNA suggest that both *Astasia* and *Khawkinea* clade with the genus *Euglena* and demonstrate that the subgenera recognized by Leedale (1967) or Bourrelly (1970) are not supported (Linton *et al.*, 1999, 2000).

Similarly, Dujardin (1841) established the genus *Phacus* to include those *Euglena*-like taxa which were rigid and slightly or markedly flattened. Unfortunately, the degree of flattening that occurs within the genus ranges from nearly leaf-like in some taxa (*Phacus longicauda*), to nearly rounded in other taxa (*Phacus pyrum*) making the use of this character questionable for defining generic limits. Similarly, many *Euglena* species (i.e. *Euglena acus*) are rigid and demonstrate bending only under stressful conditions.

Generic boundaries are also poorly defined in the genera *Trachelomonas* Ehrenberg (1833) and *Strombomonas* DeFlandre 1930. The genus *Strombomonas* was separated from *Trachelomonas* because the lorica gradually tapers to an apical pore and has no sharply defined neck or collar. However, as noted by Leedale (1967), many species of *Trachelomonas* also lack necks or collars making it difficult, if not impossible, to define generic boundaries using these characters.

The previous morphologically based examples demonstrate the need for a re-evaluation of these taxa as independent genera. This report expands previous phylogenetic studies using the SSU rDNA gene in an effort to better establish taxonomic boundaries among the photosynthetic and osmotrophic euglenoids.

METHODS

Euglenoid strains and culture conditions. The strains used in this study are listed in Table 1. Strains of *Euglena gracilis* (ASW 08025), *Euglena stellata* (ASW 08134), *Astasia comma* (ASW 08097), *Phacus aenigmaticus* (ASW 08012), *Phacus orbicularis* (ASW 08045), *Phacus parvulus* (ASW 08060), *Phacus pseudonordstedtii* (ASW 08010) and *Phacus acuminatus* (ASW 08004) were obtained from the Culture Collection of Algae at the Institute of Ecology and Conservation Biology (Kusel-Fetzmann & Schagerl, 1992), University of Vienna. The strains of *Astasia longa* (SAG 1204-17h) and *A. curvata* (SAG B 1204-5b) were obtained from the Culture Collection at the Institute of Plant Physiology, University of Göttingen (Schlösser, 1994). The strains were cultivated in biphasic media as circumscribed by Pringsheim (1954). Coloured taxa were grown at 20 °C using a 16:8 h light:dark regime (white light, 30 µM photons m⁻² s⁻¹), colourless taxa were grown in the dark at the same temperature. The sequences of all other taxa were obtained

from the GenBank database. The 18S rDNA gene sequences determined in the course of this study have been deposited in GenBank under the accession numbers given in Table 1.

Isolation of DNA and amplification. Total genomic DNA was isolated using the procedure described by Doyle & Doyle (1990) and Marin (1996), with slight modifications. PCR amplification reactions were carried out in a PTC-100 Programmable Thermal Controller (MJ Research) using the oligonucleotide primers EPA-23 > (5'-GTCATATGCTT-(CT)(GT)TTCAAGG(AG)CTAAGCC-3') and EPA-2286 < (5'-TCACCTACAGC(AT)ACCTTGTTACGAC-3'). The primers correspond to conserved regions in the published SSU rDNA sequence of *Euglena gracilis* (bases 23–50 for the forward primer EPA-23, respectively bases 2286–2272 for the reverse primer EPA-2286, GenBank accession no. M12677). A 100 µl reaction mix contained 73.5 µl ddH₂O, 10 µl 10 × reaction buffer with 15 mM MgCl₂, 8 µl 10 mM dNTPs, 2 µl of the primers each (20 pM), 2 µl template DNA (100–2000 ng µl⁻¹), 0.5 µl DNA Polymerase (BioTherm, 5U µl⁻¹), as well as 3 µl DMSO. Amplifications were carried out using the following main programme (with slight modifications for some strains): initial denaturation for 3 min at 95 °C, followed by one cycle of denaturation for 1.5 min at 95 °C, annealing for 1.5 min at 52 °C and extension for 3 min at 72 °C, followed by 25 cycles of denaturation for 1 min at 95 °C, annealing for 1.5 min at 62 °C and extension for 3 min at 72 °C. The amplification was completed by holding the reaction mixture for 10 min at 72 °C to allow complete extension of the PCR products. After amplification samples were sized by agarose electrophoresis against a 1 kb ladder (Gibco-BRL), and gel purified using a QIAquick gel extraction kit (Qiagen).

Sequencing. Three main sequencing primer pairs were used for cycle sequencing reactions: (1) EPA-23 > and EPA-468 > (5'-GGGTTTCGATCCGGAGA-3'; *Euglena gracilis* positions 468–484) together with EPA-827 < (5'-GAATTACCGCGGCTGCTG-3'; *Gyropaigne lefevri* positions 827–810), (2) EPA-679 > (5'-CGGTAATTCAGCTCC-3'; *Euglena gracilis* positions 679–694) together with EPA-1553 < (5'-ATTCCTTTAAGTTTC-3'; *Euglena gracilis* positions 1553–1539), (3) EPA-1294 > (5'-AGAGGTGAAATTCT-3'; *Euglena gracilis* positions 1294–1307) and EPA-1539 > (5'-GAACTTAAAGGAATTG-3'; *Euglena gracilis* positions 1539–1555) together with EPA-2286 <. A 10 µl cycle sequencing reaction mix contained 7 µl purified template DNA (100–400 ng µl⁻¹), 2 µl BigDye Terminator RR Mix (PE Biosystems) and 1 µl 10 pM primer. The following programme was used: denaturation for 10 s at 96 °C, annealing for 5 s at 50 °C, extension for 4 min at 60 °C (25 cycles). The purified DNA templates were sequenced in an ABI PRISM 377 DNA Sequencer (Perkin Elmer) using dye terminator labelling, as described by the manufacturer.

Sequence alignment. For editing the sequences the software programs Auto Assembler version 1.4.0 (Perkin Elmer) and DNA Strider version 1.2 (Christian Marck, Commissariat à l'Énergie Atomique) were used. The sequences were first roughly aligned by using CLUSTAL v (Higgins *et al.*, 1991), then corrected manually by using secondary structure of *Euglena gracilis* SSU rRNA as a guide for alignment. A total of 1036 nucleotides was included for the phylogenetic analyses under maximum-parsimony (MP), distance and maximum-likelihood (ML) criteria. Gaps were coded as missing data.

Phylogenetic analysis. MP, distance and ML analyses were

Table 1. Strains used in this study

Abbreviations: ASW, Algenkultur-Sammlung an der Universität Wien, Vienna, Austria; ATCC, American Type Culture Collection, Manassas, VA, USA; CBS, Centraalbureau voor Schimmelcultures, Baarn, The Netherlands; CCAP, Culture Collection of Algae and Protozoa, Cambridge, Ambleside and Oban, UK; CCMP, Provasoli-Guillard National Center for Culture of Marine Phytoplankton, McKown Point, West Boothbay Harbor, USA; SAG, Sammlung von Algenkulturen at the University of Göttingen, Göttingen, Germany; UTEX, Culture Collection of Algae, Austin, TX, USA.

Taxon	Strain designation	GenBank no.	Length (nt) complete sequence	A	C	G	T	G + C (mol%)	Length (nt) analysed dataset	A	C	G	T	G + C (mol%)
<i>Astasia comma</i> Pringsh.	ASW 08097	AF283307	3125	20	25	29	26	54	1036	24	23	30	23	53
<i>Astasia curvata</i> Klebs	SAG B 1204-5b	AY004245	2508	23	23	28	26	51	1035	25	22	30	23	52
<i>Astasia longa</i> Pringsh.	SAG 1204-17h	AF283306	2309	20	29	31	20	60	1035	24	23	30	23	53
<i>Bodo caudatus</i> Dujardin	Isolated by S. L. Hajduk	X53910	1921	25	23	27	25	50	975	27	21	29	23	50
<i>Crithidia fasciculata</i> Leger	Isolated by G. E. Kidder	Y00055	2063	25	23	27	25	50	975	27	22	28	23	50
<i>Distigma curvata</i> Pringsh.	SAG B 1216-1b	AF099081	3373	28	19	25	28	44	1035	27	22	28	23	50
<i>Distigma proteus</i> Ehrenb.	SAG B 1204-26a	AF106036	3724	23	23	26	28	49	1035	25	22	28	25	50
<i>Euglena acus</i> Ehrenb.	UTEX 1316	AF152104	2223	20	27	30	23	57	1035	25	23	29	23	52
<i>Euglena agilis</i> Carter	UTEX 1605	AF115279	2296	22	26	30	22	56	1035	24	23	30	23	53
<i>Euglena anabaena</i> Mainx	UTEX 373	AF242548	2320	22	26	29	23	55	1036	24	24	29	23	53
<i>Euglena gracilis</i> Klebs	ASW 08025	AF283309	2241	21	28	30	21	58	1035	24	23	30	23	53
<i>Euglena spirogyra</i> Ehrenb.	UTEX 1307	AF150935	2256	20	27	30	23	57	1035	24	24	30	22	54
<i>Euglena stellata</i> Mainx	UTEX 372	AF150936	2299	20	28	32	20	60	1035	25	24	30	21	54
<i>Euglena stellata</i> Mainx	ASW 08134	AF283310	2179	23	27	29	21	56	1035	25	23	30	22	53
<i>Euglena viridis</i> Ehrenb.	UTEX 85	AF112872	2265	23	27	29	21	56	1035	24	24	30	22	54
<i>Eutreptiella</i> sp. Da Cunha	CCMP 389	AF112875	2156	21	26	30	23	56	1035	25	23	28	24	51
<i>Gyropaigne lefevrei</i> Bourelly <i>et</i> Georges	CCAP 1233/1	AF110418	2904	21	26	30	23	56	1034	23	24	30	23	54
<i>Khawkinea quartana</i> (Moroff) Jahn <i>et</i> McKibben	ATCC 30895	U84732	2566	21	26	29	24	55	1035	25	23	30	22	53
<i>Lepocinclis ovata</i> Playfair	SAG B 1244-5	AF061338	2433	22	28	29	21	57	1036	24	25	29	22	54
<i>Lepocinclis ovum</i> (Ehrenb.) Lemm.	SAG B 1244-8	AF110419	2361	20	27	31	22	58	1035	24	24	30	22	54
<i>Petalomonas cantuscyni</i> Cann <i>et</i> Pennick	CCAP 1259/1	U84731	2028	24	23	29	24	52	1035	26	21	28	25	49
<i>Phacus acuminatus</i> Stokes	ASW 08004	AF283311	2114	22	26	30	22	56	1035	25	24	29	22	53
<i>Phacus aenigmaticus</i> Drez.	ASW 08012	AF283313	2182	22	27	30	21	57	1034	26	24	28	22	52
<i>Phacus orbicularis</i> Hübn.	ASW 08045	AF283315	2193	20	28	31	21	59	1034	24	24	30	22	54
<i>Phacus oscillans</i> Klebs	UTEX 1285	AF181968	2244	23	25	28	24	53	1035	25	23	29	23	52
<i>Phacus parvulus</i> Klebs	ASW 08060	AF283314	2105	22	25	29	24	54	1035	25	23	29	23	52
<i>Phacus pseudonordstedtii</i> Pochm.	ASW 08010	AF283316	2238	22	28	29	21	57	1035	24	24	29	23	53
<i>Phacus pusillus</i> Lemm.	UTEX 1282	AF190815	2244	22	25	29	24	54	1035	25	23	29	23	52
<i>Phacus pyrum</i> (E.) Stein	UTEX 2357	AF112874	2309	23	28	28	21	56	1034	24	24	29	23	53
<i>Phacus similis</i> Christen	SAG 58.81	AF119118	2234	23	24	29	24	53	1035	24	23	30	23	53
<i>Phacus splendens</i> Pochm.	UTEX 1284	AF190814	2263	22	27	30	21	57	1035	25	25	28	22	53
<i>Pseudoperamea trichophorum</i> Christen	CBS P7-13-1838	U84733	2145	26	21	27	26	48	1035	27	21	27	25	48
<i>Strombomonas</i> sp. Deflandre	Isolate from Korea	AF096994	2167	23	26	29	22	55	1035	24	24	30	22	54
<i>Trachelomonas hispida</i> (Perty) Stein	UTEX 1326	AF090377	1960	24	24	28	24	52	975	26	23	28	23	51
<i>Trachelomonas volvocina</i> Ehrenb.	Isolate from Korea	AF096995	2001	22	25	30	23	55	975	24	24	30	22	54

performed on the aligned data using PAUP*4.0b4a (Swofford, 2000) on a PowerMac G4. Each included nucleotide position was treated as an independent, unordered, multistate character of equal weight. The among-site rate variation was treated as equal rates for all sites. The transitions/transversions ratio was determined using MacClade version 3.01 (Maddison & Maddison, 1992). A parsimony heuristic search was performed using addition sequence set at random, with 1000 repetitions, ACCTRAN character state optimization, tree bisection-reconnection (TBR) branch swapping, and MULTREES on. The maximum-likelihood analyses (heuristic search with random addition and ten repetitions, TBR branch swapping, starting branch length using Rogers-Swofford method, and MULTREES on) were run using the HKY85 (Hasegawa-Kishino-Yano model; Hasegawa *et al.*, 1985). Distance trees were generated using different models provided in PAUP*4.0b4a: Kimura two-parameter (Kimura, 1980), Kimura three-parameter (Kimura, 1981), HKY85, Jukes & Cantor (Jukes & Cantor, 1969), F81 (Felsenstein, 1981) and uncorrected *p*. The left-skewness test for determination of phylogenetic signal in the dataset was performed with 1000000 randomly generated trees under the parsimony criterion (Hillis & Huelsenbeck,

1992). The χ^2 -test of PAUP*4.0b4a (Swofford, 2000) was performed to gain information about the homogeneity of the nucleotide distribution. The robustness and stability of parsimony and distance trees were estimated using bootstrapping with 1000 and 10000 replicates, respectively, and 10 replicates in the case of ML analyses. Kinetoplastids (*Crithidia fasciculata*, *Bodo caudatus*) were included because previous morphological and molecular studies have shown them to be a sister group to the euglenoids (Farmer, 1988; Simpson, 1997; Montegut-Felkner & Triemer, 1997; Linton *et al.*, 1999; Maslov *et al.*, 1999; Preisfeld *et al.*, 2000). They were used as outgroup organisms.

RESULTS

The base composition in the dataset was balanced with a mean G + C content of 55% for the aligned complete sequences and 52% for the regions used in the analyses. The left-skew ($g_1 = -1.02$, $p = 0.01$) of the random trees indicated that the data matrix contained phylogenetic signal (number of steps of the shortest random tree 3121.6, for the second shortest 3162.2).

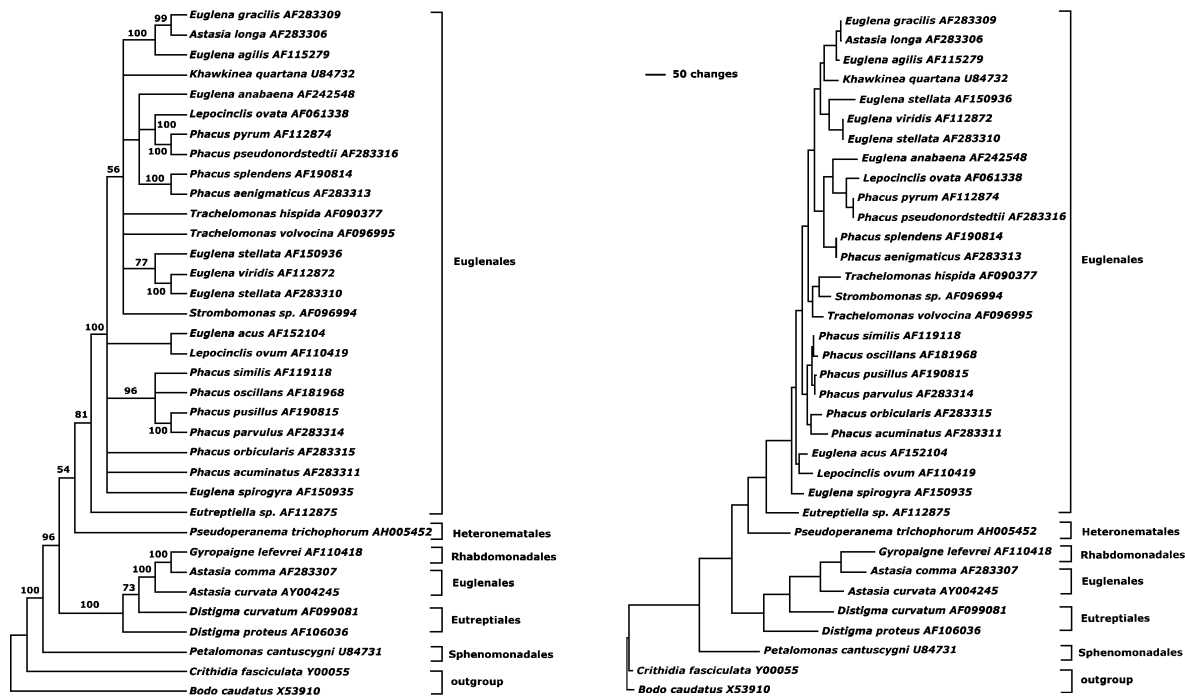


Fig. 1. (left) Strict consensus tree of the 10 most parsimonious trees obtained from the maximum-parsimony analysis of the SSU rDNA sequence dataset of 35 euglenoid strains, based on 506 parsimony-informative characters. Euglenoid orders after Leedale (1967). Numbers indicate bootstrap values of 1000 replicates (PAUP*4.0b4a). (right) One out of the two most parsimonious trees obtained from the maximum-parsimony analysis (PAUP*4.0b4a) showing a monophyletic origin of the *Trachelomonas/Strombomonas* clade.

Table 1 shows the distribution of nucleotides as well as the G + C content in 18S complete sequences and those conserved regions used in the analyses. After exclusion of highly variable sites to minimize the risk of a phylogenetic tree construction based on nucleotide distribution (Saitou & Imanishi, 1989; Hasegawa & Hashimoto, 1993; Steel *et al.*, 1993), the distribution pattern resulted in a χ^2 -test with $p = 0.99999$, allowing for phylogenetic reconstruction.

Inferred phylogenies resulting from maximum-parsimony, distance and maximum-likelihood analyses (PAUP4.0b4a; Swofford, 2000) are shown in Figs 1–3. All trees generated in this study supported the monophyletic origin of the euglenoids, with the phagotrophic *Petalomonas cantuscygni* at the base of the euglenoid lineage, the monophyly of the green euglenoids, as well as the polyphyly of the phototrophic genera *Euglena*, *Phacus* and *Lepocinclis*. *Eutreptiella* sp., representative of the Eutreptiales *sensu* Leedale (1967), was clearly separated from the Euglenales by all methods with a bootstrap value of 100.

MP analyses

Parsimony analysis of the dataset generated 10 most parsimonious trees with a length of 2240 steps based on 506 parsimony-informative characters. Fig. 1 (left) shows the strict consensus of the trees and the bootstrap support, Fig. 1 (right) shows one of the 10

most parsimonious trees obtained (one out of the two trees showing a monophyletic origin of the *Trachelomonas/Strombomonas* clade). The strict consensus tree (Fig. 1, left) does not provide a clear answer to the monophyletic or polyphyletic origin of the lorica-bearing genera *Trachelomonas* and *Strombomonas*. The relationships between most of the clades within the Euglenales in the upper part of the tree remain unresolved.

ML analyses

Maximum-likelihood (Fig. 2) showed no striking differences in comparison with the MP analyses concerning the clades resolved in the strict consensus tree. The two *Trachelomonas* species as well as the *Strombomonas* sp. do not clade together suggesting that *Trachelomonas* is not monophyletic.

Distance analyses

Distance analyses following different models (Kimura two-parameter, Kimura three-parameter, HKY85, Jukes & Cantor, F81 and uncorrected p) each resulted in one single tree, with the same topology in all different analyses (Fig. 3 shows the HKY85 tree). The overall topology in the different distance trees is in agreement with the ML tree, with few exceptions. Differences concern the position of *Trachelomonas*

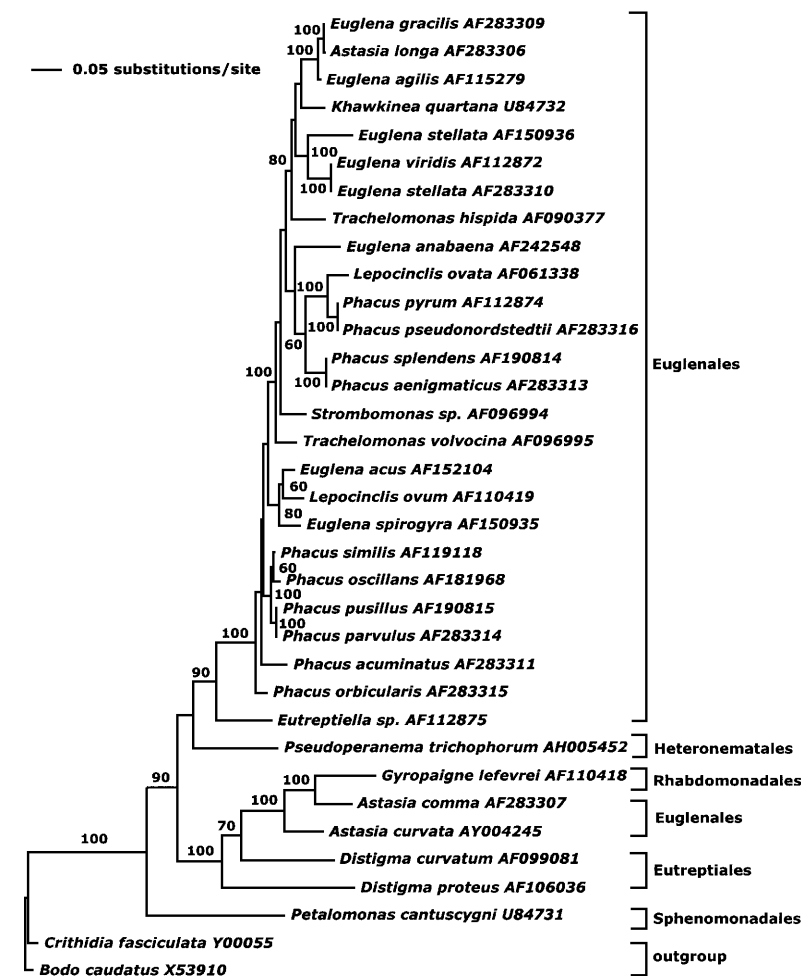


Fig. 2. Tree obtained from the maximum-likelihood analysis (PAUP*4.0b4a) following HKY85 (Hasegawa–Kishino–Yano model; Hasegawa *et al.*, 1985). Numbers indicate bootstrap values of 10 replicates.

volvocina and *Strombomonas* sp. as well as the positions of the *Phacus acuminatus* strain and *Phacus orbicularis*. In the distance tree, *Phacus acuminatus* and *Phacus orbicularis* are part of a separate clade built together with *Phacus similis*, *Phacus pusillus*, *Phacus parvulus* and *Phacus oscillans*; in the ML tree, *Phacus acuminatus* and *Phacus orbicularis* are sisters to all of the Euglenales. *Trachelomonas volvocina* and *Strombomonas* sp. either clade together (distance tree) or have separate origins (ML tree).

DISCUSSION

Implications for euglenoid classification

The monophyly of the euglenoids and the placement of the phagotrophic taxon *Petalomonas cantuscygni* at the base of the euglenoid lineage is in agreement with the results of Linton *et al.* (1999, 2000) and Preisfeld *et al.* (2000). Previous morphological (Farmer, 1988) and molecular studies (Montegut-Felkner & Triemer, 1997; Linton *et al.*, 1999, 2000; Preisfeld *et al.*, 2000) suggested that *Petalomonas cantuscygni* may be most similar to the common euglenoid ancestor. As shown here and in the previous molecular studies, *Petalomonas cantuscygni* anchors the euglenoid stem of the

tree. In agreement with our analyses, Preisfeld *et al.* (2000) have *Pseudoperanema* arising after the clade containing *Distigma* and *Gyropaigne*. This position is consistent with two possibilities: (1) the feeding apparatus arose multiple times or (2) the feeding apparatus arose once and was lost in the *Distigma* lineage. However, since *Petalomonas cantuscygni* and *Pseudoperanema trichophorum* are the only phagotrophic species that have been included in phylogenetic analyses so far, further taxon sampling is required to resolve the relationships at the base of the euglenoid lineage.

The positioning of taxa within the *Distigma*/*Astasia*/*Gyropaigne* clade and the clade itself within the tree also seems to have consequences for supra-generic classifications. The genus *Gyropaigne*, currently assigned to the Rhabdomonadales (Leedale, 1967; Farmer, 1988), appears as sister to the Euglenales and Eutreptiales (*sensu* Leedale 1967). Members of the Euglenales and Eutreptiales are divided among two separate clades within our trees, indicating that neither is a monophyletic group, thus bringing into question the currently available classification systems. Concerning the taxa in the basal clades of the tree, it seems also worth considering that both *Pseudoperanema* and *Distigma* have two flagella,

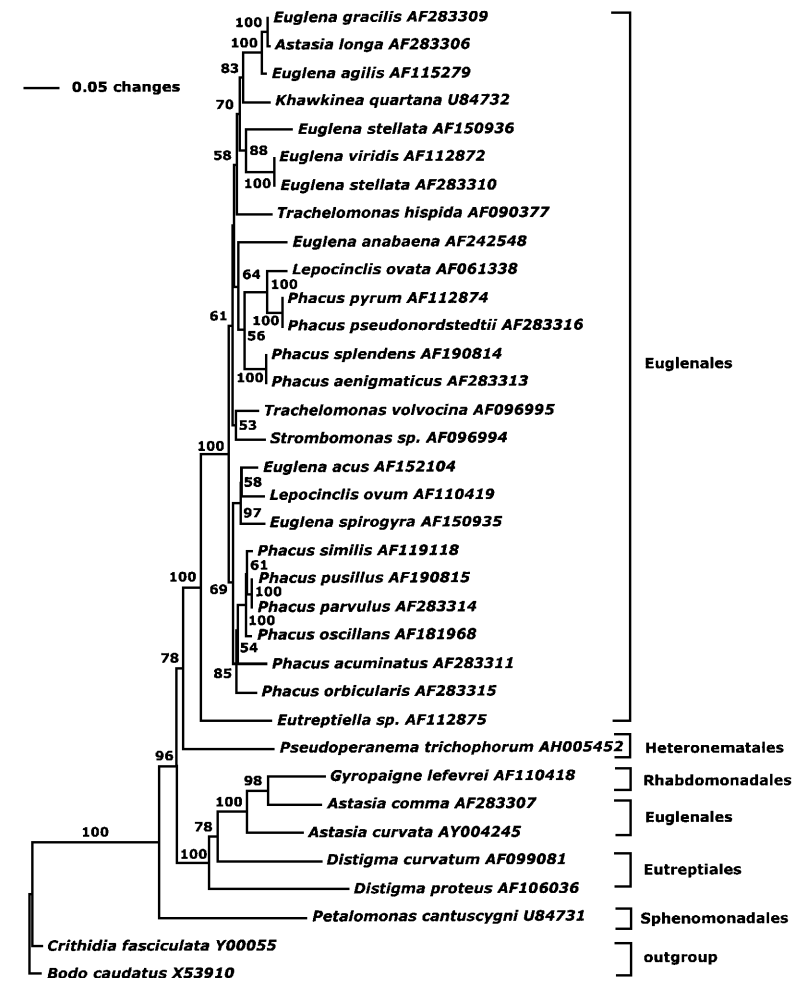


Fig. 3. Distance tree inferred from neighbour-joining analysis (PAUP*4.0b4a) following the HKY85 model. The same topology was achieved by Kimura two-parameter (Kimura, 1980), Kimura three-parameter (Kimura, 1981), Jukes & Cantor (Jukes & Cantor, 1969), F81 (Felsenstein, 1981) and uncorrected *p* analyses. Numbers indicate bootstrap values of 10000 replicates.

are very metabolic and have similar pellicle organization. In addition, the number of emergent flagella drops from two in *Distigma* to one in *Astasia* within the clade, a pattern which occurs again moving from *Pseudoperanema* and *Eutreptiella* to the rest of the Euglenales. However, from a practical point of view, systematic re-evaluations should not be carried out until more morphological, physiological and molecular data have accumulated, combined with the necessity of further taxon sampling.

Astasia. The position of the long-branched clade containing the osmotrophic taxa *Distigma curvatum*, *D. proteus*, *Astasia curvata*, *A. comma* and *Gyropaigne lefevrei* also has other interesting implications. Providing that the position of these taxa reflects true evolutionary events and comparing their position in the tree to the positions of the osmotrophs *Khawkinea quartana* and *Astasia longa*, the patterns suggest that osmotrophy arose multiple times: at least once out of the phagotrophic lineage, and twice within the photosynthetic lineage. The latter hypothesis is supported by molecular data where the presence of a plastid genome in *Astasia longa* provides evidence that the chloroplast was once present (Siemeister & Hachtel, 1990b; Gockel

et al., 1994b). In contrast to *A. longa*, no chloroplast genome has been detected in the *Astasia* species basal in the tree (*A. comma*, *A. curvata*), the *Distigma* species (*D. curvatum*, *D. proteus*) and *Gyropaigne lefevrei* so far. Christen (1963) recognizes three subgenera in *Astasia*, based on the shape of the anterior end of the cell and the nutritional state, respectively. The subgenus *Euglenoidea* (*A. longa*) comprises those species with a subapical position of the canal opening (resembling *Euglena*), the subgenus *Euastasia* (*A. curvata*, *A. comma*) comprises those with a transversally cut anterior end and an apical canal opening. The subgenus 'Phytophaga' comprises the phagotrophic species in *Astasia* with a subapical canal opening, which, like other phagotrophic euglenoids, are capable of ingesting solid prey. In our analyses the species belonging to different subgenera were separated, with the subgenus *Euastasia* emerging within the basal osmotrophic clade. Angeler (1999), by investigating different species of *Distigma* by means of transmission electron microscopy, found that some representatives, especially those belonging to the *D. proteus* group, contained intracellular bacteria. This may lend support for the capability of phagotrophic nutrition in a *D. proteus*-like ancestor.

The possibility that osmotrophy arose more than once may not be the most parsimonious solution, however, it has been shown in several cases that evolution does not necessarily follow the principle of parsimony (e.g. Gastony, 1986; Lamboy, 1994). Although ultra-structural data are scarce at the moment, the inconstancy of the osmotrophic nutrition mode could be reflected in pellicle architecture, at least in the non-rigid species. While the distigmids were shown to possess a constant number of pellicle strips throughout the entire cell length (Angeler *et al.*, 1999; Angeler, 2000b), members of the genus *Khawkinia* were shown to reduce the strip number at the anterior and posterior pole (Angeler, 2000a, b), a phenomenon which also occurs in phototrophic species (summarized in Leander & Farmer, 2000).

Considering generic classifications, the coloured genera *Euglena*, *Phacus* and *Lepocinclis* appeared to be polyphyletic in all trees generated by the different methods and models, thus confirming that the currently available classification systems (e.g. Leedale, 1967) are far from reflecting natural relationships. This seems to be true also for the osmotrophic genera. A case in point is the placement of the respective *Astasia* representatives in different clades, thus rejecting the idea of Pringsheim & Hovasse (1948), Pringsheim & Pringsheim (1952), Proviasoli *et al.* (1948), continued by Mego (1968) and Schiff & Eppstein (1968), that the genus *Astasia* as a whole evolved from an *Euglena*-like ancestor. The colourless genus *Astasia* seems to be polyphyletic, which confirms the earlier assumptions by Christen (1963) based on light microscopical investigations. Nevertheless, based on the results of this study, we feel it is useful to provide a preliminary evaluation of some of the currently available classifications for these genera.

Euglena. Seven different species of the photosynthetic genus *Euglena* used in these analyses have been organized differently in the three classification systems of Gojdics (1953), Pringsheim (1956), and Zakryś (1986) and Batko & Zakryś (1995). These authors' classifications were based mainly on morphological (chloroplast) features, therefore comparing their systems to the clading pattern of the genus *Euglena* based on molecular (SSU rDNA) data will enable an evaluation of their classification systems. *Euglena viridis* and *Euglena stellata* were assigned to the subgenus *Euglena* (Zakryś, 1986; Batko & Zakryś, 1995), 'Group A' (Gojdics, 1953) and group Radiatae (Pringsheim, 1956). These taxa form a monophyletic clade and thus appear to be a phylogenetically defensible grouping. *Euglena acus* and *Euglena spirogyra* belong to the subgenus *Discoglana* (Zakryś, 1986; Batko & Zakryś, 1995), 'Group D' (Gojdics, 1953) and group Rigidae (Pringsheim, 1956). These taxa also form a monophyletic clade and appear to be a well-supported grouping that includes *Lepocinclis ovum*. The chloroplast structure of *L. ovum* is the same as *Euglena acus* and *Euglena spirogyra* so its inclusion is supported by morphology. *Euglena gracilis*, *Euglena*

agilis and *Euglena anabaena* belong to the subgenus *Calliglana* (Zakryś, 1986; Batko & Zakryś, 1995), 'Group F' (Gojdics, 1953) group Catilliferae (Pringsheim, 1956). These are not found within a monophyletic clade but are dispersed over the upper part of the tree.

Milanowski *et al.* (2001), investigating the chloroplast small-subunit rRNA genes of 18 euglenoid species, ten belonging to the genus *Euglena*, also could not confirm the validity of the subgenera *Calliglana*, *Discoglana* and *Euglena*. With the exception of *Euglena acus*, this study comprised all *Euglena* taxa included in our study, and more. In this study only the taxa of the subgenus *Discoglana* formed a separate clade, which also contained *Lepocinclis fusiformis*. The subgenera *Euglena* and *Calliglana* appeared to be intermixed with each other as well as representatives of *Strombomonas*, *Trachelomonas* and *Colacium*.

Phacus. Ten different species of the coloured genus *Phacus* were included in this study. Pochmann (1942) in Huber-Pestalozzi (1955) assigns the taxa *Phacus similis*, *Phacus oscillans*, *Phacus pusillus*, *Phacus parvulus*, *Phacus orbicularis* and *Phacus acuminatus* to the subgenus *Chlorophacus*, section *Proterophacus* (periplast with fine strips, cells flattened). He assigns *Phacus pyrum*, *Phacus pseudonordstedtii*, *Phacus splendens* and *Phacus aenigmaticus* to section *Pleuraspis* (periplast with coarse strips, cells 'Lepocinclis-like', i.e. rounded). Comparing the branching pattern of the *Phacus* species in our trees with Pochmann's classification (which is mainly based on pellicle morphology) the two sections are indeed well separated from each other. The *Pleuraspis*-clade includes one taxon that does not belong to the genus *Phacus*, *Lepocinclis ovata*, which appears within this clade in all trees. In six out of the ten MP trees obtained, another taxon, *Euglena anabaena*, is part of the *Pleuraspis*-clade. Concerning the section *Proterophacus*, *Phacus similis*, *Phacus oscillans*, *Phacus pusillus* and *Phacus parvulus* form a stable clade in all trees. The remaining two species either appear as a polytomy (MP consensus tree), as separate divergences (ML tree), or as members of the *Proterophacus* clade (distance tree).

Trachelomonas, Strombomonas. The monophyly or polyphyly of the coloured, lorica-bearing genera *Trachelomonas* and *Strombomonas* could not be confirmed or rejected. Whereas ML and distance methods yielded a polyphyly of the lorica-bearing taxa, multiple maximum-parsimony trees gave conflicting results. In two out of the ten parsimonious trees obtained, *Trachelomonas* and *Strombomonas* appeared as a monophyletic clade, whereas in the others they were separated into different clades. In addition, it should be mentioned that the parsimony consensus trees as well as the bootstrap trees of all different methods used always resulted in a polytomy of the lorica-bearing taxa. Nevertheless our preliminary results suggest a polyphyletic origin of the genus *Trachelomonas*. More

taxa of these two genera should be included in further investigations to get a clear picture of their monophyletic or polyphyletic origin.

ACKNOWLEDGEMENTS

Financial support for this study was provided by the Oesterreichische Nationalbank (grant no. 7485), the Theodor-Körner-Fonds and a research grant from the University of Vienna. We want to thank T. F. Stuessy for logistic support and M. Schagerl and E.-L. Kusel-Fetzmann for providing euglenoid strains and help in identification.

REFERENCES

- Angeler, D. G. (1999).** *Distigma proteus* var. *longicauda* var. nov. – a new colourless euglenoid described from cultures. *Arch Hydrobiol Suppl* **127**, *Algolog Stud* **92**, 19–33.
- Angeler, D. G. (2000a).** A light microscopical and ultrastructural investigation and validation of *Khawkinia pertyi* comb. nova (Euglenophyta). *Arch Hydrobiol Suppl* **131**, *Algolog Stud* **96**, 89–103.
- Angeler, D. G. (2000b).** Taxonomy and morphology of *Distigma elegans* and *Khawkinia fritschii*, rare euglenoids rediscovered in the Iberian Peninsula. *Nova Hedwigia* **70**, 397–408.
- Angeler, D. G., Müllner, A. N. & Schagerl, M. (1999).** Comparative ultrastructure of the cytoskeleton and nucleus of *Distigma* (Euglenozoa). *Eur J Protistol* **35**, 309–318.
- Batko, A. & Zakryś, B. (1995).** Numerical proof of the subgeneric classification of *Euglena*. *Arch Hydrobiol Suppl* **111**, *Algolog Stud* **79**, 1–18.
- Bourrelly, P. (1970).** Troisième partie – Euglénophytes. In *Les Algues d'Eau Douce*, pp. 115–184. Paris: N. Boubée & Cie.
- Christen, H. R. (1963).** Zur Taxonomie der farblosen Eugleninen. *Nova Hedwigia* **4**, 437–464.
- DeFlandre, G. (1930).** *Strombomonas*, nouveau genre d'Euglénacées (*Trachelomonas* Ehrb. pro parte). *Arch Protistenkd* **69**, 551–614.
- Doyle, J. J. & Doyle, J. L. (1990).** Isolation of plant DNA from fresh tissue. *Focus* **12**, 13–15.
- Dujardin, F. (1841).** *Histoire Naturelle des Zoophytes-Infusoires*. Paris: Roret.
- Ehrenberg, C. G. (1833).** Dritter Beitrag zur Erkenntnis grosser Organisation in der Richtung des kleinsten Raumes. *Physikalische Abhandlungen der Königlichen Akademie der Wissenschaften zu Berlin* 145–336.
- Ehrenberg, C. G. (1838).** *Die Infusionstierchen als vollkommene Organismen*. Leipzig: Verlag von Leopold Voss.
- Farmer, M. A. (1988).** *A re-evaluation of the taxonomy of the Euglenophyceae based on ultrastructural characteristics*. Dissertation, University of New Jersey, Rutgers.
- Felsenstein, J. (1981).** Evolutionary trees from DNA sequences: a maximum likelihood approach. *J Mol Evol* **17**, 368–376.
- Gastony, G. J. (1986).** Electrophoretic evidence for the origin of fern species by unreduced spores. *Am J Bot* **73**, 1563–1569.
- Gockel, G., Baier, S. & Hachtel, W. (1994a).** Plastid ribosomal protein genes from the nonphotosynthetic flagellate *Astasia longa*. *Plant Physiol* **105**, 1443–1444.
- Gockel, G., Hachtel, W., Baier, S., Fliss, C. & Hencke, M. (1994b).** Genes for components of the chloroplast translational apparatus are conserved in the reduced 73-kb plastid DNA of the nonphotosynthetic euglenoid flagellate *Astasia longa*. *Curr Genet* **26**, 256–262.
- Gojdics, M. (1953).** *The Genus Euglena*. Madison: University of Wisconsin Press.
- Hall, R. P. & Jahn, T. L. (1929).** On the comparative cytology of certain euglenoid flagellates and the systematic position of the families Euglenidae Stein and Astasiidae Butschli. *Trans Am Microscop Soc* **48**, 388–405.
- Hasegawa, M. & Hashimoto, K. (1993).** Ribosomal RNA trees misleading? *Nature* **361**, 6407–6423.
- 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.
- Higgins, D. G., Bleasby, A. J. & Fuchs, R. (1991).** CLUSTAL V: improved software for multiple sequence alignment. *CABIOS* **8**, 189–191.
- Hillis, D. M. & Huelsenbeck, J. P. (1992).** Signal, noise and reliability in molecular phylogenetic analyses. *J Hered* **83**, 189–195.
- Huber-Pestalozzi, P. (1955).** *Das Phytoplankton des Süßwassers*. IV. Euglenophyceen. Edited by A. Thienemann. Stuttgart: E. Schweizerbart'sche Verlagsbuchhandlung.
- Jukes, T. H. & Cantor, C. R. (1969).** Evolution of protein molecules. In *Mammalian Protein Metabolism*, vol. 3, pp. 21–132. Edited by H. N. Munro. New York: Academic Press.
- Kimura, M. (1980).** A simple model for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *J Mol Evol* **16**, 111–120.
- Kimura, M. (1981).** Estimation of evolutionary distances between homologous nucleotide sequences. *Proc Natl Acad Sci USA* **78**, 454–458.
- Kusel-Fetzmann, E. L. & Schagerl, M. (1992).** Verzeichnis der Sammlung von Algen-Kulturen an der Abteilung für Hydrobotanik am Institut für Pflanzenphysiologie der Universität Wien. *Phyton* **32**, 209–234.
- Lamboy, W. F. (1994).** The accuracy of the maximum parsimony method for phylogeny reconstruction with morphological characters. *Syst Bot* **19**, 489–505.
- Leander, B. S. & Farmer, M. A. (2000).** Comparative morphology of the euglenid pellicle. I. Patterns of strips and pores. *J Eukaryot Microbiol* **47**, 469–480.
- Leedale, G. F. (1967).** *Euglenoid Flagellates*. Englewood Cliffs, NJ: Prentice-Hall.
- Linton, E. W., Hittner, D., Lewandowski, C., Auld, T. & Triemer, R. E. (1999).** A molecular study of euglenoid phylogeny using small subunit rDNA. *J Eukaryot Microbiol* **46**, 217–223.
- Linton, E. W., Nudelman, M. A., Conforti, V. & Triemer, R. E. (2000).** A molecular analysis of the Euglenophytes using SSU rDNA. *J Phycol* **36**, 740–746.
- Maddison, W. P. & Maddison, D. R. (1992).** *MacClade Version 3 – Analysis of Phylogeny and Character Evolution*. Sunderland, MA: Sinauer Associates.
- Marin, B. (1996).** *Die frühe Evolution der Grünalgen: Vergleichende Untersuchungen ultrastruktureller Merkmale sowie phylogenetische Sequenzanalysen von kern- und plastidkodierten ribosomalen RNA-Sequenzen*. Dissertation, University of Cologne.
- Maslov, D. A., Yasuhira, S. & Simpson, L. (1999).** Phylogenetic affinities of *Diplonema* within the Euglenozoa as inferred from the SSU rRNA gene and partial COI protein sequences. *Protist* **150**, 33–42.

- Mego, J. L. (1968).** Inhibitors of the chloroplast system in *Euglena gracilis*. In *The Biology of Euglena*, vol. 2, pp. 351–381. Edited by D. E. Buetow. London: Academic Press.
- Milanowski, R., Zakryś, B. & Kwiatowski, J. (2001).** Phylogenetic analysis of chloroplast small-subunit rRNA genes of the genus *Euglena* Ehrenberg. *Int J Syst Evol Microbiol* **51**, 773–781.
- Montegut-Felkner, A. E. & Triemer, R. E. (1997).** Phylogenetic relationships of selected euglenoid genera based on morphological and molecular data. *J Phycol* **33**, 512–519.
- Pochmann, A. (1942).** Synopsis der Gattung *Phacus*. *Arch Protistenkd* **95**, 81–252.
- Preisfeld, A., Berger, S., Busse, I., Liller, S. & Ruppel, H. G. (2000).** Phylogenetic analyses of various euglenoid taxa (Euglenozoa) based on 18S rDNA sequence data. *J Phycol* **36**, 220–226.
- Pringsheim, E. G. (1954).** *Algenreinkulturen (Pure Cultures of Algae). Ihre Herstellung und Erhaltung*. Jena: Gustav Fischer.
- Pringsheim, E. G. (1956).** Contributions towards a monograph of the genus *Euglena*. *Nova Acta Leopold* **125**, 1–168.
- Pringsheim, E. G. & Hovasse, R. (1948).** The loss of chromatophores in *Euglena gracilis*. *New Phytol* **47**, 52–87.
- Pringsheim, E. G. & Pringsheim, O. (1952).** Experimental elimination of chromatophores and eye-spot in *Euglena gracilis*. *New Phytol* **51**, 65–76.
- Proviasoli, L., Hutner, S. H. & Schatz, A. (1948).** Streptomycin-induced chlorophyll-less races of *Euglena*. *Proc Soc Exp Biol Med* **69**, 279–282.
- Reichenow, E. (1928).** In *Doflein's Lehrbuch der Protozoenkunde*, Teil II, pp. 436–864.
- Saitou, N. & Imanishi, M. (1989).** Relative efficiencies of the Fitch-Margolis, maximum parsimony, maximum likelihood, minimum evolution and neighbor-joining methods of phylogenetic tree construction in obtaining the correct tree. *J Mol Biol Evol* **6**, 514–525.
- Schiff, J. A. & Eppstein, H. T. (1968).** The continuity of the chloroplast in *Euglena*. In *The Biology of Euglena*, vol. 2, pp. 285–333. Edited by D. E. Buetow. London: Academic Press.
- Schlösser, U. G. (1994).** SAG-Sammlung von Algenkulturen at the University of Göttingen. Catalogue of strains 1994. *Bot Acta* **107**, 111–186.
- Siemeister, G. & Hachtel, W. (1990a).** Organization and nucleotide composition of ribosomal RNA genes on a circular 73 kbp DNA from the colourless flagellate *Astasia longa*. *Curr Genet* **17**, 433–438.
- Siemeister, G. & Hachtel, W. (1990b).** Structure and expression of a gene encoding the large subunit of ribulose-1,5-bisphosphate carboxylase (rbcL) in the colourless euglenoid flagellate *Astasia longa*. *Plant Mol Biol* **14**, 825–833.
- Simpson, A. G. B. (1997).** The identity and composition of the Euglenozoa. *Arch Protistenkd* **148**, 318–328.
- Steel, M. A., Lockhart, P. J. & Penny, D. (1993).** Confidence in evolutionary trees from biological sequence data. *Nature* **364**, 440–442.
- Swofford, D. L. (2000).** PAUP*. Phylogenetic Analysis Using Parsimony (*and other methods). Version 4. Sunderland, MA: Sinauer Associates.
- Zakryś, B. (1986).** Contribution to the monograph of Polish members of the genus *Euglena* Ehrenberg 1830. *Nova Hedwigia* **42**, 494–540.