The Journal of
Eukaryotic Microbiology

Published by the International Society of Protistologists

Journal of Eukaryotic Microbiology ISSN 1066-5234

ORIGINAL ARTICLE

Phylogenetic Relationships and Morphological Character Evolution of Photosynthetic Euglenids (Excavata) Inferred from Taxon-rich Analyses of Five Genes

Anna Karnkowska[a,b], Matthew S. Bennett[c], Donovan Watza[c], Jong Im Kim[d], Bożena Zakryś[a] & Richard E. Triemer[c]

a Department of Plant Systematics and Geography, Faculty of Biology, University of Warsaw, Warsaw, Poland
b Department of Parasitology, Faculty of Science, Charles University in Prague, Prague, Czech Republic
c Department of Plant Biology, Michigan State University, East Lansing, Michigan, USA
d Department of Biology, Chungnam National University, Daejeon, Korea

Keywords
Algae; Euglenea; Euglenida; *hsp90*; LSU rDNA; molecular phylogeny; protists; *psbO*; SSU rDNA.

Correspondence
A. Karnkowska, Department of Plant Systematics and Geography, University of Warsaw, Al. Ujazdowskie 4, PL-00-478 Warsaw, Poland
Telephone number: +48 225-530-554;
FAX number: +48 225-530-540;
e-mail: ankarn@biol.uw.edu.pl

Received: 13 July 2014; revised 6 October 2014; accepted October 7, 2014.

doi:10.1111/jeu.12192

ABSTRACT

Photosynthetic euglenids acquired chloroplasts by secondary endosymbiosis, which resulted in changes to their mode of nutrition and affected the evolution of their morphological characters. Mapping morphological characters onto a reliable molecular tree could elucidate major trends of those changes. We analyzed nucleotide sequence data from regions of three nuclear-encoded genes (nSSU, nLSU, *hsp90*), one chloroplast-encoded gene (cpSSU) and one nuclear-encoded chloroplast gene (*psbO*) to estimate phylogenetic relationships among 59 photosynthetic euglenid species. Our results were consistent with previous works; most genera were monophyletic, except for the polyphyletic genus *Euglena*, and the paraphyletic genus *Phacus*. We also analyzed character evolution in photosynthetic euglenids using our phylogenetic tree and eight morphological traits commonly used for generic and species diagnoses, including: characters corresponding to well-defined clades, apomorphies like presence of lorica and mucilaginous stalks, and homoplastic characters like rigid cells and presence of large paramylon grains. This research indicated that pyrenoids were lost twice during the evolution of phototrophic euglenids, and that mucocysts, which only occur in the genus *Euglena*, evolved independently at least twice. In contrast, the evolution of cell shape and chloroplast morphology was difficult to elucidate, and could not be unambiguously reconstructed in our analyses.

PHOTOSYNTHETIC euglenids (Euglenea Bütschli 1884) are a group of protists in the supergroup Excavata, which acquired a chloroplast through secondary endosymbiosis (Maruyama et al. 2011; Turmel et al. 2009; Yamaguchi et al. 2012). Müller (1786) first described a small green flagellate as *Cercaria viridis* and subsequently Ehrenberg (1830) renamed *C. viridis* as *Euglena*. Since then, many researchers have dealt with the taxonomy of photosynthetic euglenids, and taxon identification was mainly based on cell morphology (see Triemer and Farmer 2007 for review). There are 13 genera of photosynthetic euglenids. The majority of these live in freshwater, but some do live in marine or brackish habitats. They exhibit a high degree of morphological diversity in cell plasticity (from very metabolic to rigid), shape (from cylindrical to oval, from straightened to twisted) and chloroplast morphology such as their location in a cell (from axial to parietal), their number and size (from a single large chloroplast to numerous small chloroplasts) and shape (from simple disks to very complicated netlike, convoluted, or stellate forms). Morphological diversity has resulted in thousands of described taxa (Algaebase—http://www.algaebase.org reports more than 3,000 validly published names), many of which are difficult to identify based on the original descriptions and figures. The main reason for describing so many taxa is the large variation in morphological plasticity, despite the limited

number of possible diagnostic features found in these unicellular organisms. This has led to a history of taxonomic duplications and re-descriptions, as well as the formulation of artificial classification schemes. Recently, the advent of molecular phylogenies has helped to resolve some of those taxonomic confusions (Bennett et al. 2014; Karnkowska-Ishikawa et al. 2010, 2011, 2012; Kosmala et al. 2005; Linton et al. 2010; Marin et al. 2003; Milanowski et al. 2006; Triemer et al. 2006; Zakryś et al. 2002).

Molecular studies of euglenids based on nuclear encoded SSU rDNA sequences began in 1997, when Montegut-Felkner and Triemer (1997) examined phylogenetic relationships among four euglenid taxa. Subsequently, authors added more taxa (Linton et al. 1999, 2000; Marin et al. 2003; Moreira et al. 2001; Müllner et al. 2001; Nudelman et al. 2003), and the following markers: nLSU rDNA sequences (Ciugulea et al. 2008; Linton et al. 2010), chloroplast ribosomal small subunit cpSSU rDNA (Linton et al. 2010; Milanowski et al. 2001, 2006; Zakryś et al. 2002) followed by the chloroplast ribosomal large subunit cpLSU rDNA (Kim and Shin 2008). With the exception of studies on the *rbcL* gene (Thompson et al. 1995) and the genes encoding *PAR1* and *PAR2* (paraxonemal rod proteins) (Talke and Preisfeld 2002), both of which had limited taxon sampling, photosynthetic euglenid phylogenies have not incorporated protein-coding genes. However, several phylogenies for estimating deep relationships among eukaryotes (Kim et al. 2006), Excavates (Simpson et al. 2006) and Euglenoza (Simpson et al. 2004) have been performed using protein-coding sequences (*hsp90*, *hsp70*, *EF-1*α, *EF-2*, α-tubulin, and β-tubulin). *hsp90* was also used for resolving relationships among phagotrophic euglenids (Breglia et al. 2007), which suggested its utility for phototrophic euglenid phylogeny. Our data complements existing data with new protein-coding sequences in an attempt to resolve photosynthetic euglenid relationships.

In this study, we selected five molecular markers—three sequences of rDNA (both nuclear SSU and LSU rDNA, as well as chloroplast SSU rDNA) and two nuclear protein-coding sequences (*hsp*90 and *psb*O). A total of 6,915 characters were assembled, and 140 new sequences were generated. We also carefully examined and selected a set of stable and characteristic cellular features of phototrophic euglenid lineages: cell shape, chloroplast morphology, cell plasticity, paramylon grain diversity, and the presence of mucus bodies, mucilaginous stalks, and loricas. Mapping morphological features on the phylogenetic tree enabled us to understand euglenid diversity and provided insights into evolutionary trends that led to this diversity.

MATERIALS AND METHODS

Taxon sampling, cultivation and identification

Table S1 lists the 68 taxa with sources and GenBank accession numbers used in this study. Alternative name(s) in culture collections or GenBank for the taxa used in this study, due to misidentification or taxonomic changes, are also given. All taxa were obtained from culture collections or were collected from small ponds in New Jersey (NJ) and Michigan (MI), USA. NJ and MI cultures used in this study, which are not available from public culture collections, will be made available upon request. The NJ and MI cultures were collected with a plankton net (mesh size: 20 μm), and individual euglenid cells were isolated by a Pasteur capillary pipette and inoculated into sterile culture media.

All strains were cultivated in a liquid soil-water medium, enriched by a small piece of garden pea (medium 3c, Schlösser 1994) and/or in modified AF-6 medium (Watanabe et al. 2000) in a growth chamber maintained at 17–22 °C and 16:8 h Light/Dark cycle, with ca. 27 μmol photons/m^2/s provided by cool white fluorescent tubes (Philips, Amsterdam, the Netherlands).

The identity of each strain was confirmed using a Zeiss Axioskop 2 Plus microscope (Carl Zeiss Inc., Hallbergmoos, Germany) or a Nikon Eclipse E-600 microscope (Nikon, Tokyo, Japan) both equipped with differential interference contrast.

DNA and RNA isolation, amplification, and sequencing

Total genomic DNA was isolated from cultures with the DNeasy Blood &Tissue Kit (Qiagen, Valencia, CA) using the animal tissues protocol as previously described (Brosnan et al. 2003; Zakryś et al. 2002). Total RNA was extracted using the Qiagen RNeasy mini kit (Qiagen) with Qiagen QIAshredder columns (Qiagen) and with the DNAse digestion step according to the manufacturer's instructions. Synthesis of cDNA was performed using Invitrogen SuperScript III Reverse Transcriptase (Invitrogen, Carlsbad, CA) according to the manufacturer's protocol. The cDNA sample synthesized with oligo-dT was used for PCR of the *hsp*90 and *psb*O genes to avoid problems associated with aligning introns. Amplification was performed using primers for conserved regions (nSSU and nLSU rDNA, Bennett and Triemer 2012; cpSSU rDNA, Linton et al. 2010; *hsp*90, 100XF & 968XR from Simpson et al. 2002 and Table S2; *psb*O, 1F & 2R from Takahashi et al. 2007 and Table S2). For amplification of nSSU, nLSU, and cpSSU rDNA, PCR programs described by Linton et al. (2010) were used. For amplification of *hsp*90 and *psb*O, the cycling conditions started with a denaturing step at 95 °C for 5 min, followed by 35 cycles of 30 s at 95 °C, 1 min of annealing at 40–65 °C (depending on the melting temperatures of the primers), and extension at 72 °C for 1 min (7 min at 72 °C for the final cycle only). The PCR products were sized on 1% agarose gels and then purified using the Qiagen MinElute Gel Extraction Kit (Qiagen) or QIAEXII Gel Extraction Kit (Qiagen) according to the manufacturer's protocol. All amplicons were sequenced using capillary-based sequencing. DNA/cDNA sequences were quality checked and assembled in the SeqMan program from the LASERGENE package (DnaStar, Madison, WI) or using the Contig Assembly Program (CAP2) in the Mac Genetic Data Environment program (MacGDE) available at http://macgde.bio.cmich.edu/.

Data analysis

To exclude pseudogenes from the protein-coding gene dataset, sequences obtained from cDNAs were translated into the corresponding amino acids to check for stop codons and individual gene trees were compared to ensure similar topologies, as suggested by Song et al. (2008). The alignment of sequences was obtained using MUSCLE (Edgar 2004), with default options. The sequences for the nSSU, nLSU, and cpSSU rDNAs were manually aligned according to secondary structure of *Euglena gracilis* Klebs as a guide (Kjer 1995; Wuyts et al. 2002) in MacGDE. Ambiguous sites in the alignment were excluded from analyses. All five gene sequences used in this study were derived mainly from the same culture (Table S1). However, when this was not possible due to culturing or sequencing difficulties, alternative sources were used. Taxon identities were confirmed by sequence comparisons of the first ~700 bases of the nSSU rDNA sequences. Individual gene trees were checked for inconsistencies in topologies, and a combined dataset (DNA sequences of five genes) was used in the final phylogenetic analysis. The alignment is available in TreeBASE (S12291).

The Bayesian Information Criterion (BIC) analysis using MODELTEST 3.7 (Posada and Crandall 1998) was used to determine the best model for each dataset (nSSU, nLSU, cpSSU, *hsp90,* and *psbO*), and a partitioned dataset was used with both the Maximum-likelihood (ML) and Bayesian analyses. The ML analysis was carried out using a general-time-reversible model incorporating invariable sites and a discrete gamma distribution (GTR+I+Γ) in RAxML version 7.5.4 (Stamatakis 2006). The Maximum-liklihood tree was determined based on 20 distinct starting trees and 1,000 bootstrap replicates—both utilizing random number seeds. The Bayesian analysis was performed using a mixed model in MrBayes version 3.2 (Ronquist et al. 2012). A gamma correction with eight categories and proportion of invariable sites was used, and two independent analyses were run with four Markov chains (default temperature parameter value). In each case, a total of 1,000,000 generations were calculated with trees sampled every 100 generations discarding the first 25% of trees. Convergence was confirmed via the sump command. A majority-rule consensus tree was created from the remaining 7,500 trees. For both analyses, trees were rooted, post analysis, using *Eutreptia viridis*, *Eutreptiella braarudii,* and *Eutreptiella pomquetensis* as outgroup taxa.

Character evolution

Morphological characteristics were selected based on the most common diagnostic features used in generic and species descriptions from the photosynthetic euglenid literature (Table 1). Ancestral character state analyses and transitions were reconstructed on the Bayesian phylogenies tree (Fig. 1) using ML and MP methods implemented in Mesquite v.2.75 (Maddison and Maddison 2011). ML optimizations were done using the Markov k-state one-parameter model (Lewis 2001). The best estimate of the character state at each node was determined using the likelihood ratio test implemented in Mesquite v.2.75. If the log likelihoods of two states differed by 2.0 or more units, the state with the lower likelihood was rejected, and the alternate state was considered the best estimate for that branch with strong statistical support (Pagel 1999). Results presented as character gains and losses of characters were mapped on a simplified phylogenetic tree, and phylogenetic lineages which shared all analyzed characters were collapsed (Fig. 2).

Table 1. List of eight morphological characters and 21 states used in the ancestral state reconstruction analysis

Characters	Morphological states
1. Cell shape	0: Fusiform (Fig. S1a)
	1: Cylindrical (Fig. S1b)
	2: Oval (Fig. S1c)
	3: Oblate (Fig. S1e)
	4: Ellipsoidal (Fig. S1d)
2. Metaboly	0: Metabolic (Fig. S2a)
	1: Slightly metabolic or rigid (Fig. S2b)
3. Large paramylon grains	0: Absent (Fig. S3a)
	1: Present (Fig. S3b–e)
4. Chloroplast morphology	0: Discoid (Fig. S4e)
	1: Lobed (Fig. S4a–d)
	2: Spherical (Fig. S4f)
	3: Stellate (Fig. S4 g)
5. Pyrenoids	0: Absent (Fig. S4f)
	1: Present (Fig. S4a–d)
6. Mucocysts	0: Absent (Fig. S5a)
	1: Present (Fig. S5b)
7. Lorica	0: Absent
	1: Present (Fig. S6)
8. Mucilaginous stalk	0: Absent
	1: Present (Fig. S7)

RESULTS

Molecular phylogenetic tree based on concatenated rDNA sequences and protein-coding sequences

We sampled sequence data from 68 strains belonging to 59 species across the photosynthetic euglenids, including data already available from GenBank (Table S1). Our novel data for this study represented ~41% of the total data set (140 sequences of a total of 340). The concatenated data set consisted of 6,915 characters with relative contributions of 1,290 characters from nSSU, 2,026 characters from nLSU, 1,330 characters from cpSSU, 1,885 characters from *hsp*90, and 384 characters from *psb*O.

The best-fit model of evolution for these data using the BIC criterion for model selection in ModelTest was the GTR+I+G (nSSU, nLSU and cpSSU rDNA), TIM+I+G (*hsp*90), and SYM+I+G (*psb*O). Bayesian analysis (Fig. 1) and maximum-likelihood analysis (not shown) recovered the same tree topology.

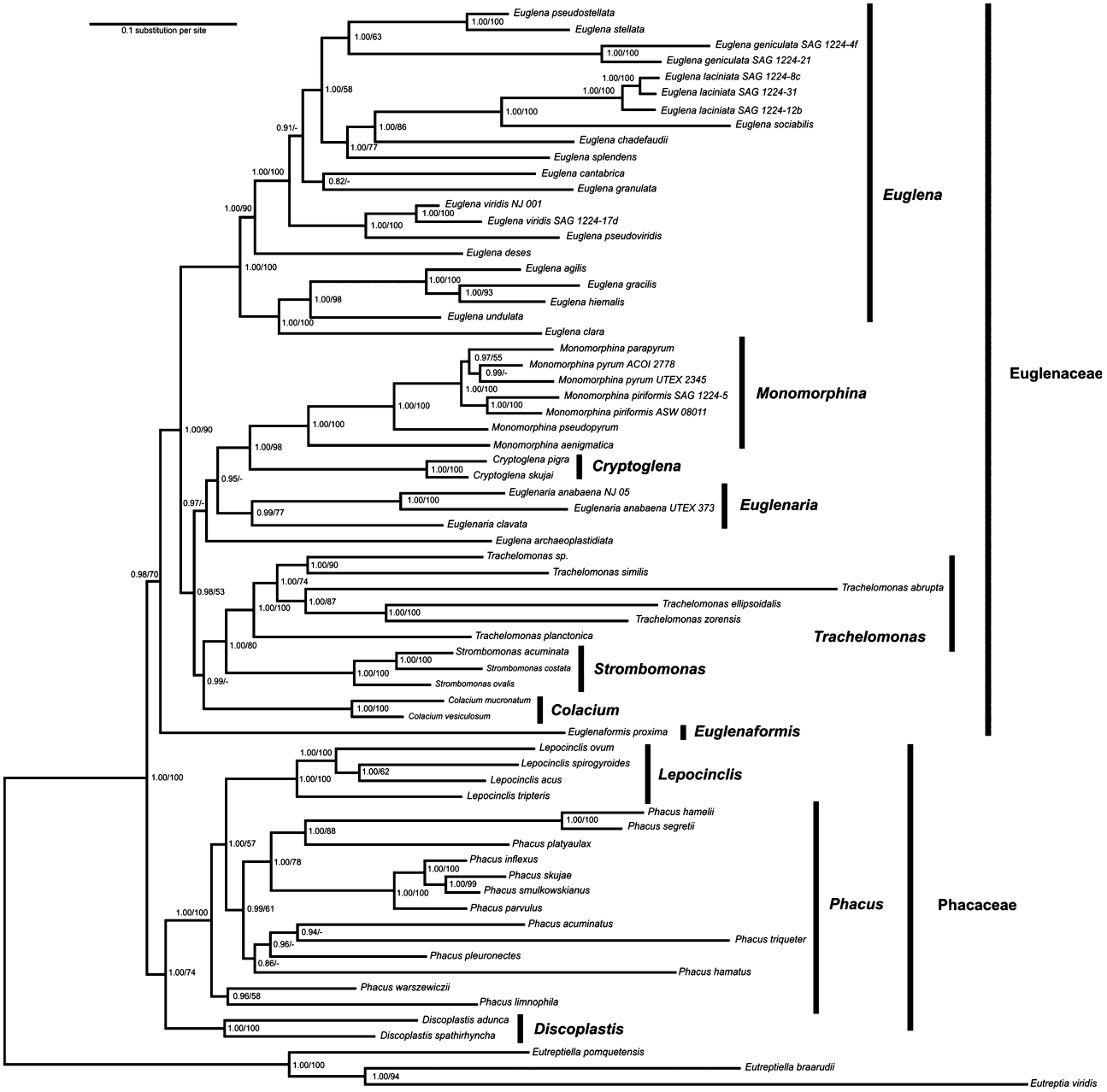

Figure 1 Bayesian tree inferred from 6,915 characters (*hsp*90, *psb*O, nSSU, nLSU and cpSSU) obtained from 68 taxa (maximum-likelihood tree has the same topology). Numbers at internal nodes represent posterior probability (pp) values (left) and bootstrap support (bs) values (right). Pp values below 0.75 and bs values below 50 are marked with dashes.

Two families of freshwater photosynthetic euglenids, Euglenaceae (0.98/70 = posterior probability [pp]/bootstrap support [bs]), and Phacaceae (1.00/74), were recovered as monophyletic groups. Most of the genera within the families were monophyletic with strong support, except for polyphyletic *Euglena* and paraphyletic *Phacus* (Fig. 1).

Within the family Euglenaceae the genus *Euglena* was split into two clades, of which the vast majority of species were grouped in one major, well-supported clade (1.00/100). Only *Euglena archaeoplastidiata* was located outside of the main *Euglena* clade, however, its basal position to the *Euglenaria* and *Monomorphina*/*Cryptoglena* clades was supported only in the Bayesian analyses (0.97). *Monomorphina* and *Cryptoglena* taxa shared a common ancestor and formed sister clades (1.00/98) as two well-supported monophyletic lineages (1.00/100). Sister to them was the genus *Euglenaria,* represented by three species that formed a moderately supported clade

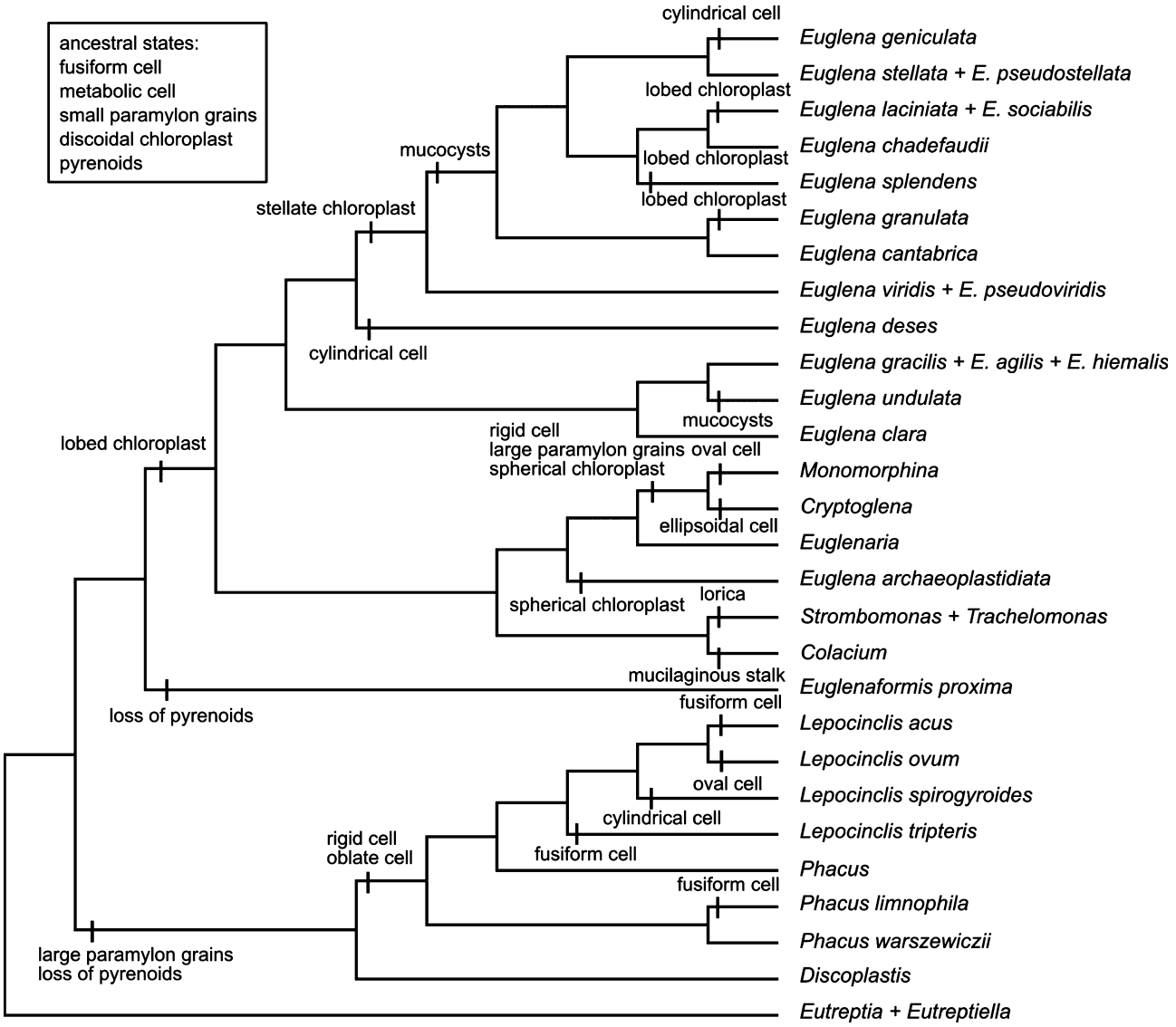

Figure 2 Trait mapping of photosynthetic euglenid characters. Morphological characters were numerically coded (Table 1) and mapped onto the phylogenetic tree (Fig. 1).

(0.99/77). The sister position of *Euglenaria* and the *Monomorphina/Cryptoglena* clade was supported only in the Bayesian analyses (0.95). The genus *Colacium,* represented only by two species, was well supported (1.00/100) and branched off prior to the clade uniting *Trachelomonas* and *Strombomonas* (1.00/80). This position of the *Colacium* clade was supported only in the Bayesian analyses (0.99). Finally, the only known species in the genus *Euglenaformis—Euglenaformis proxima* (0.98/70), was positioned at the base of the family Euglenaceae. *Lepocinclis*, *Phacus,* and *Discoplastis* genera made up the family Phacaceae, and formed one clade (1.00/74). *Discoplastis* formed a well-supported clade (1.00/100) sister to the *Phacus/Lepocinclis* clade (1.00/100). However, *Lepocinclis* was positioned within the *Phacus* clade (1.00/57). This position indicates the paraphyly of the genus *Phacus*, but only with moderate support.

Character evolution

A total of 21 states from eight characters were included in the morphological matrix (Table 1, Fig. S1–S7). All characters and states were optimized across the tree using parsimony and maximum likelihood ancestral state reconstruction, and superimposed on the Bayesian phylogeny (Fig. 2).

Cell shape

A fusiform cell shape is the most common across the photosynthetic euglenid lineages, and cylindrical, ellipsoi-

dal, oblate, and oval cell shapes were distributed sporadically across the tree.

Metaboly

Rigid cells were characteristic for four clades (*Phacus, Lepocinclis, Monomorphina,* and *Cryptoglena*), while the rest of the clades, as well as the ancestor of the freshwater photosynthetic euglenids (the Eutreptiales), were characterized by metabolic cells. Ancestral state reconstruction suggested that rigid cells arose twice, once in the common ancestor of *Phacus* and *Lepocinclis* and once in the common ancestor of *Monomorphina* and *Cryptoglena* (Fig. 2).

Large (dimorphic) paramylon grains

Large paramylon grains are found in all members of the Phacaceae, and also in the *Monomorphina*/*Cryptoglena* clade (Fig. 2). Ancestral state reconstruction suggested that this character evolved twice, once in the common ancestor of the family Phacaceae, and once in the common ancestor of *Monomorphina* and *Cryptoglena*.

Chloroplasts

Ancestral state reconstruction suggested that ancestral chloroplasts of freshwater autotrophic euglenids were numerous and discoid. This type of chloroplast characterized the family Phacaceae as well as *E. proxima*, the most basal species of the family Euglenaceae. Ancestral state reconstruction suggested that lobed chloroplasts evolved once after *E. proxima* split from the main line of the Euglenaceae. Spherical chloroplasts evolved independently many times as indicated by their scattered occurrence in the tree. Evolution of stellate chloroplasts was not clear based on ancestral state reconstruction. Their scattered occurrence might be explained by independent origins or they may have appeared only once and then were lost in several lineages.

Pyrenoids

At least some of the marine photosynthetic euglenids (*Eutreptia* and *Eutreptiella*) have pyrenoids, and ancestral state reconstruction suggests that the common ancestor of the freshwater photosynthetic euglenids had pyrenoids, however, they were lost independently twice, once in the common ancestor of the family Phacaceae, and second time in *E. proxima*.

Mucocysts

The "true mucocysts" (sensu Kosmala et al. 2009) are characteristic only for some representatives of *Euglena*. Ancestral state reconstruction indicated that the common ancestor of *Euglena* (Fig. 2) had no such mucocysts, so they probably appeared independently at least two times in the genus *Euglena*.

Loricas

Loricas occurred only in two sister genera, *Trachelomonas* and *Strombomonas*, and appeared in the common ancestor of those two genera (Fig. 2).

Mucilaginous stalks

These were characterized in only one clade, and were a synapomorphy for the genus *Colacium* (Fig. 2).

DISCUSSION

Evolutionary relationships

The phylogenetic reconstruction based on concatenated sequences of five genes: three rDNA sequences (nSSU, nLSU and cpSSU), and two protein-coding sequences (*hsp*90 and *psb*O) is consistent with previous results obtained on the basis of rDNA genes alone (Kim et al. 2010; Linton et al. 2010). Our molecular phylogenetic study supports the monophyly of the families Euglenaceae and Phacaceae, and most genera. Euglenaceae was divided into eight clades that represent genera and single taxon *E. archaeoplastidiata*. The earliest described and most problematic genus is *Euglena*. The main clade consisting of most of the *Euglena* species was recognized in previous studies with strong support (Kim et al. 2010; Linton et al. 2010; Marin et al. 2003; Milanowski et al. 2006; Moreira et al. 2001; Triemer et al. 2006). However, in addition to the main *Euglena* clade, there were always several separated taxa that have recently been renamed and established as new genera—*Discoplastis* (Triemer et al. 2006), *Euglenaria* (Linton et al. 2010) and *Euglenaformis* (Bennett et al. 2014). Our analysis supported the monophyly of those recently described genera. The phylogenetic position of *E. archaeoplastidiata*, which diverged prior to the *Monomorphina*/*Cryptoglena* and *Euglenaria* clade, was the same in both analyses, however, with no support in the ML analysis. Prior to this study, the phylogenetic position of this species was only studied once, and it branched off at the base of the genus *Euglena* (Kim et al. 2010). *Euglena archaeoplastidiata* has a single, parietal chloroplast similar to those observed in *Monomorphina* and *Cryptoglena*, but some characters (e.g. the presence of only small paramylon grains, diplopyrenoids and metaboly) are more similar to the characters of *Euglena* or *Euglenaria*. Since only one taxon is available, taxonomic decisions cannot be inferred, but this species seems to be crucial for interpretation of the evolution of characters because of the presence of a single chloroplast in this species.

Our analysis supported the monophyly of *Monomorphina*, *Cryptoglena*, *Strombomonas*, *Trachelomonas,* and *Colacium*. Also, relationships among those genera were supported, and were generally consistent with, previous studies (Kim et al. 2010; Linton et al. 2010; Marin et al. 2003; Milanowski et al. 2006; Triemer et al. 2006). The topological position of the genus *Colacium* was the same as what has been seen in most previous studies, namely sister to the *Trachelomonas* and *Strombomonas* clade (Ciugulea et al. 2008; Kim et al. 2010; Triemer et al. 2006), but received no bootstap support in the ML tree.

The family Phacaceae was described recently by Kim et al. (2010), and our analyses also indicated its monophyly. The diagnosis for the genus *Phacus* was emended

by Marin et al. (2003) to clearly separate taxa within that genus from taxa in *Cryptoglena* and *Monomorphina*. *Phacus* sensu Marin et Melkonian formed a monophyletic sister clade to *Lepocinclis* in all previous studies (Kim et al. 2010; Linton et al. 2010; Marin et al. 2003; Milanowski et al. 2006; Triemer et al. 2006) and was clearly morphologically distinct from taxa in the genus *Lepocinclis*. However, in the phylogeny of Linton et al. (2010), three additional taxa—*Euglena limnophila* [*Phacus limnophila*], *Phacus warszewiczii,* and *Lepocinclis salina* [*Phacus salina*], were added to the analysis. *Euglena limnophila* [*Phacus limnophila*] and *P. warszewiczii* formed a well-supported sister group to the *Phacus* sensu Marin et Melkonian clade (1.00/96) and *L. salina* [*Phacus salina*] grouped within the *Phacus* clade. This presented a problem because, morphologically, *E. limnophila* [*Phacus limnophila*], and *L. salina* [*Phacus salina*] did not fit into the emended diagnosis of *Phacus* by Marin et al. (2003). Consequently, the generic description of *Phacus* was revised to include ovoid and spindle-shaped species (Linton et al. 2010), which meant that a flattened cell shape was no longer a defining characteristic that separated *Phacus* from *Lepocinclis*. In our analyses, the genus *Phacus* sensu Linton et Karnkowska was paraphyletic. One moderately supported (1.00/57) clade was comprised of a subclade containing members of the genus *Lepocinclis*, and another subclade containing *Phacus* sensu Marin et Melkonian taxa. The other clade was comprised of *P. limnophila* and *P. warszewiczii*, and was basal to the clade grouping *Lepocinclis* and *Phacus* sensu Marin et Melkonian. This result suggested that *Phacus* sensu Linton et Karnkowska could be split into two genera, *Phacus* sensu Marin et Melkonian and a new genus containing *P. limnophila* and *P. warszewiczii.* However, detailed morphological analyses support the position of *P. warszewiczii* as among the earliest diverging *Phacus* species (Esson and Leander 2010). The results presented in these phylogenetic analyses do not necessarily contest the assignment of these two taxa to the genus *Phacus*, but rather demonstrate that an increase in taxon sampling within this group is necessary to either support or reject *Phacus* sensu Linton et Karnkowska and ultimately may require the amalgamation of *Phacus* and *Lepocinclis* into a single genus.

Our results suggest that simply adding more molecular markers to a dataset does not necessarily resolve all taxonomical problems, and that better results may be achieved by adding new species to the analysis. This has been shown in previous analyses of photosynthetic euglenids, in which an increase in species representation on phylogenetic trees resulted in the disclosure of new evolutionary lines (Linton et al. 2010; Marin et al. 2003; Triemer et al. 2006). One major issue with this finding is that over last decade almost all known strains from culture collections have been sequenced and included in phylogenetic analyses. However, new approaches such as multiple displacement amplification using just a few cells (Bennett and Triemer 2012) or single-cell approach (Lax and Simpson 2013) have recently been shown to be a valuable tool to study the biodiversity and phylogeny of autotrophic, as well as phagotrophic, euglenids.

Character evolution

The common opinion is that microorganisms do not show the variability in morphological features found in multicellular eukaryotes. This is not true for the unicellular eukaryotic organisms that have undergone endosymbiotic episodes in their history (Leander et al. 2007). The transition from a heterotrophic to an autotrophic lifestyle has significantly influenced the evolution of morphological features in many lineages. Such a situation was also observed in photosynthetic euglenids whose direct ancestors were phagotrophic (Leander et al. 2007). Among the most diverse characters are: (1) shape and cell plasticity (metaboly), (2) chloroplast number and its morphology (including presence of pyrenoids), and (3) the presence of large paramylon grains and mucus bodies.

Most euglenids had rounded cells (not flat) and were spindle-shaped (broad-fusiform, narrow fusiform, cylindrical-fusiform, etc.). Rarely was the cell body flattened and oblate (leaf-shaped) (Fig. S1). This latter form occurs only in *Phacus,* and appeared only once in the course of evolution. However, studies have shown that "flattening" is not characteristic of the whole genus, because a few representatives had cylindrical (*P. limnophila*) or rounded (*P. salina*) cells (Linton et al. 2010). Spherical and cylindrical cells were found multiple times and appeared independently in several genera. The evolution of cell shape was not closely linked to the phylogeny of photosynthetic euglenids, and cell shape, as a primary diagnostic feature, was applicable only in a few cases.

Cell plasticity (metaboly) was a feature imprecisely described in the literature using terms like "very metabolic" or "low metaboly" (Fig. S2). It could be, however, closely linked to rigidity of the periplast depending on its structure, for example, the number of pellicle strips, or their size and/or shape. Trends in the evolution of the euglenid pellicle were described in a series of detailed studies (Esson and Leander 2008; Leander and Farmer 2000, 2001a,b; Leander et al. 2001b). Because these data were available for only a limited number of the taxa included in our study, these ultrastructural features could not be incorporated into our character evolution analyses.

The most metabolic forms were described in the literature as having "euglenoid movement" (Harris 1969). This metabolic movement was a characteristic feature for *Discoplastis*, *Euglena*, *Euglenaria,* and *Euglenaformis*, as well as for the marine photosynthetic euglenids (*Eutreptiales*) and their phagotrophic ancestors. This evidence suggested that strong cell plasticity (and the resulting way of moving) originated in the phagotrophic ancestors. With the advent of photosynthesis, it was possible that the evolutionary pressure on the cytoskeleton involved in cell movement and nutrition changed, because most species of photosynthetic euglenids did not have the ability to undergo the gliding motility found in many heterotrophic forms (Leander et al. 2007). The structure of the periplast of certain

genera caused their cells to be strongly rigid (*Phacus*, *Lepocinclis*, *Monomorphina*, *Cryptoglena*). According to Leander et al. (2007), evolutionary changes in the cytoskeleton, leading to a stiffening of the cells, took place at least three times: in the ancestor of the *Monomorphina*/*Cryptoglena* clade, the *Phacus* clade, and the *Lepocinclis* clade, while the common ancestor of *Lepocinclis* and *Phacus* probably still had a flexible periplast similar to the modern representatives of the genus *Discoplastis*. Ancestral state reconstruction suggested that the common ancestor of *Phacus* and *Lepocinclis* had rigid cells, and representatives of both genera inherited this trait. Undoubtedly, this feature appeared more than once during the evolution of photosynthetic euglenids, and it was correlated with a reduction in the total number of pellicle strips. The morphology and organization of the pellicular strips in phototrophs might be an adaptation for diffusing the light used in photosynthesis and for protection from predators (Leander 2004).

Paramylon grains occur in the cytoplasm of all photosynthetic euglenids, and are often numerous in the area of pyrenoids. However, paramylon grains are not directly related to the presence of chloroplasts as evidenced by their presence in heterotrophic euglenids such as *Peranema* (Leander et al. 2001a). This suggests that the accumulation of carbon in the form of paramylon (β-1,3 glucan) is an ancestral feature of all phototrophic euglenids and appeared before the secondary endosymbiotic event (Leander et al. 2001a). β-1,3 glucans are also the storage material for phototrophic dinoflagellates or chrysophyta (Kiss et al. 1987; Kivic and Walne 1984; Leedale 1967), however, they are not identical polymers. Chrysolaminarin, the β-1,3 glucan found in diatoms and haptophytes, has occasional branching at the C2 and C6 positions, has a much lower degree of polymerization and is water soluble. The β-1,3 glucan of haptophytes is a branched polymer with β-1,3 and β-1,6 linkages and the storage product may either be soluble and low molecular weight, as in *Phaeocystis globosa*, or may accumulate as a granule in the cytoplasm as in *Pavlova lutheri* (Hirokawa et al. 2008). However, although the storage product of *Pavlova* is crystalline, it has been shown that the internal structure of the crystal is very different from paramylon (Kiss and Triemer 1988). So far, it is not known whether the biosynthetic pathway for paramylon synthesis in euglenids was inherited from the ancestor, or evolved independently in this lineage. One could speculate that the pathway had been transferred from an algal donor with the ability to produce a β-1,3 glucan storage product, but it could also have been obtained through transfer from a food source such as yeast (e.g. *Saccharomyces cerevisiae*), which produce an insoluble β-1,3 glucan polymer (Triemer 1997).

The simultaneous occurrence of both small and large paramylon grains (dimorphism of paramylon grains—Fig. S3b–e) were found in two clades: the Phacaceae and the *Monomorphina*/*Cryptoglena* clade (Fig. 2). Character evolution analyses suggested that the ancestor of photosynthetic euglenids did not have this trait, and that it arose independently in these two lineages. It is difficult to determine the role of large paramylon grains. Perhaps their absence in cells is associated with metaboly—it is easy to imagine that large grains could hinder metabolic movements.

Characteristics associated with the morphology of chloroplasts are difficult to observe and relatively difficult to describe. Mapping features showed that numerous, small, parietal, disk-shape chloroplasts without pyrenoids (characteristic for the family Phacaceae and genera *Eutreptia* and *Euglenaformis*—Fig. S4e), probably existed in the common ancestor of the photosynthetic euglenids as well, while the single, parietal, spherical (or almost spherical) chloroplast characteristic for *Monomorphina*, *Cryptoglena* (Kosmala et al. 2007), and *E. archaeoplastidiata* is derived (Fig. S4f). In the case of the *Monomorphina*/*Cryptoglena* clade (Fig. 2), it is likely that their common ancestor possessed a single large chloroplast. Axial, stellate chloroplasts are present only in several species of *Euglena* (Kosmala et al. 2009; Fig. S4g). On the basis of character evolution, we cannot unambiguously reconstruct the evolution of this type of chloroplast. However, we propose that such a complicated structure more likely evolved only once and then disappeared in some lineages, rather than appeared independently many times.

Nevertheless, there appears to be a correlation between the shapes of chloroplasts and metabolic behavior. Euglenid species with rigid cells typically have numerous parietal chloroplasts as in *Phacus* and *Lepocinclis* (Ciugulea and Triemer 2010; Pochmann 1942), or a single, spherical chloroplast, as in *Monomorphina* (Kosmala et al. 2007). In both cases, the entire surface of the chloroplast has a uniform exposure to light. In contrast, the most metabolic genera (*Euglena*, *Euglenaria*) have a much more complicated chloroplast structure (Fig. S4a–d), often with the center of the chloroplasts located deep in the cytoplasm and with long bands reaching beneath the surface of the cell (Fig. S4d) (Ciugulea and Triemer 2010; Karnkowska-Ishikawa et al. 2013; Kosmala et al. 2009). These complicated structures could mean that their total chloroplast surface area may be even greater, but the exposure to the light is smaller. Perhaps this is related to protection against excessive exposure to light as often can occur in shallow water bodies.

One of the characteristic features for euglenids is the presence of "paramylon caps" adjacent to the pyrenoid on the outside of the chloroplast (Fig. S4b–d), or the accumulation of small paramylon grains in the immediate vicinity of pyrenoids outside of the chloroplasts ("paramylon center") (Fig. S4g). The presence of paramylon caps indirectly indicates the presence of pyrenoids. For many photosynthetic euglenids, pyrenoids are an important diagnostic feature, although the presumption of their presence is based solely on the presence or absence of paramylon caps (Kusel-Fetzmann and Weidinger 2008; Pringsheim 1956). Thus, in the literature, there is a great deal of contradictory information. Modern transmission electron microscopy revealed the presence of pyrenoids in many species for which they were thought to be absent, for example *Monomorphina* (Nudelman et al. 2006), *Colacium*

or *Trachelomonas* (Brown et al. 2003) as well as many representatives of *Euglena* (Zakryś et al. 2001; Zakryś and Walne 1998) (Fig. S4a). The presence of pyrenoids in *Cryptoglena* is still debatable—they have never been reported within this genus, but the electron micrographs by Rosowski and Lee (1978) suggest the possibility of their occurrence. The literature also lacks clear and reliable information regarding *Eutreptiella* and *Eutreptia*—it is believed that some species have them while others do not. Numerous reports in the literature concerning pyrenoids in photosynthetic euglenids refer mainly to *E. gracilis*. It was shown that this species, when kept in the dark, did not appear to form pyrenoids. However, after transferring cells into a mineral medium, the pyrenoids appear in rudimentary form (Osafune et al. 1988). Consequently, it seems that the development of pyrenoids is only possible in the presence of light (Kiss et al. 1987). The presence of pyrenoids may also be related to ontogeny. Kiss et al. (1987) showed that RubisCO was concentrated in pyrenoids during the growth phase, but just before and during division, they became undetectable by the antibodies. Therefore, the lack of pyrenoids could be a state of transition. However, pyrenoids have not been observed in *Phacus*, *Lepocinclis,* and *Discoplastis* (Kim et al. 2010), as well as in the genus *Euglenaformis* (Bennett et al. 2014). A lack of pyrenoids was proposed as a diagnostic feature for the Phacaceae (Kim et al. 2010), but due to problems with their detection, as well as their most likely homoplastic character, the absence of pyrenoids was not a very effective diagnostic indicator. Mapping of characters has also shown that the common ancestor of photosynthetic euglenids probably had pyrenoids, so the lack of pyrenoids in those four genera would be a secondary feature. Additional evidence in favor of this scenario is the presence of pyrenoids in *Rapaza viridis*, a recently discovered taxon which is believed to be a transitional stage between heterotrophic and autotrophic euglenids (Yamaguchi et al. 2012).

Several types of cellular structures responsible for the production of mucus have been reported in euglenids. Mucus primarily facilitates the euglenoid movement (swimming or gliding), is the main building material for stalks in *Colacium*, provides protection for division stages (palmellas) and is the primary component of loricas (Hausmann and Mignot 1977; Hilenski and Walne 1983; Mignot 1966). The presence and shape of some kinds of mucocysts ("true mucocysts"—Kosmala et al. 2009) can be a very good diagnostic feature for species in the genus *Euglena* (Karnkowska-Ishikawa et al. 2013; Kosmala et al. 2009), where mucocysts are always of a uniform shape (spindle-shaped, or spherical) and are small (1–3 μm) bodies that lie just below the surface of the periplast (Fig. S5b). In most species mucocysts are not easily visible without neutral red staining. This type of mucosyst is found exclusively in some *Euglena* species. Earlier study suggested that mucocysts appeared only once in *Euglena* (Milanowski et al. 2006), and our character evolution analyses suggested that it happened twice in the genus *Euglena*. Another possibility is that the mucocysts are homologous to the extrusomes present in phagotrohic euglenids (Gojdics 1953), but were lost in the majority of the autotrophic euglenid lineages, except in the genus *Euglena*.

CONCLUSIONS

Five molecular markers and a sampling of 59 taxa provided the most comprehensive tree of photosynthetic euglenids to date. Two nuclear-encoded genes (*hsp90, psbO*) had not previously been used in phylogenetic analyses with photosynthetic euglenids. The resultant topology has strong nodal support for most of the branches and is consistent with all previous phylogenies. The monophyletic status of two families and 11 genera is well supported. However, the position of *E. archaeoplastidiata*, although somewhat variable, always remains outside of the main *Euglena* clade, which makes this genus polyphyletic. In addition, the clade containing *P. limnophila* and *P. warszewiczii* was paraphyletic to the rest of the *Phacus* taxa—though this relationship was only moderately supported.

We also analyzed character evolution in photosynthetic euglenids using our well-resolved phylogenetic tree and morphological traits. On the basis of the inferred ancestral state reconstructions, we identified many well-supported clades defined by apomorphic morphological characters, however, we also recognized homoplastic traits.

ACKNOWLEDGMENTS

Financial support was provided by Polish Ministry of Science and Higher Education grant no. N N303 796040 and by the National Science Foundation PEET program, grant no. DEB4-21348. A.K. was supported by the European Social Fund and the state budget of the Czech Republic. Project no. CZ.1.07/2.3.00/30.0061.

LITERATURE CITED

Bennett, M. S. & Triemer, R. E. 2012. A new method for obtaining nuclear gene sequences from field samples and taxonomic revisions of the photosynthetic euglenoids *Lepocinclis* (*Euglena*) *helicoideus* and *Lepocinclis* (*Phacus*) *horridus* (Euglenophyta). *J. Phycol.*, 48:254–260.

Bennett, M., Wiegert, K. E. & Triemer, R. E. 2014. Characterization of new genus *Euglenaformis* and the chloroplast genome of *Euglenaformis* [*Euglena*] *proxima* (Euglenophyta). *Phycologia* 53: 66–73.

Breglia, S. A., Slamovits, C. H. & Leander, B. S. 2007. Phylogeny of phagotrophic euglenids (Euglenozoa) as inferred from hsp90 gene sequences. *J. Eukaryot. Microbiol.*, 52:86–94.

Brosnan, S., Shin, W., Kjer, K. M. & Triemer, R. E. 2003. Phylogeny of the photosynthetic euglenophytes inferred from the nuclear SSU and partial LSU rDNA. *Int. J. Syst. Evol. Microbiol.*, 53:1175–1186.

Brown, P. J. P., Zakryś, B. & Farmer, M. A. 2003. Plastid morphology, ultrastructure, and development in *Colacium* and the loricate Euglenophytes (Euglenophyceae). *J. Phycol.*, 39:115–121.

Ciugulea, I., Nudelman, M. A., Brosnan, S. & Triemer, R. E. 2008. Phylogeny of the euglenoid loricate genera *Trachelomonas* and

Strombomonas (Euglenophyta) inferred from nuclear SSU and LSU rDNA. *J. Phycol.*, 44:406–418.

Ciugulea, I. & Triemer, R. E. 2010. A Color Atlas of Photosynthetic Euglenoids. Michigan State University Press, East Lansing, 204 p.

Edgar, R. C. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.*, 32:1792–1797.

Ehrenberg, C. G. 1830. Neue Beobachtungen über blutartige Erscheinungen in Aegypten, Arabien und Sibirien, nebst einer Uebersicht und Kritik der früher bekannten. *Annalen der Physik und Chemie*, 8:477–514.

Esson, H. J. & Leander, B. S. 2008. Visualizing the complex substructure of euglenid pellicle strips with SEM. *Phycologia*, 47:529–532.

Esson, H. J. & Leander, B. S. 2010. Evolution of distorted pellicle patterns in rigid photosynthetic Euglenids (*Phacus* Dujardin). *J. Eukaryot. Microbiol.*, 57:19–32.

Gojdics, M. 1953. The Genus Euglena. The University of Wisconsin Press, Madison. 268 p.

Harris, J. 1969. Microscopical observations of vast numbers of animalcules seen in water. *Philos. Trans. R. Soc. London Ser. B*, 19: 254–259.

Hausmann, K. & Mignot, J. P. 1977. Unterschungen an den mucocysten von Euglena splendens Dangeard 1901. *Protistologica*, 13:213–217.

Hilenski, L. L. & Walne, P. L. 1983. Ultrastructure of ejectile mucocysts in *Peranema trichophorum* (Euglenophyceae). *J. Protozool.*, 30:491–496.

Hirokawa, Y., Fujiwara, S., Suzuki, M., Akiyama, T., Sakamoto, M., Kobayashi, S. & Tsuzuki, M. 2008. Structural and physiological studies on the storage β-polyglucan of haptophyte *Pleurochrysis haptonemofera*. *Planta*, 227:589–599.

Karnkowska-Ishikawa, A., Milanowski, R., Kwiatowski, J. & Zakryś, B. 2010. Taxonomy of the *Phacus oscillans* (Euglenaceae) and its close relatives – balancing morphological and molecular features. *J. Phycol.*, 46:172–182.

Karnkowska-Ishikawa, A., Milanowski, R., Triemer, R. E. & Zakryś, B. 2012. Taxonomic revisions of morphologically similar species from two genera: *Euglena* (*E. granulata* and *E. velata*) and *Euglenaria* (*Eu. anabaena, Eu. caudata, Eu. clavata*). *J. Phycol.*, 48:729–739.

Karnkowska-Ishikawa, A., Milanowski, R., Triemer, R. E. & Zakryś, B. 2013. A rediscription of morphologically similar species from the genus *Euglena*: *E. laciniata*, *E. sanguinea*. *E. sociabilis* and *E. splendens*. *J. Phycol.*, 49:616–626.

Karnkowska-Ishikawa, A., Milanowski, R. & Zakryś, B. 2011. The species *Euglena deses* revisited: new morphological and molecular data. *J. Phycol.*, 47:653–661.

Kim, J. I. & Shin, W. 2008. Phylogeny of the Euglenales inferred from plastid LSU rDNA sequences. *J. Phycol.*, 44:994–1000.

Kim, J. I., Shin, W. & Triemer, R. E. 2010. Multigene analyses of photosynthetic euglenoids and new family Phacaceae (Euglenales). *J. Phycol.*, 46:1278–1287.

Kim, E., Simpson, A. G. B. & Graham, L. E. 2006. Evolutionary relationships of apusomonads inferred from taxon-rich analyses of 6 nuclear encoded genes. *Mol. Biol. Evol.*, 23:2455–2466.

Kiss, J. Z. & Triemer, R. E. 1988. A comparative study of the storage carbohydrate granules from *Euglena* (Euglenida) and *Pavlova* (Prymnesiida). *J. Protozool.*, 35:237–241.

Kiss, J. Z., Vasconcelos, A. C. & Triemer, R. E. 1987. Structure of the euglenoid storage carbohydrate, Paramylon. *Am. J. Bot.*, 74:877–882.

Kivic, P. A. & Walne, P. L. 1984. An evaluation of possible phylogenetic relationship between the Euglenophyta and Kinetoplastida. *Origins Life*, 13:269–288.

Kjer, K. M. 1995. Use of rRNA secondary structure in phylogenetic studies to identify homologous positions: an example of alignment and data presentation from the frogs. *Mol. Phylogenet. Evol.*, 4:314–330.

Kosmala, S., Karnkowska, A., Milanowski, R., Kwiatowski, J. & Zakryś, B. 2005. The phylogenetic and taxonomic position of *Lepocinclis fusca* comb. nova (=*Euglena fusca*) (Euglenaceae). Morphological and molecular justification". *J. Phycol.*, 41:258–267.

Kosmala, S., Karnkowska-Ishikawa, A., Milanowski, R., Kwiatowski, J. & Zakryś, B. 2009. Phylogeny and systematics of species from the genus *Euglena* (Euglenaceae) with axial, stellate chloroplasts based on morphological and molecular data – new taxa, emended diagnoses and epitypifications. *J. Phycol.*, 45:464–481.

Kosmala, S., Milanowski, R., Brzóska, K., Pękala, M., Kwiatowski, J. & Zakryś, B. 2007. Phylogeny and systematics of the genus *Monomorphina* (Euglenaceae) based on morphological and molecular data. *J. Phycol.*, 43:171–185.

Kusel-Fetzmann, E. & Weidinger, M. 2008. Ultrastructure of five *Euglena* species positioned in the subdivision Serpentes. *Protoplasma*, 233:209–222.

Lax, G. & Simpson, A. G. B. 2013. Combining molecular data with classical morphology for uncultured phagotrophic Euglenids (Excavata): a single-cell approach. *J. Eukaryot. Microbiol.*, 60:615–625.

Leander, B. S. 2004. Did trypanosomatid parasites have photosynthetic ancestors? *Trends Microbiol.*, 12:251–258.

Leander, B. S., Esson, H. J. & Breglia, S. A. 2007. Macroevolution of complex cytoskeletal systems in euglenids. *BioEssays*, 29:987–1000.

Leander, B. S. & Farmer, M. A. 2000. Comparative morphology of the euglenid pellicle. I. Patterns of strips and pores. *J. Eukaryot. Microbiol.*, 47:469–479.

Leander, B. S. & Farmer, M. A. 2001a. The evolution of *Phacus* (Euglenozoa) as inferred from pellicle morphology and SSU rDNA. *J. Phycol.*, 37:1–17.

Leander, B. S. & Farmer, M. A. 2001b. Comparative morphology of the euglenid pellicle. II. Diversity of strip substructure. *J. Eukaryot. Microbiol.*, 48:202–217.

Leander, B. S., Triemer, R. E. & Farmer, M. A. 2001a. Character evolution in heterotrophic euglenoids. *Eur. J. Protistol.*, 37:337–356.

Leander, B. S., Witek, R. P. & Farmer, M. A. 2001b. Trends in the evolution of the euglenid pellicle. *Evolution*, 55:2215–2235.

Leedale, G. F. 1967. Euglenoid Flagellates. Prentice-Hall, Englewood Cliffs, NJ.

Lewis, P. O. 2001. A likelihood approach to estimating phylogeny from discrete morphological character data. *Syst. Biol.*, 50:913–925.

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., Karnkowska-Ishikawa, A., Kim, J. I., Shin, W., Bennett, M., Kwiatowski, J., Zakryś, B. & Triemer, R. E. 2010. Reconstructing Euglenoid evolutionary relationships using three genes: nuclear SSU and LSU, and chloroplast 16S rDNA sequences and the description of *Euglenaria* gen. nov. (Euglenophyta). *Protist*, 161:603–619.

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. 2011. Mesquite: a modular system for evolutionary analysis. Version 2.75. http://mesquiteproject.org. (accessed 10 July 2014).

Marin, B., Palm, A., Klingberg, M. & Melkonian, M. 2003. Phylogeny and taxonomic revision of plastid-containing Euglenophytes based on SSU rDNA sequence comparisons and synapomorphic signatures in the SSU rRNA secondary structure. *Protist*, 154:99–145.

Maruyama, S., Suzaki, T., Weber, A. P. M., Archibald, J. M. & Nozaki, H. 2011. Eukaryote-to-eukaryote gene transfer gives rise to genome mosaicism in euglenids. *BMC Evol. Biol.*, 11:105.

Mignot, J. P. 1966. Structure et ultrastructure de quelques Euglenomonadines. *Protistologica*, 2:51–117.

Milanowski, R., Kosmala, S., Zakryś, B. & Kwiatowski, J. 2006. Phylogenetic analysis of photosynthetic Euglenophytes based on combined chloroplast and cytoplasmic SSU rDNA sequences. *J. Phycol.*, 42:721–730.

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.

Moreira, D., López-García, P. & Rodriguez-Valera, F. 2001. New insights into the phylogenetic position of diplonemids: G+C content bias, differences of evolutionary rate and a new environmental sequence. *Int. J. Syst. Evol. Microbiol.*, 51:2211–2219.

Müller, O. F. 1786. Animalcula infusoria fluviatilia et marina quæ detexit, systematice descripsit et ad vivum delineari curavit Otho Fridericus Müller; sistet opus hoc posthumum, quod cum tabulis Aeneis L. in lucem tradit vidua ejus nobilissima cura Othonis Fabricii. Hauniæ [Copenhagen]: Typis N. Mölleri. pp. i-lvi, 1–367.

Müllner, A. N., Angeler, D. G., Samuel, R., Linton, E. W. & Triemer, R. E. 2001. Phylogenetic analysis of phagotrophic, phototrophic and osmotrophic euglenoids by using the nuclear 18S rDNA sequence. *Int. J. Syst. Evol. Microbiol.*, 51:783–791.

Nudelman, M. A., Leonardi, P. I., Conforti, V., Farmer, M. A. & Triemer, E. R. 2006. Fine structure and taxonomy of *Monomorphina aenigmatica* comb. nov. (Euglenophyta). *J. Phycol.*, 42:194–202.

Nudelman, M. A., Rossi, M. S., Conforti, V. & Triemer, R. E. 2003. Phylogeny of Euglenophyceae based on small subunit rDNA sequences: taxonomic implications. *J. Phycol.*, 39:226–235.

Osafune, T, Ehara, T., Sumida, S. & Hase, E. 1988. Light-independent and dependent processes of the formation of pyrenoid and thylakoids in dark-grown cells of *Euglena gracilis* Z. Proceedings of the IVth Asia-Pacific Conference and Workshop on Electron Microscopy. Bangkok, p. 519–520.

Pagel, M. 1999. The maximum likelihood approach to reconstructing ancestral character states of discrete characters on phylogenies. *Syst. Biol.*, 48: 612–622.

Pochmann, A. 1942. Synopsis der Gattung Phacus. *Arch. Protistenk.*, 95:81–252.

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

Pringsheim, E. G. 1956. Contributions towards a monograph of the genus *Euglena*. *Nova Acta Leopold.*, 18:1–168.

Ronquist, F., Teslenko, M., van der Mark, P., Ayres, D., Darling, A., Höhna, S., Larget, B., Liu, L., Suchard, M. A. & Huelsenbeck, J. P. 2012. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst. Biol.*, 61:539–542.

Rosowski, J. R. & Lee, K. W. 1978. *Cryptoglena pigra*: a euglenoid with one chloroplast. *J. Phycol.*, 14:160–166.

Schlösser, U. G. 1994. SAG – Sammlung von Algenkulturen at the University of Göttingen. Cataloge of strains 1994. *Bot. Acta*, 107:113–187.

Simpson, A. G. B., Gill, E. E., Callahan, H. A., Litaker, R. W. & Roger, A. J. 2004. Early evolution within Kinetoplastids (Euglenozoa), and the late emergence of Trypanosomatids. *Protist*, 155:407–422.

Simpson, A. G. B., Inagaki, Y. & Roger, A. J. 2006. Comprehensive multi-gene phylogenies of excavate protists reveal the evolutionary positions of 'primitive' eukaryotes. *Mol. Biol. Evol.*, 23:615–625.

Simpson, A. G. B., Lukes, J. & Roger, A. J. 2002. The evolutionary history of kinetoplastids and their kinetoplasts. *Mol. Biol. Evol.*, 19:2071–2083.

Song, H., Buhay, J. E., Whiting, M. F. & Crandall, K. A. 2008. Many species in one: DNA barcoding overestimates the number of species when nuclear mitochondrial pseudogenes are coamplified. *Proc. Natl Acad. Sci. USA*, 105:13486–13491.

Stamatakis, A. 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics*, 22:2688–2690.

Takahashi, F., Okabe, Y., Nakada, T., Sekimoto, H., Ito, M., Kataoka, H. & Nozaki, H. 2007. Origins of the secondary plastids of Euglenophyta and Chlorarachniophyta as revealed by an analysis of the plastid-targeting, nuclear-encoded gene psbO. *J. Phycol.*, 43:1302–1309.

Talke, S. & Preisfeld, A. 2002. Molecular evolution of euglenozoan paraxonemal rod genes par1 and par2 coincides with phylogenetic reconstruction based on small subunit rDNA data. *J. Phycol.*, 38:995–1003.

Thompson, M. D., Copertino, D. W., Thompson, E., Favreau, M. & Hallick, R. B. 1995. Evidence for the late origin of introns in chloroplast genes from an evolutionary analysis of the genus *Euglena*. *Nucleic Acids Res.*, 23:4745–4752.

Triemer, R. E. 1997. Feeding in *Peranema trichophorum* revisited (Euglenophyta). *J. Phycol.*, 33:649–654.

Triemer, R. E. & Farmer, M. A. 2007. A decade of euglenoid molecular phylogenetics. *In*: Brodie, J. & Lewis, J. (ed.), Unravelling the Algae: The Past, Present and Future of Algal Systematics. Systematics Association Series. CRC Press, Boca Rato, FL. p. 315–330.

Triemer, R. E., Linton, E., Shin, W., Nudelman, A., Monfils, A., Bennett, M. & Brosnan, S. 2006. Phylogeny of the Euglenales based upon combined SSU and LSU rDNA sequence comparisons and description of *Discoplastis* gen. nov. (Euglenophyta). *J. Phycol.*, 42:731–740.

Turmel, M., Gagnon, M. C., O'Kelly, C. J., Otis, C. & Lemieux, C. 2009. The chloroplast genomes of the green algae *Pyramimonas*, *Monomastix*, and *Pycnococcus* shed new light on the evolutionary history of prasinophytes and the origin of the secondary chloroplasts of euglenids. *Mol. Biol. Evol.*, 26:631–648.

Watanabe, M. M., Kawachi, M., Hiroki, M. & Kasai, F. 2000. NIES Collection List of Strains. Sixth Edition, 2000, Microalgae and Protozoa. Microbial Culture Collections, National Institute for Environmental Studies, Tsukuba, Japan, 159 pp.

Wuyts, J., van de Peer, Y., Winkelmans, T. & De Wachter, R. 2002. The European database on small subunit ribosomal RNA. *Nucleic Acids Res.*, 30:183–185.

Yamaguchi, A., Yubuki, N. & Leander, B. S. 2012. Morphostasis in a novel eukaryote illuminates the evolutionary transition from phagotrophy to phototrophy: description of *Rapaza viridis* n. gen. et sp. (Euglenozoa, Euglenida). *BMC Evol. Biol.*, 12:29.

Zakryś, B., Jaume-Sanchez, J. & Walne, P. L. 2001. Chloroplast ultrastructure of *Euglena cantabrica* Pringsheim, *E. cuneata*

Pringsheim, *E. deses* Ehrenberg, *E. mutabilis* Schmitz (Euglenophyceae): taxonomic significance. *Acta Protozool.*, 40:161–167.
Zakryś, B., Milanowski, R., Empel, J., Borsuk, P., Gromadka, R. & Kwiatowski, J. 2002. Two different species of *Euglena*, *E. geniculata* and *E. myxocylindracea* (Euglenophyceae), are virtually genetically and morphologically identical. *J. Phycol.*, 38:1190–1199.
Zakryś, B. & Walne, P. L. 1998. Comparative ultrastructure of chloroplast in the Subgenus *Euglena* (Euglenophyta): taxonomic significance. *Cryptogamie*, 9:3–18.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article:

Table S1. Taxa used with culture collection and GenBank accession codes.
Table S2. Primers newly designed for this study.
Figure S1. Cell shape. (**a**) fusiform (*Lepocinclis acus*), (**b**) cylindrical (*Euglena deses*), (**c**) oval (*Monomorphina pyrum*), (**d**) ellipsoidal (*Cryptoglena skujae*), (**e**) oblate (*Phacus caudatus*). Scale bars 10 μm.
Figure S2. Metaboly. (**a**) metabolic (*Euglena splendens*), (**b**) slightly metabolic or rigid (*Lepocinclis tripteris*).
Figure S3. Large paramylon grains. (**a**) absent, (**b**–**e**) present (b: ring, c: long rods, d: plates, e: disks).
Figure S4. Chloroplast morphology. (**a**) lobed with pyrenoids (arrow) (*Euglena deses*); (**b**–**d**) lobed with pyrenoids and "paramylon caps" (arrows) (b: *Euglenaria clavata*, c: *Euglenaria anabaena*, d: *Euglena sanguinea*), (**e**) discoid (*Phacus orbicularis*), (**f**) spherical (*Monomorphina aenigmatica*; with irregular, small holes visible—arrow), (**g**) stellate with "paramylon center" (arrow) (*Euglena stellata*).
Figure S5. Mucocysts. (**a**) absent, (**b**) present (rows of spindle-shaped mucocysts visible in the cell of *Euglena sanguinea*).
Figure S6. Lorica present (arrow). (**a**: *Trachelomonas hispida* var. *coronata,* **b**: *Strombomonas verrucosa*). Scale bars 10 μm.
Figure S7. Mucilaginous stalks present (arrow). Scale bar 10 μm.