Molecular Phylogenetics and Evolution 177 (2022) 107607

Contents lists available at ScienceDirect

Molecular Phylogenetics and Evolution

journal homepage: www.elsevier.com/locate/ympev

# Redefining Chlorobotryaceae as one of the principal and most diverse lineages of eustigmatophyte algae

Dovilė Barcytė [a,*], Martina Zátopková [a], Yvonne Němcová [b], Michal Richtář [a], Tatiana Yurchenko [a], Karin Jaške [a], Karen P. Fawley [c], Pavel Škaloud [b], Tereza Ševčíková [a], Marvin W. Fawley [c], Marek Eliáš [a,*]

[a] Department of Biology and Ecology, Faculty of Science, University of Ostrava, Chittussiho 10, 710 00 Ostrava, Czech Republic
[b] Department of Botany, Faculty of Science, Charles University, Benátská 2, 128 00 Prague, Czech Republic
[c] Division of Science and Mathematics, University of the Ozarks, Clarksville, AR 72830, USA

ARTICLE INFO

*Keywords:*
*Chlorobotrys*
Eustigmatophyceae
Phylogeny
Plastid genomes
Taxonomy
Zoospores

ABSTRACT

Eustigmatophyceae is one of the ~17 classes of the vast algal phylum Ochrophyta. Over the last decade, the eustigmatophytes emerged as an expansive group that has grown from the initially recognized handful of species to well over 200 genetically distinct entities (potential species). Yet the majority of eustigs, remain represented by unidentified strains, or even only metabarcode sequences obtained from environmental samples. Moreover, the formal classification of the group has not yet been harmonized with the recently uncovered diversity and phylogenetic relationships within the class. Here we make a major step towards resolving this issue by addressing the diversity, phylogeny and classification of one of the most prominent eustigmatophyte clades previously informally called the "Eustigmataceae group". We obtained 18S rDNA and *rbcL* gene sequences from four new strains from the "Eustigmataceae group", and from several additional eustig strains, and performed the most comprehensive phylogenetic analyses of Eustigmatophyceae to date. Our results of these analyses confirm the monophyly of the "Eustigmataceae group" and define its major subclades. We also sequenced plastid genomes of five "Eustigmataceae group" strains to not only improve our understanding of the plastid gene content evolution in eustigs, but also to obtain a robustly resolved eustigmatophyte phylogeny. With this new genomic data, we have solidified the view of the "Eustigmataceae group" as a well-defined family level clade. Crucially, we also have firmly established the genus *Chlorobotrys* as a member of the "Eustigmataceae group". This new molecular evidence, together with a critical analysis of the literature going back to the 19th century, provided the basis to radically redefine the historical concept of the family Chlorobotryaceae as the formal taxonomic rubric corresponding to the "Eustigmataceae group". With this change, the family names Eustigmataceae and Characiopsidaceae are reduced to synonymy with the Chlorobotryaceae, with the latter having taxonomic priority. We additionally studied in detail the morphology and ultrastructure of two Chlorobotryaceae members, which we describe as *Neustupella aerophytica* gen. et sp. nov. and *Lietzensia polymorpha* gen. et sp. nov. Finally, our analyses of partial genomic data from several Chlorobotryaceae representatives identified genes for hallmark flagellar proteins in all of these strains. The presence of the flagellar proteins strongly suggests that zoosporogenesis is a common trait of the family and also occurs in the members never observed to produce flagellated stages. Altogether, our work paints a rich picture of one of the most diverse principal lineages of eustigmatophyte algae.

## 1. Introduction

Eustigmatophytes, or eustigs for short, form a neatly defined taxon of microalgae formally recognized as the class, Eustigmatophyceae, in the vast algal radiation referred to as stramenopile algae or Ochrophyta (Eliáš et al., 2017). They are unicellular coccoid algae exhibiting a unique suite of cytological and biochemical characteristics that are distinct from other yellow-green algae historically placed in the class Xanthophyceae. These characteristics led to the recognition of the Eustigmatophyceae as a class separate from the Xanthophyceae

* Corresponding authors.
*E-mail addresses:* dovile.barcyte@gmail.com (D. Barcytė), marek.elias@osu.cz (M. Eliáš).

https://doi.org/10.1016/j.ympev.2022.107607
Received 19 May 2022; Received in revised form 11 July 2022; Accepted 5 August 2022
Available online 11 August 2022

(Hibberd and Leedale, 1970, 1972; Whittle and Casselton, 1975). Eustigs predominantly thrive in freshwater and terrestrial habitats, although the best-known representatives of the class, the genera *Nannochloropsis* and *Microchloropsis*, are nearly exclusively marine. The two aforementioned genera in particular, but more increasingly other eustigs, have been the subjects of extensive research aimed primarily at establishing and exploiting these organisms as biotechnology resources (Stoyneva-Gärtner et al., 2019). Comparatively less attention has been historically given to the diversity, ecology, and evolution of eustigmatophyte algae, although recent investigations have provided fundamental insights that paint a new picture of the group (Eliáš et al., 2017).

The major change in our perception of the Eustigmatophyceae concerns the diversity of the class. After its establishment, the number of species and genera included in the class has long remained modest. Hibberd's seminal work summarizing the first decade of eustigmatophyte research counted only 12 species in six genera as positively identified eustigs (Hibberd, 1981). As of 2017, the census had grown to ~30 formally recognised species in ~15 genera as a result of re-evaluating additional "xanthophytes" as well as discovery and description of truly novel taxa (Eliáš et al., 2017). Additional eustigmatophyte phylogenetic lineages have been recently uncovered by reassessing the genus *Characiopsis*, which had previously been placed in the Xanthophyceae (Amaral et al., 2020, 2021). Moreover, a systematic screening of algal cultures with the aim to reveal potential new eustigmatophyte taxa has indeed yielded numerous strains that are genetically distinct from previously characterized eustigs (as evaluated by sequencing the 18S rRNA and/or *rbcL* genes). These strains represent additional eustig species and genera yet to be identified or described (Amaral et al., 2020; Fawley et al., 2014, 2021; Kryvenda et al., 2018; Wolf et al., 2018).

However, we are certain that currently available eustig cultures represent only a fraction of the actual eustigmatophyte diversity. Numerous genotypes identified in a single east African lake by analysing amplicons of the 18S rRNA gene obtained by eustig-specific primers (Villanueva et al., 2014) differ from genotypes of cultured eustigmatophytes and even define novel higher-order clades lacking any cultured representatives (Amaral et al., 2020). A recent broad environmental DNA survey targeting eustigmatophyte *rbcL* sequences in freshwater samples has yielded 184 haplotypes, with only 10 of them matching *rbcL* sequences from cultured strains (Fawley et al., 2021). Hence, eustigs are far more specious and phylogenetically diverse than appears from textbooks, floras, or on-line taxonomic compendia currently in use (Graham et al., 2016; Lee, 2018; Ott et al., 2015; Guiry and Guiry, 2022). Furthermore, the scope of the Eustigmatophyceae has been potentially broadened by the recent demonstration of the existence of a diverse clade sister to all currently known members of the class and most likely responsible for the production of long chain 1,13- and 1,15-diols (lipids of enigmatic origin present in marine and lacustrine environments; Rampen et al., 2022). These novel putative eustigs remain documented only by partial rDNA sequences (the nuclear 18S and plastidial 23S), except for a full-length 18S rDNA sequence from an apparently misidentified algal isolate, for which no details have been published (GenBank accession number KY980400.1).

One of the major tasks of eustigmatophyte research is to obtain a detailed and robustly reconstructed phylogeny of the group and to develop a formal classification that appropriately captures the diversity and evolutionary history of these algae. Prior to the advent of molecular phylogenetics, Hibberd (1981) proposed a scheme classifying all then known eustigs into a single order, Eustigmatales, and four families distinguished by different combinations of morphological traits concerning vegetative cells and zoospores. Three of these families were newly established (Eustigmataceae, Pseudocharaciopsidaceae, Monodopsidaceae), whereas the fourth, Chlorobotryaceae Pascher, was adopted from previous literature and formally transferred to Eustigmatophyceae from Xanthophyceae (Hibberd, 1981). Much later a fifth family, the Loboceae, was proposed to accommodate a newly described eustig, *Pseudotetraëdriella kamillae*, because the morphology of this organism did not fit into any of the four existing families (Hegewald et al., 2007). However, molecular phylogenetic analyses presented in the same study indicated *P. kamillae* is nested within the Monodopsidaceae. The placement of *P. kamillae* in the monotypic family Loboceae would thus render the family Monodopsidaceae paraphyletic. Furthermore, the name "Loboceae" is invalid, as it does not conform to the International Code of Nomenclature for algae, fungi, and plants (ICN; article 18: "The name of a family … is formed from the genitive singular of a name of an included genus by replacing the genitive singular inflection … with the termination -aceae"; Turland et al., 2018). However, as elaborated below, the discrepancies between the historically developed classification of eustigmatophytes and their phylogeny are much more profound.

Phylogenetic relationships within the group have been addressed primarily by employing two molecular genetic markers, the nuclear 18S rRNA gene and of the plastid *rbcL* gene. Phylogenetic analyses of these two markers provide broadly consistent results in terms of the major splits in the eustig phylogenetic tree, although disagreement is observed at some nodes and statistical support is lacking for many branches (Amaral et al., 2020, 2021). Trees inferred from the 18S rRNA gene strongly support the existence of two principal lineages, one matching to the single formally established eustigmatophyte order, Eustigmatales, and the other comprised of more recently recognized eustigs and described under the PhyloCode as the clade *Goniochloridales* (Amaral et al., 2020; Fawley et al., 2014). Analyses of *rbcL* sequences support the split between these two groups, except for the genus *Paraeustigmatos*, which is resolved by analysis of *rbcL* sequence data as a third lineage separate from the Eustigmatales and the *Goniochloridales* (Fawley et al., 2019, 2021). However, 18S rDNA analysis or analysis of the combined 18S and *rbcL* data place *Paraeustigmatos* as a basal lineage within the Eustigmatales (Fawley et al., 2019). Crucially, phylogenetic analyses of concatenated sequences of plastid genome-encoded proteins provided full support for the Eustigmatales / *Goniochloridales* split, with *Paraeustigmatos* (as strain Mont 10/10-1w) branching with Eustigmatales rather than as an independent lineage (Ševčíková et al., 2019). It would thus seem natural to divide Eustigmatophyceae into two orders, but resolution of the status of the clade *Goniochloridales* is hampered by technical issues of formal taxonomy (Fawley et al., 2014) and has to be addressed in the future.

Putting *Paraeustigmatos* aside, cultured representatives of Eustigmatales are neatly divided into three major clades by both 18S rRNA and *rbcL* phylogenies (Amaral et al., 2020, 2021; Fawley et al., 2014). One clade corresponds to Hibberd's family Monodopsidaceae, expanded by the inclusion of *P. kamillae* (see above). The second clade, recognized for the first time by Fawley et al. (2014), did not correspond to any of the taxa defined in Hibberd's classification and was designated the "*Pseudellipsoidion* group" based on one of its constituent representatives. More recently this clade has been formalized as a new family named Neomonodaceae (Amaral et al., 2020). However, it was brought to our attention that the correct derivation of a family name from the genus name *Neomonodus* (considering the actual form of its genitive singular, see ICN article 18 and section 4.4 below) is Neomonodontaceae. We will use the corrected orthography in the rest of the paper. The third major Eustigmatales clade was dubbed the "Eustigmataceae group" by Fawley et al. (2014) to account for the fact that it includes representatives of multiple eustigmatophyte families (Chlorobotryaceae and Pseudocharaciopsidaceae, in addition to Eustigmataceae). Apart from eustigs assigned to particular species or genera, each major clade of the Eustigmatales contains lineages represented solely by unidentified strains (Fawley et al., 2014, 2021; Kryvenda et al., 2018). The richest clade in this regard is the "Eustigmataceae group", with several taxonomically unassigned subclades, including the deeply diverged "Clade Ia" comprising multiple genus-level lineages (Fawley et al., 2014, 2021). Thus, the "Eustigmataceae group" constitutes one of the major areas of eustig systematic biology for further research.

Here we address several key open questions concerning the diversity,

phylogeny, and taxonomy of the "Eustigmataceae group". One of the challenges to cope with, pertinent to eustigmatophyte research in general, is the need to harmonize the findings of modern investigations on eustig strains with the large body of work from the "pre-molecular" era of microalgal research, which led to the description of an impressive number of generally poorly documented taxa. As a result, it is often extremely difficult to decide whether an unidentified alga represents a previously described species or whether it is a novel organism. We gathered morphological, ultrastructural, and DNA sequence data from several previously reported, as well as novel eustig strains that clarified the phylogenetic position of the genus *Chlorobotrys* and provided strong evidence for the recognition of two new eustig species and genera. We also obtained complete plastid genome sequences from five representatives of the "Eustigmataceae group", including the first representative of the prominent but poorly characterised Clade Ia, that helped us to corroborate the monophyly of the group and to better define its internal relationships. The improved sampling of the eustig plastid genomes further illuminated the evolution of their salient traits, most notably of the peculiar six-gene *ebo* operon that was previously shown to occur in plastid genomes of certain eustigs as a result of horizontal gene transfer (HGT) from a lineage of bacterial endosymbionts (Yurchenko et al., 2016, 2018). The genomic data we generated also helped illuminate the occurrence of zoosporogenesis in the group. Finally, our careful analysis of the literature from the 19th century on led us to clarify the convoluted taxonomic history of eustigmatophyte algae and to solidify their classification by re-evaluating the "Eustigmataceae group" as the redefined family Chlorobotryaceae. Altogether, more than a century after it was conceived, we finally put the concept of the family Chlorobotryaceae on firm ground and construe this taxon as one of the principal and phylogenetically most diverse lineages of Eustigmatophyceae.

## 2. Material and methods

### 2.1. Isolation and cultivation of algal strains

The strain *Chlorobotrys* sp. FD2 was isolated in 2014 from a peatbog (47°51′15.696′'N and 6°41′8.916′'E; water pH 5.4, conductivity 29 μS/cm) located close to Lake Étang de la Goutte in France. *Chlorobotrys* sp. B2 and 2E5 were isolated in 2014 from the peatbog Černohorská rašelina located in the Czech Republic, and from the peatbog Klín located in Slovakia, respectively. The details on the origin of an additional *Chlorobotrys* strain, UP3 5/31–7 m, are provided in Fawley et al. (2021). The three European *Chlorobotrys* strains were cultivated in both prefiltered and sterilized bog water and in the modified DY IV liquid medium (Andersen et al., 1997). The strain *Chlorobotrys* sp. UP3 5/31–7 m was cultivated in liquid modified Woods Hole Medium (WH+; Fawley et al., 1990; Fawley and Fawley, 2017). Two unidentified eustigmatophyte strains SAG 2217 and SAG 2220 (Kryvenda et al., 2018) were obtained from the Culture Collection of Algae at the University of Göttingen, Germany (SAG), the unidentified eustigmatophyte strain CAUP Q 801 was obtained from the Culture Collection of Algae of Charles University, Prague, and the strain CCALA 278 (*Pleurochloris meiringensis*) was purchased from the Culture Collection of Autotrophic Organisms (CCALA), Institute of Botany of the AS CR, Třeboň, Czech Republic. These strains were cultivated in liquid Bold's Basal Medium (BBM; Bischoff and Bold, 1963). In addition, SAG 2220 was also cultivated on agar slants of WH + medium. CAUP Q 801 was also cultivated both in the liquid and on the agarised BBM medium in order to evaluate and describe the morphological plasticity and induce zoospores. For the latter purpose, cultivation in nitrogen (N)-free liquid BBM medium (withdrawing $NaNO_3$ from the recipe) was also applied. Both *N*-rich and *N*-poor liquid and *N*-rich agar cultures were exposed to different lengths of light–dark periods. An additional approach, a combination of the two, was also designed: first growing the strain on *N*-rich agarised BBM medium in Petri dish exposed to the continuous light and eventually switching to a dark chamber prior to supplying with the same liquid medium.

### 2.2. Light, electron, and confocal microscopy

The algal strains were studied under light and differential interface contract (DIC) microscopy with an Olympus BX53 (Tokyo, Japan). Microphotographs were taken with an Olympus DP73 (Tokyo, Japan) digital camera. Additional images were taken with either a Nikon N*i*U or Nikon E600 microscope with an Olympus SC180 camera. Olympus cellSens Imaging Software v1.6 was used to process images and obtain morphometric measurements of the cells. For transmission electron microscopy (TEM), samples were fixed for 2 h at 5 °C in a 2% solution of glutaraldehyde in BBM, post-fixed for 2 h at 5 °C in 1% osmium tetroxide in 0.05 mol/L phosphate buffer and at 5 °C in 1% uranyl acetate in 50% methanol overnight. After dehydration through ethanol series (70%, 96%, 100%), cells were embedded in Spurr's medium (Spurr, 1969) via butan-1-ol. Ultrathin sections, cut with a diamond knife on an Ultracut E (Reichert-Jung, Wien, Austria), were post-stained with lead citrate and examined using a JEOL 1011 TEM (JEOL Ltd., Tokyo, Japan). Microphotographs were obtained using a Veleta CCD camera equipped with image analysis software (Olympus Soft Imaging Solution GmbH). For confocal microscopy, a Leica TCS SP8 laser scanning confocal microscope (Leica Microsystems, Wetzlar, Germany) equipped with an Argon-Krypton laser was used. We applied a 488 nm excitation line passing 488/561/633 main beam splitter. Emitted light between 600 and 750 nm was collected using the Internal Hybrid Detector. The autofluorescence of chlorophyll was exploited for visualization of the plastid structure. The algal cells were embedded in agarized medium to minimize their moving, and scanned by an HC PL APO CS2 63x/1.20 water immersion objective. A series of optical sections through plastids were captured and used for 3-dimensional reconstruction of their morphology. The plastid reconstructions were produced by the ImageJ 1.34p program (Abràmoff et al., 2004), using the "Volume viewer" plugin.

### 2.3. DNA isolation and sequencing

Total genomic DNA from the cultures of SAG 2217 and SAG 2220 was extracted using an Invisorb® Spin Plant Mini Kit (STRATEC Molecular GmbH, Berlin, Germany), following the manufacturer's instructions; from the cultures of CAUP Q 801 and *Chlorobotrys* sp. FD2 by the modified Dellaporta et al. (1983) protocol, including additional steps of RNAse and proteinase K treatments; from the culture of WTwin 8/9 T-6m6.8 (details reported in Fawley et al., 2021) using the procedure described in Fawley and Fawley (2004). The isolated DNA samples were sent to Macrogen, Inc. (Seoul, South Korea) for library construction with TruSeq Nano DNA Kit (insert size 350 bp) and Illumina NovaSeq 6000 platform sequencing, resulting in total of 43,022,716 (SAG 2217), 40,515,452 (SAG 2220), 46,603,880 (CAUP Q 801), 54,717,852 (FD2), and 41,906,188 (WTwin 8/9 T-6m6.8) 150 bp paired-end reads.

For the purpose of sequencing individual phylogenetic markers, genomic DNA was additionally isolated from the three *Chlorobotrys* sp. strains: from UP3 5/31-7m as described in Fawley and Fawley (2004) and from B2 and 2E5 using InstaGene matrix (Bio-Rad Laboratories, Inc.), following the manufacturer's instructions. DNA isolation from the CCALA 278 strain was done as described in Přibyl et al. (2012). The nuclear 18S rDNA of *Chlorobotrys* sp. UP3 5/31-7m was amplified and sequenced with the primers NS1-X and 18L-X (Phillips and Fawley, 2000), NS5 (White et al., 1990), and the new primer, 18–600 (5′-CGAAATCCAACTACGAGC-3′). 18S rDNA of *Chlorobotrys* sp. B2 and 2E5 and the CCALA 278 strain was amplified and sequenced using universal eukaryotic primers 18SF and 18SR (Katana et al., 2001) in combination with eustigmatophyte specific primers EustigF1 (5′-GACAATAAATAACAATGCCGG-3′) and EustigR1 (5′-GTTATAAACTCGTTGAACGCA-3′), including additional internal sequencing primers (Katana et al., 2001).

The plastid *rbcL* gene of the latter two *Chlorobotrys* sp. strains was amplified with the primer pair EU-rbcL-F3 (5′-ACGTTATGAATCAGGTGTAATC-3′) and EU-rbcL-R3 (5′-CTGTATCAGTTGATGAGTAGTTG-3′), and additionally sequenced with internal primer pair EUSrbcL-sF2 (5′-ACAAATGCACCAATTACTTAA-3′) and EUSrbcL-sR1 (5′-AACGCATGAAWGGTTGWGAGTT-3′) specific for eustigmatophytes. In addition, using a previously reported DNA prep (Yurchenko et al., 2018) and primers EU-rbcL-F1 and EU-rbcL-R1 (Fawley et al., 2015) we amplified the *rbcL* gene from *Pseudostaurastrum* sp. 10174 and sequenced it using the extra internal primers EUSrbcL-sF1 (5′-AACTCWCAACCWTTCATGCGTT-3′) and EUSrbcL-sR1. All PCR reactions were done using MyTaq Red Reaction Buffer and MyTaq Red DNA Polymerase (Bioline). The obtained PCR products were purified with Gel/PCR DNA Fragments Extraction Kit (Geneaid). The assembled sequences were deposited in the GenBank under the accession numbers ON924315–ON924322 (18S rDNA) and ON920848–ON920850 (*rbcL*).

### 2.4. Genome assembly and analyses

The raw Illumina reads were trimmed with Trimmomatic v0.39 (Bolger et al., 2014) and assembled using SPAdes v3.13.0 (Bankevich et al., 2012). The assemblies, together with the one obtained previously for *Characiopsis acuta* ACOI 456 as part of the plastid genome sequencing of this species (Ševčíková et al., 2019), were deposited at Figshare (https://doi.org/10.6084/m9.figshare.20286678.v1). Scaffolds derived from plastid genomes were identified with BLAST (Altschul et al., 1997) using standard plastid genes as queries. A conventional plastid genome architecture was considered and confirmed by the detection of an inverted repeat region of a double read coverage. Illumina reads were subsequently mapped to the assembled genomes using Bowtie2 v.2.3.4.1 (Langmead and Salzberg, 2012) and the result was inspected in Tablet v.1.14.04.10 (Milne et al., 2013) to check the accuracy of the assembled sequences. Validated genome sequences were annotated with MFannot (https://megasun.bch.umontreal.ca/cgi-bin/mfannot/mfannotInterface.pl) using the genetic code 11 (Bacterial, Archaeal and Plant Plastid). The obtained annotation was checked manually and issues were fixed, including correction of misidentified coding sequence starts and identification of some of the *orf*s not recognised by MFannot as homologs of standard plastid genes. The tRNA genes assigned by MFannot as *trnM* were checked and reannotated as initiator fMet-tRNA, elongator eMet-tRNA and Ile-tRNA (AUA-decoding, with the CAU anticodon modified by lysidinylation). A few noncoding RNA genes completely missed by MFannot, including *ssrA* and an intron-containing *trnL(uaa)*, were annotated manually. The borders of the annotated genes were validated comparing them with homologs from previously studied eustigmatophyte plastomes (Ševčíková et al., 2019). All four unique plastid genomes (only one for the SAG 2220/2017 strain pair with exactly identical plastomes) were deposited in GenBank under accession numbers ON929294–ON929296 and ON938208. Graphical maps of the new plastid genomes were prepared using OGDRAW v.1.3.1 (https://chlorobox.mpimp-golm.mpg.de/OGDraw.html; Greiner et al., 2019).

The total genome assemblies (including the newly obtained ones and the previously generated one from *C. acuta* ACOI 456) that contain incomplete nuclear genome sequences of varying contiguity and read coverage were searched with TBLASTN to identify genes corresponding to the Mas family proteins (Hee et al., 2019) and heavy chains of flagellar dyneins (Kollmar, 2016). For the Mas genes the corresponding regions were extracted and the exon–intron structure of the genes was deduced manually, facilitated by sequence conservation and guided by comparison with homologs from *Vischeria* sp. C74 and *Monodopsis* sp. C73 and C141, which were annotated as part of the respective genome sequencing projects (Yang et al., 2021) and verified by transcriptome data. For the dyneins only the presence of the different paralogs in the respective genome assembly was monitored, with the identity of the genes verified by reciprocal BLASTX searches against a previously published reference dataset including dynein heavy chains from a broad sample of eukaryotes (Kollmar, 2016) expanded by the addition of manually curated dynein heavy chain protein sequences from *Vischeria* sp. C74.

### 2.5. Phylogenetic analyses

An 18S rDNA sequence alignment was built by combining the newly obtained data with the Eustigmatales and *Goniochloridales* sequences previously gathered by Amaral et al. (2020, 2021), including a slight modification, i.e. using our assembled complete 18S rDNA sequence of the SAG 2220 instead of the one published by Kryvenda et al. (2018). A *rbcL* alignment (hereafter rbcL-ref) was built by expanding the reference dataset from Fawley et al. (2021) by adding the new individually determined sequences and those extracted from the plastome assemblies reported here. To further improve the taxonomic sampling of the *rbcL* gene, we also used sequences extracted from our unpublished plastomes of *Pseudostaurastrum enorme* SAG 11.85, *Pseudostaurastrum limneticum* SAG 14.94, and *Pseudotetraëdriella kamillae* SAG 2056; the sequences were deposited to GenBank with accession numbers ON920851–ON920853. The sequences of the two separate datasets were aligned by MAFFT v7 available online (https://mafft.cbrc.jp/alignment/software/) employing the Auto strategy (Katoh and Standley 2013). The resulting alignments were trimmed manually in BioEdit v7.2.5 (Hall 1999). The final 18S rDNA alignment contained 120 sequences and 1760 aligned positions, the final rbcL-ref alignment contained 155 sequences and 1428 aligned positions.

Maximum likelihood (ML) analyses were conducted with IQ-TREE multicore version 2.0.3 (Minh et al., 2020) applying the best evolutionary models for the data selected by the program using ModelFinder (Kalyaanamoorthy et al., 2017). The selected models were as follows: TN + F + R3 for the 18S rDNA dataset, and GTR + F + I + G4 for the first, TNe + R3 for the second, and TIM3 + F + I + G4 for the third codon position for the rbcL-ref alignment. Statistical support values for tree topologies were assessed applying nonparametric bootstrapping with 100 replications. Bayesian inference (BI) was performed with MrBayes v3.2.7 (Ronquist and Huelsenbeck, 2003) using the GTR + Γ + I and GTR + Γ models for the rbcL-ref and 18S rDNA datasets, respectively. Two Markov chain Monte Carlo runs (MCMC) for 2,000,000 generations with trees sampled every 100 generations were performed, with the first 25% of generated trees discarded as burn-in. The obtained posterior probabilities were used to assess branch support.

A second *rbcL* sequence alignment, hereafter rbcL-env, was built by expanding the rbcL-ref alignment with partial (370 bp) *rbcL* sequences obtained by Fawley et al. (2021) as community metabarcodes with eustig-targeting amplification primers. Metabarcode sequences that constituted suspiciously long branches or turned out to occupy an unstable position in different preliminary trees were examined to detect possible chimeric sequences missed in the chimera check by Fawley et al. (2021). To this end the sequences were compared by BLASTN against the NCBI nr nucleotide database and in parallel to the in-house *rbcL* sequence database containing also the metabarcode sequences. The pairwise alignments of the query with the non-self best hits were inspected to find out if different regions of the alignment exhibit pronounced differences in the degree of similarity of the sequences aligned. When this was the case, such regions were used as separate queries in BLASTN searches. Eighteen metabarcode sequences proved to contain regions (in the 5′ or 3′ part) that were more similar to *rbcL* from non-eustigmatophytes (different algae and even bacteria); in most cases these regions were highly similar to or identical with *rbcL* sequences from a particular non-eustigmatophyte taxon. Eleven additional metabarcode sequences were identified as apparently embracing regions of the *rbcL* gene derived from different major eustig subgroups. These 29 sequences (with further details listed in Table S1) are putative chimeras and were thus removed from final analyses. The final rbcL-env alignment, constructed in the same way as the rbcL-ref alignment, contained 438 sequences and 1428 aligned positions. Tree inference was

conducted with IQ-TREE using ultrafast bootstrapping with 1000 replications and GTR + F + R5 as the best-fit model for the dataset.

Alignments of plastid genome-encoded proteins originally built by Ševčíková et al. (2019) and updated by Barcytė et al. (2021) were further expanded by incorporating the newly sequenced organisms, resulting in 69 conserved protein-coding genes coming from 57 taxa, including a total of 21 eustigmatophyte algae. Single-gene datasets were aligned using MAFFT with the E-INS-I method and trimmed with trimAL (Capella-Gutiérrez et al., 2009) with the -gappyout mode. The resulting alignments were concatenated by FASconCAT-G_v1.05 (Kück and Longo, 2014), yielding a final supermatrix of 17,753 aligned positions. The ML phylogenetic tree was inferred using the IQ-TREE with the LG + F + I + G4 model and 100 nonparametric bootstraps. The ML phylogeny of the Mas protein family was obtained by preparing the alignment and running the analysis in the same manner as described for plastid proteins. The final alignment consisted of 343 aligned positions, and the model WAG + I + G4 was chosen as best-fitting for the dataset.

All resulting and here presented phylogenetic trees were visualized

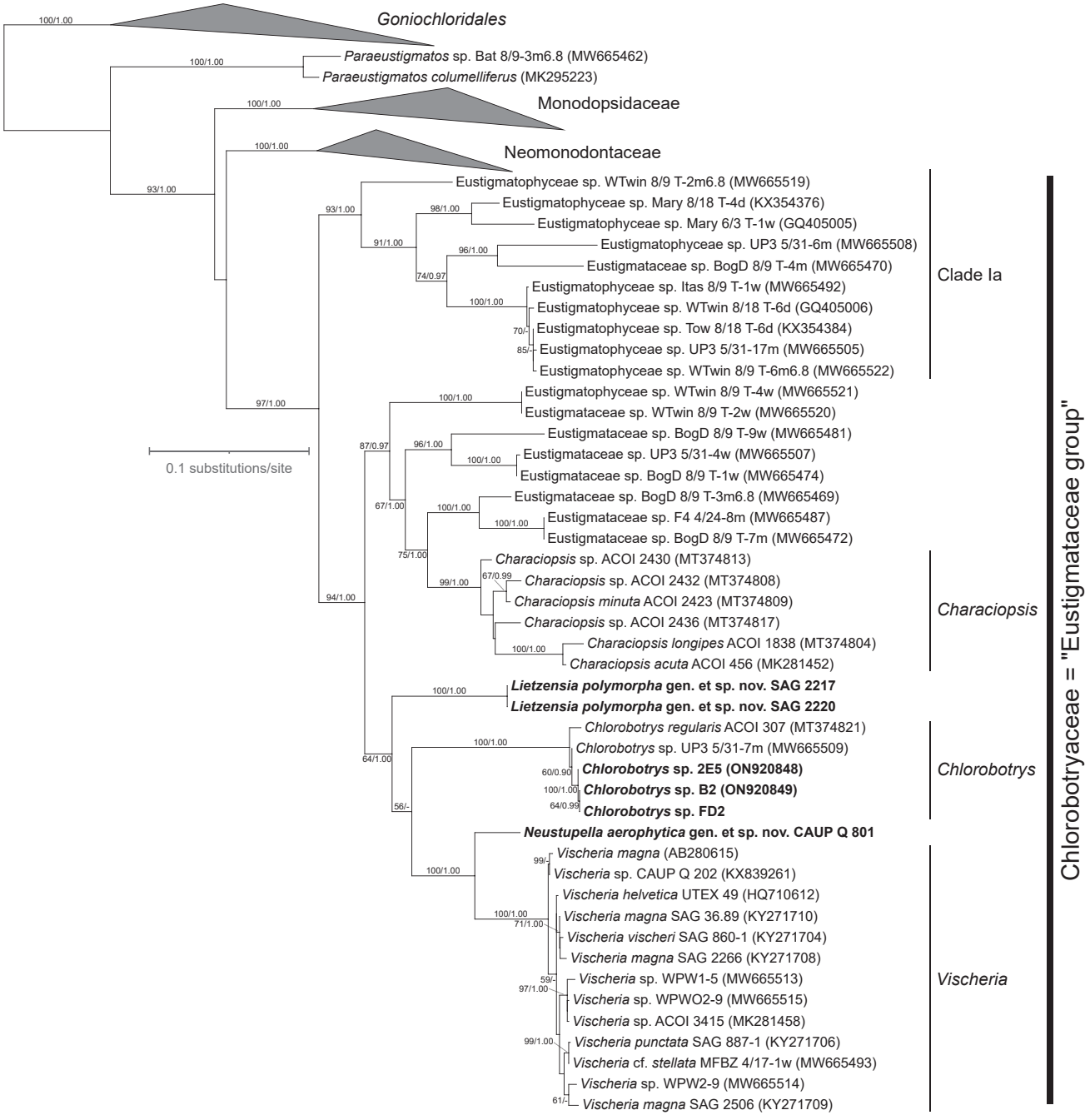


**Fig. 1.** Maximum likelihood phylogenetic tree inferred from eustigmatophyte *rbcL* sequences. The alignment used for the tree inference consisted of 1428 aligned nucleotide positions, the substitution model employed was GTR + F + I + G4 for the first, TNe + R3 for the second, and TIM3 + F + I + G4 for the third codon position. The root is arbitrarily placed between Eustigmatales and the clade *Goniochloridales* based on the previous conclusions and results of the phylogenetic analysis of plastome-encoded proteins (see Fig. 3). Taxa for which the *rbc*L sequence is newly reported in this study are in bold. Sequences with no accession number indicated were extracted from the respective full plastid genome sequences (listed in Table S5). The numbers at branches correspond to bootstrap support values (shown when ≥ 50%) and Bayesian posterior probabilities (shown when > 0.85). The triangles represent larger non-focal clades collapsed for the sake of simplicity. The full version of the tree is available as Fig. S1.

with the Interactive Tree of Life (iTOL) v6 (https://itol.embl.de/; Letunic and Bork, 2021) and postprocessed with Inkscape v1.0.1. Sequence alignments used to infer the trees are available at Figshare (https://doi.org/10.6084/m9.figshare.20286723.v1).

## 3. Results

### 3.1. Phylogeny of the "Eustigmataceae group" based on rbcL and 18S rRNA gene sequences

The ML and BI phylogenies based on the single-gene datasets of *rbcL* (Fig. 1, Fig. S1) and 18S rDNA (Fig. 2, Fig. S2) sequences confidently placed all the studied strains within the previously delimited "Eustigmataceae group" of the order Eustigmatales. The branching order of the main "Eustigmataceae" clades was incongruent between the two datasets but it was in agreement with the previous studies employing similar datasets and showing the same inconsistencies (Amaral et al., 2020, 2021). All *Chlorobotrys rbcL* sequences formed a fully supported clade. The same clade was well supported by the 18S rDNA sequence analysis; however, an additional 18S rDNA sequence from the *Chlorobotrys regularis* strain CCAP 810/1 (reported by Fawley et al., 2014) together with a nearly identical sequence from the strain CAUP Q 801 was resolved as a lineage sister to the genus *Vischeria*. The *rbcL* tree supported this relationship of CAUP Q 801 and *Vischeria*; however, the *rbcL* sequence from *C. regularis* CCAP 810/1 is unavailable. The position of the "main" *Chlorobotrys* clade within the "Eustigmataceae group" differed between the two phylogenies with no statistical support in either analysis. The phylogenetic position of the pair of unidentified eustigmatophytes SAG 2217 and SAG 2220 was likewise variable between the two trees and lacked significant support values to firmly consider their sister relationship with any of the main "Eustigmataceae" clades. These two strains exhibited an identical *rbcL* (and the whole plastid genome, see below) sequence, whereas the previously published 18S rDNA sequences (Kryvenda et al., 2018) that we used in the tree differ by a few mismatches (including single-nucleotide indels). These differences apparently resulted from sequencing errors in the SAG 2220 sequence (KY271668.1), as confirmed by the comparison of the sequence to our partial genome assembly from the strain (see below). Finally, the

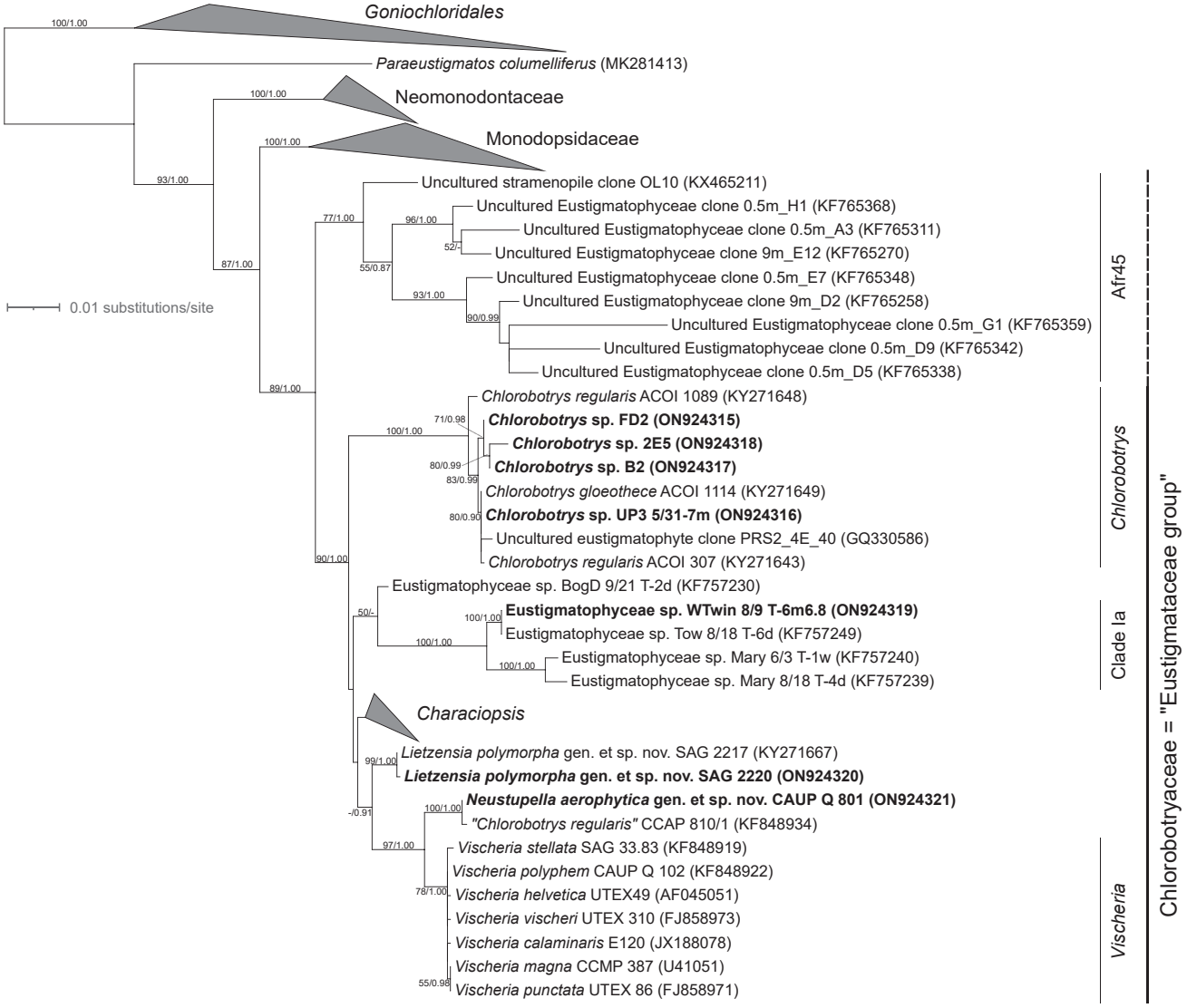


**Fig. 2.** Maximum likelihood phylogenetic tree inferred from eustigmatophyte 18S rRNA gene sequences. The alignment used for the tree inference consisted of 1760 aligned nucleotide positions, the substitution model employed was TN + F + R3. Taxa for which the 18S rDNA sequence is newly reported in this study are in bold. The rooting and display conventions are the same as for the tree in Fig. 1. The clade denoted "Afr45" consists of environmental DNA clones from African lakes and comprises groups 4 and 5 as delimited by Villanueva et al. (2014) plus the "uncultured stramenopile clone OL10" reported by Luo et al. (2017). The full version of the tree is available as Fig. S2.

unidentified strain WTwin 8/9 T-6m6.8 was placed by both markers among a suite of other morphologically uncharacterized eustigs previously denoted as Clade Ia (Fawley et al., 2014).

Several additional aspects of the *rbcL* and 18S rDNA trees are worth mentioning beyond the relationships in the "Eustigmataceae group". Firstly, a clade consisting solely of environmental 18S rDNA clones from east African lakes (Luo et al., 2017; Villanueva et al., 2014) and denoted here Afr45 emerged as the immediate sister group of the "Eustigmataceae group" in its original scope (Fig. 2). Secondly, the newly determined *rbcL* sequence from *Pseudotetraëdriella kamillae* placed this species into the family Monodopsidaceae as a sister lineage of a *Monodopsis/Monodus* clade (Fig. S1), consistent with the results of 18S rDNA-based analyses (Fig. S2) and further invalidating the view of *P. kamillae* as a representative of a separate family (see Introduction). Thirdly, we for the first time included in the phylogenetic analysis of the *rbc*L gene sequences from the genus *Pseudostaurastrum*. The result (Fig. S1) agreed with the outcome of the 18S rDNA-based phylogeny (Fig. S2) and confirmed the monophyly of all three representatives sampled, as well as the position of *Pseudostaurastrum* as a *Goniochloridales* lineage separate from the Clades IIa, IIb, and IIc.

A recent metabarcoding (environmental DNA-based) study has uncovered a tremendous diversity of eustigmatophyte *rbcL* genotypes in nature (Fawley et al., 2021), and we were interested to see whether any of the previously unidentified genotypes can be matched to the eustig strains newly studied here. To this end we expanded the *rbcL* dataset analysed by Fawley et al. (2021) by adding the previously missing sequences from the CAUP Q 801 and SAG 2220/2017 strains as well as from *P. kamillae* and the representatives of the genus *Pseudostaurastrum*. On the other hand, we excluded 29 metabarcode sequences from the dataset of Fawley et al. (2021), since we recognized them as obvious or putative chimeras (combining regions from different eustig lineages or even including segments derived from other organismal groups; Table S1). In contrast to the previous study, which reconstructed the *rbcL* phylogeny by trimming the length of the alignment analysed to the length of the metabarcode amplicons (370 bp), we kept the alignment at the length of the reference *rbcL* sequences (1428 bp).

The resulting tree (Fig. S3) is generally congruent with the tree reported by Fawley et al. (2021), but provides a much stronger evidence for the notion that the group of environmental genotypes referred to by these authors as "possible Eustigmatophyceae" indeed belongs to the class (the monophyly of Eustigmatophyceae including the aforementioned environmental sequence group receiving bootstrap support of 80%). To facilitate communication, we label this group the "clade X", whereas the second cluster consisting purely from environmental genotypes, in the previous study designated "uncertain Eustigmatales", is here called "clade Y" (Fig. S3). All the other eustigmatophyte metabarcode or metagenomic *rbcL* genotypes fall into the major established clades (Monodopsidaceae, Neomonodontaceae, the "Eustigmataceae group", *Goniochloridales*) or are affiliated to the stand-alone genus *Paraeustigmatos*. Clade X is placed sister to all (other) eustigmaotphytes, whereas clade Y is nested within Eustigmatales. Notably, one of the previously unidentified genotypes, UVA_Wise_retention_pond_1991_ASV 546, has now emerged to belong to the genus *Pseudostaurastrum* as a close relative of *P. enorme* SAG 11.85. The other newly added sequences from cultured algae, however, neither match nor are closely related to any of the environmental genotypes. These results further underscore the previously noticed surprisingly low overlap between the eustigs studies based on cultured material and those documented by metabarcoding (Fawley et al., 2021).

### 3.2. Plastid genome evolution and multigene phylogeny of the "Eustigmataceae group"

The statistical support for the monophyly of the "Eustigmataceae group" provided by the single-gene phylogenies is high but not full (Figs. 1 and 2). To further test the robustness of this clade by a multigene phylogenetic analysis, we used Illumina HiSeq to obtain complete plastid genomes, as well as partial nuclear genome data (and mitochondrial genome sequences not studied here), from five eustigmatophyte strains: CAUP Q 801, SAG 2217, SAG 2220, *Chlorobotrys* sp. FD2, and WTwin 8/9 T-6m6.8. All five newly determined plastomes exhibit the conventional circular-mapping architecture with inverted repeats of varying length (Table 1). Maps showing identified and annotated genes and other genomic features are provided as Fig. S4. The plastome sequences of SAG 2217 and SAG 2220 were completely identical. Thus, we refer in the subsequent discussion only to one of the strains, SAG 2220.

The gene content of the newly sequenced plastomes does not depart in any unexpected way from the highly conserved set of eustigmatophyte plastid genes defined by previous studies (see Ševčíková et al., 2019). All four plastomes have a nearly identical set of non-coding RNA genes, the only difference being the absence of the *trnL(caa)* gene from the strain WTwin 8/9 T-6m6.8. The number of protein-coding genes varies between 130 and 135. One cause of the variation is the exclusive presence of a hypothetical open reading frame (*orf187*, potentially encoding a protein 187 amino acid residues long) inserted between the genes *thiS* and *rbcL* in the plastome of the strain WTwin 8/9 T-6m6.8 (Fig. S4). We could not identify any homology of the hypothetical protein product of this *orf* to other proteins or protein domains even when using the highly sensitive tool HHpred (Söding, 2005), precluding to make any conclusions of the function and even functionality of the *orf*. The other reason for the gene content difference is the differential retention of the *psbZ* gene, which is present only in WTwin 8/9 T-6m6.8 among the four plastomes analysed, and of the *acpP* gene retained only by SAG 2220. Finally, of the newly sequenced eustigs only CAUP Q 801 and *Chlorobotrys* sp. FD2 exhibit the so-called *ebo* operon, a six-gene cluster previously identified in plastomes of the genera *Vischeria* and *Monodopsis* and acquired by horizontal gene transfer from the lineage of eustigmatophyte-specific endosymbiontic bacteria, *Candidatus* Phycorickettsia (Ševčíková et al., 2019; Yurchenko et al., 2016, 2018).

We expanded the previously employed dataset of conserved plastome-encoded proteins (Ševčíková et al., 2019) by including the newly generated plastome sequence data and used the resulting alignment to infer a multigene phylogeny with the most comprehensive

**Table 1**
Features of the newly sequenced eustigmatophyte plastid genomes.

| | Strain CAUP Q 801 | Strain SAG 2220 (=SAG 2217) | *Chlorobotrys* sp. FD2 | Strain WTwin 8/9 T-6m6.8 |
|---|---|---|---|---|
| size (bp) | 125957 | 120721 | 126383 | 119531 |
| inverted repeat (bp) | 9641 | 9844 | 9644 | 9858 |
| LSC region (bp) | 62004 | 55723 | 62340 | 55105 |
| SSC region (bp) | 44671 | 45310 | 44755 | 44710 |
| total GC content (%) | 32.99 | 33.97 | 33.39 | 32.44 |
| gene content (total*) | 167 | 162 | 167 | 162 |
| common conserved plastid protein-coding genes | 127 | 128 | 127 | 128 |
| conserved group-specific genes (*ycf95, orf1_eust*) | 2 | 2 | 2 | 2 |
| *ebo* operon genes | 6 | 0 | 6 | 0 |
| ORFs without homologs | 0 | 0 | 0 | 1 |
| rRNA genes | 3 | 3 | 3 | 3 |
| tRNA genes | 28 | 28 | 28 | 27 |
| other noncoding RNA genes (only *ssrA*) | 1 | 1 | 1 | 1 |
| number of genes in inverted repeat | 12 | 12 | 12 | 12 |

* Genes present in inverted repeat (IR) are counted just once. The *clpC* gene split into two separate open reading frames in eustigmatophytes (*clpC_A* and *clpC_B*) is counted as one gene. Abbreviations: LSC, long single-copy; SSC, short single-copy; ORFs, open-reading frames.

representation of eustigmatophytes to date (Fig. 3A). The monophyly and the internal topology of the eustigmatophyte subtree, fully supported with the exception of a single branch commented on below, confirms the previous result and divides the Eustigmatophyceae into two deeply diverged clades, *Goniochloridales* and the order Eustigmatales. Four principal lineages can be recognized within Eustigmatales: the most deeply diverged lineage represented by *Paraeustigmatos cellumeliferus*; the Neomonodontaceae represented by *Pseudellipsoidion edaphicum*; and two sister subclades that correspond to the Monodopsidaceae and the "Eustigmataceae group". Within the "Eustigmataceae group", the strain CAUP Q 801 is resolved as sister to the genus *Vischeria*, with both united into a higher-order grouping with *Chlorobotrys* sp. FD2 and the pair of SAG strains. *Characiopsis acuta* is then sister to this larger grouping, and WTwin 8/9 T-6m6.8, representing Clade Ia, branches as the deepest lineage of the "Eustigmataceae group". Monophyly of the Eustigmatophyceae and the relationships within this class received maximal support in the analysis, except for the branching order of the *Vischeria* plus CAUP Q 801 clade, *Chlorobotrys* sp. FD2 and the SAG isolates, with the latter two forming a clade with only moderate bootstrap support (75%).

Beyond the phylogeny of the eustigmatophytes, our results illuminate relationships of other ochrophytes. Above all, our analysis recovers with high bootstrap support the sisterhood of the Pinguiophyceae and *Olisthodiscus luteus* (Olisthodiscophyceae), which was observed with weak statistical support in a previous analysis based on a much smaller set of plastid genes (Barcytė et al., 2021). In addition, this is the first full plastome-based analysis to include *Schizocladia ischiensis*, the sole known representative of the ochrophyte class Schizocladiophyceae. *Schizocladia* is placed as a sister lineage of brown algae (Phaeophyceae) in our tree (Fig. 3A), confirming the results of previous analyses based on much more limited sequence data (Kawai et al., 2003; Yang et al., 2012).

The inferred relationships and the expanded dataset of sequenced plastomes allow us to refine the picture of the changes in the plastid gene content along the phylogeny of the "Eustigmataceae group" (Fig. 3B). After the divergence of Clade Ia (represented here by the WTwin 8/9 T-

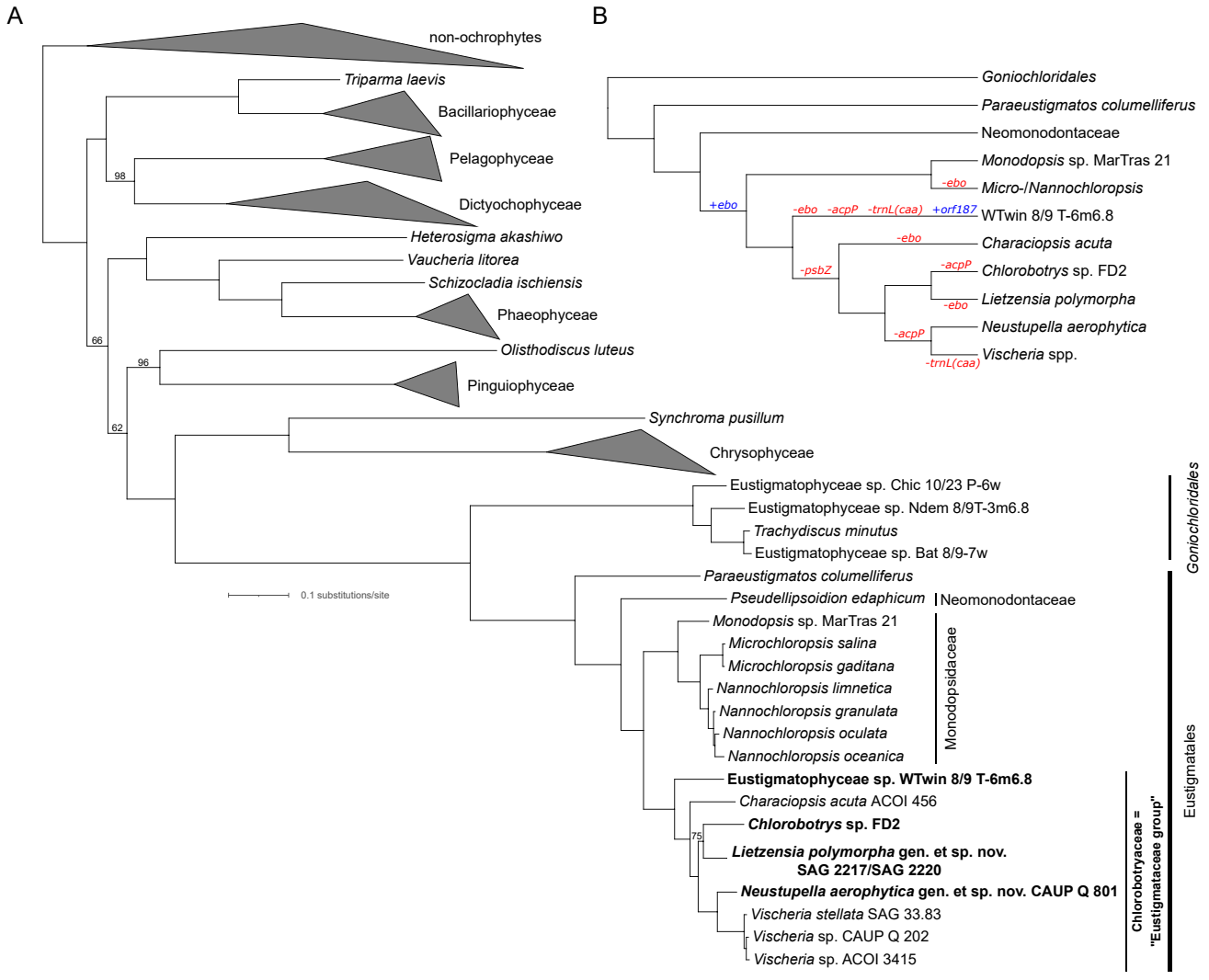


**Fig. 3.** (A) Maximum likelihood phylogenetic tree inferred from a concatenated dataset of 69 plastome-encoded proteins from eustigmatophyte and selected other ochrophytes representing the principal lineages of the group. The alignment used for the tree inference consisted of 17,753 amino acid positions, the substitution model employed was LG + F + I + G4. Bootstrap support values are 100 for all branches if not indicated otherwise in the tree. The root is placed between an outgroup (haptophytes, cryptophytes, a red alga, and a glaucophyte) and ochrophytes. Taxa with the plastid genome newly sequenced in this study are in bold. The full version of the tree is provided as Fig. S8, sources of sequence data used to generate the tree are provided in Table S5. (B) Plastid gene gain and loss mapped onto a schematic phylogeny of eustigmatophytes (based on the tree shown in panel A), with the focus on Chlorobotryaceae (=the "Eustigmataceae group"). Gains are in blue, losses in red, "*ebo*" refers to a cluster of six genes (*eboA* to *eboF*) gained or lost as a whole. Gains/losses within the lineages comprised of multiple species (*Goniochloridales*, Neomonodontaceae, *Micro-/Nannochloropsis*, *Vischeria* spp.) are not indicated for simplicity.

6m6.8 strain), *psbZ* was lost once, whereas *acpP* seems to have been lost independently-three times in the "Eustigmataceae group" (with three additional loss events inferred previously for other eustig branches; Ševčíková et al., 2019). The propensity of the gene for loss is explained by the existence in all eustigmatophytes of a presumably functionally redundant nucleus-encoded version of the protein gained by HGT from *Candidatus* Phycorickettsia (Ševčíková et al., 2019). Our expanded analysis also increases the minimal number of independent losses invoked to explain the distribution of the *ebo* operon in eustig plastomes from the previously inferred two (Ševčíková et al., 2019) to four (Fig. 3B). As the physiological function of the *ebo* genes remains undefined, the evolutionary factors behind the recurrent *ebo* operon loss cannot be deduced at present. Finally, the absence of the *trnL(caa)* gene in the strain WTwin 8/9 T-6m6.8 implies one additional independent loss of this gene in addition to the three inferred previously (in *Vischeria* and two lineages outside the "Eustigmataceae group"; Ševčíková et al., 2019). Disregarding the gene losses and the gain of *orf187* in the WTwin 8/9 T-6m6.8 lineage, there are no additional changes in the gene order among the plastomes of different members of the "Eustigmataceae group". Moreover, comparisons to plastomes of the sister clade Monodopsidaceae indicate that the gene order has remained the same at least since the common ancestor of the two clades.

### 3.3. Morphology and ultrastructure of the newly studied eustigmatophyte strains

Vegetative cells of the strain CAUP Q 801 exhibited a mostly circular to sometimes oval or somewhat triangular outline (Fig. 4A, B), and were solitary or rarely in groups of two (Fig. 4C). The cell size ranged from 5.0 to 12.5 μm in diameter (n = 100) when grown in liquid medium. In contrast, the maximum diameter for cells cultivated on agar reached 16 μm. Some of the cells grown on agar appeared to have an undulating pattern to the cell surface giving them an angular look (Fig. 4D). The cell wall of CAUP Q 801 was smooth, without ornamentation, protrusions, or mucilage. Young cells contained one or two massive plastids that were parietal and trough-, bowl-, or cup- shaped. The plastids of mature cells typically were deeply lobed with irregular incisions, giving the false appearance of multiple plastids. However, the lobes were always connected at the base of the plastid (Fig. 4E). One to two irregularly polyhedral pyrenoids protruded from each of the plastids (Fig. 4A; arrowheads). Some of the mature cells also contained a prominent vacuole (Fig. 4A, F). Reddish globules, appearing either as a solid unit (Fig. 4G), or distinctly composed of several separate droplets (Fig. 4H), were visible in some cells. Lamellate vesicles (Fig. 4I; arrows) and unidentified granular material (Fig. 4A) were noticeable as well. In addition, numerous shiny crystals (possibly guanine crystals), were evident especially when observed with DIC optics (Fig. 4J). Reproduction typically occurred by 2–4 autospores (Fig. 4K, L). After their release from the mother cell wall, autospores ranged in size from approximately 3.5–5.0 μm in diameter.

Zoospores were never observed for the CAUP Q 801 strain when cultivated in liquid or agarised BBM. However, if we cultivated the alga on agar in a Petri dish under light, and then added liquid medium and incubated the culture in the dark for 1–2 days, we observed the release of zoospores in the liquid medium. The zoospores were elongated and cylindrical in shape. They contained a single trough-shaped plastid without a pyrenoid that occupied two-thirds of the cell volume and an extraplastidial eyespot located at the anterior extremity (Fig. 4M–O).

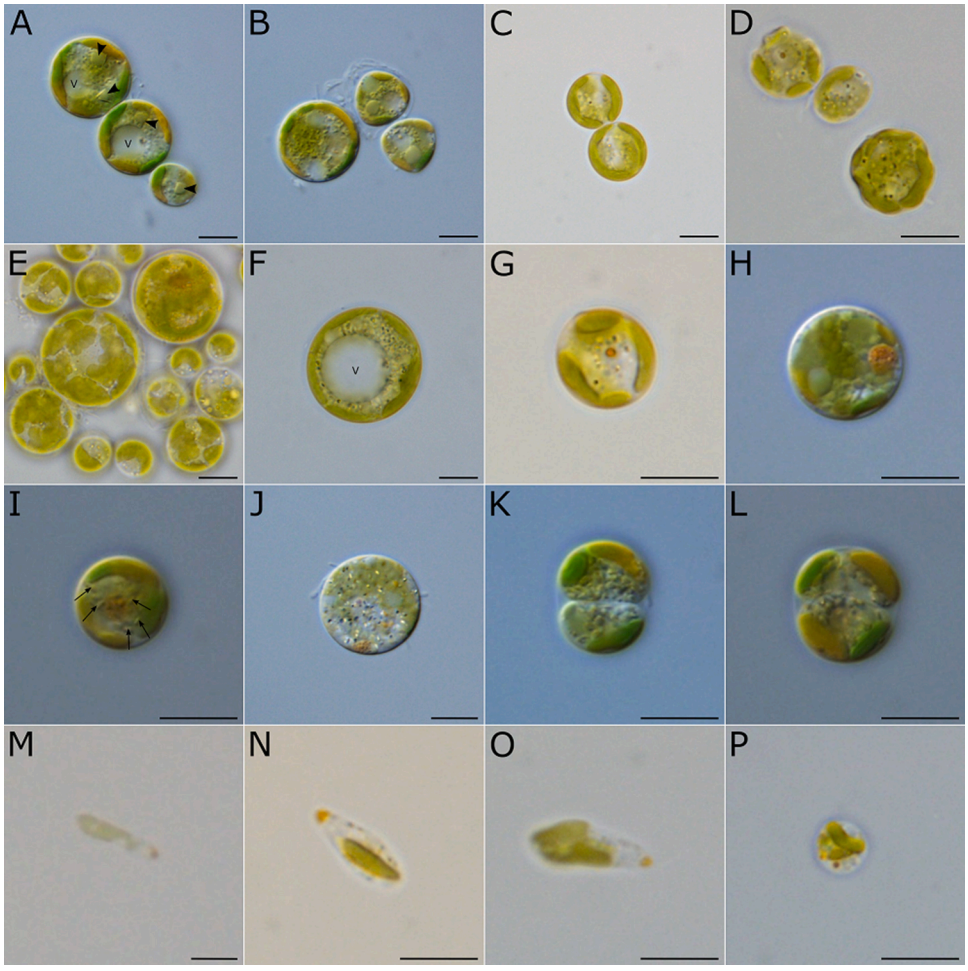


**Fig. 4.** Light micrographs of *Neustupella aerophytica* gen. et sp. nov., strain CAUP Q 801. (A) Vegetative spherical cells of different sizes with prominent pyrenoids (arrowheads) and vacuoles (v); (B) Young triangular cells; (C) Temporary grouping of two cells; (D) Vegetative cells with an undulating cell outline; (E) Plastids with deep incisions and irregular cracks; (F) A prominent vacuole filling the half of the cell's volume; (G) Small reddish globule appearing as a solid unit; (H) Reddish globule composed of several separate droplets; (I) Lamellate vesicles (arrows); (J) Shining crystals under DIC; (K) Asexual reproduction by two autospores; (L) Reproduction by four autospores; (M) Elongated-cylindrical zoospore before being exposed to bright light; (N—O) Spindle-shaped zoospores shortly after being exposed to light with prominent extraplastidial eyespots; (P) Rearranged zoospore with a dismantled eyespot and spherical outline. Scale bars = 5 μm.

Their size reached up to approximately 12.0 μm in length and 3.0 μm in width. Zoospores moved very quickly and possessed a single visible flagellum. When the zoospores were exposed to brighter light, the protoplast was immediately reduced in size. Initially this change made the cells appear spindle-shaped or broadly ellipsoidal, after which the zoospores transitioned to the spherical form (Fig. 4P). The zoospores lost their flagella, the eyespot disassembled, and the plastid rearranged to the normal shape found in vegetative cells.

The vegetative cells of the strain SAG 2220 were also predominantly circular when grown in liquid medium (Fig. 5A). Cells were solitary with occasional groups of two (Fig. 5B). The cell diameter ranged from 3.5 μm for autospores to 8.0 μm for mature vegetative cells. When cultivated on agar, some of the cells of the SAG 2220 also were elongated to pyriform (Fig. 5C, D) and the cells were also slightly larger, with a maximum length of 10 μm. In addition, bud-like structures on the cell wall were common when grown on agar (Fig. 5E, F). No other cell wall protrusions or ornamentations were present. The cell interior contained one to two parietal plastids. They were either plate- or cup-shaped, lying on one side of the cell or forming a prominent ring (Fig. 5H). The plastids of mature cells exhibited only shallow incisions resulting in tiny lobes and there were almost no crevices present (Fig. 5I). Even though not always conspicuous, the polyhedral pyrenoid could be discerned either protruding outside the plastid (Fig. 5A; arrowhead), or embedded within it (Fig. 5J; arrowhead). A reddish globule was almost always noticeable. In addition, old cells, in particular, exhibited abundant accumulation of lipid droplets (Fig. 5K). Reproduction by only two autospores was observed (Fig. 5L). Fig. 5L also shows that cellular material, including the reddish globule, is often ejected from the cell during autospore production. No flagellated cells were observed despite the effort to induce them.

The ultrastructure of vegetative cells of the CAUP Q 801 and SAG 2220 strains was very similar, showing one to two plastids with prominent pyrenoids surrounded by flattened lamellate vesicles and attached to the plastid by a narrow stalk (though not always obvious), a single nucleus (not typically discernible with light microscopy), several mitochondrial profiles, and a thick multi-layered cell wall (Fig. 6). Additional fine structural features, such as the lack of the girdle lamellae and absence of continuity between the outermost membranes of the nuclear and plastid envelopes (Fig. 6D), as well as the pyrenoid matrix being free of plastid thylakoids (Fig. 6C), re-confirmed the features typical for the class Eustigmatophyceae. In contrast to CAUP Q 801, where plastids were deeply lobed (Fig. 6A, B), the plastids of SAG 2220 were confirmed to lack prominent incisions (Fig. 6D, E). In addition, the reddish globule was composed of numerous different sized compartments enclosed by the single membrane in the CAUP Q 801 (Fig. 6B), whereas in SAG 2220 it was typically made of the single compartment (Fig. 6E). Finally, lamellate vesicles were also demonstrated to occur freely in the cytoplasm without being associated with the pyrenoid in SAG 2220 (Fig. 6F).

Confocal microscopy affirmed the differences in the plastid morphology of the two strains, especially in mature vegetative cells, where the plastid(s) contained more and typically deeper incisions in CAUP Q 801 (Fig. 6G) than in SAG 2220 (Fig. 6H). In addition, the plastid of the SAG 2220 strain exhibited a ring or crescent architecture with the lumen often being traversed only by the single prominent plastid lobe (Fig. 6H, middle section). Meanwhile the cell lumen of CAUP Q 801 was nearly always occupied by several smaller protruding plastid lobes (Fig. 6G) with the exception in the autospores. Finally, the incisions were evident also in the marginal parts of the plastid(s) in CAUP Q 801 (Fig. 6G), whereas in SAG 2220, if present, they typically occurred in the central part (Fig. 6H). A summary of morphological and ultrastructural features of the strains CAUP Q 801 and SAG 2220, and their comparison to a selection of previously described algae, is provided in Table S2.

### 3.4. *Genomic evidence for flagellated stages being common in the "Eustigmataceae group"*

A recent study reported on the finding of flagellum-related genes in the nuclear genomes of the genus *Monodopsis*, a taxon never seen to produce flagellated stages (Yang et al., 2021). This prompted us to ask if the genomic data we obtained from the SAG 2217/2220 strain pair, *Chlorobotrys* sp. FD2, and WTwin 8/9 T-6m6.8, all eustigs where zoosporogenesis has not been observed or studied, may provide signatures of a (cryptic) ability to form zoospores. We focused on the Mas family of

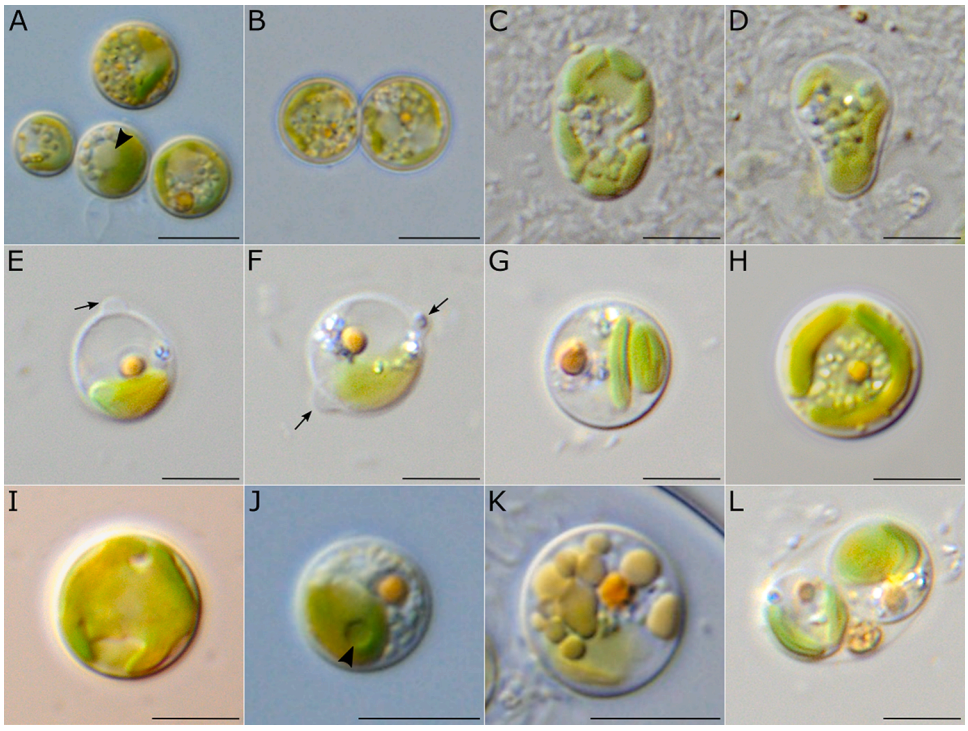


**Fig. 5.** Light micrographs of *Lietzensia polymorpha* gen. et sp. nov., strain SAG 2220. (A) Vegetative spherical cells cultivated in liquid medium. A pyrenoid (arrowhead) is visible; (B) Temporary grouping of two young cells; (C, D) Elongated to pyriform shapes of the cells when cultivated on agar slants; (E, F) A single parietal plastid and bud-like structures formed on the cell walls; (G) Cell containing two plastids and a prominent reddish globule; (H) Ring-shaped arrangement of the plastids; (I) Shallow incisions of the plastid; (J) Pyrenoid (arrow) embedded within the plastid; (K) Numerous lipid droplets in old cells; (L) Asexual reproduction by two autospores. Scale bars = 5 μm.

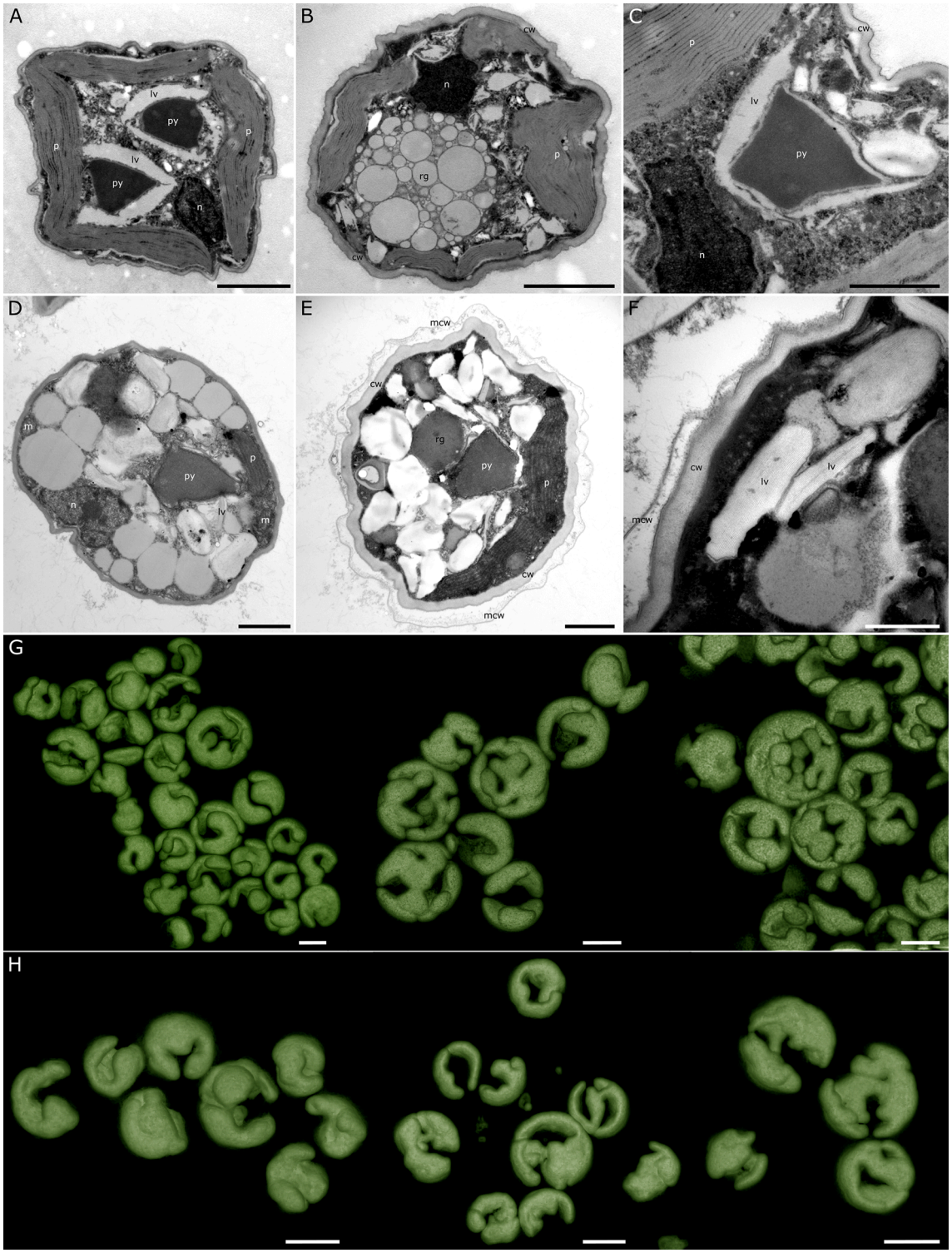


**Fig. 6.** Ultrastructure and chloroplast morphology of *Neustupella aerophytica* gen. et sp. nov., strain CAUP Q 801 (A–C; G) and *Lietzensia polymorpha* gen. et sp. nov., strain SAG 2220 (D–F; H). (A) Vegetative cell showing two plastids (one of them being deeply lobed) and two stalked pyrenoids. The pyrenoids are surrounded by lamellate vesicles. (B) Vegetative cells with a thick multi-layered cell wall and a reddish globule composed of numerous different sized compartments enclosed by a single membrane. (C) Close-up look at the fine structure of the pyrenoid. (D) Vegetative cell containing a single plastid with a stalked pyrenoid. The cell volume is completely filled by the storage material. The lack of continuity between the plastid and nuclear envelopes is obvious. (E) Prominent multi-layered cell wall surrounded by a mother-cell wall, and a reddish globule composed of the single compartment. (F) Lamellate vesicles scattered in the cytoplasm. (G) Confocal laser scanning microscopy of *N. aerophytica*. (H) Confocal laser scanning microscopy of *L. polymorpha*. Abbreviations: cw – cell wall; lv – lamellate vesicle; m – mitochondrion; mcw – mother cell wall; n – nucleus; p – plastid; py – pyrenoid; rg – red globule. Scale bars = 2 μm (A, B), 1 μm (C–E), 0,5 μm (F), 5 μm (G, H).

proteins that are the constituents of the Stramenopiles-specific tripartite mastigonemes on the anterior flagellum (Hee et al., 2019; Honda et al., 2007; Yamagishi et al., 2009). The family consists of three paralogs denoted Mas1 to Mas3, whose origin predates the split between oomycetes and ochrophytes (Hee et al., 2019). Interestingly, all eustigs analysed for the presence of the Mas genes, including the presumably azoosporic ones, exhibit orthologs of each of them (Table S3).

The assignment of the eustig genes to the three Mas subgroups is clearly supported by phylogenetic analysis (Fig. 7). Within each Mas clade eustig sequences form a monophyletic subclade, with the sequences from *Monodopsis* (two closely related strains; Yang et al., 2021) being sister to sequences from the "Eustigmataceae group". The relationships of the individual Mas genes from the "Eustigmataceae group" generally follow the relationships of the organisms themselves, and there is no indication from the tree that any of the species would exhibit an increased substitution rate of the Mas sequences potentially indicative of relaxed selective constraints (loss of function) or functional shifts (deployment of the Mas proteins for a novel, possibly flagellum-independent role). Interestingly, whereas only a single Mas3 gene is present in the "Eustigmataceae group", the Mas3 gene appears to have duplicated in the *Monodopsis* lineage, thus paralleling a Mas3 gene duplication previously encountered in the chrysophyte *Ochromonas danica* (Yamagishi et al., 2009). The genome assembly available for *Chlorobotrys* sp. FD2 is extremely fragmented due to insufficient read coverage, and all three Mas genes are represented in the assembly by multiple short pieces. Hence, this species was not included in the phylogenetic analysis presented in Fig. 7, but the assignment of the gene pieces to the Mas paralogs still appears robust based on sequence similarity comparisons.

To provide further evidence that the Mas genes have been retained by the eustigs to produce flagellar mastigonemes instead of serving an unknown flagellum-unrelated role, we additionally searched the genome assemblies for genes encoding heavy chains of flagellar dyneins (Kollmar, 2016). Indeed, all eustigs analysed possess genes for all major types, including the axonemal dyneins forming both the outer and the inner dynein arm, as well as DHC2 implicated in intra-flagellar transport (Table S4).

## 4. Discussion

### 4.1. Redefined family Chlorobotryaceae as the appropriate formal taxon embracing the "Eustigmataceae group"

By inferring phylogenies from highly sampled individual molecular phylogenetic markers (*rbcL* and 18S rRNA gene) and more sparsely sampled plastid multigene datasets, we confirm the existence of a eustigmatophyte clade, the "Eustigmataceae group," first recognized by Fawley et al. (2014) based on an analysis of the 18S rRNA gene. The robust support for the "Eustigmataceae group" from independent phylogenetic datasets and the degree of its separation from other major

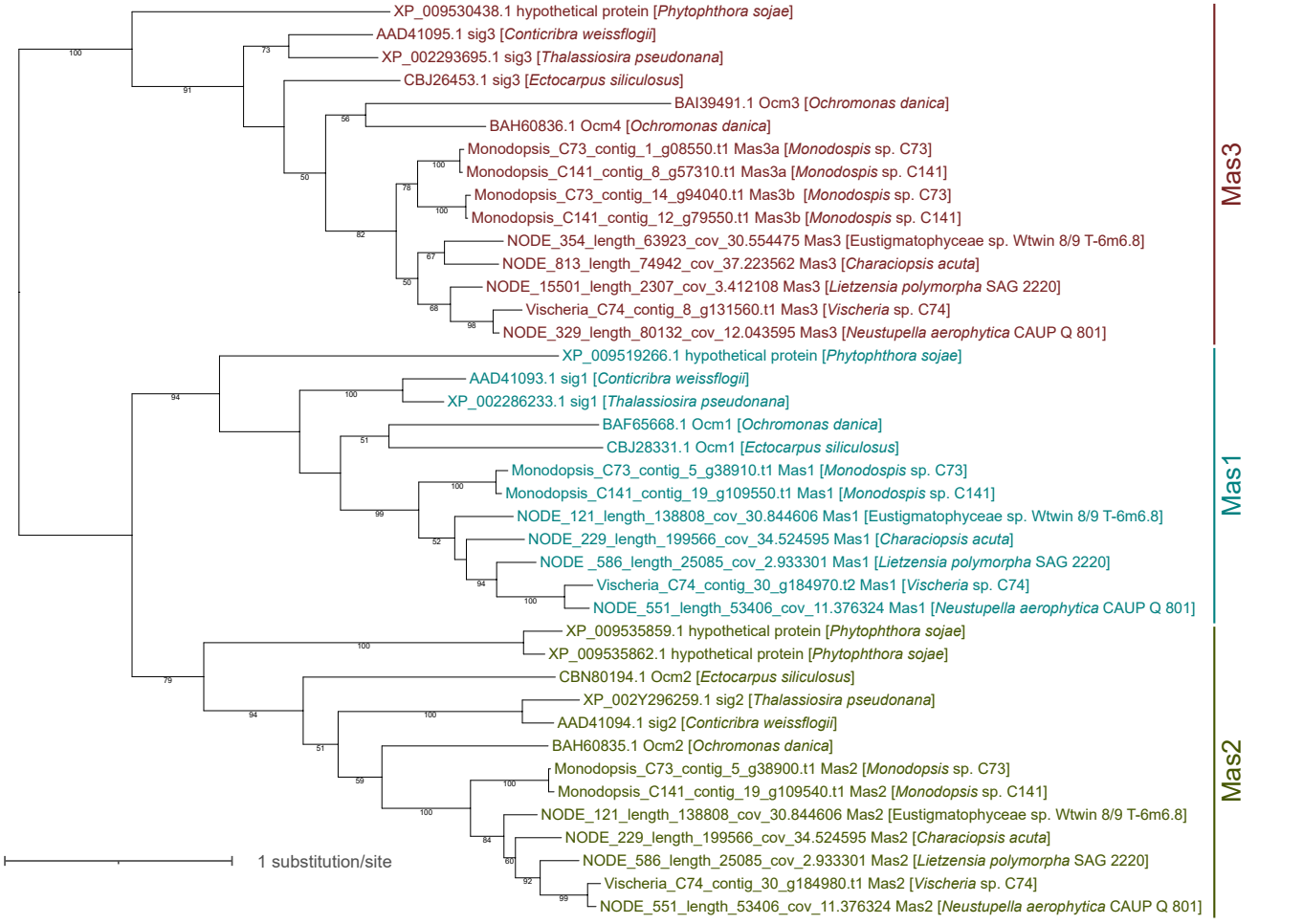


**Fig. 7.** Phylogenetic tree of the Mas protein family. The tree was inferred from a multiple alignment comprised of 343 amino acid position using the model WAG + I + G4. The tree clearly shows three separate subgroups (paralogs) in the Mas family, with the root placed arbitrarily between Mas3 and the other two paralogs. Sequence IDs correspond to GenBank accession numbers (non-eustigmatophytes), protein model IDs from genome annotations reported by Yang et al. (2021; *Vischeria* and *Monodopsis* spp.), and scaffold IDs from genome assemblies reported here (other eustigs).

eustigmatophyte clades, call for this group to be formally recognized as a taxon. Based on existing phylogenetic analyses of DNA sequence data, the "Eustigmataceae group" is comparable to two other major clades that have been recognized as families: Monodopsidaceae, introduced by Hibberd (1981), and the family Neomonodontaceae described by Amaral et al. (2020; with the incorrect spelling "Neomonodaceae", see Introduction). Hence, in the context of the existing taxonomy of Eustigmatophyceae, the "Eustigmataceae group" is best considered as a family of the order Eustigmatales.

An open question remains as to whether this family should also include the clade of environmental 18S rRNA gene sequences from Africa lakes (Afr45) that constitutes a sister group of what has been originally defined as the "Eustigmataceae group" by Fawley et al. (2014). However, except for the sequence KX465211.1, the other Afr45 sequences are short (522 bp) and hence the position of this clade should be treated with caution. The proper classification of the Afr45 clade can only be resolved by identifying and studying the organisms behind the sequences, and employing additional phylogenetic markers including *rbcL*. It is possible that the Afr45 clade may be equivalent to the lineage of represented by the *rbcL* metabarcode Big_Cherry_at_dam_ASV 733, which is placed sister to (other) Chlorobotryaceae in the tree presented in Fig. S3. However, the position of this metabarcode sequence is unstable and in other trees (employing a different algorithm or substitution model) it is nested within Chlorobotryaceae (data not shown). Alternatively, the organisms behind the Afr45 clade might be the same as those corresponding to the *rbcL* "clade Y", given its branching within Eustigmatales (Fig. S3); this would support a separate family status for the Afr45 clade. The *rbcL* "clade X" is placed in a position sister to sequences from other eustigs (Fig. S3), and it is tempting to speculate that it overlaps with the uncultured algal group documented by partial environmental sequences of two other markers (plastidial 23S rDNA and nuclear 18S rDNA) and positioned as a sister lineage of known eustigmatophytes (Rampen et al., 2022).

The organisms belonging the "Eustigmataceae group" have been previously included in several different families. The family Eustigmataceae, established by Hibberd (1981) to accommodate the genera *Eustigmatos* and *Vischeria*, corresponds to an extremely narrow lineage with little genetic divergence among its members (Figs. 1 and 2). These two genera have recently been merged into the single genus *Vischeria* (*Eustigmatos* being a junior heterotypic synonym; Kryvenda et al., 2018). The genus *Characiopsis* typifies a separate family Characiopsidaceae described by Pascher (1937–1938). Different sets of other genera were included in the family by Pascher and later by Ettl (1978), some of which may eventually be confirmed as being related to *Characiopsis* and hence part of the "Eustigmataceae group" (e.g. *Dioxys*; see Amaral et al., 2021). However, other genera that have been placed in Characiopsidaceae are clearly unrelated to the "Eustigmataceae group", including the genus *Harpochytrium* now known to be a fungus (Dee et al., 2015). The genus *Pseudocharaciopsis* has also been placed in the "Eustigmataceae group" based on molecular data from the type species *Pseudocharaciopsis texensis*, later called *Pseudocharaciopsis minuta*) (Fawley et al. 2014). The genus *Pseudocharaciopsis* typifies the monogeneric family Pseudocharaciopsidaceae (K.W.Lee & Bold ex Hibberd), which could be applied to the "Eustigmataceae group". However, a thorough study of the genus *Characiopsis* concluded that *Pseudocharaciopsis* is a junior synonym of *Characiopsis* (Amaral et al., 2021), which implies that Pseudocharaciopsidaceae is synonymous with Characiopsidaceae.

The final presently known member of the "Eustigmataceae group" previously formally placed into a family is the genus *Chlorobotrys*, which according to Hibberd's classification scheme represents a separate family Chlorobotryaceae, with Pascher indicated as the authority for the family (Hibberd, 1981). The family Chlorobotryaceae, with the original orthography "Chlorobotrydaceae", was introduced into the literature by Pascher in 1912, but the respective publication (Pascher 1912) includes only the family name with a list of genera included. Hence, the use of the name was not "accompanied by a description or diagnosis", which is a requirement of a valid publication of the family name according the ICN (article 38). This requirement was, however, apparently fulfilled in a subsequent publication by the same author (Pascher, 1915), which lists "Chlorobotrydaceae" as part of a classification scheme for the suborder Chlorobotrydinae. By explicitly referring to a characteristic feature of the Chlorobotrydaceae ("freilebend", i.e. free-living) and thus contrasting it with another family included in the same suborder, the Chlorotheciaceae (described as "festsitzend", i.e. sessile), Pascher (1915) provided a diagnosis of Chlorobotrydaceae. Therefore, the correct authority of the family name in our opinion is "Pascher 1915". Indeed, Silva (1979) apparently reached the same conclusion. In contrast, the opinion presented in AlgaeBase (Guiry and Guiry, 2022) considers the name Chlorobotrydaceae to be published later ("Pascher, 1925: 48").

As follows from the analysis above, the family name Chlorobotrydaceae has priority over any of the family names Eustigmataceae, Characiopsidaceae, or Pseudocharaciopsidaceae. However, the spelling of the name needs to be changed to "Chlorobotryaceae", as introduced by Hibberd (1981). This name change is based on article 18 of the ICN (see Introduction), and the fact that the genitive singular of the root term "botrys" (βότρῦς) of the type genus name is "botryos" (βότρῦος). As the correction of spelling does not change authorship and publication date of the family name, Chlorobotryaceae is a prime candidate for the oldest described family name applicable to the "Eustigmataceae group" if it is to be recognized as a single family. The validity of this conclusion, however, depends on two conditions: (1) the group really includes the type of the family name, i.e. the genus *Chlorobotrys* Bohlin 1901, which is typified by the species *Chlorobotrys regularis* (West) Bohlin, and originally described as *Chlorococcum regulare* West (1892); (2) none of the presently unknown or unidentified members of the "Eustigmataceae group" typifies a family name with a valid publication preceding that of Chlorobotryaceae.

Sequences of the 18S rRNA or *rbcL* genes assigned to *C. regularis* or *Chlorobotrys* in general were reported prior to this study (Amaral et al., 2021; Fawley et al., 2014, 2021; Kryvenda et al., 2018). All of these organisms, as well as three additional isolates presented in this study can be confidently placed in the "Eustigmataceae group" based on the analyses of the DNA sequence data (Figs. 1 and 2). However, the actual biological sources of the *Chlorobotrys* sequences published prior to this study were not documented in any detail by the authors. Thus, we provide micrographs of our four *Chlorobotrys* sp. isolates (FD2, B2, 2E5, UP3 5/31-7m; Fig. S5). All of these isolates closely match the morphology of *Chlorobotrys* including *C. regularis* as described in previous sources including Pascher (1937–1938) and are readily distinguished from other genera of eustigmatophytes or xanthophytes. These descriptions include the presence of lamellate mucilage or mucilaginous masses in which pairs or groups of spherical or nearly spherical cells are embedded (Hibberd, 1974; Ettl, 1978). The consistency of our identification of the *Chlorobotrys* isolates with the previous "tradition" is evident from the comparison to the strain CCAP 810/1 studied by Hibberd (1974), and identified by him as *Chlorobotrys regularis*. Although this strain is lost and its exact position in the eustigmatophyte phylogeny cannot be verified by molecular data (see Note S1), the ultrastructural details reported by Hibberd (1974) match closely the ultrastructure of *Chlorobotrys* sp. B2, including specifically a protruding pyrenoid, fibrous layers of mucilage separated by tripartite membrane-like structures, and lamellate vesicles lying freely in the cytoplasm (Fig. S6). It is interesting to note that in the original description of *Chlorococcum regulare*, West (1892) explicitly mentioned that the mucilage is non-lamellate, which is in fact true for strains kept in cultures for a long time, where the extensive production of mucilage can be lost (as we have observed, for example, in the strain FD2). However, the other characteristics provided by the author fit nicely into the current concept of the alga. It is highly significant that West noted the presence of a "red dot in every cell", apparently corresponding to the reddish globule so characteristic for eustigmatophytes.

The molecular sequence data from the multiple *Chlorobotrys* strains indicate a degree of genetic differentiation that suggests the existence of multiple species within the group. We refrain here from addressing the species-level classification and taxonomy of *Chlorobotrys*, which requires a deeper study at both the morphological and genetic level. Ideally, a broader set of isolates, including ones from the type localities of *C. regularis* (Harrop Tarn and Bowness in the Lake District, North West England; West, 1892), will be utilized in that future study to assess the identity and validity of the additional historically described, currently accepted *Chlorobotrys* species and to select the best candidate for epitypification of *C. regularis*. Regardless, the identification of *Chlorobotrys* as a particular lineage of the "Eustigmataceae group" is unambiguous and may thus serve as the basis for family-level classification of the whole group. An issue directly related to the question of the identity and phylogenetic position of the genus *Chlorobotrys* is the existence of a partial 18S rRNA gene sequence (KF848934.1) that was previously reported for the strain *Chlorobotrys regularis* CCAP 810/1 (Fawley et al., 2014). This sequence is not affiliated with the 18S rRNA sequences from the *Chlorobotrys* strains discussed above, but is virtually identical (barring a single nucleotide substitution) to the 18S rRNA sequence from the strain CAUP Q 801 described below as a new species in a new genus. As we explain in Note S1, this sequence was with near certainty derived from the strain CAUP Q 801 contaminating the CCAP 810/1 culture at the time of sequencing, and the actual CCAP 810/1 culture no longer contains the alga studied as *C. regularis* by Hibberd (1974).

What about the possibility that the "Eustigmataceae group" is eventually shown to include a member that typifies a family validly described earlier than Chlorobotryaceae? To address this question, we checked candidates for presently unknown members of the "Eustigmataceae group", and considered algae that may belong in this lineage based on morphological similarity to presently known members or whose relationship to the verified representatives of the group has been suggested by historical classifications. As detailed in Note S2, no plausible alternative family name applicable to the "Eustigmataceae group" emerged from the literature review. Chlorotheciaceae and Chlorosaccaceae are two families that are older than Chlorobotryaceae and considered previously to include some of the presently known "Eustigmataceae group" members. These families are typified by genera (*Chlorothecium* and *Chlorosaccus*, respectively) that are unlikely to belong to the "Eustigmataceae group" or even eustigmatophytes as a whole based on morphological features (Note S2). Two additional families, Gloeobotrydaceae and Pleurochloridaceae, historically overlap with the "Eustigmataceae group", but were established only after Chlorobotryaceae.

Altogether, the evidence gathered from our *Chlorobotrys* sp. isolates (FD2, B2, 2E5, UP3 5/31-7m), together with a careful re-evaluation of the taxonomic history of the taxa concerned, lead us to the conclusion that the appropriate family name for the "Eustigmataceae group" is Chlorobotryaceae. We note that Nakayama et al. (2015), when reporting on the newly discovered eustigmatophyte *Vacuoliviride crystalliferum*, included in the paper a phylogenetic tree based on the 18S rRNA gene with the clade corresponding to the "Eustigmataceae group" annotated as Chlorobotryaceae in the respective figure. Interestingly, the authors did not comment on their rationale to use this taxonomic name for the respective clade and did not address the taxonomy of this eustigmatophyte subgroup in any detail (note that *V. crystalliferum* belongs to the clade Goniochloridales). Apparently, the authors assigned the name Chlorobotryaceae to that clade assuming that it contains the type, *Chlorobotrys regularis*, but the only evidence at that time was based on the 18S rRNA sequence attributed to *C. regularis* CCAP 810/1, which we suggest (Note S1) is a mistake due to contamination of the *C. regularis* culture by CAUP Q 801. Hence, curiously Nakayama et al. arrived at a correct conclusion, but based on misinformation.

### 4.2. Diversity and new taxa of Chlorobotryaceae

With the revised circumscription the family Chlorobotryaceae now includes three previously defined genera, *Chlorobotrys*, *Vischeria*, and *Characiopsis* (including the synonyms of the latter two, i.e. *Eustigmatos* and *Pseudocharaciopsis*, respectively). However, molecular phylogenetic evidence from unidentified cultured strains as well as from environmental DNA surveys indicates a much higher genus-level diversity in the family (Fig. 1; Fig. S3). While acknowledging the limited resolution provided by the short metabarcode *rbcL* sequences and the fact that the delimitation of separate genera may not necessarily be dictated by the degree of divergence in a single phylogenetic marker, it seems that Chlorobotryaceae encompasses $>$ 15 genera other than the three currently known ones. Some may correspond to previously established genera that are yet to be identified as Chlorobotryaceae members. A good candidate is the monotypic genus *Botryochloropsis*, certainly a eustigmatophyte based on its ultrastructural characteristics and pigment composition (Preisig and Wilhelm, 1989) but not investigated yet by means of molecular taxonomy. Interestingly, preliminary observations suggest that this genus may find its home in the Clade Ia as one of its sublineages including the strain WTwin 8/9 T-6m6.8, for which we here report a complete plastome sequence (data to be published elsewhere). However, it is certain there is a substantial true novelty at the genus level in Chlorobotryaceae (like in eustigmatophytes in general), as there are simply not enough previously established candidates that could accommodate the diversity apparent from molecular phylogenetic studies.

In this study we have directly addressed the generic identity of two Chlorobotryaceae lineages represented by the strains CAUP Q 801 and SAG 2220. Based on the general morphological characteristics of its vegetative stage as well as zoospores, CAUP Q 801 highly resembles members of the most closely related genus, i.e. *Vischeria*, in particular those species characterized by a smooth cell wall and previously classified in the separate genus *Eustigmatos* (Kryvenda et al., 2018). The specific relationship of CAUP Q 801 to *Vischeria* is additionally consistent with the similar ultrastructure of the reddish globule, being composed of numerous variably sized compartments enclosed by a single membrane as demonstrated in different species of *Vischeria* (Eliáš, 2017; Gärtner et al., 2012; Trzcińska et al., 2014). Other eustigs known so far to have a somewhat compartmentalized reddish globule are *Chlorobotrys* (Fig. S6A) and the distantly related *Neomonodus ovalis* (Amaral et al., 2020). However, additional TEM investigations of other eustig lineages are needed to firmly consider the taxonomic significance of this specific structure. Finally, all so far known *Vischeria* isolates represent terrestrial species, same as CAUP Q 801 (isolated from subaerial algal growth on decaying bare wood in a tropical forest in Singapore), while the rest of the Chlorobotryaceae are freshwater algae. However, there is no doubt that CAUP Q 801 represents a species different from any of the historically described species of *Vischeria* and *Eustigmatos*: these are firmly placed within the redefined *Vischeria* genus by molecular evidence or their morphology is clearly different from CAUP Q 801 (the latter applies to certain *Vischeria* species with a highly characteristic cell shape or cell wall sculpting; Ettl, 1978). While classifying CAUP Q 801 as a new species of the genus *Vischeria* is formally possible, we argue that the depth of the phylogenetic divergence between CAUP Q 801 and known *Vischeria* species as evidenced by both nuclear (18S rRNA; Fig. 2) and plastid markers (*rbcL*, concatenated plastome-encoded proteins; Fig. 1, Fig. 3) is such that CAUP Q 801 is more appropriately considered to represent a separate genus. The strain SAG 2220 (together with SAG 2217) is even more deeply separated from any of the three established Chlorobotryaceae genera in all molecular phylogenies (Figs. 1-3).

Hence, if none of the presently defined Chlorobotryaceae genus can accommodate the strain CAUP Q 801 or SAG 2220, can they be identified as representatives of genera, or even species, that were described before but not yet tied by phylogenetic evidence to Chlorobotryaceae?

Straight away we exclude *Botryochloropsis* and its single species *B. similis*, as it differs from both CAUP Q 801 and SAG 2220 by the lack of a pyrenoid and other features (such as the zoospore morphology when compared to CAUP Q 801), and as mentioned above, algae resembling it much closer are found among the Clade Ia members. From the traditional perspective (Ettl, 1978; Pascher, 1937–1939), both CAUP Q 801 and SAG 2220 would be most likely considered as members of the formally xanthophyte genera *Pleurochloris* and *Chloridella*. Both genera were described by Pascher (in 1925 and 1932, respectively), with the distinguishing feature being the apparent lack of zoospores in the latter. As CAUP Q 801 produces zoospores, it is natural to ask if it may be identified with any of the previously described *Pleurochloris* species. In this context it is notable that two species originally described as belonging to the genus *Pleurochloris* were later transferred to the genus *Eustigmatos* (Hibberd, 1981) to be eventually recombined as *Vischeria polyphem* and *Vischeria magna* (Kryvenda et al., 2018), so they are indeed close relatives of CAUP Q 801. Nevertheless, as further discussed below, zoospore production by coccoid algae may require specific clues not commonly present or easy to introduce in laboratory conditions, questioning the validity of its apparent absence as a reliable taxonomic criterion. Hence, we do not consider the fact that CAUP Q 801 is capable of zoospore production as a reason to *a priori* exclude its potentially affiliation to *Chloridella*. In the opposite regard, the fact that we did not observe zoospore production by SAG 2220, does not preclude the possibility that this strain belongs to the genus *Pleurochloris*, particularly when considering the genomic evidence for the capability to build flagella by this alga (see below).

Critical for assessing possible affinities between CAUP Q 801 or SAG 2220 and the genus *Pleurochloris* is defining the phylogenetic position of the type species of the genus, *Pleurochloris commutata*. It was described by Pascher (1937–1938) as a common soil alga lacking a pyrenoid and with biflagellate, broadly oval or pyriform zoospores. There has been no modern account on this species and, as discussed by Hibberd (1981), it is even not clear whether it is a xanthophyte or an eustigmatophyte. Crucially, the morphological features of *P. commutata* make it unlikely to be specifically related to CAUP Q 801 or SAG 2220. Both strains are distinguished by the presence of a pyrenoid, and the zoospore morphology of CAUP Q 801 does not match the zoospores of *P. commutata*. Although we cannot directly compare the morphology of the putative zoospores of SAG 2220, the fact that zoospores are not readily produced is by itself a potentially relevant diagnostic feature, and additionally, as a planktic alga (isolated from a lake in Germany) SAG 2220 differs from *P. commutata* ecologically.

The possibility that CAUP Q 801 or SAG 2220 represents an undescribed *Chloridella* species ultimately depends on the phylogenetic position of the type species of the genus, *Chloridella neglecta* (Pascher & Geitler) Pascher (basionym: *Chlorobotrys neglectus* Pascher & Geitler). However, this species is characterized by the presence of multiple smaller, visually well separated plastids, which is a situation untypical for eustigmatophytes yet common in xanthophytes. In fact, *C. neglecta* is very similar, perhaps indistinguishable, from vegetative cells of certain *Pleurochloris* species characterized by multiple smaller plastids, exemplified by *Pleurochloris meiringensis*, which has been confirmed as a xanthophyte by molecular evidence from its authentic culture (Andreoli et al., 1999). Interestingly, some culture collections include subcultures of the strain V.216 (isolated by Vischer in 1940) that are listed as *Chloridella neglecta* (CCAP 813/1, UTEX B 431), whereas other collections have subcultures of the same strain listed as *Pleurochloris meiringensis* (CCALA 278, SAG 813-1), with the information that they were previously identified as *Chloridella neglecta*. Morphologically this alga indeed resembles the original description of *C. neglecta*, as documented here in Fig. S7 using the subculture CCALA 278. However, as we found out by sequencing its 18S rRNA gene, this alga is closely related, if not conspecific to the xanthophyte *P. meiringensis*, as it differs in a single nucleotide substitution from the sequence reported for the authentic strain of this species (AF109728.1). This is consistent with the fact that both the CCALA and SAG collections presently identify the Vischer's strain V.216 as *P. meiringensis*. Another strain (SAG 48.84) originally referred to in the culture collections as *Chloridella neglecta* was previously investigated by Kryvenda et al. (2018), who showed that it is firmly nested in the genus *Vischeria* based on 18S rRNA sequence data. The authors concluded that the strain is misidentified, which is a conclusion we endorse upon checking the light microscopy images of the strain provided by the authors. Thus, *Chloridella neglecta*, and hence the genus it typifies, is best interpreted as a xanthophyte.

Based on the arguments above we discount the possibility that *Pleurochloris* or *Chloridella* could be considered as a taxonomic home of any of the strains CAUP Q 801 and SAG 2220. However, as the phylogenetic coherence of the genera *Pleurochloris* and *Chloridella* is uncertain, we need to check the possibility that CAUP Q 801 or SAG 2220 correspond to some of the previously described *Pleurochloris* and *Chloridella* (non-type) species. However, our careful comparison, summarized in Table S2, did not point to any of the *Pleurochloris* species as sufficiently similar to CAUP Q 801 to hypothesize their conspecificity. Some of these species can be excluded as CAUP Q 801 relatives on the grounds of a plastidial eyespot having been clearly documented in their zoospores, a feature indicating they most likely belong to Xanthophyceae (confirmed by molecular data for *P. meiringensis*, see above). Most of the other species differ from CAUP Q 801 by the lack of a pyrenoid or the zoospore morphology. For example, *Pleurochloris pyrenoidosa* possesses a eustigmatophyte-like plastid and a pyrenoid, but its zoospores are pear-shaped with two flagella and lack a noticeable eyespot, while CAUP Q 801 contains elongated zoospores with a single emergent flagellum and a prominent extraplastidial eyespot. Possible identification of CAUP Q 801 as any of the previously described *Chloridella* species (other than *C. neglecta*) is hampered by the fact that none of them was described as possessing a pyrenoid and each exhibits a combination of the plastid morphology, the cell size range, and cell wall features further distinguishing them from CAUP Q 801 (Table S2).

A similar review of morphological features of the existing *Pleurochloris* and *Chloridella* species indicates that none of them is similar enough to the SAG 2220 strain (Table S2). The plastid features in SAG 2220 somewhat resemble the ones described for *P. pyrenoidosa*, i.e., mostly lining only one side of the cell, often remarkably small and very delicate and sometimes with a very regular, deeply lobed outline. However, the cell size measured for SAG 2220 is much smaller (3.5–8.0 μm) as noted for *P. pyrenoidosa* (8.0–12.0 μm) and the plastids are usually only shallowly lobed (Fig. 6H). Our culture of the SAG 2217 strain was overgrown by another eustigmatophyte alga (*Monodopsis* sp.) before we could carefully investigate its morphology and compare it to previously described species, but the observations by Kryvenda et al. (2018) show it is highly alike SAG 2220. Furthermore, the molecular sequence data, including complete plastome sequences, indicate beyond any doubt that the two strains are conspecific, which is also consistent with the fact they were isolated in different years from the same locality.

The analyses presented above build a solid case for both CAUP Q 801 and SAG 2220 being different from any previously described algal species, and do not point to any candidate genera that could accommodate these strains as their new species. Hence, below we provide a formal description of CAUP Q 801 as *Neustupella aerophytica* gen. et sp. nov. and of SAG 2220 as *Lietzensia polymorpha* gen. et sp. nov.

### 4.3. Trait evolution in Chlorobotryaceae

While the phylogenetic coherence of the redefined Chlorobotryaceae is obvious, there are presently no unique shared traits (synapomorphies) – beyond specific substitutions in gene sequences (the source of the phylogenetic signal underpinning the Chlorobotryaceae monophyly in molecular phylogenetic reconstructions) – that could be used to define the family in the "traditional" perspective of biological classification. Indeed, the members of the family vary in various traits historically considered of taxonomic significance for eustigmatophyte classification,

which explains why representatives of Chlorobotryaceae as redefined here were distributed in three different families by Hibberd (1981). Thus, these organisms differ in the number of flagella (one or two) of zoospores, the ability to produce a stipe, the cell shape (spherical to fusiform), or the presence versus absence of cell groups embedded in lamellate mucilaginous envelopes (Eliáš et al., 2017). There also seems to be variation in the occurrence and shape of the pyrenoid in Chlorobotryaceae. Whereas pyrenoids are common in most members of the Chlorobotryaceae (Amaral et al., 2021; Hibberd, 1981; this study), we have not observed pyrenoids in members of the Clade Ia (unpublished observations). The presence or absence of pyrenoids has already been shown to vary in the Monodopsidaceae (Eustigmatales). Species of *Monodopsis* possess pyrenoids, whereas those of *Nannochloropsis* and *Microchloropsis* are devoid of them (Santos and Leedale, 1995). However, as the molecular underpinnings of the pyrenoid formation in eustigs are not known, it is not certain whether gain or loss of pyrenoid formation should be viewed as a major evolutionary step or a volatile trait depending on relatively minute genetic changes. Despite this possible disparity of the redefined Chlorobotryaceae, it is notable that by broadening its circumscription we are paradoxically making it closer to its original conception. Thus, the basionym of the type of the genus *Vischeria*, i.e. *Vischeria stellata*, is *Chlorobotrys stellatus* Chodat, and the genus *Characiopsis* was considered a member of the family when Chlorobotryaceae was initially mentioned in the literature (Pascher, 1912).

It is additionally noteworthy that the taxa classified in the redefined Chlorobotryaceae may be less disparate than generally thought when it comes to one particular characteristic considered by Hibberd (1981) as important for differentiating eustigmatophytes at the family level: zoosporogenesis or the lack thereof. Thus, the production of zoospores by members of Hibberd's families Eustigmataceae and Pseudocharaciopsidaceae was contrasted to the absence of this reproduction mode in Monodopsidaceae and the narrowly defined Chlorobotryaceae (restricted to *Chlorobotrys* in Hibberd's scheme). However, it has been well established that zoosporogenesis in coccoid algae may be an occasional phenomenon under control of rarely occurring specific biotic and abiotic clues or their combinations (Agrawal, 2012). Moreover, the zoosporogenesis-inducing conditions may also be species- or strain-specific. For example, zoospores in the freshwater eustigs *Trachydiscus minutus* and *Goniochloris sculpta* (*Goniochloridales*) were induced by transferring synchronized cultures into fresh media and kept in the dark for 3–4 days (Přibyl et al., 2012). A similar approach, however, did not work when we tested it for *Neustupella aerophytica*. Considering the terrestrial nature of our isolate, we tried to imitate the actual conditions in nature: the period of drought (agar exposed to the continuous light) with the following rain and a night period (providing liquid medium and transferring to the dark). And indeed, zoospore production started within 1–2 days. Interestingly, we then transferred old agar clumps containing one of our terrestrial *Vischeria* sp. isolates to the liquid media, in two days' time we could observe several swimming zoospores produced even in the light conditions (data not shown). Considering the fact that *Lietzensia polymorpha* thrives in a water body, it was not surprising that a similar approach did not induce zoospores. That said, the presence of enough water for zoospores to swim in (or sudden exposure to wet conditions) and absence of light are likely the key factors triggering zoosporogenesis in terrestrial eustigmatophyte algae, whereas additional clues may be needed in eustigs that live continuously exposed to free water.

Interestingly, signatures for the ability to produce zoospores can in principle be deduced from genome data. Indeed, recent investigations identified genes for hallmark flagellar components in genomes of phylogenetically diverse coccoid algae that have never been observed to produce flagellated stages, such as the pelagophyte *Aureococcus anophagefferens* or the chlorophytes *Prasinoderma coloniale* and *Monoraphidium neglectum* (Eliáš et al., 2016; Li et al., 2020). Most pertinent for this study was the recent finding of such genes in the genomes of two different strains of the genus *Monodopsis* in the family Monodopsidaceae, indicating that this eustig lineage may be zoosporogenic after all (Yang et al., 2021). Inspired by this we probed the genomic data from the strains studied here for the presence of genes for two categories of flagellar proteins, the mastigoneme-constituting Mas family and flagellar dynein heavy chains, and found them in all Chlorobotryaceae members. These include *L. polymorpha* (SAG 2220), where we failed to observe zoospores despite trying to induce their production by various treatments, as well as *Chlorobotrys* sp. FD2, which represents a eustig genus that is considered azoosporic based on previous studies (Hibberd, 1974, 1981). No observation of zoospores has been made for the strain WTwin 8/9 T-6m6.8, but we have documented them in very closely related strains (results to be published elsewhere).

It is beyond the scope of this study to carry out a detailed reconstruction of the complement of the flagellar genes in Chlorobotryaceae, as transcriptome data for most species and strains are lacking, making accurate prediction of gene models (exon–intron structures) difficult. In addition, the quality of the genomic data for some of the strains is not sufficient for more extensive comparative genomic analyses. The genome assembly is particularly poor in the case of *Chlorobotrys* sp. FD2 and hence we could not ascertain whether the sequences homologous to Mas- and dynein-encoding genes of the other eustigs actually represent structurally intact and potentially functional genes. We thus cannot formally rule out the possibility that *Chlorobotrys* sp. FD2 has a decaying set of flagellar genes that does not allow for the production of functional flagella. Indeed, such a situation has been documented at the infraspecific level in the haptophyte *Gephyrocapsa* (=*Emiliania*) *huxleyi*, which includes lineages "locked" in the aflagellated diploid stage and exhibiting loss and pseudogenization of genes important for flagellum biogenesis (von Dassow et al., 2015). However, if this was the case also in *Chlorobotrys* sp. FD2, the loss of the ability to produce flagellated stages in this genus would still be a relatively recent event.

The genomic evidence for the ability of an organism to build flagella does not by itself define the form of the putative flagellated stage, so we cannot rule out the possibility that some Chlorobotryaceae may produce flagellated stages functioning only as gametes (as is the case of, e.g., centric diatoms; Poulíčková and Mann, 2019) rather than producing zoospores. The sexual process has not been directly observed in any eustigmatophyte, but homologs of meiotic components are encoded by various eustigs including the Chlorobotryaceae representative *Vischeria* (Yang et al., 2021), so the production of gametes as the sole flagellated cell type in some eustigs certainly is a possibility that needs to be examined. The physiology of and molecular mechanisms governing zoosporogenesis in eustigmatophytes is an important topic for future research.

### 4.4. Formal taxonomy

Based on the observations presented in the Results section and on the preceding discussion, we emend or describe several eustigmatophyte taxa at a different taxonomic level, following the rules and recommendations of The International Code of Nomenclature (ICN) for algae, fungi, and plants (Turland et al., 2018).

**Chlorobotryaceae** Pascher emend. Barcytė & M.Eliáš.

**Description**: Unicellular, free living or epibiontic; solitary or in loose groups. Cell shapes vary from spherical or oval to spindle-shaped. Cell wall smooth, sculptured or covered in mucilage (including several layers of them). Stipe (stalk) formed in some species. Uninucleate. Reddish globules either homogenous or composed of several different sized compartments. Lamellate vesicles present. Plastids one to several; with or without stalked pyrenoids. Reproduction by autospores and bi- or uniflagellate zoospores. Terrestrial or freshwater.

**Type genus**: *Chlorobotrys* Bohlin.

**Included genera**: *Chlorobotrys* Bohlin, *Vischeria* Pascher, *Characiopsis* Borzi, *Neustupella* gen. nov., *Lietzensia* gen.nov.

***Neustupella*** Barcytė, Němcová & M.Eliáš gen. nov.

**Description:** Cells spherical or subspherical to somewhat irregular in outline. Cell wall robust and smooth. One to two plastids with pyrenoids. Nucleus single. Reproduction by autospores and zoospores. The genus represents a separate phylogenetic lineage within the Chlorobotryaceae related to *Vischeria* Pascher.

**Etymology:** The name is a combination of the surname "Neustupa" and the diminutive suffix "-ella". The genus name is proposed to honour our colleague and friend prof. Jiří Neustupa, an acknowledged scholar in the field of phycology and an inspiring person; in 2008 he isolated the strain CAUP Q 801, based on which the genus is described.

**Type species:** *Neustupella aerophytica* sp. nov.

***Neustupella aerophytica*** Barcytė, Němcová & M.Eliáš sp. nov. (Fig. 4).

**Description:** Cells solitary, rarely and temporarily in groups of two, spherical to oval or somewhat triangular in outline, ranging from 3.5 to 12.5 (-16.0) μm in diameter. Cell wall rigid, multilayered, and smooth without sculpture or mucilage. Cells uninucleate. One to two parietal plastids. In young cells the plastid(s) bowl- or trough-shaped. In mature cells the plastid(s) cup-shaped, divided into several unequal lobes by deep incisions. Polyhedral pyrenoid attached to the plastid by a stalk. A huge vacuole may be present. Reddish globule present. Crystals may be present. Reproduction by two or four autospores or zoospores. Zoospores cylindrical to spindle-shaped, reaching up to 12 μm in length and up to 3 μm in width, with a single apparent flagellum, an extraplastidial eyespot at the anterior extremity, and a single plastid. Terrestrial.

**Holotype (designated here):** TEM resin block H–CAUP Q 801 of the strain CAUP Q 801 at the Culture Collection of Algae of Charles University (CAUP), Prague, Czech Republic.

**Reference strain (ex-type):** CAUP Q 801.

**DNA sequences:** ON924321 (18S rDNA), ON929296 (plastid genome).

**Type locality:** Central Catchment Nature Reserve, Singapore (1°21′12.56′'N and 103°48′43.09′'E).

**Etymology:** The species epithet reflects the aerophytic nature of the holotype.

***Lietzensia*** Barcytė, Němcová, K.P.Fawley, M.W.Fawley & M.Eliáš gen. nov.

**Description:** Cells spherical to somewhat irregular in outline. Cell wall robust and smooth without sculpture. One to two plastids with pyrenoids. Nucleus single. Reproduction by autospores. The genus represents a separate phylogenetic lineage within the Chlorobotryaceae.

**Etymology:** The generic name *Lietzensia* refers to the Lake Lietzensee, where the alga was found.

**Type species:** *Lietzensia polymorpha* sp. nov.

***Lietzensia polymorpha*** Barcytė, Němcová, K.P.Fawley, M.W.Fawley & M.Eliáš sp. nov. (Fig. 5).

**Description:** Cells solitary, rarely and temporarily in groups of two, spherical to pyriform. The latter only when grown on agar. Cell size range from 3.5 to 8.0 (-10.0) μm in diameter. Cell wall robust and smooth but additional buds might be present. Cells uninucleate. One to two parietal plastids; in young cells plate- or bowl-shaped shaped, lining one side of the cell; in mature cells ring-shaped surrounding the periphery; might be slightly lobed. Indistinct pyrenoid present. Reddish globule conspicuous. Oil droplets produced. Only reproduction by two autospores observed, but genomic evidence suggests zoospores can be formed, too. Freshwater.

**Holotype (designated here):** TEM resin block H–SAG 2220 of the strain SAG 2220 at the Culture Collection of Algae of Charles University (CAUP), Prague, Czech Republic.

**Reference strain (ex-type):** SAG 2220.

**DNA sequences:** ON924320 (18S rDNA), ON938208 (plastid genome).

**Type locality:** Lake Lietzensee, Berlin, Germany.

**Etymology:** The species epithet reflects the morphological plasticity of the species.

**Note:** The species is represented by another living strain SAG 2217 housed at the Culture Collection of Algae at the University of Göttingen, Germany (SAG).

**Neomonodontaceae** R.Amaral, K.P.Fawley, Němcová, T.Ševčíková, Lukešová, M.W.Fawley, L.M.A.Santos et M.Eliáš, emended name.

**Note:** The root term of *Neomonodus*, the type of the family, is "odus" (ὀδούς), the genitive singular of which is "odontos" (ὀδόντος), so the original spelling of the family, name Neomonodaceae (Amaral et al., 2020), was incorrect.

### 4.5. Concluding remarks

Our work has resolved one of the long-standing issues of the higher-level classification of eustigmatophyte algae and further improved the knowledge of the biodiversity of the group by carefully documenting and formally describing two new genera and species. It is, however, obvious that further research is required on many fronts if a truly comprehensive understanding of the diversity and evolution of eustigmatophytes is the goal. Focusing specifically on the family Chlorobotryaceae, additional careful taxonomic studies are required to resolve the identity of the number of phylogenetically diverse unidentified cultures belonging to this group. In addition, the evidence from *rbcL*-based metabarcoding for the existence of numerous additional lineages in Chlorobotryaceae currently not represented by cultured organisms (Fig. S3) calls for a continuing sampling and isolation effort. The latter is also critical for uncovering the organismal identity of the mysterious Afr45 clade and resolving its formal taxonomic status in relation to Chlorobotryaceae. An obvious future research direction is not only to further expand the list of sequenced eustig plastid genomes, which we have shown provide a highly valuable resource for determining the phylogenetic relationships within eustigmatophytes, but also to exploit other segments of the eustig genetic blueprint for phylogenetic studies, including mitochondrial and nuclear genomes. These will be instrumental for resolving problematic areas of the eustig phylogeny, such as the branching order of *Chlorobotrys*, *Lietzensia*, and the *Vischeria/Neustupella* clade, and will lay a basis for possible further elaborations on the formal taxonomy of Chlorobotryaceae, such as delimiting subfamilies to properly capture the deep divergencies in the family.

## CRediT authorship contribution statement

**Dovilė Barcytė:** Conceptualization, Data curation, Formal analysis, Investigation, Visualization, Writing – original draft, Writing – review & editing. **Martina Zátopková:** Investigation, Resources, Writing – review & editing. **Yvonne Němcová:** Investigation, Writing – review & editing. **Michal Richtář:** Data curation, Formal analysis, Writing – review & editing. **Tatiana Yurchenko:** Data curation, Formal analysis, Writing – review & editing. **Karin Jaške:** Investigation, Resources, Writing – review & editing. **Karen P. Fawley:** Investigation, Resources, Writing – review & editing. **Pavel Škaloud:** Investigation, Supervision, Writing – review & editing. **Tereza Ševčíková:** Investigation, Supervision, Writing – review & editing. **Marvin W. Fawley:** Investigation, Resources, Writing – review & editing. **Marek Eliáš:** Conceptualization, Formal analysis, Funding acquisition, Investigation, Supervision, Validation, Writing – original draft, Writing – review & editing.

## Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Acknowledgements

We thank Dr. Maike Lorenz (Experimental Phycology and Culture Collection of Algae (SAG), Albrecht-von-Haller-Institute for Plant Sciences, Georg August University Göttingen) for her assistance in

obtaining the strains SAG 2217 and SAG 2220, and Lubor Kysučan (University of Ostrava) for his advice on the ancient Greek grammar. This work was supported by the Czech Science Foundation project 20-27648S (to ME), the project CePaViP (CZ.02.1.01/0.0/0.0/16_019/0000759) provided by ERD Funds (to ME), and a grant from the National Science Foundation (grant number DEB114514) to KPF and MWF. This project was also supported by the Arkansas INBRE program, with a grant from the National Institute of General Medical Sciences, (NIGMS), P20 GM103429 from the National Institutes of Health (to KPF).

## Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ympev.2022.107607.

## References

Abràmoff, M.D., Magalhães, P.J., Ram, S.J., 2004. Image processing with ImageJ. Biophotonics Int. 11, 36–42.
Agrawal, S.C., 2012. Factors controlling induction of reproduction in algae—review: the text. Folia Microbiol. 57, 387–407. https://doi.org/10.1007/s12223-012-0147-0.
Altschul, S.F., Madden, T.L., Schäfer, A.A., Zhang, J., Zhang, Z., Miller, W., Lipman, D.J., 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 25, 3389–3402. https://doi.org/10.1093/nar/25.17.3389.
Amaral, R., Fawley, K.P., Němcová, Y., Ševčíková, T., Lukešová, A., Fawley, M.W., Santos, L.M.A., Eliáš, M., 2020. Toward modern classification of eustigmatophytes, including the description of Neomonodaceae fam. nov. and three new genera. J. Phycol. 56, 630–648. https://doi.org/10.1111/jpy.12980.
Amaral, R., Ševčíková, T., Eliáš, M., Santos, L.M.A., 2021. *Characiopsis* Borzì belongs to the Eustigmatophyceae. Eur. J. Phycol. 56, 186–202. https://doi.org/10.1080/09670262.2020.1808712.
Andersen, R.A., Morton, S.L., Sexton, J.P., 1997. Provasoli-Guillard National Center for Culture of Marine Phytoplankton 1997 list of strains. J. Phycol. 33 (suppl.), 1–75.
Andreoli, C., Moro, I., La Rocca, N., Rigoni, F., Valle, L.D., Bargelloni, L., 1999. *Pseudopleurochloris antarctica* gen. et sp. nov., a new coccoid xanthophycean from pack-ice of Wood Bay (Ross Sea, Antarctica): ultrastructure, pigments and 18S rRNA gene sequence. Eur. J. Phycol. 34, 149–159. https://doi.org/10.1080/09670269910001736202.
Bankevich, A., Nurk, S., Antipov, D., Gurevich, A.A., Dvorkin, M., Kulikov, A.S., Lesin, V. M., Nikolenko, S.I., Pham, S., Prjibelski, A.D., Pyshkin, A.V., Sirotkin, A.V., Vyahhi, N., Tesler, G., Alekseyev, M.A., Pevzner, P.A., 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J. Comput. Biol. 19, 455–477. https://doi.org/10.1089/cmb.2012.0021.
Barcytė, D., Eikrem, W., Engesmo, A., Seoane, S., Wohlmann, J., Horák, A., Yurchenko, T., Eliáš, M., 2021. *Olisthodiscus* represents a new class of Ochrophyta. J. Phycol. 57, 1094–1118. https://doi.org/10.1111/jpy.13155.
Bischoff, H.W., Bold, H.C., 1963. Phycological Studies IV. Some soil algae from Enchanted Rock and related algal species. Univ. Texas Publ. 6318, 1–95.
Bohlin, K., 1901. Étude sur la flore algologique d'eau douce des Açores. Bihang till Kongl. Svenska vetenskaps-akademiens handlingar 27 Afd. III(4), 1–85.
Bolger, A.M., Lohse, M., Usadel, B., 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30, 2114–2120. https://doi.org/10.1093/bioinformatics/btu170.
Capella-Gutiérrez, S., Silla-Martínez, J.M., Gabaldón, T., 2009. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics 25, 1972–1973. https://doi.org/10.1093/bioinformatics/btp348.
Dee, J.M., Mollicone, M., Longcore, J.E., Roberson, R.W., Berbee, M.L., 2015. Cytology and molecular phylogenetics of Monoblepharidomycetes provide evidence for multiple independent origins of the hyphal habit in the Fungi. Mycologia 107, 710–728. https://doi.org/10.3852/14-275.
Dellaporta, S.L., Wood, J., Hicks, J.B., 1983. A plant DNA minipreparation: version II. Plant Mol. Biol. Rep. 1, 19–21. https://doi.org/10.1007/BF02712670.
Eliáš, M., et al., 2017. Eustigmatophyceae. In: Archibald, J.M., Simpson, A.G., Slamovits, C.H. (Eds.), Handbook of the Protists. Springer International Publishing, Cham, pp. 367–406.
Eliáš, M., Klimeš, V., Derelle, R., Petrželková, R., Tachezy, J., 2016. A paneukaryotic genomic analysis of the small GTPase RABL2 underscores the significance of recurrent gene loss in eukaryote evolution. Biol. Direct. 11, 5. https://doi.org/10.1186/s13062-016-0107-8.
Ettl, H., 1978. Xanthophyceae. 1.Teil, in: Ettl, H., Gerloff, J., Heynig, H. (Eds.), Süßwasserflora von Mitteleuropa, Bd. 3. Gustav Fischer, Stuttgart & New York.
Fawley, M.W., Douglas, C.A., Stewart, K.D., Mattox, K.R., 1990. Light-harvesting pigment-protein complexes of the Ulvophyceae (Chlorophyta): characterization and phylogenetic significance. J. Phycol. 26, 186–195. https://doi.org/10.1111/j.0022-3646.1990.00186.x.
Fawley, K.P., Eliáš, M., Fawley, M.W., 2014. The diversity and phylogeny of the commercially important algal class Eustigmatophyceae, including the new clade *Goniochloridales*. J. Appl. Phycol. 26, 1773–1782. https://doi.org/10.1007/s10811-013-0216-z.
Fawley, M.W., Fawley, K.P., 2004. A simple and rapid technique for the isolation of DNA from microalgae. J. Phycol. 40, 223–225. https://doi.org/10.1111/j.0022-3646.2004.03-081.x.
Fawley, M.W., Fawley, K.P., 2017. Rediscovery of *Tetraëdriella subglobosa* Pascher, a member of the Eustigmatophyceae. Fottea 17, 96–102. https://doi.org/10.5507/fot.2016.018.
Fawley, M.W., Jameson, I., Fawley, K.P., 2015. The phylogeny of the genus *Nannochloropsis* (Monodopsidaceae, Eustigmatophyceae), with descriptions of *N. australis sp. nov*. and *Microchloropsis gen. nov*. Phycologia 54, 545–552. https://doi.org/10.2216/15-60.1.
Fawley, M.W., Němcová, Y., Fawley, K.P., 2019. Phylogeny and characterization of *Paraeustigmatos columelliferus*, gen. et sp. nov., a member of the Eustigmatophyceae that may represent a basal group within the Eustigmatales. Fottea 19 (2), 107–114. https://doi.org/10.5507/fot.2019.002.
Fawley, M.W., Fawley, K.P., Cahoon, A.B., Lane, C., 2021. Finding needles in a haystack–Extensive diversity in the eustigmatophyceae revealed by community metabarcode analysis targeting the *rbc*L gene using lineage-directed primers. J. Phycol. 57 (5), 1636–1647. https://doi.org/10.1111/jpy.13196.
Gärtner, G., Stoyneva, M.P., Uzunov, B.A., Mancheva, A.D., Ingolić, E., 2012. Ultrastructure of vegetative cells and autospores of an aerophytic strain of *Vischeria stellata* (Chodat ex Poulton) Pascher (Eustigmatophyceae) from Bulgaria. Fottea 12, 273–280. https://doi.org/10.5507/fot.2012.019.
Graham, L.E., Graham, J.M., Wilcox, L.W., Cook, M.E., 2016. Algae, 3rd ed. LJLM Press.
Greiner, S., Lehwark, P., Bock, R., 2019. OrganellarGenomeDRAW (OGDRAW) version 1.3.1: expanded toolkit for the graphical visualization of organellar genomes. Nucleic Acids Res. 47, W59–W64. https://doi.org/10.1093/nar/gkz238.
Guiry, M.D., Guiry, G.M., 2022. AlgaeBase. World-wide electronic publication, National University of Ireland, Galway. https://www.algaebase.org; searched on 21/02/2022.
Hall, T.A., 1999. BioEdit: A user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symp. Ser. 41, 95–98.
Hee, W.Y., Blackman, L.M., Hardham, A.R., 2019. Characterisation of Stramenopile-specific mastigoneme proteins in *Phytophthora parasitica*. Protoplasma 256, 521–535. https://doi.org/10.1007/s00709-018-1314-1.
Hegewald, E., Padisák, J., Friedl, T., 2007. *Pseudotetraëdriella kamillae*: taxonomy and ecology of a new member of the algal class Eustigmatophyceae (Stramenopiles). Hydrobiologia 586, 107–116. https://doi.org/10.1007/s10750-006-0566-5.
Hibberd, D.J., 1974. Observations on the cytology and ultrastructure of *Chlorobotrys regularis* (West) Bohlin with special reference to its position in the Eustigmatophyceae. Br. Phycol. J. 9 (1), 37–46. https://doi.org/10.1080/00071617400650051.
Hibberd, D.J., 1981. Notes on the taxonomy and nomenclature of the algal classes Eustigmatophyceae and Tribophyceae (synonym Xanthophyceae). Bot. J. Linn. Soc. 82, 93–119. https://doi.org/10.1111/j.1095-8339.1981.tb00954.x.
Hibberd, D.J., Leedale, G.F., 1970. Eustigmatophyceae – a new algal class with unique organization of the motile cell. Nature 225 (5234), 758–760. https://doi.org/10.1038/225758b0.
Hibberd, D.J., Leedale, G.F., 1972. Observations on the cytology and ultrastructure of the new algal class. Eustigmatophyceae. Ann. Bot. 36, 49–71. https://doi.org/10.1093/oxfordjournals.aob.a084577.
Honda, D., Shono, T., Kimura, K., Fujita, S., Iseki, M., Makino, Y., Murakami, A., 2007. Homologs of the sexually induced gene 1 (*sig1*) product constitute the stramenopile mastigonemes. Protist 158, 77–88. https://doi.org/10.1016/j.protis.2006.08.004.
Kalyaanamoorthy, S., Minh, B.Q., Wong, T.K.F., von Haeseler, A., Jermiin, L.S., 2017. ModelFinder: fast model selection for accurate phylogenetic estimates. Nat. Methods 14, 587–589. https://doi.org/10.1038/nmeth.4285.
Katana, A., Kwiatowski, J., Spalik, K., Zakryś, B., Szalacha, E., Szymańska, H., 2001. Phylogenetic position of *Koliella* (Chlorophyta) as inferred from nuclear and chloroplast small subunit rDNA. J. Phycol. 37, 443–451. https://doi.org/10.1046/j.1529-8817.2001.037003443.x.
Katoh, K., Standley, D.M., 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol. Biol. Evol. 30, 772–780. https://doi.org/10.1093/molbev/mst010.
Kawai, H., Maeba, S., Sasaki, H., Okuda, K., Henry, E.C., 2003. *Schizocladia ischiensis*: a new filamentous marine chromophyte belonging to a new class, Schizocladiophyceae. Protist 154, 211–228. https://doi.org/10.1078/143446103322166518.
Kollmar, M., 2016. Fine-tuning motile cilia and flagella: evolution of the dynein motor proteins from plants to humans at high resolution. Mol. Biol. Evol. 33, 3249–3267. https://doi.org/10.1093/molbev/msw213.
Kryvenda, A., Rybalka, N., Wolf, M., Friedl, T., 2018. Species distinctions among closely related strains of Eustigmatophyceae (Stramenopiles) emphasizing ITS2 sequence-structure data: *Eustigmatos* and *Vischeria*. Eur. J. Phycol. 53, 471–491. https://doi.org/10.1080/09670262.2018.1475015.
Kück, P., Longo, G.C., 2014. FASconCAT-G: extensive functions for multiple sequence alignment preparations concerning phylogenetic studies. Front. Zool. 11, 81. https://doi.org/10.1186/s12983-014-0081-x.
Langmead, B., Salzberg, S., 2012. Fast gapped-read alignment with Bowtie 2. Nat. Methods 9, 357–359. https://doi.org/10.1038/nmeth.1923.
Lee, R.E., 2018. Phycology, 5th ed. Cambridge University Press.
Letunic, I., Bork, P., 2021. Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. Nucleic Acids Res. 49, W293–W296. https://doi.org/10.1093/nar/gkab301.
Li, L., Wang, S., Wang, H., Sahu, S.K., Marin, B., Li, H., Xu, Y., Liang, H., Li, Z., Cheng, S., Reder, T., Çebi, Z., Wittek, S., Petersen, M., Melkonian, B., Du, H., Yang, H., Wang, J., Wong, G.K., Xu, X., Liu, X., Van de Peer, Y., Melkonian, M., Liu, H., 2020.

The genome of *Prasinoderma coloniale* unveils the existence of a third phylum within green plants. Nat. Ecol. Evol. 4, 1220–1231. https://doi.org/10.1038/s41559-020-1221-7.

Luo, W., Li, H., Kotut, K., Krientitz, L., 2017. Molecular diversity of plankton in a tropical crater lake switching from hyposaline to subsaline conditions: Lake Oloidien, Kenya. Hydrobiologia 788, 205–229. https://doi.org/10.1007/s10750-016-2998-x.

Milne, I., Stephen, G., Bayer, M., Cock, P.J., Pritchard, L., Cardle, L., Shaw, P.D., Marshall, D., 2013. Using Tablet for visual exploration of second-generation sequencing data. Brief Bioinform. 14, 193–202. https://doi.org/10.1093/bib/bbs012.

Minh, B.Q., Schmidt, H.A., Chernomor, O., Schrempf, D., Woodhams, M.D., von Haeseler, A., Lanfear, R., 2020. IQ-TREE 2: New models and efficient methods for phylogenetic inference in the genomic era. Mol. Biol. Evol. 37 (5), 1530–1534. https://doi.org/10.1093/molbev/msaa015.

Nakayama, T., Nakamura, A., Yokoyama, A., Shiratori, T., Inouye, I., Ishida, K., 2015. Taxonomic study of a new eustigmatophycean alga, *Vacuoliviride crystalliferum* gen. et sp. nov. J Plant Res 128, 249–257. https://doi.org/10.1007/s10265-014-0686-3.

Ott, D.W., Oldham-Ott, C.K., Rybalka, N., Friedl, T., 2015. Xanthophyte, Eustigmatophyte, and Raphidophyte Algae. In: Wehr, J.D., Sheath, R.G., Kociolek, J. P. (Eds.), Freshwater Algae of North America: Ecology and Classification. Academic Press, Amsterdam, pp. 485–536.

Pascher, A., 1912. Zur Gliederung der Heterokonten. Hedwigia 53, 6–22.

Pascher, A., 1915. Über Halosphaera. Ber. Deutsch. Bot. Ges. 33, 488–492.

Pascher, A., 1925. Heterokontae, in: Pascher, A. (Ed.), Die Susswasser-Flora Deutschlands, Osterreich und der Schweiz, Heft 11: Heterokontae, Phaeophyta, Rhodophyta, Charophyta. G. Fischer, Jena, pp. 1–118.

Pascher, A., 1937–1939. Heterokonten, in: L. Rabenhorst's Kryptogamen-Flora von Deutschland, Osterreich und der Schweiz. Aufl. 2. Band XI. Akad. Verlagsges, Leipzig.

Phillips, K.A., Fawley, M.W., 2000. Diversity of coccoid algae in shallow lakes during winter. Phycologia 39, 498–506. https://doi.org/10.2216/i0031-8884-39-6-498.1.

Poulíčková, A., Mann, D.G., 2019. Diatom Sexual Reproduction and Life Cycles. In: Seckbach, J., Gordon, R. (Eds.), Diatoms: Fundamentals and Applications. Scrivener Publishing LLC, pp. 245–272.

Přibyl, P., Eliáš, M., Cepák, V., Lukavský, J., Kaštánek, P., 2012. Zoosporogenesis, morphology, ultrastructure, pigment composition, and phylogenetic position of *Trachydiscus minutus* (Eustigmatophyceae, Heterokontophyta). J. Phycol. 48, 231–242. https://doi.org/10.1111/j.1529-8817.2011.01109.x.

Rampen, S.W., Friedl, T., Rybalka, N., Thiel, V., 2022. The Long chain Diol Index: A marine palaeotemperature proxy based on eustigmatophyte lipids that records the warmest seasons. Proc. Natl. Acad. Sci. U. S. A. 119 https://doi.org/10.1073/pnas.2116812119.

Ronquist, F., Huelsenbeck, J.P., 2003. MRBAYES 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19, 1572–1574. https://doi.org/10.1093/bioinformatics/btg180.

Santos, L.M.A., Leedale, G.F., 1995. Some notes on the ultrastructure of small azoosporic members of the algal class Eustigmatophyceae. Nova Hedwigia 60, 219–225.

Ševčíková, T., Yurchenko, T., Fawley, K.P., Amaral, R., Strnad, H., Santos, L.M.A., Fawley, M.W., Eliáš, M., Katz, L.A., 2019. Plastid genomes and proteins illuminate the evolution of eustigmatophyte algae and their bacterial endosymbionts. Genome Biol. Evol. 11 (2), 362–379. https://doi.org/10.1093/gbe/evz004.

Silva, P.C., 1979. Review of the taxonomic history and nomenclature of the yellow-green algae. Arch. Protistenkunde 121 (1–2), 20–63. https://doi.org/10.1016/s0003-9365(79)80008-1.

Söding, J., 2005. Protein homology detection by HMM-HMM comparison. Bioinformatics 21, 951–960. https://doi.org/10.1093/bioinformatics/bti125.

Spurr, A.R., 1969. A low viscosity epoxy resin embedding medium for electron microscopy. J. Ultrastruct. Res. 26 (1–2), 31–43. https://doi.org/10.1016/S0022-5320(69)90033-1.

Stoyneva-Gärtner, M., Uzunov, B., Gärtner, G., Borisova, C., Draganova, P., Radkova, M., Stoykova, P., Atanassov, I., 2019. Current bioeconomical interest in stramenopilic Eustigmatophyceae: a review. Biotechnol. Biotechnol. Equip. 33, 302–314. https://doi.org/10.1080/13102818.2019.1573154.

Trzcińska, M., Pawlik-Skowrońska, B., Krokowski, D., Watanabe, S., 2014. Genetic and morphological characteristics of two ecotypes of *Eustigmatos calaminaris* sp. nov. (Eustigmatophyceae) inhabiting Zn- and Pb-loaded calamine mine spoils. Fottea 14, 1–13. https://doi.org/10.5507/fot.2014.001.

Turland, N.J., Wiersema, J.H., Barrie, F.R., Greuter, W., Hawksworth, D.L., Herendeen, P.S., Knapp, S., Kusber, W.H., Li, D.Z., Marhold, K., May, T.W., McNeill, J., Monro, A. M., Prado, J., Price, M.J., Smith, G.F., 2018. International Code of Nomenclature for algae, fungi, and plants (Shenzhen Code). Regnum Vegetabile 159. Koeltz Botanical Books, Glashütten.

Villanueva, L., Besseling, M., Rodrigo-Gámiza, M., Rampen, S.W., Verschuren, D., Sinningh Damsté, J.S., 2014. Potential biological sources of long chain alkyl diols in a lacustrine system. Org. Geochem. 68, 27–30. https://doi.org/10.1016/j.orggeochem.2014.01.001.

von Dassow, P., John, U., Ogata, H., Probert, I., Bendif el, M., Kegel, J.U., Audic, S., Wincker, P., Da Silva, C., Claverie, J.M., Doney, S., Glover, D.M., Flores, D.M., Herrera, Y., Lescot, M., Garet-Delmas, M.J., de Vargas, C., 2015. Life-cycle modification in open oceans accounts for genome variability in a cosmopolitan phytoplankton. ISME J. 9, 1365–1377. https://doi.org/10.1038/ismej.2014.221.

West, W., 1892. Algae of the English Lake District. J. R. Microsc. Soc. 1892, 713–748.

White, T.J., Bruns, T.D., Lee, S.B., Taylor, J.W., 1990. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In: Innis, M.A., Gelfand, D.H., Sninsky, J.J., White, T.J. (Eds.), PCR Protocols: A Guide to Methods and ApplicAtions. Academic Press Inc., New York, pp. 315–322.

Whittle, S.J., Casselton, P.J., 1975. The chloroplast pigments of the algal classes Eustigmatophyceae and Xanthophyceae. I. Eustigmatophyceae. Br. Phycol. J. 10 (2), 179–191. https://doi.org/10.1080/00071617500650171.

Wolf, B.M., Niedzwiedzki, D.M., Magdaong, N.C.M., Roth, R., Goodenough, U., Blankenship, R.E., 2018. Characterization of a newly isolated freshwater eustigmatophyte alga capable of utilizing far-red light as its sole light source. Photosynth. Res. 135, 177–189. https://doi.org/10.1007/s11120-017-0401-z.

Yamagishi, T., Motomura, T., Nagasato, C., Kawai, H., 2009. Novel proteins comprising the stramenopile tripartite mastigoneme in *Ochromonas danica* (Chrysophyceae). J. Phycol. 45, 1110–1115. https://doi.org/10.1111/j.1529-8817.2009.00722.x.

Yang, E.C., Boo, G.H., Kim, H.J., Cho, S.M., Boo, S.M., Andersen, R.A., Yoon, H.S., 2012. Supermatrix data highlight the phylogenetic relationships of photosynthetic stramenopiles. Protist 163, 217–231. https://doi.org/10.1016/j.protis.2011.08.001.

Yang, H.P., Wenzel, M., Hauser, D.A., Nelson, J.M., Xu, X., Eliáš, M., Li, F.W., 2021. *Monodopsis* and *Vischeria* genomes shed new light on the biology of eustigmatophyte algae. Genome Biol. Evol. 13, evab233. https://doi.org/10.1093/gbe/evab233.

Yurchenko, T., Ševčíková, T., Strnad, H., Butenko, A., Eliáš, M., 2016. The plastid genome of some eustigmatophyte algae harbours a bacteria-derived six-gene cluster for biosynthesis of a novel secondary metabolite. Open Biol. 6 (11), 160249. https://doi.org/10.1098/rsob.160249.

Yurchenko, T., Ševčíková, T., Přibyl, P., El Karkouri, K., Klimeš, V., Amaral, R., Zbránková, V., Kim, E., Raoult, D., Santos, L.M.A., Eliáš, M., 2018. A gene transfer event suggests a long-term partnership between eustigmatophyte algae and a novel lineage of endosymbiotic bacteria. ISME J. 12, 2163–2175. https://doi.org/10.1038/s41396-018-0177-y.