2021

A PHYLOGENOMIC APPROACH TO EXPLORE PHOTOSYNTHETHIC STRAMENOPILE EVOLUTION

Kristina Xenia Terpis
University of Rhode Island, kristina.terpis@gmail.com

A PHYLOGENOMIC APPROACH TO EXPLORE PHOTOSYNTHETHIC STRAMENOPILE EVOLUTION

BY

KRISTINA XENIA TERPIS

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE

REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

IN

BIOLOGICAL AND ENVIRONMENTAL SCIENCES

UNIVERSITY OF RHODE ISLAND

2021

MASTER OF SCIENCE THESIS

OF

KRISTINA X TERPIS

APPROVED:

Thesis Committee:

Major Professor Christopher Lane

Bethany Jenkins

Roxanne Beinart

Brenton DeBoef
DEAN OF THE GRADUATE SCHOOL

UNIVERSITY OF RHODE ISLAND
2021

ABSTRACT

Stramenopiles, Alveolates, and Rhizaria (SAR) are environmentally ubiquitous and form one of the most genetically diverse groups of eukaryotes on earth. Combined, these phyla encompass nearly half of all known eukaryotic genomic diversity. Stramenopiles comprise both non-photosynthetic taxa including important pathogens of animals and plants such as the gastric parasite *Blastocystis* and oomycetes like *Phytophthora*, the causative agent of the Irish Potato Famine. Additionally, photosynthetic members from the Phylum Ochrophyta, like the charismatic seaweeds such as giant kelp, create important ecosystems while diverse microbial phytoplankton including diatoms are fundamental to biogeochemical cycles and form the base of the food web. Despite the known diversity within the stramenopiles, the lack of obvious morphological characteristics for distinguishing taxa has led to some classes becoming a dumping ground for difficult to place taxa. The combination of largely understudied, yet genetically diverse taxa with few reliable distinguishing characteristics has complicated stramenopile systematics for centuries. The absence of a resolved phylogeny representing the diversity of the group significantly hinders the ability to make meaningful inferences about ochrophyte evolution. Here we expand the taxonomic sampling for ochrophytes by generating twenty-eight novel transcriptomes, focusing on classes lacking genome-scale data and species of uncertain taxonomic affiliation. By combining these data with other publicly available data, we constructed the most taxonomically comprehensive phylogeny of ochrophytes comprised of nineteen

out of the twenty-one previously described classes. Our analyses robustly resolve relationships between stramenopiles and support four main clades of ochrophytes. We placed historically difficult to resolve taxa. A rapid radiation in the diversification of ochrophytes remains problematic for resolving the root of ochrophytes and deep internal branches. These outstanding questions will likely require identification and sequencing of additional, and even novel ochrophyte taxa or methodological advances to extract a reliable phylogenetic signal from these ancient events.

ACKNOWLEDGMENTS

This research was made possible due to the support and encouragement of my advisor Dr. Chris Lane who enabled me to pursue my degree. I would also like to thank my committee members Dr. Roxanne Beinart and Dr. Bethany Jenkins for their support and advice. This work would also not have been possible without the expertise of my collaborators Dr. J. Craig Bailey, Dr. Matt Brown, Dr. Marek Eliaš, Dr. Martin Koliskó, and Dr. Heroen Verbruggen. I would like to thank the past and present members of the Lane lab for their friendship and advice. This research was supported in part by a Grant-in-Aid of Research awarded to me by The Phycological Society of America. And a special thank you to Dr. Salomaki, for everything

PREFACE

This thesis is presented in manuscript format. Chapter 1 is prepared for submission to Journal of Phycology as a literature review. Chapter 2 is in prepared for submission as a research article to Molecular Biology and Evolution.

TABLE OF CONTENTS

ABSTRACT ii

ACKNOWLEDGEMENTS iv

PREFACE v

TABLE OF CONTENTS vi

LIST OF TABLES vii

LIST OF FIGURES viii

CHAPTER 1 1

CHAPTER 2 19

LIST OF TABLES

TABLE PAGE

Table 1. Taxonomy, sources, and BUSCO completeness based on the Eukaryota ODB10 reference gene set for data used in this study 46

LIST OF FIGURES

FIGURE PAGE

Chapter 1

Figure 1. Timeline showing important events in stramenopile classification ..12

Chapter 2

Figure 1. Phylogeny of SAR 40

Figure 2. Phylogeny of ochrophytes + Pseudofungi.....42

Figure 3. Effects of removing fast evolving and heterotacheous sites, and random gene subsampling on apicomplexan relationships.....43

Figure 4. Possible root placement for the root of ochrophytes.....45

CHAPTER 1

This chapter is prepared for submission to Journal of Phycology

A Brief History of Stramenopiles

Introduction

The term Chromophyta was first introduced by Linnæus (1751) to refer to colored plants. The term “colored plants” was not to mean the opposite of colorless, but rather any color other than green. In 1813, Lamouroux was the first to distinguish algae based on their color. He established groups for the red, the brown, and the green algae called the Fucacées, Floridées, and Ulvaceés, along with two additional groups, one for *Codium* and the other for *Dictyota* (Lamouroux, 1813). Around the turn of the century, researchers started to utilize additional morphological features for algal classification including the flagella (or cilium) as a distinguishing characteristic. For example, Luther separated the yellow-green algae (todays Xanthophytes, specifically *Tribonema*) from the other green algae (Ulvaceés) due to differences in their flagella, and called this new group the Heterokontae (Luther, 1899). A few years later, Chodat suggested that all algal groups with a flagellated life stage should be placed into two groups, Chlorophyceés (green algae) and Phéophyceés, comprised of Luther’s Heterokontae, as well as diatoms, dinoflagellates, and euglenoids (Chodat, 1909). Over time, this definition changed as more ultrastructural studies revealed greater diversity among this group.

In 1952, the Stockholm Congress made the -phyta ending the standard terminology for divisions in the International Code of Botanical Nomenclature

and the term Chromophyta was later introduced by Bourrelly (1957). In 1960, Chadefaud updated the names of the algal groups to include -phyco- in them, to designate them as non-vascular plants (Chadefaud, 1960). The three groups that contained all eukaryotic algae were now known as Rhodophycophyta, Chlorophycophyta, and Chromophycophyta. Since red algae are also colored but lack flagella, Christensen (1962) introduced the group Contophora to encompass both the rhodophycophytes and the chromophycophytes, with the distinguishing characteristic of presence of a flagellum was present. Christensen (1962) gave the designation of Chromophyta (*nomen nudum)* as a group compromising taxa that contain chlorophylls *a* and *c*, lack chlorophyll *b*, and are flagellated. This included the current ochrophytes, haptophytes, cryptophytes, Dinophyta, and the choanoflagellates, while the euglenoids were transferred to the Chlorophyta. In 1986, Cavalier-Smith introduced the term Chromista consisting of three groups: the Heterokonts, Haptophytes, and Cryptomonads (Cavalier-Smith, 1986). The term ‘Stramenopile’ was later introduced by Patterson (1989) to encompass the clade of protists that shared the synapomorphy of tripartite hairs because the meaning of Heterokont had changed throughout time and became ambiguous. Stramenopiles (Patterson, 1989), also known as Heterokonts (Cavalier-Smith, 1986) remain as one of the most diverse eukaryotic clades.

Generally, stramenopiles have two flagella that are used for swimming. One is longer and smooth, while the other is shorter and has tripartite hairs

attached. Most vegetative cells of the Chrysophyceae, Bolidophyceae, Raphidophyceae, Pinguiophyceae, and Synurophyceae have two typical flagella (Andersen, 1989; Guillou et al., 1999; Heywood, 1990; Hibbered, 1976; Honda & Inouye, 2002). Similarly, the zoospores of some members of the Chrysomerophyceae, Eustigmatophyceae, Schizocladiophyceae, Phaeothamniophyceae, and Xanthophyceae also retain two flagella (Andersen et al., 1998; Hibbered, 1990; Kawai et al., 2003; O'Kelly, 1989). One of these flagella is smooth and shorter, while the other is longer and has tripartite hairs. The orientation of these flagella vary, and in some species the flagella can be pointing in the same direction, while in others, the flagella may be on opposite sides of the biflagellated cell. Further complicating this, some stramenopiles retain only a single flagellum in parts of their lifecycle and others have life stages that lack flagella. For example, some species of the Dictyochophyceae have only a single flagellum during their skeletal stage (Chang, 2015). Similarly, the male gametes of diatoms have only a single flagellum, while the vegetative state of diatoms have lost their flagella completely. The presence of tripartite hairs on the shorter flagellum is the only known synapomorphy of stramenopiles, uniting the heterotrophic stramenopiles with ochrophytes.

Throughout the twentieth century, pigment analysis was primarily used to characterize distinct groups of photosynthetic stramenopiles. It was clear that groups could be defined by the lack or presence of chlorophyll *b*, and differences in pigment profiles. It was believed at that time that all chromophytes contain chlorophylls *a* and *c* plus various accessory pigments,

with the most prevalent being fucoxanthin, which gives them their characteristic golden-brown color. Carotenoid distribution within the ochrophytes also varies. For example, fucoxanthin is always the predominate carotenoid in the Phaeophyceae, with the next major pigment being violaxanthin. When the class Chrysophyceae was established in 1914 by Pascher, he placed all taxa that were not diatoms or brown algae into this new class and created orders for each of the different cellular types such as coccoid, filamentous, and capsoid (Andersen et al., 1993). This class was then modified in the 1950's to contain all taxa that were not raphidophytes, xanthophytes, diatoms, or brown algae. In the 1970's six different species of what was then considered chrysophytes *sensu strictu* were examined for their carotenoid composition and it was found that the profiles varied wildly, causing confusion and belief that the chrysophytes were highly diverse in their pigment retention (Bjørnland, 1989). When the diversity of this group was eventually examined using electron microscopy coupled with these pigment analyses, this group was split into several different groups. Today, pigment composition correlates well with clades resolved using phylogenies from molecular data.

However, losses of these pigments have led to certain taxa being incorrectly assigned to other groups. For example, members of the Eustigmatophyceae and Xanthophyceae, which appear to be bright green in color due to the loss of fucoxanthin, were originally placed in the Chlorophyceae. Further pigment analysis and ultrastructural examination determined that they belong within the ochrophytes. As new taxa have been

described, two ochrophyte classes, the Eustigmatophyceae and the monotypic Aureanophyceae, have each independently lost chlorophyll *c*. In the Raphidophyceae, freshwater taxa resemble the Xanthophyceae in carotenoid composition, whereas marine taxa resemble those of the Chrysophyceae and Phaeophyceae. Christensen (1980) attempted to resolve this difference in taxa that contain or lack fucoxanthin, by placing fucoxanthin containing raphidophytes into the Chrysophyceae and merging those lacking it into the Xanthophyceae. However, these relationships did not hold up with ultrastructural characteristics.

Oomycetes were originally grouped with Fungi due to similarities in morphology, however with biochemical and ultrastructural information it was discovered that they share a common ancestor with other stramenopiles. Other studies further corroborated the relationship between the photosynthetic and heterotrophic stramenopiles. For example, Manton et al. (1951) examined the morphological similarities in the zoospores of *Saprolegnia* spp. (Oomycota) and the coenocytic *Vaucheria* spp. (Ochrophyta) and Parker et al. (1963) identified biochemical parallels between these two groups. Additionally, heterotrophic Bicosoecides resemble some sessile stages of photosynthetic chrysophytes. Ultrastructural data has confirmed the presence of two flagella with the shorter of the two having tripartite tubular hairs. This feature was found to unite the heterotrophic and photosynthetic taxa, however, the relationships between different groups of stramenopiles remained ambiguous

and clear morphological features demonstrating evolutionary histories were lacking.

As mentioned above, morphological similarities were able to unite heterotrophic and photosynthetic stramenopiles. For example, the oomycetes share a handful of structural similarities with the Xanthophyceae and Eustigmatophyceae such as a closed pattern of mitosis, which is rare in chromophyte algae (Beakes 1989). This relationship was further supported by the presence of electron-dense encystment vesicles around the outside of the zoospore (Beakes, 1983; Ott & Brown, 1974b, 1975) Additionally, oomycetes share similarities of the flagellar root system with Phaeophyceae and some members of the Chrysophyceae, yet it remained unclear how the group evolved and were specifically related. It was believed that the oomycetes shared a common ancestor with the xanthophytes, but the oomycetes secondarily lost the ability to photosynthesis and accordingly also lost their plastid. Alternative scenarios were proposed, including that chloroplasts were gained independently at least twice, or that a chromophytic stramenopile lost chloroplasts, and then regained them after the evolution within the plastid-less lineages. Li and Volcani (1987) proposed the latter to explain how apocholoric diatoms arose.

More recently, the placement of *Symbiomonas scintillans* and *Picophagus flagellatus* within the stramenopiles proved to be difficult. Both species are hetereotrophs that possess key stramenopile ultrastructural features, including mitochondria with tubular cristae and tripartite hairs on their

shorter flagellum, yet lack ultrastructural evidence for a plastid. Molecular phylogenetics were required to conclusively place these taxa within the phylum. Surprisingly, *Picophagus* was placed sister to the Chrysophyceae and Synurophyceae, indicating it was a heterotrophic ochrophyte (Guillou et al., 1999). Detailed ultrastructural analyses demonstrated that *P. flagellatus* has not only lost the ability to photosynthesize, but has apparently lost its plastid entirely (Guillou et al., 1999). Without molecular data these taxa would not have been included within the ochrophytes, but rather as a heterotrophic stramenopiles. The use of molecular phylogenetics made it clear that photosynthesis arose once in stramenopiles resulting from a secondary endosymbiotic event, dividing ochrophytes from heterotrophic stramenopiles (Andersen et al., 1999; Cavalier-Smith et al., 1995; Keeling, 2013; Sibbald & Archibald, 2020).

Throughout the evolution of stramenopiles there have been multiple independent loses of either photosynthetic capabilities, or the plastid entirely. While the most obvious role of the chloroplasts is photosynthesis, plastids also perform other cellular functions such as biosynthesis of fatty acids, isoprenoids, iron sulfur clusters, and/or haem (Salomaki & Kolisko, 2019; Sibbald & Archibald, 2020). The production of these, and other compounds, relies heavily on proteins encoded in the nucleus and targeted to the plastid, with many fewer of these proteins being encoded on the plastid genome itself. Without establishing other mechanisms to replace these essential metabolic

functions, the plastid remains essential, and is retained even without performing photosynthesis.

Several ochrophytes have independently lost photosynthetic ability, which has confounded their placement within the group. For example, members of the Synchromophyceae are mixotrophic and capable of acquiring nutrients without photosynthesis. While *Synchroma* spp. and *Chlamydomyxa* spp. are mixotrophic and retain photosynthetic capabilities, *Leukarachnion* is solely heterotrophic and retains a cryptic plastid whose function remains unknown. Although the Synchromophyceae exhibit a wide range of plastid morphologies and functions, they are united by their amoeboid (rhizopodial) morphology and lack of flagella in their vegetative state. The chrysophytes and synurophytes are two classes that are trophically diverse including obligate phototrophs, mixotrophs, and obligate heterotrophic lineages. Photosynthesis has been lost multiple times in the chrysophytes, and mixotrophy is very common. Dorrell et al. (2019) investigated the plastid of *Paraphysomonas* and found that, in contrast to all other nonphotosynthetic chrysophytes, *Paraphysomonas* spp. do not retain a plastid genome, but still retain the organelle that functions in haem biosynthesis.

While the chrysophytes, synurophytes, and synchromophytes are close relatives, diatoms represent a more distant ochrophyte lineage that has also undergone the loss of photosynthesis. There are at least twenty or so free-living species of *Nitzschia* that have lost the ability to photosynthesize. However, the nuclear encoded, plastid targeted genes, are still retained and

provided important metabolic capabilities (Kamikawa et al., 2015, 2017). Recent phylogenomic analysis suggests that there was a single loss of photosynthesis within *Nitzschia*, giving rise to a monophyletic grouping within the autotrophic *Nitzschia* spp. (Onyshchenko et al., 2019). The origin of the plastid enabled the diversification of ochrophytes and allowed them to occupy new ecological niches, however secondary losses of photosynthesis has further allowed this group to expand its nutrition acquisition strategies.

Both autotrophic and heterotrophic stramenopiles play significant roles in diverse ecosystems. Diatoms are major contributors to the global carbon cycle and the brown algae (kelp) create large ecosystems in the inter and subtidal. Oomycetes include notable plant pathogens such as *Phytophthora infestans*, the causative agent of the Irish potato famine, and *Blastocystis* is a parasite of humans and other animals. Meanwhile there is a scarcity of information for other taxa belonging to this phylum and with the increased use of molecular data for taxonomy novel lineages are continually found. The use of increasingly available molecular tools provides the ability to explore the evolution of ochrophytes and provide a greater understanding of their adaptation to unique environments. Recent studies have focused on resolving the evolutionary relationships within ochrophytes. Some studies have used few genes to explore many classes (Cavalier-Smith & Scoble, 2013; Wetherbee et al., 2018; Yang et al., 2012), while others used a phylogenomic approach using many genes, but did not have deep taxonomic sampling (Derelle et al., 2016; Leonard et al., 2018; Strassert et al., 2019). Research

presented in this thesis adds transcriptomic data for 7 ochrophyte classes previously lacking genome-scale data. This work investigates the evolution of ochrophytes using the most taxonomically comprehensive phylogenomic study of stramenopiles, to date.

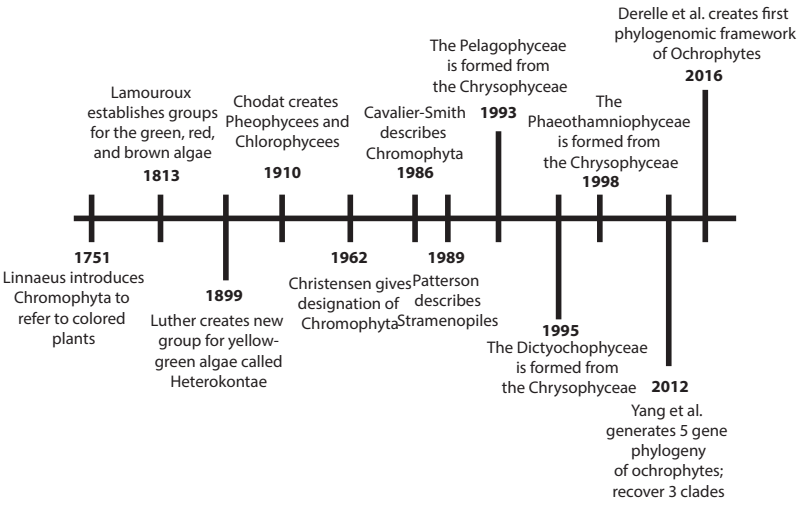

Figure 1. **Timeline showing important events in stramenopile classification**. Our understanding of the evolutionary relationships of this diverse group has changed considerably overtime with morphological, pigment, and molecular data.

BIBLIOGRAPHY

Andersen, R.A. 1989. The Synurophyceae and their relationship to other golden algae. In J. Kristiansen, G. Cronberg, and U. Geissler [eds.], Chrysophytes developments and perspectives. *Nova Hedwigia Beiheft*. 95:143–59.

Andersen, R.A., de Peer, Y.V., Potter, D., Sexton, J.P., Kawachi, M. & LaJeunesse, T. 1999. Phylogenetic analysis of the SSU rRNA from members of the Chrysophyceae. *Protist*. 150:71–84.

Andersen, R.A., Potter, D., Bidigare, R.R., Latasa, M., Rowan, K. & O'Kelly, C.J. 1998. Characterization and phylogenetic position of the enigmatic golden alga *Phaeothamnion confervicola*: Ultrastructure, pigment composition and partial SSU rDNA sequence. *Journal of Phycology*. 34:286–98.

Andersen, R.A., Saunders, G.W., Paskind, M.P. & Sexton, J.P. 1993. Ultrastructure and 18S rRNA gene sequence for *Pelagomonas calceolata* gen. et sp. nov. and the description of a new algal class, the Pelagophyceae classis nov. *Journal of Phycology*. 29:701–15.

Beakes, G.W. 1983. A comparative account of cyst coat ontogeny in saprophytic and fish-lesion (pathogenic) isolates of the *Saprolgenia diclina-parasitica* complex. *Canadian Journal of Botany*. 603–25.

Beakes, G.W 1989. Oomycete fungi: their phylogeny and relationship to chromphyte algae. *The chromophyte algae: problems and perspectives*. 38:325–42.

Bjørnland, T. 1989. Distribution patterns of carotenoids in relation to chromophyte phylogeny and systematics. *The chromophyte algae: problems and perspectives*. 37–61.

Bourrelly, P. 1957. Recherches sur les Chrysophycées. *Revue Algologique, Mémoire Hors-Série*. 1.

Cavalier-Smith, T. 1986. The kingdom Chromista: origin and systematics. *Progress in Phycologial Research*. 4:309–47.

Cavalier-Smith, T., Chao, E.E. & Allsopp, M.T.E.P. 1995. Ribosomal RNA evidence for chloroplast loss within Heterokonta: Pedinellid relationships and a revised classification of ochristan algae. *Archiv für Protistenkunde*. 145:209–20.

Cavalier-Smith, T. & Scoble, J.M. 2013. Phylogeny of Heterokonta: *Incisomonas marina*, a uniciliate gliding opalozoan related to Solenicola (Nanomonadea), and evidence that Actinophryida evolved from raphidophytes. *European Journal of Protistology*. 49:328–53.

Chadefaud, M. 1960. Les végétaux non vascularies (Cryptogamie). *Traité de botanique systématique*.

Chang, F.H. 2015. Cell morphology and life history of *Dictyocha octonaria* (Dictyochophyceae, Ochrophyta) from Wellington Harbour, New Zealand. *Phycological Research*. 63:253–64.

Chodat, R. 1909. Étude critique et expérimentale sur le polymorphisme des algues. *George, Geneva*.

Christensen, T. 1962. Alger. *Botanik, Db. 2, Systembatisk Botanik, Nr. 2*. 1–178.

Christensen, T. 1980. Algae, a taxonomic survey. *Fasc. 1. AiO Tryk, Odense*.

Derelle, R., López-García, P., Timpano, H. & Moreira, D. 2016. A phylogenomic framework to study the diversity and evolution of stramenopiles (=heterokonts). *Mol Biol Evol*. 33:2890–8.

Dorrell, R.G., Azuma, T., Nomura, M., Audren de Kerdrel, G., Paoli, L., Yang, S., Bowler, C. et al. 2019. Principles of plastid reductive evolution illuminated by nonphotosynthetic chrysophytes. *Proceedings of the National Academy of Sciences*. 116:6914–23.

Guillou, L., Chrétiennot-Dinet, M.-J., Boulben, S., Moon-van der Staay, S.Y. & Vaulot, D. 1999. *Symbiomonas scintillans* gen. et sp. nov. and *Picophagus flagellatus* gen. et sp. nov. (Heterokonta): Two new heterotrophic flagellates of picoplanktonic size. *Protist*. 150:383–98.

Heywood, P. 1990. Phylum Raphidophyta. *L. Margulis, J. O. Corliss, M. Melkonian, and D. J. Chapman [eds.], Handbook of Protoctista*. 318–25.

Hibbered, D.J. 1976. The ultrastructure and taxonomy of the Chrysophyceae and Prymnesiophyceae (Haptophyceae): a survey with some new observations on the ultrastructure of the Chrysophyceae. *Botanical Journal of the Linnean Society*. 72:55–80.

Hibbered, D.J. 1990. Phylum Eustigmatophyta. *L. Margulis, J. O. Corliss, M. Melkonian, and D. J. Chapman [eds.], Handbook of Protoctista*. 326–33.

Honda, D. & Inouye, I. 2002. Ultrastructure and taxonomy of a marine photosynthetic stramenopile *Phaeomonas parva* gen. et sp. nov. (Pinguiophyceae) with emphasis on the flagellar apparatus architecture. *Phycological Research*. 75–89.

Kamikawa, R., Moog, D., Zauner, S., Tanifuji, G., Ishida, K.-I., Miyashita, H., Mayama, S. et al. 2017. A non-photosynthetic diatom reveals early steps of reductive evolution in plastids. *Molecular Biology and Evolution*. 34:2355–66.

Kamikawa, R., Yubuki, N., Yoshida, M., Taira, M., Nakamura, N., Ishida, K., Leander, B.S. et al. 2015. Multiple losses of photosynthesis in *Nitzschia* (Bacillariophyceae). *Phycological Research*. 63:19–28.

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*. 154:211–28.

Keeling, P.J. 2013. The number, speed, and impact of plastid endosymbioses in eukaryotic evolution. *Annual Review of Plant Biology*. 64:583–607.

Lamouroux, J.V.F. 1813. Essai sur les genres de la famille des thalassiophytes non articulées. *Annales du Muséum d'Histoire Naturelle*. 20:21–47, 115–39, 267–93.

Leonard, G., Labarre, A., Milner, D.S., Monier, A., Soanes, D., Wideman, J.G., Maguire, F. et al. 2018. Comparative genomic analysis of the 'pseudofungus' *Hyphochytrium catenoides* . *Open Biology*. 8:170184.

Li, C.W. & Volcani, B.E. 1987. Four new apochlorotic diatoms. *Journal of Protozoology*. 37–58.

Linnæus, C. 1751. Philosophia botanica.

Luther, A. 1899. Ueber Cholorsaccus eine neue Gattung der Süsswasseralgen, nebst einigen Bemerkungen zur Systematik vermandter Algen. *Bihang till Kungliga Svenska Vetenskapsakademiens Handlingar*. 24:1–22.

Manton, I., Clarke, B. & Greenwood, A.D. 1951. Observations with the electron microscope on a species of Saprolegnia. *Journal of Experimental Botany*. 2:321–31.

O'Kelly, C.J. 1989. The evolutionary origin of the brown algae: information from studies of motile cell structure. *The chromophyte algae: problems and perspectives*.

Onyshchenko, A., Ruck, E.C., Nakov, T. & Alverson, A.J. 2019. A single loss of photosynthesis in the diatom order Bacillariales (Bacillariophyta). *American Journal of Botany*. 106:560–72.

Ott, Donald W. & Brown, R.W. 1975. Developmental cytology of the genus *Vaucheria* III. Emergence, settlement and germination of the mature zoospore of *V*. *fontinalis*. *British Phycological Journal*. 10:49–56.

Ott, D.W. & Brown, R.M. 1974b. Developmental cytology of the genus *Vaucheria* II. Sporogenesis in *V. fontinalis* (L). *British Phycological Journal*. 9:333–51.

Parker, B.C. 1963. Studies of the structure and chemical composition of the cell walls of Vaucheriaceae and the Saprolegniaceae. *Proceedings of the Royal Society*. B158:435–44.

Patterson, D.J. 1989. Stramenopiles: Chromophytes from a protistan perspective. *Chromophyte Algae Problems and Perspectives*. 357–79.

Salomaki, E.D. & Kolisko, M. 2019. There is treasure everywhere: reductive plastid evolution in Apicomplexa in light of their close relatives. *Biomolecules*. 9.

Sibbald, S.J. & Archibald, J.M. 2020. Genomic insights into plastid evolution. *Genome Biol Evol*. 12:978–90.

Strassert, J.F.H., Jamy, M., Mylnikov, A.P., Tikhonenkov, D.V. & Burki, F. 2019. New phylogenomic analysis of the enigmatic phylum Telonemia further resolves the eukaryote tree of life. *Mol Biol Evol*. 36:757–65.

Wetherbee, R., Jackson, C.J., Repetti, S.I., Clementson, L.A., Costa, J.F., van de Meene, A., Crawford, S. et al. 2018. The golden paradox – a new heterokont lineage with chloroplasts surrounded by two membranes. *Journal of Phycology*.

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*. 163:217–31.

CHAPTER 2

This manuscript is prepared for submission to Molecular Biology and Evolution

A Phylogenomic Approach to Explore Photosynthetic Stramenopile Evolution

Abstract

Stramenopiles, Alveolates, and Rhizaria (SAR) are environmentally ubiquitous and form one of the most genetically diverse groups of eukaryotes on earth. Combined, these phyla encompass nearly half of all known eukaryotic genomic diversity. Stramenopiles are comprised of lineages that are non-photosynthetic including important pathogens of animals and plants such as the gastric parasite *Blastocystis* and oomycetes like *Phytophthora*, the causative agent of the Irish Potato Famine. Additionally, there is a singular photosynthetic lineage (Ochrophyta) which includes the charismatic seaweeds (e.g., giant kelp) which create important ecosystems and diverse microbial phytoplankton (e.g., diatoms) that are fundamental to biogeochemical cycles and form the base of the food web. Despite the known diversity within the stramenopiles, the lack of obvious morphological characteristics for distinguishing taxa has led to some classes becoming a dumping ground for difficult to place taxa. The combination of largely understudied, yet genetically diverse taxa with few reliable distinguishing characteristics has complicated stramenopile systematics for centuries. The absence of a resolved phylogeny representing the diversity of the group significantly hinders the ability to make meaningful inferences about ochrophyte evolution. Here we expand the taxonomic sampling for ochrophytes by generating twenty-eight novel transcriptomes, focusing on classes lacking genome-scale data and species of

uncertain taxonomic affiliation. By combining these data with other publicly available data, we constructed the most taxonomically comprehensive phylogeny of ochrophytes comprising nineteen out of the twenty-one previously described classes. Our analyses robustly resolve relationships between stramenopiles and support four main clades of ochrophytes and placed historically difficult to resolve taxa. A rapid radiation in the diversification of ochrophytes remains problematic for resolving the root of ochrophytes and deep internal branches. These outstanding questions will likely require identification and sequencing of additional, and even novel, ochrophyte taxa or methodological advances to extract a reliable phylogenetic signal from these ancient events.

Introduction

The stramenopiles, alveolates, and Rhizaria form one of the most diverse major eukaryotic clades on earth and is commonly referred to as SAR (Grattepanche et al., 2018). Some members of SAR have been intensely studied, such as the parasitic *Blastocystis*, oomycetes (stramenopiles), or *Plasmodium* (Alveolata). Photosynthetic lineages such as diatoms and kelp (stramenopiles), and dinoflagellates (alveolates) have also been extensively studied due to their importance as primary producers in marine ecosystems. However, many members of this supergroup remain largely ignored (Lin et al., 2012; Massana et al., 2004).

Photosynthesis has a complex evolutionary history across SAR exemplified by multiple, independent, plastid acquisitions across the supergroup (Delwiche, 1999; Derelle et al., 2016; Dorrell et al., 2017; Keeling, 2013; Ševčíková et al., 2016; Strassert et al., 2021). Stramenopiles include a single photosynthetic clade, termed Ochrophyta (Cavalier-Smith et al., 1995), that branch sister to numerous heterotrophic lineages. The ochrophytes are an incredibly diverse clade of stramenopiles, including members ranging in complexity from 30-m long multicellular kelp, to unicellular protists <20-µm in size. They are found in marine, freshwater, and terrestrial habitats and exhibit a wide range of morphological, life history, and nutritional strategies. Surveys of marine environments have revealed that stramenopiles make up a large fraction of the poorly studied microscopic marine eukaryotes (Logares et al., 2014; Pernice et al., 2016).

Despite diverse morphologies, a characteristic that remains consistent among all stramenopiles is that motile cells possess two unequal flagella: one long flagellum oriented towards the anterior of the cell with tripartitie tubular hairs (mastigonemes), and a shorter and generally posterior and smooth flagellum. The underlying synapomorphy of ochrophytes is a plastid of red algal origin acquired via an endosymbiotic event (Dorrell et al., 2017). Current evidence indicates a single plastid acquisition in the ancestor of ochrophytes (Dorrell et al., 2017; Ševčíková et al., 2015; Stiller et al., 2014), has been followed by a few independent losses of photosynthesis (Dorrell et al., 2019; Grant et al., 2009; Kamikawa et al., 2017; Onyshchenko et al., 2019). Most

ochrophyte plastids are characterized by the presence of chlorophylls *a* and *c*, with fucoxanthin as the predominant xanthophyll (Andersen, 2004). However, photosynthetic pigment profiles vary considerably, and differential losses of accessory pigments have led to the incorrect assignment of some species to other stramenopile classes, or even other phyla (Ott et al., 2015).

Currently, 17 classes of Ochrophyta are recognized, many of which have been described in the last 30 years, based solely on small-subunit (SSU) rRNA genes, or other single gene phylogenies. Single gene phylogenies have been efficient for placing newly discovered taxa in known or novel classes when distinguishing morphology is lacking, but phylogenies employing the SSU are unable to resolve deep relationships (Gentekaki et al., 2014; Hampl et al., 2009). Furthermore, clades with weak statistical support are unstable and prone to move with the addition of more, or different data. The lack of informative morphological characteristics for distinguishing ochrophyte classes has had historical consequences for inferring evolutionary relationships (Derelle et al., 2016). For example, prior to molecular techniques, characteristically biflagellated brown cells were often classified within Chrysophyceae (Bailey et al., 1998). From careful examination of these cells, new classes were formed based on light microscopy, ultrastructural, and DNA sequence data, such as the Phaeothamniophyceae (Bailey et al., 1998) Pelagophyceae (Andersen et al., 1993), and Chrysomerophyceae (Cavalier-Smith et al., 1995).

In an attempt to resolve the evolutionary relationships among ochrophytes, Yang et al. (2012), used a five-gene alignment, recovering three main clades that they termed SI, SII, and SIII. Shortly after, the Marine Microbial Eukaryotic Sequencing Project (MMETSP) represented a fundamental turning point in data availability for stramenopiles. Over 650 novel transcriptomes were sequenced from across the tree of life (Keeling et al., 2014) and nearly half of the data generated (269 transcriptomes), were from stramenopiles. This focus was intentional, due to the low abundance of available data relative to the diversity of the phylum. However, even these data were heavily biased towards diatoms (Diatomeae *sensu lato*), due to their importance in global carbon cycling. Despite taxonomic biases, these data were combined with additional new transcriptomes to produce the first phylogenomic analysis of ochrophytes (Derelle et al., 2016). While there was robust support for the monophyly and relationships among the classes represented, data for only 9 of the 17 known classes ochrophytes of were included. Subsequently, the expanded application of high-throughput sequencing enabled phylogenomic studies to significantly improve the resolution of relationships within and among stramenopile classes (Dorrell et al., 2019; Jackson et al., 2017; Leonard et al., 2018; Noguchi et al., 2016).

The range of morphology, life history, ecology, and overall diversity of stramenopiles, provides enormous scope for studying the evolution of traits. However, the lack of a resolved ochrophyte phylogeny representing the breadth of diversity of the group significantly hinders the ability to make

meaningful inferences about their evolution. Here, we expand the taxonomic sampling for ochrophytes, focusing on classes lacking genome-scale data, and species of uncertain taxonomic affiliation. We constructed the most taxonomically comprehensive ochrophyte phylogeny to date, comprised of 17 previously described classes. Within this robustly resolved evolutionary framework, we place historically difficult to resolve taxa in three different classes and transfer two others to different ochrophyte classes. Ultimately, we have improved class level taxonomic representation and phylogenetic resolution among ochrophyte lineages. An apparent rapid radiation early in the evolutionary history of ochrophytes remains problematic for resolving deep branches within the clade and will likely require additional taxa and methodological advances to extract a reliable phylogenetic signal from these ancient events.

Results/Discussion

Phylogenomic Analyses

Our phylogenomic dataset consists of 242 genes from 114 taxa (Table 1), resulting in a concatenated alignment of 75,752 amino acid (AA) sites with 18.71% missing data across the entire dataset. As in previous stramenopile phylogenies, we recover full support for Ochrophyta, consistent with a single plastid origin in stramenopiles (Dorrell et al., 2017; Sibbald & Archibald, 2020). Our 28 newly sequenced transcriptomes were combined with publicly available data to robustly place five ochrophyte classes previously lacking

molecular data (Phaeothamniophyceae, Phaeosacciophyceae, Chrysoparadoxophyceae, Picophagophycae, and Olisthodiscophyceae) and two taxa that were considered *incertae sedis* (*Botrydiopsis pyrenoidosa* SAG31.83 and *Phaeobotrys solitaria* SAG15.95) (Figure 1). The topology resolves a fully supported Diatomista and Chrysista, similar to the results presented in Derelle et al. (2016). Our dataset also provides full support for the monophyly of all classes, including the Phaeothamniophyceae, Phaeosacciophyceae, Eustigmatophyceae, and the Pinguiophyceae. Additionally, we confidently place the classes that are monotypic in our dataset: Picophagea, Chrysoparadoxophyceae, and Schizocladiophyceae (Figure 1).

In our phylogeny, Diatomeae and Bolidophyceae form a clade (Khakista), with the Dictyochophyceae and Pelagophyceae branching together to form a sister clade (Hypogyrista) (Figure 1). These four classes comprise the Diatomista (*sensu* Cavalier-Smith), which has also been called SIII by Yang et al. (2012). In agreement with the SII clade recovered by Yang et al. (2012), we recover a clade containing the Chrysophyceae, Synurophyceae, and Synchromophyceae. The Synurophyceae has been established as a class separate from the Chrysophyceae, although previous analyses suggest that the Synurophyceae are an internal branch of the Chrysophyceae (Dorrell et al., 2019; Noguchi et al., 2016). Dorrell et al. (2019) demonstrated the paraphyly of the Chrysophyceae *sensu stricto*, by focusing on these two classes in their study, where *Paraphysomonas* spp. (Chrysophyceae) branch

sister to a clade containing the remaining Chrysophyceae and the Synurophyceae. We recover this same relationship with full support (Figure 1). We are unable to resolve this paraphyly with the data presented here, however, increased taxonomic sampling of chrysophyte lineages may help resolve these relationships.

We recover the Pinguiophyceae and the recently described class, Olisthodiscophyceae, branching sister to the SII clade. *Olisthodiscus luteus* was classified as a raphidophyte (Hara & Chihara, 1987; Inouye et al., 1992) before Cavalier-Smith & Scoble (2013) weakly recovered this taxon as sister to the Pelagophyceae and Dictyochophyceae in their 18S rDNA phylogeny. Our study agrees with the recent results of a plastid gene-based phylogenetic analysis, that *O. luteus* is a deeply divergent sister lineage to the Pinguiophyceae, and thus constitutes a separate class, Olisthodiscophyceae (Barcyté et al., 2021). With the addition of our new data, we expand the taxonomic breadth of the SII clade recovered by Yang et al. (2012) and Wetherbee et al. (2018). However, this topology places the pinguiophytes as early branching members of the SII clade, as opposed to their inclusion as members of the Diatomista by Adl et al. (2019).

The remaining ochrophytes form a clade sister to the SII group with the Eustigmatophyceae being the earliest diverging lineage. The Raphidophyceae then branch sister to a fully supported clade, widely known as the PX clade, that is comprised of Phaeophyceae, Xanthophyceae, and several smaller classes (Figure 1). Other studies also recover strong support for the grouping

of the Raphidophyceae with the PX clade and this has been termed SI (Yang et al., 2012). While other studies have recovered strong support for the PX clade, the relationships within the group were always poorly resolved (Kawachi et al., 2002; Kawai et al., 2003; Wetherbee et al., 2018; Yang et al., 2012). We recover full support for all taxa contained within the PX clade, except for the node containing Phaeothamniophyceae and Phaeophyceae, which only has 87% bootstrap support (Figure 1).

Resolving Contentious Ochrophyte Relationships

The eustigmatophytes have historically proven difficult to place within the ochrophyte phylogeny, with affinities being linked to clades SI, SII, or branching as sister to all other ochrophytes depending on genes and taxa used in the analyses (Bailey et al., 1998; Kai et al., 2008; Kawachi et al., 2002; Ševčíková et al., 2016; Yang et al., 2012). However, in these studies the placement of the Eustigmatophyceae is rarely well-supported. Although phylogenomic analyses have provided the ability to resolve historically difficult relationships, the first study dedicated to investigating the relationships among ochrophytes did not include the Eustigmatophyceae (Derelle et al., 2016). Subsequent phylogenomic studies that included the Eustigmatophyceae within their phylogenomic datasets were unable to robustly place this class (Leonard et al., 2018; Noguchi et al., 2016). Here, we recover the eustigmatophytes branching sister to the SI clade with full statistical support in ML and from Bayesian analyses (Figure 1).

Although our analyses provided full support for the Eustigmatophyceae forming a sister lineage to the SI clade, there were several very short, internal branches present in the ochrophyte early divergence. Such rapid radiations have proven difficult to resolve (Giarla & Esselstyn, 2015; Mitchell et al., 2017; Rothfels et al., 2017; Whitfield & Lockhart, 2007), and given the variety of previously published results, we decided to further investigate the deep evolutionary history. We repeated our ML analyses after removing all distant, long-branching outgroups (Alveolata, Rhizaria, Bigyra), leaving Pseudofungi, the closest known relatives of ochrophytes with publicly available genome-scale data, as our outgroup taxa. In this analysis, the relationships between most classes remains unchanged. However, we now recover a clade comprised of the SII and SIII groups, with SI branching sister to them (Figure 2) as opposed to SI and SII branching together (Figure 1). Our analyses after removing long branching outgroup taxa places the Eustigmatophyceae as the earliest diverging class of ochrophytes (Figure 2). This placement is uncommon in phylogenetic analyses, however, this relationship is supported by mitochondrial data. The Eustigmatophyceae has previously been reported as the only ochrophyte lineage to encode the *atp*1 gene on its mitochondrion genome (Ševčíková et al., 2016). Moreover, in their phylogeny of the *atp*1 gene, the authors recover the eustigmatophytes among heterotrophic stramenopiles, which also encode *atp*1 on the mitochondrion genome (Ševčíková et al., 2016). This placement suggests that retention of the *atp*1 gene on the mitochondrion genome may be an ancestral trait of stramenopiles

which was lost or transferred to the nuclear genome early in the ochrophyte radiation. Accordingly, the topology showing the Eustigmatophyceae as the earliest branching ochrophyte lineage could indicate the transfer of the gene occurred after they diverged from other ochrophytes.

As noted above, our phylogeny also fully supports the relationship of *Olisthodiscus* with the Pinguiophyceae as members of the SII clade. However, the placement of the pinguiophytes within SII, rather than the Diatomista (SII), as described in the recent taxonomic updates by Adl et al. (2019), warrants further investigation. Interestingly, Barcyté et al. (2021) recover *Olisthodiscus* as sister to the pinguiophytes in their concatenated plastid encoded dataset, and 18S and 28S concatenated Bayesian tree, but not in the 18S and 28S concatenated ML tree. The relationship between these lineages is therefore also worth further analyses. This is especially true, given the fact that both *Olisthodiscus* and pinguiophytes represent long branches within the ochrophyte phylogeny, suggesting that their close relationship may be an artifact of long branch attraction (LBA).

In an effort to resolve the true root of ochrophytes, as well as the relationship of Pinguiophyceae and Olisthodiscophyceae, we ran Approximately Unbiased (AU) tests on the competing topologies, and tested the possible impact of systematic errors (e.g., LBA) and gene selection as drivers of support for either of the topologies. We performed serial fast-evolving and heterotacheous site removal (FSR and HSR, respectively) on both of our datasets, with the 5000 fastest or most heterotacheous sites being

removed at each step and assessed the support for several clades of interest in each newly generated dataset (Figure 3). In the SAR dataset, for both the fast site removal and heterotachy analyses, the support of eustigmatophytes belonging to the SI quickly lost support. Interestingly, the placement of the eustigmatophytes with SII is supported in the fast site removal and heterotachy once 40,752 sites remained, but this is roughly the same point once the monophyly of SII begins to fall apart. The AU test on the SAR dataset rejected the topology of eustigmatophytes branching first (P=0.039) and eustigmatophytes being a member of the SII clade (P=0.003) with the only topology not rejected being eustigmatophytes with SI (P=0.799).

For the ochrophytes + Pseudofungi dataset, the placement of eustigmatophytes with SI is never supported in either fast site removal or heterotachy analysis. The support for eustigmatophytes being sister to all other ochrophytes remained strong until 31,189 sites remained, when eustigmatophytes branched with SII for just one step with 60% bootstrap support before the relationship switched back. The placement of the eustigmatophytes branching sister to all other ochrophytes was the only supported topology in the heterotachy analysis with good support (90ML) (Figure 3). The AU test on the ochrophyte + Pseudofungi dataset rejected the placement of eustigmatophytes as a part of the SII clade (P=0.001) but was unable to reject the eustigmatophytes branching with the SI clade (P=0.117), or eustigmatophytes as the earliest branching ochrophytes (P=0.872).

A comparison of random gene subsampling for the SAR dataset and the ochrophyte + Pseudofungi dataset provided a striking difference in support for the placement of the Eustigmatophyceae within ochrophytes. In the 40%, 60%, and 80% subsets of the SAR dataset, Eustigmatophyceae branching with the SI clade was strongly supported. Only in the 20% dataset does support for this grouping waver, but the Eustigmatophyceae are never well supported as sister to the remaining ochrophytes, or branching with the SII clade. When the more divergent outgroup taxa are removed from the dataset and only Pseudofungi are used as an outgroup in the analyses, the Eustigmatophyceae are strongly supported as a sister lineage of all other ochrophytes, with this topology receiving the highest support in the 40% and 80% datasets. In the 20% and 60% datasets, support for the Eustigmatophyceae branching sister to ochrophytes, or the SI was highly variable. These results demonstrate that the placement of the Eustigmatophyceae with the SI clade is strongly impacted by gene selection and highlights the difficultly of resolving relationships at the ochrophyte radiation. This conflict may be a result of genes have differing rates of evolution which could be producing these differing topologies. Certain genes that are included may not contain enough phylogenetically informative signal to resolve deeper or shallower relationships resulting in different results. This is especially true when trying to resolve deep evolutionary relationships, as adequate phylogenetic signal in the selected genes may have been lost.

The relationship between the Pinguiophyceae and Olisthodiscophyceae was strongly supported in both the SAR and ochrophyte + Pseudofungi dataset for FSR, HSR, and random gene subsampling datasets (Figure 3). Throughout the analyses, a relationship for the Pinguiophyceae + SIII, or Pinguiophyceae + Olisthodiscophyceae + SIII was never supported. In both datasets, topologies where pinguiophytes were branching with SIII, and pinguiophytes + *Olisthodiscus* + SIII were rejected. This confirms the finding of Barcyté et al. (2021) that *Olisthodiscus* is a deeply divergent taxon, related to the Pinguiophyceae, and rejects the affinity of the Pinguiophyceae with SIII as summarized in Adl et al (2019). We also never recover any support for a clade containing *Leukarachion*, *Chylamdomyxa*, *Synchroma*, and *Picophagus*, demonstrating that *Picophagus* and the Synchromophyceae are separate classes.

Although it is still uncertain where the root of ochrophytes sits, our topologies, both with and without distant ancestors as outgroups, identify Eustigmatophyceae diverging from the members of the SI, SII, and SIII clades as described by Yang et al. (2012). Based on this pattern, the Eustigmatophyceae represent an independent early diverging lineage of Ochrophyta that cannot be considered belonging to the SI, SII, or SIII clades. Further supporting this hypothesis, eustigmatophytes are the only known ochrophyte class to have preserved the *atp*1 gene in the mitochondrial genome. All other ochrophytes have a nuclear encoded mitochondrial targeted *atp*1 protein. Our Pseudofungi + ochrophyte topology supports a scenario in

which the *atp*1 gene was retained in the mitochondrion in the common ancestor of ochrophytes, and then transferred to the nucleus in the common ancestor of the remaining ochrophytes. In both of our phylogenies we recovered full support for four distinct clades of ochrophytes, SI, SII, SIII, and eustigmatophytes, with the key difference being the root placement (Figure 4). Fully resolving the placement of the ochrophyte root will likely require the identification and inclusion of molecular data from more deep branching ochrophyte taxa and perhaps novel heterotrophic stramenopiles, or the creation of evolutionary models better suited to identify phylogenetic signal in short internal branches, as suggested by Di Franco et al. (2021).

Conclusions

Here we have the most complete reconstruction of ochrophyte evolutionary history to date with 17 known classes. We resolved the placement of *Olisthodiscus luteus* with the pinguiophytes as early branching members of the SII clade. The placement of the root of ochrophytes remains a challenge that may require increased sampling of early branching photosynthetic stramenopiles. Finally, our analyses indicate that the Eustigmatophyceae should be considered as an independent ochrophyte lineage rather than a member of any of the previous clades.

Methods

Material sources and culturing

Cultures were purchased from culture collections based on availability. The cultures were grown in 250mL glass flasks under the culture collections recommended culturing conditions in a Percival incubator. Culture growth was monitored by fluorometer and once each culture reached mid-exponential phase, cells were harvested over a 2-micron 25-millimeter poly carbonate filter and flash frozen in liquid nitrogen and stored in a -80º C freezer until RNA was isolated.

RNA isolation, sequencing, and assembly

RNA was extracted using Machery Nagel plant RNA kit following manufacturers protocol with the exception of cell lysis. This was carried out by bead beating for 5 minutes using a mixture of 0.1 and 0.5mm zirconia/silica beads (BioSpec) in a BioSpec Mini-Beadbeater.

Sequencing libraries were constructed using the NEB Next Ultra RNA prep kit and sequenced on the Illumina HiSeq (4000) in paired-end mode (2x150bp). Adapters and low-quality regions were removed using Trimmomatic (Bolger et al., 2014). Trimmed reads were assembled using rnaSPAdes v3.13 (Bushmanova et al., 2018) and transcripts were translated to protein sequences using Transdecoder. Completeness of all genomes or transcriptomes after removal of low coverage sequences in this study were assessed using BUSCO with the Eukaryotic gene set (Table 1) Prior to the phylogenomic dataset construction, individual transcriptomes were clustered at 99% using CDHit (Li & Godzik, 2006).

Phylogenomic dataset construction

The previously published dataset from (Leonard et al., 2018) was used as a starting point for this study because of its broad representation of photosynthetic and heterotrophic stramenopiles. This dataset was generated using OrthoMCL (Li et al., 2003) to identify candidate genes from a broad eukaryotic taxon sampling. All orthogroups were examined and groups that had at least 80% retained data were kept. These orthogroups were examined further for monophyly of eukaryotic groups and removed genes that contained too many missing lineages, were too paralogous, or were too unreliable for phylogenomics. A subset of those genes were used in the curation of the Leonard et al. dataset based on the presence of stramenopiles.

In addition to the 22 transcriptomes generated for this study, 20 publicly available photosynthetic stramenopile transcriptomes were gathered to increase taxonomic sampling of ochrophytes (Table 1).

Representative sequences from *Vaucheria litorea*, *Chrysophyceae* sp., *Phaeomonas parva*, and *Thalassiosira pseudonana* were selected to be used as BLAST queries for each gene in the dataset. All potential ortholog candidates were identified using a BLASTP search (evalue 1e-10, BLAST v2.7.1). Sequences represented by low coverage (<5) were removed from the blast results and then the top 5 remaining BLAST hits for each taxon being added to the dataset were selected as potential orthologs and added to the dataset for subsequent single gene tree construction and manual curation of orthologs.

Single gene datasets were aligned using MAFFT (--globalpair --maxiterate 1000) (Katoh & Standley, 2013), alignment errors were filtered using DIVVIER (--partial -mincol 4 -divvygap) (Ali et al., 2019) and trimmed with trimAL (Capella-Gutiérrez et al., 2009) using a gap threshold of 0.2. Trees were built using RAxML (Stamatakis, 2014) under the LG model with 100 rapid bootstrap replicates. For each single gene dataset, a total of three rounds of manual inspection were performed to detect and remove contaminants and paralogs. The first round was to remove deep paralogs, a second for the removal of midparalogs and a third to address inparalogs. After removal of paralogs and contaminants during each round of manual curation, the datasets were realigned, trimmed and trees were reconstructed. During the second and third rounds of manual curation, filtered using Divvier (-partial -mincol 4) prior to trimming in trimAL.

The final ortholog-only dataset for each of the 242 genes was masked using Prequal under the default settings (Whelan et al., 2018), aligned using MAFFT (--globalpair --maxiterate 1000), and filtered using Divvier (--partial -mincol 4 -divvygap). Single gene alignments were concatenated, and genes were partitioned using Partition Finder 2 (Lanfear et al., 2017) under the LG+G model using the corrected Aikaike Information Criterion (AICc). A Maximum likelihood (ML) tree was inferred using IQ Tree v 1.6.7 (Nguyen et al., 2015) under the LG+C60+F+G4 mixture model. The resulting ML tree was used as a guide tree under the same model to estimate the PMSF profiles (Wang et al., 2018). Support values were then estimated using 500 nonparametric bootstrap

replicates under the LG+C60+F+G4+PMSF model. Bayesian analyses were inferred using Phylobayes-MPI v. 1.5 (Lartillot et al., 2013) with the CAT+GTR+G4 model and constant sites removed (-dc). Four independent Markov Chain Monte Carlo (MCMC) chains were run for >10,000 generations (all sampled).

For the fast site removal (FSR) analysis, per site evolutionary rates were estimated and sites in the alignment were sorted from fastest to slowest evolving. This was followed by sequential removal of 5,000 fastest sites at a time generating increasingly smaller alternative datasets at each step. Similarly, for the Heterotacheous sites removal (HSR), the most heterotacheous sites were removed in a stepwise fashion 5000 sites at a time, producing iteratively smaller datasets until no further sites could be removed. Random subsampling of 20%, 40%, 60%, or 80% of genes that comprise the complete phylogenomic dataset was conducted under the default 95% confidence interval setting as in Brown et al. 2018 (Brown et al., 2018). The generation of these datasets were performed using scripts in the Phylofisher software package (https://github.com/TheBrownLab/PhyloFisher). For each of the FSR, HSR, and randomly resampled gene analyses, each step or subsample was analyzed using 100 fast bootstrap replicates in IQ-Tree under the LG+C60+F+G+PMSF model.

Both datasets used in the phylogenomic analyses were subjected to constrained ML tree reconstruction in IQTree (LG+C60+F+G) using the same partitioning scheme utilized in our initial analyses. Trees were constrained to

reflect the following relationships: Eustigmatophyceae + SI, Eustigmatophyceae + SII, Eustigmatophyceae outside, Pinguiophyceae + SIII, and Pinguiophyceae + *Olisthodiscus* + SIII. For each of the constrained topologies, ML trees were reconstructed in IQTree under the LG+C60+F+G model of evolution. Gene concordance factors (gCF) for the best ML tree for our initial phylogenomic analysis and the constrained topologies were calculated in IQTree. Additionally, the approximately unbiased (AU) test (Minh et al., 2020) was conducted on both datasets, the constrained trees, and 100 distinct local topologies saved during the initial ML analysis (-wt option in IQTree).

Maximum likelihood analyses were conducted under LG+C60+F+G-PMSF in IQTree.

To further investigate the role of systematic errors contributing to the observed topologies, we also ran phylogenomic analyses using the less complex LG model to strengthen a possible artifactual result compared to the partitioned LG+C60+F+G used in our previous analyses.

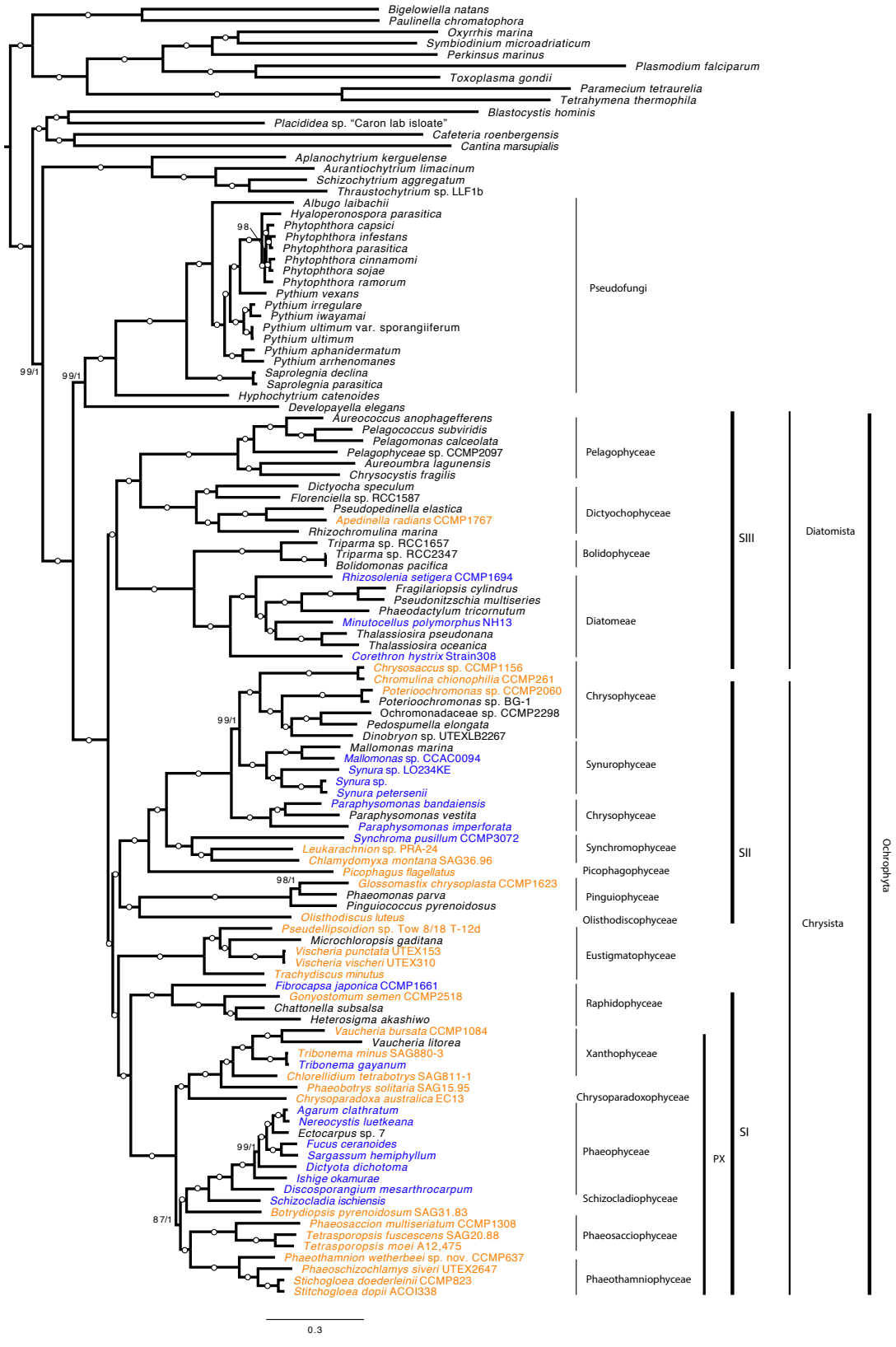
Bigelowiella natans
Paulinella chromatophora
Oxyrrhis marina
Symbiodinium microadriaticum
Perkinsus marinus
Plasmodium falciparum
Toxoplasma gondii
Paramecium tetraurelia
Tetrahymena thermophila
Blastocystis hominis
Placididea sp. "Caron lab isloate"
Cafeteria roenbergensis
Cantina marsupialis
Aplanochytrium kerguelense
Aurantiochytrium limacinum
Schizochytrium aggregatum
Thraustochytrium sp. LLF1b
Albugo laibachii
Hyaloperonospora parasitica
Phytophthora capsici
Phytophthora infestans
Phytophthora parasitica
Phytophthora cinnamomi
Phytophthora sojae
Phytophthora ramorum
Pythium vexans
Pythium irregulare
Pythium iwayamai
Pythium ultimum var. sporangiiferum
Pythium ultimum
Pythium aphanidermatum
Pythium arrhenomanes
Saprolegnia declina
Saprolegnia parasitica
Hyphochytrium catenoides
Developayella elegans
Aureococcus anophagefferens
Pelagococcus subviridis
Pelagomonas calceolata
Pelagophyceae sp. CCMP2097
Aureoumbra lagunensis
Chrysocystis fragilis
Dictyocha speculum
Florenciella sp. RCC1587
Pseudopedinella elastica
Apedinella radians CCMP1767
Rhizochromulina marina
Triparma sp. RCC1657
Triparma sp. RCC2347
Bolidomonas pacifica
Rhizosolenia setigera CCMP1694
Fragilariopsis cylindrus
Pseudonitzschia multiseries
Phaeodactylum tricornutum
Minutocellus polymorphus NH13
Thalassiosira pseudonana
Thalassiosira oceanica
Corethron hystrix Strain308
Chrysosaccus sp. CCMP1156
Chromulina chionophilia CCMP261
Poterioochromonas sp. CCMP2060
Poterioochromonas sp. BG-1
Ochromonadaceae sp. CCMP2298
Pedospumella elongata
Dinobryon sp. UTEXLB2267
Mallomonas marina
Mallomonas sp. CCAC0094
Synura sp. LO234KE
Synura sp.
Synura petersenii
Paraphysomonas bandaiensis
Paraphysomonas vestita
Paraphysomonas imperforata
Synchroma pusillum CCMP3072
Leukarachnion sp. PRA-24
Chlamydomyxa montana SAG36.96
Picophagus flagellatus
Glossomastix chrysoplasta CCMP1623
Phaeomonas parva
Pinguiococcus pyrenoidosus
Olisthodiscus luteus
Pseudellipsoidion sp. Tow 8/18 T-12d
Microchloropsis gaditana
Vischeria punctata UTEX153
Vischeria vischeri UTEX310
Trachydiscus minutus
Fibrocapsa japonica CCMP1661
Gonyostomum semen CCMP2518
Chattonella subsalsa
Heterosigma akashiwo
Vaucheria bursata CCMP1084
Vaucheria litorea
Tribonema minus SAG880-3
Tribonema gayanum
Chlorellidium tetrabotrys SAG811-1
Phaeobotrys solitaria SAG15.95
Chrysoparadoxa australica EC13
Agarum clathratum
Nereocystis luetkeana
Ectocarpus sp. 7
Fucus ceranoides
Sargassum hemiphyllum
Dictyota dichotoma
Ishige okamurae
Discosporangium mesarthrocarpum
Schizocladia ischiensis
Botrydiopsis pyrenoidosum SAG31.83
Phaeosaccion multiseriatum CCMP1308
Tetrasporopsis fuscescens SAG20.88
Tetrasporopsis moei A12,475
Phaeothamnion wetherbeei sp. nov. CCMP637
Phaeoschizochlamys siveri UTEX2647
Stichogloea doederleinii CCMP823
Stitchogloea dopii ACOI338
98
99/1
99/1
99/1
98/1
99/1
87/1
Pseudofungi
Pelagophyceae
Dictyochophyceae
Bolidophyceae
Diatomeae
Chrysophyceae
Synurophyceae
Chrysophyceae
Synchromophyceae
Picophagophyceae
Pinguiophyceae
Olisthodiscophyceae
Eustigmatophyceae
Raphidophyceae
Xanthophyceae
Chrysoparadoxophyceae
Phaeophyceae
Schizocladiophyceae
Phaeosacciophyceae
Phaeothamniophyceae
SIII
SII
SI
PX
Diatomista
Chrysista
Ochrophyta
0.3

Figure 1. Phylogeny of SAR. Maximum likelihood and Bayesian phylogeny of ochrophytes inferred using IQTree (LG+C60+F+Γ4) and PhyloBayes (CAT-GTR+Γ4) based on our initial dataset including heterotrophic stramenopiles, Rhizaria, and alveolates as outgroup taxa. This dataset was comprised of 242 genes and 75,752 sites. Statistical support is derived from non-parametric PMSF based bootstrap support (n=500) (BS) and Bayesian posterior probabilities (PP). Ochrophyte transcriptomes newly sequenced as part of this study are colored orange and transcriptomes added to increase taxonomic sampling are colored blue. Branch supports correspond to ML BS (right) and Bayesian posterior probabilities (left). Branches fully supported in each analysis are shown with hollow circles (100 BS/1.0 PP).

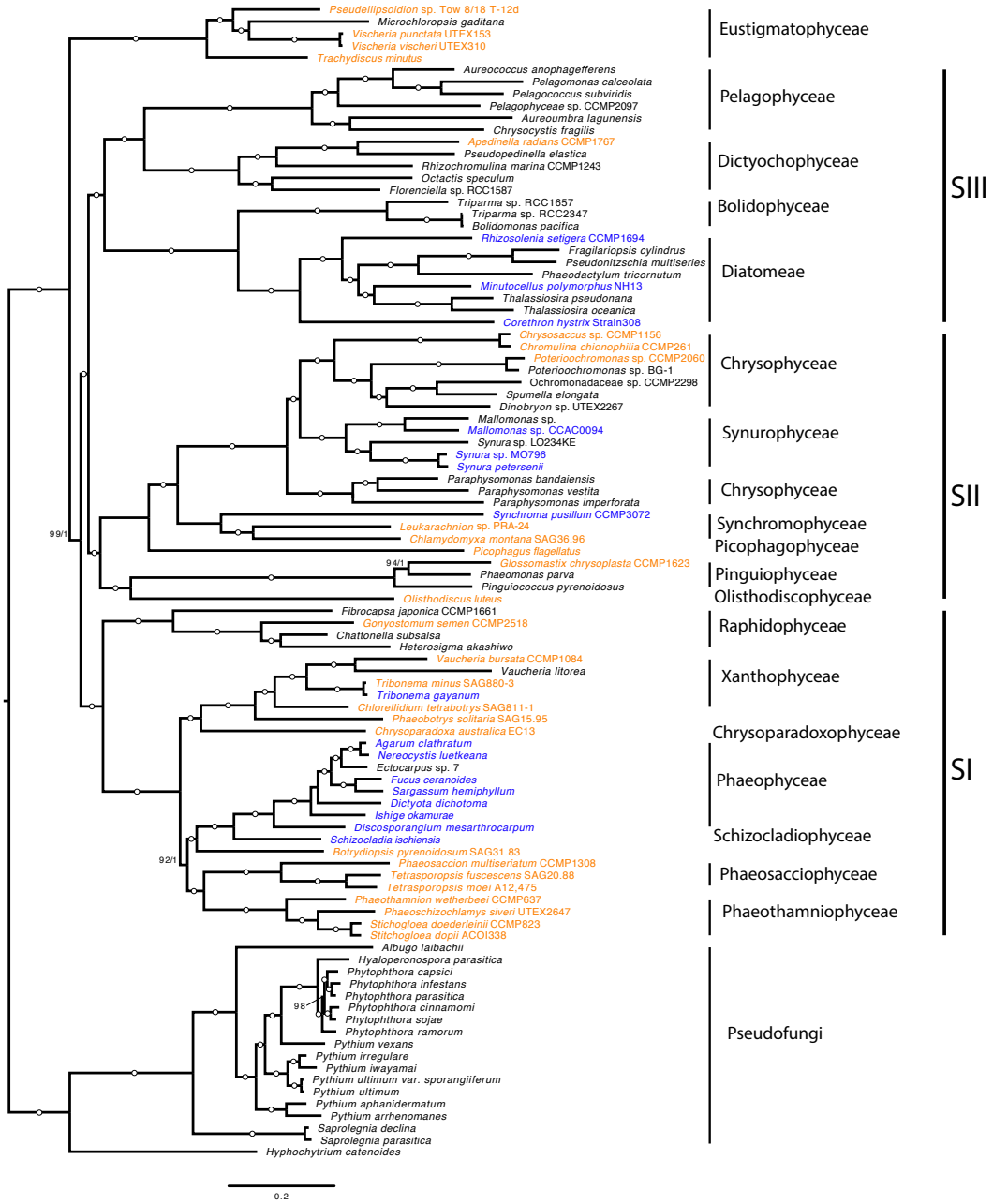

Figure 2. Phylogeny of Ochrophytes + Pseudofungi Maximum likelihood phylogeny of Ochrophytes + Pseduofungi inferred using IQTree (LG+C60+F+Γ4) and PhyloBayes (CAT-GTR+Γ4). This dataset was comprised of 242 genes and 76,189 sites. Statistical support is derived from non-parametric PMSF based bootstrap support (n=500) (BS) and Bayesian posterior probabilities (PP). Ochrophyte transcriptomes newly sequenced as part of this study are colored orange and transcriptomes added to increase taxonomic sampling are colored blue. Branch supports correspond to ML BS (right) and Bayesian posterior probabilities (left). Branches fully supported in each analysis are shown with hollow circles (100 BS/1.0 PP).

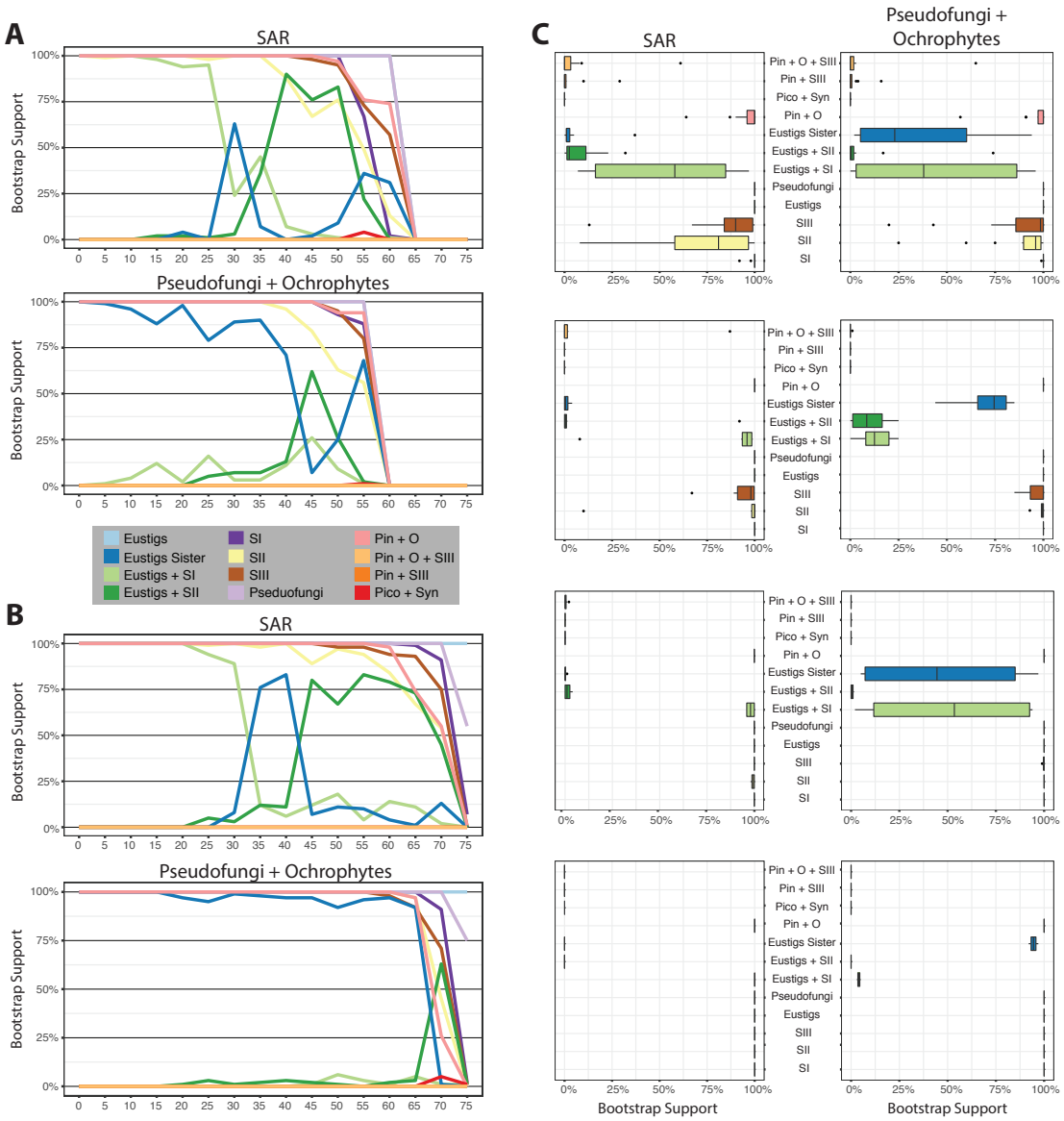

Figure 3. Effects of removing fast evolving and heterotacheous sites, and random gene subsampling on apicomplexan relationships. (A) Graphs plotting support for bipartitions of interest for the SAR and Pseudofungi + Ochrophytes datasets after the stepwise removal of the 5,000 fastest evolving sites until all sites are removed from each dataset. Non-parametric bootstrap support (n=100) values are on the y-axis and number of sites removed, measured in thousands, is shown on the x-axis. (B) Graphs plotting support for bipartitions of interest for the SAR and Pseudofungi + Ochrophytes datasets after the stepwise removal of the 5,000 most heterotacheous sites until all sites are removed from each dataset. Non-parametric bootstrap support (n=100) values are on the y-axis and number of sites removed, measured in thousands, is shown on the x-axis. (C) Box-and-whisker plots showing support for randomly sampled subsets of genes from

each total dataset (SAR dataset on the left and ochrophytes + pseudofungi on the right). Non-parametric bootstrap support (n=100) values are on the y-axis and subsample percentage is shown on the x-axis. The number of individual datasets required to sample every gene with 95% probability of sampling every gene for each percentage of genes being sampled is shown in parentheses. Abbreviated groups of interest are: SI = Raphidophyceae + Xanthophyceae + Chrysoparadoxophyceae + Phaeophyceae + Schizocladophyceae + Phaeosacciophyceae + Phaeothamniophyceae; SII = Olisthodiscophyceae + Pinguiophyceae + Picophagophyceae + Synchromophyceae + Chrysophyceae + Synurophyceae + Other Chrysophyceae; SIII = Pelagophyceae + Dictyochophyceae + Bolidophyceae + Diatomae; Pseudofungi = Pseudofungi; Eustigs = Eustigmatophyceae; Eustigs Sister = Eustigmatophyceae are earliest diverging ochrophytes; Eustigs + SI = Eustigmatophyceae + SI taxa; Eustigs + SII = Eustigmatophyceae + SII taxa; Pin + O = Pinguiophyceae + Olisthodiscophyceae; Pin + O + SIII = Pinguiophyceae + Olisthodiscophyceae + SIII taxa; Pin + SIII = Pinguiophyceae+ SIII taxa; Pico + Syn = Picophagophyceae + Synchromophyceae.

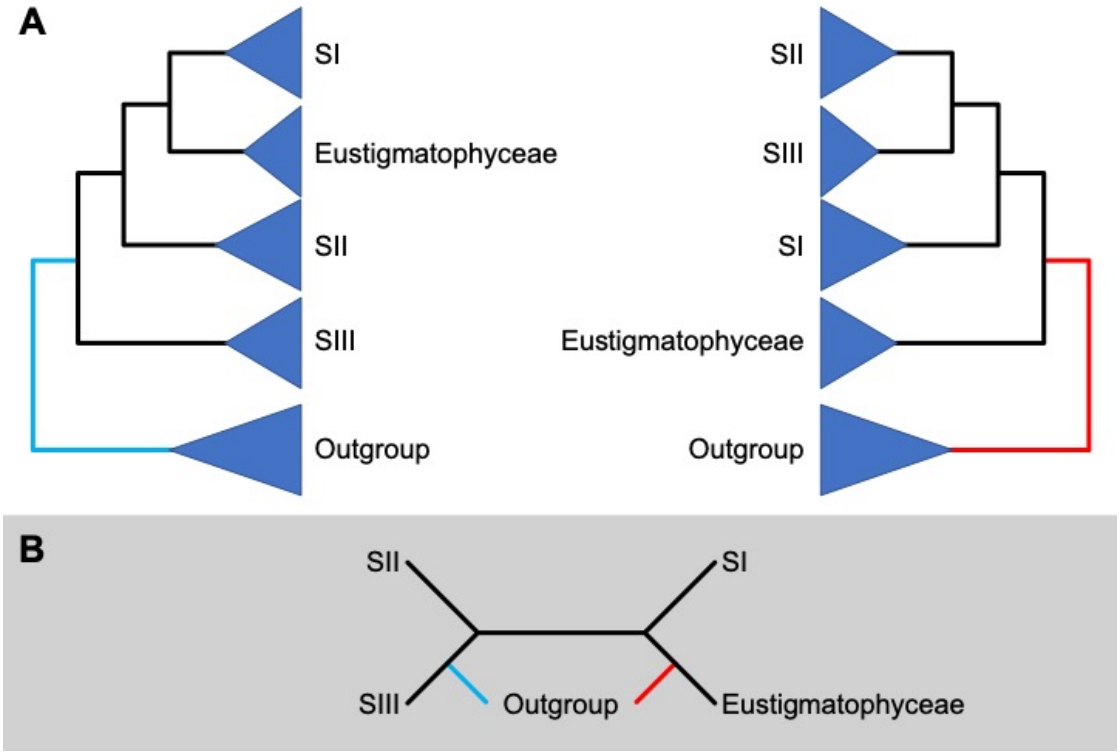

Figure 4. Possible placement for the root of ochrophytes. A) Two topologies recovered in ML analyses utilizing different outgroup taxa in this study. The SAR dataset is shown on the left with a blue line that places the root closest to the SIII clade. The ochrophytes + pseudofungi dataset is shown on the right with a red line placing the root closest to the Eustigmatophyceae. B) A star tree showing that the relationships between the ochrophytes lineages are the same in both datasets, with the placement of the root differing between the two.

Class	Taxa	Data Source	BUSCO Genes (%)		
			Complete	Fragment	Missing
Apicomplexa	*Plasmodium falciparum*	GCF_000002765.4	56.5	3.5	40
Apicomplexa	*Toxoplasma gondii*	GCF_000006565.2	61.2	8.6	30.2
Bolidophyceae	*Triparma* sp. RCC1657	http://datacommons.cyverse.org/browse/iplant/home/shared/imicrobe/projects/104/samples/2106/MMETSP1321_clean.pep.fa	40.8	18	41.2
Bolidophyceae	*Triparma* sp. RCC2347	https://datacommons.cyverse.org/browse/iplant/home/shared/imicrobe/projects/104/samples/2103/MMETSP1320_clean.pep.fa	43.9	15.7	40.4
Bolidophyceae	*Bolidomonas pacifica*	https://de.cyverse.org/anon-files//iplant/home/shared/imicrobe/projects/104/samples/2102/MMETSP1319_clean.pep.fa	48.6	19.2	32.2
Cercozoa	*Bigelowiella natans*	GCF_000002455.1	73	11.4	15.6
Cercozoa	*Paulinella chromatophora*	SRR3221671	80	10.6	9.4
Chrysoparadoxophyceae	*Chrysoparadoxa australica* EC13	This study	61.6	24.7	13.7
Chrysophyceae	*Chrysosaccus* sp. CCMP1156	This study	75.3	11.4	13.3
Chrysophyceae	*Poterioochromonas* sp. BG-1	https://de.cyverse.org/anon-files//iplant/home/shared/imicrobe/pr	49.5	11.4	39.1

		ojects/104/samples/2176/MMETSP1105_clean.pep.fa			
Chrysophyceae	*Poterioochromonas* sp. CCMP2060	This study	48.2	21.6	30.2
Chrysophyceae	*Chromulina chionophilia* CCMP261	This study	70.2	13.3	16.5
Chrysophyceae	Ochromonadaceae sp. CCMP2298	SRR485951	67.9	12.2	19.9
Chrysophyceae	*Dinobryon sp. UTEX LB 2267*	https://de.cyverse.org/anon-files//iplant/home/shared/imicrobe/projects/104/samples/1691/MMETSP0019_clean.pep.fa & https://de.cyverse.org/anon-files//iplant/home/shared/imicrobe/projects/104/samples/1692/MMETSP0020_clean.pep.fa	61.6	8.6	29.8
Chrysophyceae	*Paraphysomonas bandaiensis*	MMETSP1103	81.1	6.7	12.2
Chrysophyceae	*Paraphysomonas imperforata*	MMETSP 0103 & 0104	60.4	9	30.6
Chrysophyceae	*Paraphysomonas vestita*	https://de.cyverse.org/anon-files//iplant/home/shared/imicrobe/projects/104/samples/2178/MMETSP1107_clean.pep.fa	15.3	9	75.7
Ciliophora	*Paramecium tetraurelia*	GCF_000165425.1	71.4	3.9	24.7

Ciliophora	*Tetrahymena thermophila*	GCF_000189635.1	76.4	3.9	19.7
Developea	*Developayella elegans*	DRR049556	85.8	8.2	6
Diatomaea	*Phaeodactylum tricornutum*	GCF_000150955.2	78.9	5.9	15.2
Diatomeae	*Fragilariopsis cylindrus*	GCA_001750085.1	73.4	10.2	16.4
Diatomeae	*Corethron hystrix* Strain 308	http://datacommons.cyverse.org/browse/iplant/home/shared/imicrobe/projects/104/samples/1674/MMETSP0010_clean.pep.fa	49	20.4	30.6
Diatomeae	*Rhizosolenia setigera* CCMP1694	http://datacommons.cyverse.org/browse/iplant/home/shared/imicrobe/projects/104/samples/2039/MMETSP0789_clean.pep.fa	76.5	8.2	15.3
Diatomeae	*Minutocellus polymorphus* NH13	http://datacommons.cyverse.org/browse/iplant/home/shared/imicrobe/projects/104/samples/1900/MMETSP1070_clean.pep.fa	75.3	10.6	14.1
Diatomeae	*Pseudo-nitzschia multiseries*	https://genome.jgi.doe.gov/portal/Psemu1/download/Psemu1_GeneCatalog_proteins_20111011.aa.fasta.gz	80.4	6.3	13.3
Diatomeae	*Thalassiosira oceanica*	https://genome.jgi.doe.gov/portal/Thaoce1/download/Thaoce1_GeneCatalog_proteins_20171026.aa.fasta.gz	53.7	13.7	32.6
Diatomeae	*Thalassiosira pseudonana*	GCF_000149405.2	71.4	10.6	18

Dictyochophyceae	*Apedinella radians* CCMP1767	This study	33.3	32.9	33.8
Dictyochophyceae	*Octactis speculum*	https://de.cyverse.org/anon-files//iplant/home/shared/imicrobe/projects/104/samples/2496/MMETSP1174_clean.pep.fa	62.4	13.7	23.9
Dictyochophyceae	*Florenciella sp. RCC1587*	https://de.cyverse.org/anon-files//iplant/home/shared/imicrobe/projects/104/samples/2114/MMETSP1324_clean.pep.fa	36.9	17.6	45.5
Dictyochophyceae	*Pseudopedinella elastica*	https://de.cyverse.org/anon-files//iplant/home/shared/imicrobe/projects/104/samples/1899/MMETSP1068_clean.pep.fa	62	17.6	20.4
Dictyochophyceae	*Rhizochromulina sp. CCMP1243*	https://de.cyverse.org/anon-files//iplant/home/shared/imicrobe/projects/104/samples/2120/MMETSP1173_clean.pep.fa	68.3	9.4	22.3
Dinoflagellata	*Oxyrrhis marina*	SRR1296901	72.9	11	16.1
Dinoflagellata	*Symbiodinium microadriaticum*	GCA_001939145.1	46.3	15.3	38.4
Eustigmatophyceae	*Pseudellipsoidion* sp. Tow 8/18 T-12d	This study	83.6	7.1	9.3
Eustigmatophyceae	*Microchloropsis gaditana*	GCF_000240725.1	14.9	11.8	73.3
Eustigmatophyceae	*Trachydiscus minutus*	This study	76.1	14.5	9.4

Eustigmatophyceae	*Vischeria punctata* UTEX153	This study	73.3	14.5	12.2
Eustigmatophyceae	*Vischeria vischeri* UTEX310	This study	70.5	19.6	9.9
Hyphochytriales	*Hyphochytrium catenoides*	https://github.com/guyleonard/hyphochytrium/blob/master/gene_predictions/proteins/2016/second_pass.all.maker.proteins.fasta	4.3	3.1	92.6
Incertae sedis	*Phaeobotrys solitaria* SAG15.95	This study	85.1	7.1	7.8
Incertae sedis	*Botrydiopsis pyrenoidosum* SAG31.83	This study	77.6	12.9	9.5
Olisthodiscophyceae	*Olisthodiscus luteus*	This study	57.2	23.1	19.7
Oomycota	*Albugo laibachii*	ftp://ftp.ensemblgenomes.org/pub/release-41/protists/fasta/albugo_laibachii/pep/	88.2	3.5	8.3
Oomycota	*Hyaloperonospora parasitica*	https://genome.jgi.doe.gov/portal/Hyaar1/download/Hyaar1_GeneCatalog_proteins_20210111_promoters_1k.fa.gz	80.8	9	10.2
Oomycota	*Phytophthora capsici*	https://genome.jgi.doe.gov/portal/Phyca11/download/Phyca11_filtered_proteins.fasta.gz	92.2	3.1	4.7

Oomycota	*Phytophthora cinnamomi*	https://genome.jgi.doe.gov/portal/Phyci1/download/Phyci1_GeneCatalog_proteins_20120612.aa.fasta.gz	92.2	2.7	5.1
Oomycota	*Phytophthora infestans*	http://ftp.ebi.ac.uk/ensemblgenomes/pub/release-51/protists/fasta/phytophthora_infestans/pep/Phytophthora_infestans.ASM14294v1.pep.all.fa.gz	90.2	4.7	5.1
Oomycota	*Phytophthora parasitica*	GCF_000247585.1	94.5	2	3.5
Oomycota	*Phytophthora ramorum*	http://ftp.ebi.ac.uk/ensemblgenomes/pub/release-51/protists/fasta/phytophthora_ramorum/pep/Phytophthora_ramorum.ASM14973v1.pep.all.fa.gz	89.8	5.5	4.7
Oomycota	*Phytophthora sojae*	https://genome.jgi.doe.gov/portal/Physo3/download/Physo3_GeneCatalog_proteins_20110401.aa.fasta.gz	86.6	3.9	9.5
Oomycota	*Pythium ultimum var sporangiiferum*	https://fungidb.org/fungidb/app/record/organism/NCBITAXON_1223559	61.6	20.4	18
Ooomycota	*Pythium aphanidermatum*	ftp://ftp.ensemblgenomes.org/pub/protists/release-51/fasta/pythium_aphanidermatum	87.9	6.7	5.4
Ooomycota	*Pythium arrhenomanes*	ftp://ftp.ensemblgenomes.org/pub/protists/release-51/fasta/pythium_arrhenomanes	81.2	9.8	9

Ooomycota	*Pythium irregulare*	ftp://ftp.ensemblgenomes.org/pub/release-41/protists/fasta/pythium_irregulare/pep/	93	3.5	3.5
Ooomycota	*Pythium iwayamai*	ftp://ftp.ensemblgenomes.org/pub/protists/release-51/fasta/pythium_iwayamai	86.7	6.3	7
Ooomycota	*Pythium ultimum*	ftp://ftp.ensemblgenomes.org/pub/protists/release-51/fasta/pythium_ultimum	92.1	2.7	5.2
Ooomycota	*Pyuthium vexans*	ftp://ftp.ensemblgenomes.org/pub/protists/release-51/fasta/pythium_vexans	89.8	4.3	5.9
Ooomycota	*Saprolegnia diclina*	GCF_000281045.1	93.7	1.2	5.1
Ooomycota	*Saprolegnia parasitica*	https://genome.jgi.doe.gov/portal/Sappar1/download/Sappar1_GeneCatalog_proteins_20180131.aa.fasta.gz	87.8	3.5	8.7
Opalozoa	*Placididea* sp. "Caron lab isolate"	https://de.cyverse.org/anon-files//iplant/home/shared/imicrobe/projects/104/samples/2175/MMETSP1104_clean.pep.fa	64.3	11.8	23.9
Opalozoa	*Cafeteria roenbergensis*	https://de.cyverse.org/anon-files//iplant/home/shared/imicrobe/projects/104/samples/2026/MMETSP0942_clean.pep.fa	46.3	10.6	43.1
Opalozoa	*Cantina marsupialis*	DRR030401	76.5	7.1	16.4

Oplaozoa	*Blastocystis* sp. subtype 7 strain B	GCF_000151665.1	42.7	18.8	38.5
Pelagophyceae	*Aureococcus anophagefferens*	https://genome.jgi.doe.gov/portal/Auran1/download/proteins.Auran1_FilteredModels3.fasta.gz	52.9	11.4	35.7
Pelagophyceae	*Aureoumbra lagunensis*	https://de.cyverse.org/anon-files//iplant/home/shared/imicrobe/projects/104/samples/1975/MMETSP0890_clean.pep.fa & https://de.cyverse.org/anon-files//iplant/home/shared/imicrobe/projects/104/samples/1976/MMETSP0891_clean.pep.fa & https://de.cyverse.org/anon-files//iplant/home/shared/imicrobe/projects/104/samples/1980/MMETSP0892_clean.pep.fa & https://de.cyverse.org/anon-files//iplant/home/shared/imicrobe/projects/104/samples/1981/MMETSP0893_clean.pep.fa	54.1	10.2	35.7
Pelagophyceae	*Chrysocystis fragilis*	https://de.cyverse.org/anon-files//iplant/home/shared/imicrobe/projects/104/samples/2100/MMETSP1165_clean.pep.fa	40	14.5	45.5
Pelagophyceae	Pelagophyceae sp. CCMP2097	SRR485867	58.8	5.9	35.3

Pelagophyceae	*Pelagomonas calceolata*	https://de.cyverse.org/anon-files//iplant/home/shared/imicrobe/projects/104/samples/2093/MMETSP0886_clean.pep.fa & https://de.cyverse.org/anon-files//iplant/home/shared/imicrobe/projects/104/samples/2094/MMETSP0887_clean.pep.fa & https://de.cyverse.org/anon-files//iplant/home/shared/imicrobe/projects/104/samples/1973/MMETSP0888_clean.pep.fa & https://de.cyverse.org/anon-files//iplant/home/shared/imicrobe/projects/104/samples/1974/MMETSP0889_clean.pep.fa	61.2	11.8	27
Pelagophyceae	*Pelagococcus subviridis*	https://de.cyverse.org/anon-files//iplant/home/shared/imicrobe/projects/104/samples/2077/MMETSP0882_clean.pep.fa & https://de.cyverse.org/anon-files//iplant/home/shared/imicrobe/projects/104/samples/2078/MMETSP0883_clean.pep.fa & https://de.cyverse.org/anon-files//iplant/home/shared/imicrobe/projects/104/samples/2086/MMETSP0884_clean.pep.fa & https://de.cyverse.org/anon-	52.2	19.6	28.2

		files//iplant/home/shared/imicrobe/projects/104/samples/2087/MMETSP0885_clean.pep.fa			
Perkinsozoa	*Perkinsus marinus*	GCF_000006405.1	57.2	13.7	29.1
Phaeophyceae	*Agarum clathratum*	SRR3711333	70.6	18.4	11
Phaeophyceae	*Dictyota dichotoma*	SRR5088950	36.4	20	43.6
Phaeophyceae	*Discosporangium mesarthrocarpum*	http://www.research.kobe-u.ac.jp/rcis-ku-macc/e.p.folder/e.i.folder/download.html	16.9	31.8	51.3
Phaeophyceae	*Ectocarpus* sp. 7	GCA_000310025.1	84.7	7.8	7.5
Phaeophyceae	*Fucus ceranoides*	ERR1161612	60.4	5.1	34.5
Phaeophyceae	*Ishige okamurae*	https://ftp.cngb.org/pub/gigadb/pub/10.5524/100001_101000/100627/assemblies/APTP-Ishige_okamurae/APTP-SOAPdenovo-Trans-assembly.fa.gz	69.5	18.8	11.7
Phaeophyceae	*Nereocystis luetkeana*	SRR3711301	72.1	16.1	11.8
Phaeophyceae	*Sargassum hemiphyllum*	ERR2041177	58	20.4	21.6
Phaeosacciophyceae	*Tetrasporopsis moei* A12,475	This study	84.3	7.1	8.6
Phaeosacciophyceae	*Phaeosaccion multiseriatum* CCMP1308	This study	70.2	14.5	15.3

Phaeosacciophyceae	*Tetrasporopsis fuscescens* SAG20.88	This study	76.9	10.6	12.5
Phaeothamniophyceae	*Stitchogloea dopii* ACOI338	This study	73.7	11.8	14.5
Phaeothamniophyceae	*Phaeothamnion wetherbeei* sp. nov. CCMP637	This study	66.3	19.2	14.5
Phaeothamniophyceae	*Stichogloea doederleinii* CCMP823	This study	78	14.9	7.1
Phaeothamniophyceae	*Phaeoschizochlamys siveri* UTEX2647	This study	86.2	8.2	5.6
Picophagophyceae	*Picophagus flagellatus*	This study	75.7	10.2	14.1
Pinguiophyceae	*Pinguiococcus pyrenoidosus*	https://de.cyverse.org/anon-files//iplant/home/shared/imicrobe/projects/104/samples/2158/MMETSP1160_clean.pep.fa	51.8	4.7	43.5
Pinguiophyceae	*Phaeomonas parva*	http://datacommons.cyverse.org/browse/iplant/home/shared/imicrobe/projects/104/samples/2098/MMETSP1163_clean.pep.fa	49.4	18.8	31.8
Pinguiophyceae	*Glossomastix chrysoplasta* RCC625	This study	52.9	20.4	26.7
Raphiodphyceae	*Fibrocapsa japonica* CCMP1661	http://datacommons.cyverse.org/browse/iplant/home/shared/imicrobe/projects/104/samples/2520/MMETSP1339_clean.pep.fa	37.7	17.3	45

Raphiodphyceae	*Gonyostomum semen* CCMP2518	This study	88.3	3.5	8.2
Raphiodphyceae	*Chattonella subsalsa*	https://de.cyverse.org/anon-files//iplant/home/shared/imicrobe/projects/104/samples/1923/MMETSP0947_clean.pep.fa & https://de.cyverse.org/anon-files//iplant/home/shared/imicrobe/projects/104/samples/1924/MMETSP0948_clean.pep.fa & https://de.cyverse.org/anon-files//iplant/home/shared/imicrobe/projects/104/samples/1929/MMETSP0949_clean.pep.fa &https://de.cyverse.org/anon-files//iplant/home/shared/imicrobe/projects/104/samples/1930/MMETSP0950_clean.pep.fa	72.1	10.2	17.7
Raphiodphyceae	*Heterosigma akashiwo*	https://de.cyverse.org/anon-files//iplant/home/shared/imicrobe/projects/104/samples/1800/MMETSP0292_clean.pep.fa & https://de.cyverse.org/anon-files//iplant/home/shared/imicrobe/projects/104/samples/1801/MMETSP0294_clean.pep.fa & https://de.cyverse.org/anon-files//iplant/home/shared/imicrobe/projects/104/samples/1802/MMETSP	58.8	21.2	20

		0295_clean.pep.fa & https://de.cyverse.org/anon-files//iplant/home/shared/imicrobe/projects/104/samples/1803/MMETSP0296_clean.pep.fa			
Sagenista	*Aplanochytrium kerguelense*	https://genome.jgi.doe.gov/portal/Aplke1/download/Aplke1_GeneCatalog_proteins_20121220.aa.fasta.gz	91.4	3.1	5.5
Sagenista	*Aurantiochytrium limacinum*	https://genome.jgi.doe.gov/portal/Aurli1/download/Aurli1_GeneCatalog_proteins_20120618.aa.fasta.gz	85.9	6.3	7.8
Sagenista	*Schizochytrium aggregatum*	JGI Project: 402022, Schag1	80	8.2	11.8
Sagenista	*Thraustochytrium* sp. LLF1b	https://de.cyverse.org/anon-files//iplant/home/shared/imicrobe/projects/104/samples/1765/MMETSP0198_clean.pep.fa & https://de.cyverse.org/anon-files//iplant/home/shared/imicrobe/projects/104/samples/1766/MMETSP0199_clean.pep.fa	81.1	5.5	13.4
Schizocladiophyceae	*Schizocladia ischiensis*	http://www.research.kobe-u.ac.jp/rcis-ku-macc/e.p.folder/e.i.folder/download.html	6.3	26.3	67.4
Synchromophyceae	*Synchroma pusillum* CCMP3072	http://datacommons.cyverse.org/browse/iplant/home/shared/imicrobe/pr	55.3	16.9	27.8

		ojects/104/samples/2283/MMETSP1452_clean.pep.fa			
Synchromophyceae	*Leukarachnion* sp. PRA-24	This study	91.4	2.4	6.2
Synchromophyceae	*Chlamydomyxa montana* SAG36.96	This study	43.9	28.6	27.5
Synurophyceae	*Mallomonas marina*	https://de.cyverse.org/anon-files//iplant/home/shared/imicrobe/projects/104/samples/2105/MMETSP1167_clean.pep.fa	40.8	20.8	38.4
Synurophyceae	*Mallomonas* sp.	https://ftp.cngb.org/pub/gigadb/pub/10.5524/100001_101000/100627/assemblies/BOGT-Mallomonas_sp/BOGT-SOAPdenovo-Trans-assembly.fa.gz	63.6	18.4	18
Synurophyceae	*Synura petersenii*	https://ftp.cngb.org/pub/gigadb/pub/10.5524/100001_101000/100627/assemblies/DBYD-Synura_petersenii/DBYD-SOAPdenovo-Trans-assembly.fa.gz	40.4	28.2	31.4
Synurophyceae	*Synura* sp. LO234KE	HAGD00000000.1	15.7	16.5	67.8
Synurophyceae	*Synura* sp.	https://ftp.cngb.org/pub/gigadb/pub/10.5524/100001_101000/100627/assemblies/VKVG-Synura_sp/VKVG-SOAPdenovo-Trans-assembly.fa.gz	38.1	26.7	35.2
Synurophyceae?	*Pedospumella elongata*	https://de.cyverse.org/anon-files//iplant/home/shared/imicrobe/pr	70.6	7.1	22.3

		ojects/104/samples/2171/MMETSP1098_clean.pep.fa			
Xanthophyceae	*Vaucheria bursata* CCMP1084	This study	77.3	12.5	10.2
Xanthophyceae	*Chlorellidium tetrabotrys* SAG811-1	This study	81.2	9.8	9
Xanthophyceae	*Tribonema minus* SAG880-3	This study	69.8	13.3	16.9
Xanthophyceae	*Tribonema gayanum*	SRR7470091	55.7	13.3	31
Xanthophyceae	*Vaucheria litorea*	https://de.cyverse.org/anon-files//iplant/home/shared/imicrobe/projects/104/samples/1916/MMETSP0945_clean.pep.fa & https://de.cyverse.org/anon-files//iplant/home/shared/imicrobe/projects/104/samples/1917/MMETSP0946_clean.pep.fa	60	12.2	27.8

Table 1. Taxonomy, sources, and BUSCO completeness based on the Eukaryota ODB10 reference gene set for data used in this study.

BIBLIOGRAPHY

Adl, S. M., Bass, D., Lane, C. E., Lukeš, J., Schoch, C. L., Smirnov, A., Agatha, S., Berney, C., Brown, M. W., Burki, F., et al. (2019). Revisions to the classification, nomenclature, and diversity of eukaryotes. *Journal of Eukaryotic Microbiology*, *66*(1), 4–119. https://doi.org/10.1111/jeu.12691

Ali, R. H., Bogusz, M., & Whelan, S. (2019). Identifying clusters of high confidence homologies in multiple sequence alignments. *Molecular Biology and Evolution*, *36*(10), 2340–2351. https://doi.org/10.1093/molbev/msz142

Andersen, R. A. (2004). Biology and systematics of heterokont and haptophyte algae. *American Journal of Botany*, *91*(10), 1508–1522. https://doi.org/10.3732/ajb.91.10.1508

Andersen, R. A., Saunders, G. W., Paskind, M. P., & Sexton, J. P. (1993). Ultrastructure and 18S rRNA gene sequence for *Pelagomonas Calceolata* gen. et sp. nov. and the description of a new algal class, the Pelagophyceae classis nov. *Journal of Phycology*, *29*(5), 701–715. https://doi.org/10.1111/j.0022-3646.1993.00701.x

Bailey, J. C., Bidigare, R. R., Christensen, S. J., & Andersen, R. A. (1998). Phaeothamniophyceae classis vova: A new lineage of chromophytes based upon photosynthetic pigments, rbcL sequence analysis and

ultrastructure. *Protist*, *149*(3), 245–263. https://doi.org/10.1016/S1434-4610(98)70032-X

Barcyté, D., Eikrem, W., Engesmo, A., Seoane, S., Wohlmann, J., Horák, A., Yurchenko, T., & Eliáš, M. (2021). Olisthodiscus represents a new class of Ochrophyta. *Journal of Phycology*, *n/a*(n/a). https://doi.org/10.1111/jpy.13155

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics*, *30*(15), 2114–2120. https://doi.org/10.1093/bioinformatics/btu170

Brown, M. W., Heiss, A. A., Kamikawa, R., Inagaki, Y., Yabuki, A., Tice, A. K., Shiratori, T., Ishida, K.-I., Hashimoto, T., Simpson, A. G. B., & Roger, A. J. (2018). Phylogenomics places orphan protistan lineages in a novel eukaryotic super-group. *Genome Biology and Evolution*, *10*(2), 427–433. https://doi.org/10.1093/gbe/evy014

Bushmanova, E., Antipov, D., Lapidus, A., & Przhibelskiy, A. D. (2018). rnaSPAdes: A *de novo* transcriptome assembler and its application to RNA-Seq data. *BioRxiv*, 420208. https://doi.org/10.1101/420208

Capella-Gutiérrez, S., Silla-Martínez, J. M., & Gabaldón, T. (2009). trimAl: A tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics*, *25*(15), 1972–1973. https://doi.org/10.1093/bioinformatics/btp348

Cavalier-Smith, T., Chao, E. E., & Allsopp, M. T. E. P. (1995). Ribosomal RNA evidence for chloroplast loss within Heterokonta: Pedinellid

relationships and a revised classification of ochristan algae. *Archiv Für Protistenkunde*, *145*(3–4), 209–220. https://doi.org/10.1016/S0003-9365(11)80316-7

Cavalier-Smith, Thomas, & Scoble, J. M. (2013). Phylogeny of Heterokonta: *Incisomonas marina*, a uniciliate gliding opalozoan related to *Solenicola* (Nanomonadea), and evidence that Actinophryida evolved from raphidophytes. *European Journal of Protistology*, *49*(3), 328–353. https://doi.org/10.1016/j.ejop.2012.09.002

Delwiche, C. (1999). Tracing the thread of plastid diversity through the tapestry of life. *The American Naturalist*, *154*(S4), S164–S177. https://doi.org/10.1086/303291

Derelle, R., López-García, P., Timpano, H., & Moreira, D. (2016). A phylogenomic framework to study the diversity and evolution of stramenopiles (=heterokonts). *Molecular Biology and Evolution*, *33*(11), 2890–2898. https://doi.org/10.1093/molbev/msw168

Di Franco, A., Baurain, D., Glöckner, G., Melkonian, M., & Philippe, H. (2021). Lower statistical support with larger datasets: insights from the Ochrophyta radiation. *BioRxiv*, 2021.01.14.426536. https://doi.org/10.1101/2021.01.14.426536

Dorrell, R. G., Azuma, T., Nomura, M., Audren de Kerdrel, G., Paoli, L., Yang, S., Bowler, C., Ishii, K., Miyashita, H., Gile, G. H., & Kamikawa, R. (2019). Principles of plastid reductive evolution illuminated by

nonphotosynthetic chrysophytes. *Proceedings of the National Academy of Sciences*, *116*(14), 6914–6923. https://doi.org/10.1073/pnas.1819976116

Dorrell, R. G., Gile, G., McCallum, G., Méheust, R., Bapteste, E. P., Klinger, C. M., Brillet-Guéguen, L., Freeman, K. D., Richter, D. J., & Bowler, C. (2017). Chimeric origins of ochrophytes and haptophytes revealed through an ancient plastid proteome. *ELife*, *6*, e23717. https://doi.org/10.7554/eLife.23717

Gentekaki, E., Kolisko, M., Boscaro, V., Bright, K. J., Dini, F., Di Giuseppe, G., Gong, Y., Miceli, C., Modeo, L., Molestina, R. E., Petroni, G., Pucciarelli, S., Roger, A. J., Strom, S. L., & Lynn, D. H. (2014). Large-scale phylogenomic analysis reveals the phylogenetic position of the problematic taxon Protocruzia and unravels the deep phylogenetic affinities of the ciliate lineages. *Molecular Phylogenetics and Evolution*, *78*, 36–42. https://doi.org/10.1016/j.ympev.2014.04.020

Giarla, T. C., & Esselstyn, J. A. (2015). The challenges of resolving a rapid, recent radiation: Empirical and simulated phylogenomics of Philippine shrews. *Systematic Biology*, *64*(5), 727–740. https://doi.org/10.1093/sysbio/syv029

Grant, J., Tekle, Y. I., Anderson, O. R., Patterson, D. J., & Katz, L. A. (2009). Multigene evidence for the placement of a heterotrophic amoeboid lineage *Leukarachnion* sp. among photosynthetic stramenopiles. *Protist*, *160*(3), 376–385. https://doi.org/10.1016/j.protis.2009.01.001

Grattepanche, J.-D., Walker, L. M., Ott, B. M., Pinto, D. L. P., Delwiche, C. F., Lane, C. E., & Katz, L. A. (2018). Microbial diversity in the eukaryotic SAR clade: Illuminating the darkness between morphology and molecular data. *BioEssays*, *40*(4), 1700198. https://doi.org/10.1002/bies.201700198

Hampl, V., Hug, L., Leigh, J. W., Dacks, J. B., Lang, B. F., Simpson, A. G. B., & Roger, A. J. (2009). Phylogenomic analyses support the monophyly of Excavata and resolve relationships among eukaryotic "supergroups." *Proceedings of the National Academy of Sciences of the United States of America*, *106*(10), 3859–3864. https://doi.org/10.1073/pnas.0807880106

Hara, Y., & Chihara, M. (1987). Morphology, ultrastructure and taxonomy of the raphidophycean alga *Heterosigma akashiwo*. *The Botanical Magazine = Shokubutsu-Gaku-Zasshi*, *100*(2), 151–163. https://doi.org/10.1007/BF02488320

Inouye, I., Hara, Y., & Chihara, M. (1992). Further observations on *Olisthodiscus luteus* (Raphidophyceae, Chromophyta): The flagellar apparatus ultrastructure. *Japanese Journal of Phycology*, *40*, 333–348.

Jackson, C., Salomaki, E. D., Lane, C. E., & Saunders, G. W. (2017). Kelp transcriptomes provide robust support for interfamilial relationships and revision of the little known Arthrothamnaceae (Laminariales). *Journal of Phycology*, *53*(1), 1–6. https://doi.org/10.1111/jpy.12465

Kai, A., Yoshii, Y., Nakayama, T., & Inouye, I. (2008). Aurearenophyceae classis nova, a new class of Heterokontophyta based on a new marine unicellular alga *Aurearena cruciata* gen. et sp. nov. inhabiting sandy beaches. *Protist*, *159*(3), 435–457. https://doi.org/10.1016/j.protis.2007.12.003

Kamikawa, R., Moog, D., Zauner, S., Tanifuji, G., Ishida, K.-I., Miyashita, H., Mayama, S., Hashimoto, T., Maier, U. G., Archibald, J. M., & Inagaki, Y. (2017). A non-photosynthetic diatom reveals early steps of reductive evolution in plastids. *Molecular Biology and Evolution*, *34*(9), 2355–2366. https://doi.org/10.1093/molbev/msx172

Katoh, K., & Standley, D. M. (2013). MAFFT Multiple sequence alignment software version 7: improvements in performance and usability. *Molecular Biology and Evolution*, *30*(4), 772–780. https://doi.org/10.1093/molbev/mst010

Kawachi, M., Inouye, I., Honda, D., O'Kelly, C. J., Bailey, J. C., Bidigare, R. R., & Andersen, R. A. (2002). The Pinguiophyceae classis nova, a new class of photosynthetic stramenopiles whose members produce large amounts of omega-3 fatty acids. *Phycological Research*, *50*(1), 31–47. https://doi.org/10.1046/j.1440-1835.2002.00260.x

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*(2), 211–228. https://doi.org/10.1078/143446103322166518

Keeling, P. J. (2013). The number, speed, and impact of plastid endosymbioses in eukaryotic evolution. *Annual Review of Plant Biology*, *64*(1), 583–607. https://doi.org/10.1146/annurev-arplant-050312-120144

Keeling, P. J., Burki, F., Wilcox, H. M., Allam, B., Allen, E. E., Amaral-Zettler, L. A., Armbrust, E. V., Archibald, J. M., Bharti, A. K., Bell, C. J., et al. (2014). The marine microbial eukaryote transcriptome sequencing project (MMETSP): Illuminating the functional diversity of eukaryotic life in the oceans through transcriptome sequencing. *PLOS Biology*, *12*(6), e1001889. https://doi.org/10.1371/journal.pbio.1001889

Lanfear, R., Frandsen, P. B., Wright, A. M., Senfeld, T., & Calcott, B. (2017). PartitionFinder 2: New methods for selecting partitioned models of evolution for molecular and morphological phylogenetic analyses. *Molecular Biology and Evolution*, *34*(3), 772–773. https://doi.org/10.1093/molbev/msw260

Lartillot, N., Rodrigue, N., Stubbs, D., & Richer, J. (2013). PhyloBayes MPI: Phylogenetic reconstruction with infinite mixtures of profiles in a parallel environment. *Systematic Biology*, *62*(4), 611–615. https://doi.org/10.1093/sysbio/syt022

Leonard, G., Labarre, A., Milner, D. S., Monier, A., Soanes, D., Wideman, J. G., Maguire, F., Stevens, S., Sain, D., Grau-Bové, X., et al. (2018). Comparative genomic analysis of the ‘pseudofungus’ *Hyphochytrium catenoides*. *Open Biology*, *8*(1), 170184. https://doi.org/10.1098/rsob.170184

Li, W., & Godzik, A. (2006). Cd-hit: A fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics (Oxford, England)*, *22*(13), 1658–1659. https://doi.org/10.1093/bioinformatics/btl158

Lin, Y.-C., Campbell, T., Chung, C.-C., Gong, G.-C., Chiang, K.-P., & Worden, A. Z. (2012). Distribution patterns and phylogeny of marine stramenopiles in the North Pacific Ocean. *Applied and Environmental Microbiology*, *78*(9), 3387–3399. https://doi.org/10.1128/AEM.06952-11

Logares, R., Audic, S., Bass, D., Bittner, L., Boutte, C., Christen, R., Claverie, J.-M., Decelle, J., Dolan, J. R., Dunthorn, Met al. (2014). Patterns of rare and abundant marine microbial eukaryotes. *Current Biology*, *24*(8), 813–821. https://doi.org/10.1016/j.cub.2014.02.050

Massana, R., Castresana, J., Balagué, V., Guillou, L., Romari, K., Groisillier, A., Valentin, K., & Pedrós-Alió, C. (2004). Phylogenetic and ecological analysis of novel marine stramenopiles. *Applied and Environmental Microbiology*, *70*(6), 3528–3534. https://doi.org/10.1128/AEM.70.6.3528-3534.2004

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. *Molecular Biology and Evolution*, *37*(5), 1530–1534. https://doi.org/10.1093/molbev/msaa015

Mitchell, N., Lewis, P. O., Lemmon, E. M., Lemmon, A. R., & Holsinger, K. E. (2017). Anchored phylogenomics improves the resolution of evolutionary relationships in the rapid radiation of Protea L. *American Journal of Botany*, *104*(1), 102–115. https://doi.org/10.3732/ajb.1600227

Nguyen, L.-T., Schmidt, H. A., von Haeseler, A., & Minh, B. Q. (2015). IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Molecular Biology and Evolution*, *32*(1), 268–274. https://doi.org/10.1093/molbev/msu300

Noguchi, F., Tanifuji, G., Brown, M. W., Fujikura, K., & Takishita, K. (2016). Complex evolution of two types of cardiolipin synthase in the eukaryotic lineage stramenopiles. *Molecular Phylogenetics and Evolution*, *101*, 133–141. https://doi.org/10.1016/j.ympev.2016.05.011

Onyshchenko, A., Ruck, E. C., Nakov, T., & Alverson, A. J. (2019). A single loss of photosynthesis in the diatom order Bacillariales (Bacillariophyta). *American Journal of Botany*, *106*(4), 560–572. https://doi.org/10.1002/ajb2.1267

Ott, D. W., Oldham-Ott, C. K., Rybalka, N., & Friedl, T. (2015). Chapter 11—Xanthophyte, Eustigmatophyte, and Raphidophyte Algae. In J. D. Wehr, R. G. Sheath, & J. P. Kociolek (Eds.), *Freshwater Algae of North America (Second Edition)* (pp. 485–536). Academic Press. https://doi.org/10.1016/B978-0-12-385876-4.00011-6

Pernice, M. C., Giner, C. R., Logares, R., Perera-Bel, J., Acinas, S. G., Duarte, C. M., Gasol, J. M., & Massana, R. (2016). Large variability of

bathypelagic microbial eukaryotic communities across the world's oceans. *The ISME Journal*, *10*(4), 945–958. https://doi.org/10.1038/ismej.2015.170

Rothfels, C. J., Pryer, K. M., & Li, F.-W. (2017). Next-generation polyploid phylogenetics: Rapid resolution of hybrid polyploid complexes using PacBio single-molecule sequencing. *New Phytologist*, *213*(1), 413–429. https://doi.org/10.1111/nph.14111

Ševčíková, T., Horák, A., Klimeš, V., Zbránková, V., Demir-Hilton, E., Sudek, S., Jenkins, J., Schmutz, J., Přibyl, P., Fousek, J., et al. (2015). Updating algal evolutionary relationships through plastid genome sequencing: Did alveolate plastids emerge through endosymbiosis of an ochrophyte? *Scientific Reports*, *5*, 10134. https://doi.org/10.1038/srep10134

Ševčíková, T., Klimeš, V., Zbránková, V., Strnad, H., Hroudová, M., Vlček, Č., & Eliáš, M. (2016). A comparative analysis of mitochondrial genomes in eustigmatophyte algae. *Genome Biology and Evolution*, *8*(3), 705–722. https://doi.org/10.1093/gbe/evw027

Sibbald, S. J., & Archibald, J. M. (2020). Genomic insights into plastid evolution. *Genome Biology and Evolution*, *12*(7), 978–990. https://doi.org/10.1093/gbe/evaa096

Stamatakis, A. (2014). RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, *30*(9), 1312–1313. https://doi.org/10.1093/bioinformatics/btu033

Stiller, J. W., Schreiber, J., Yue, J., Guo, H., Ding, Q., & Huang, J. (2014). The evolution of photosynthesis in chromist algae through serial endosymbioses. *Nature Communications*, *5*(1), 5764. https://doi.org/10.1038/ncomms6764

Strassert, J. F. H., Irisarri, I., Williams, T. A., & Burki, F. (2021). A molecular timescale for eukaryote evolution with implications for the origin of red algal-derived plastids. *Nature Communications*, *12*(1), 1879. https://doi.org/10.1038/s41467-021-22044-z

Wang, H.-C., Susko, E., & Roger, A. J. (n.d.). The relative importance of modeling site pattern heterogeneity versus partition-wise heterotachy in phylogenomic inference. *Systematic Biology*. https://doi.org/10.1093/sysbio/syz021

Wetherbee, R., Jackson, C. J., Repetti, S. I., Clementson, L. A., Costa, J. F., van de Meene, A., Crawford, S., & Verbruggen, H. (2018). The golden paradox – a new heterokont lineage with chloroplasts surrounded by two membranes. *Journal of Phycology*. https://doi.org/10.1111/jpy.12822

Whelan, S., Irisarri, I., & Burki, F. (2018). PREQUAL: Detecting non-homologous characters in sets of unaligned homologous sequences. *Bioinformatics*, *34*(22), 3929–3930. https://doi.org/10.1093/bioinformatics/bty448

Whitfield, J. B., & Lockhart, P. J. (2007). Deciphering ancient rapid radiations. *Trends in Ecology & Evolution*, *22*(5), 258–265. https://doi.org/10.1016/j.tree.2007.01.012

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*(2), 217–231. https://doi.org/10.1016/j.protis.2011.08.001