The evolution of red algal organellar genomes and the origin of red complex plastids

Fabian van Beveren

université PARIS-SACLAY

The evolution of red algal organellar genomes and the origin of red complex plastids

L'évolution des génomes organellaires des algues rouges et l'origine des plastides complexes rouges

Thèse de doctorat de l'Université Paris-Saclay

École doctorale n°577, Structure et Dynamique des Systèmes Vivants (SDSV)
Spécialité de doctorat: Évolution
Graduate School : Sciences de la vie et santé
Référent : Faculté des Sciences d'Orsay

Thèse préparée dans l'unité de recherche Ecologie Systématique et Evolution (Université Paris Saclay, CNRS, AgroParisTech), sous la direction de **David MOREIRA**, directeur de recherche, la co-direction de **Laura EME**, chargée de recherche, la co-direction de **Purificación LÓPEZ-GARCÍA**, directrice de recherche

Thèse soutenue à Paris-Saclay, le 28 Novembre 2022, par

Fabian VAN BEVEREN

Composition du jury

Didier CASANE Professeur, Université Paris Cité	Président
Fabien BURKI Professeur, Uppsala University	Rapporteur & Examinateur
Ingrid LAFONTAINE Professeure, Sorbonne Université	Rapportrice & Examinatrice
Richard DORRELL Chargé de recherche, École Normale Supérieure	Examinateur
David MOREIRA Directeur de recherche, Université Paris-Saclay	Directreur de thèse

THESE DE DOCTORAT

NNT : 2022UPASL073

université PARIS-SACLAY

ÉCOLE DOCTORALE
Structure et dynamique des systèmes vivants (SDSV)

Title: The evolution of red algal organellar genomes and the origin of red complex plastids

Keywords: algae, endosymbiosis, plastids, phylogenomics, eukaryotes, Proteorhodophytina

Abstract: Algae are a diverse group of eukaryotic photosynthesizers found in many lineages across the tree of eukaryotes. They are responsible for most of the primary production on Earth, yet the origin of algal diversity remains unclear. Glaucophytes, red algae, and green plants (green algae and land plants) have cyanobacterial-derived plastids, probably obtained from a single endosymbiosis event before their last common ancestor. In contrast, many other eukaryotic lineages have acquired their so-called "complex" plastids not directly from a cyanobacterium but by an endosymbiotic event with a green or red alga. The host and endosymbiont have been identified in the case of green complex plastids (in euglenids and chlorarachniophytes), but it remains unclear how the variety of red complex plastid-bearing lineages has originated. Four divergent lineages harbor red complex plastids: Myzozoa, Ochrophyta, Cryptophyta, and Haptophyta. They include many ecologically significant groups such as diatoms, kelp, and coccolithophores.

The red algal origins of their plastids may have remained unclear due to the poor taxon sampling available for unicellular red algae, which are more likely to be related to the red algal ancestor of red complex plastids. The recently erected superphylum Proteorhodophytina contains many unicellular and filamentous algae that have not yet been included in plastid phylogenomic analyses. Moreover, their plastid genomes are unique among red algae as they are highly dynamic due to the proliferation of self-splicing introns. We have sequenced and analyzed organellar genomes of 25 unicellular and filamentous red algae. We detect multiple intron proliferation events resulting in plastid and mitochondrial genome expansion. The sister lineages *Rhodella* and *Corynoplastis* contain the largest plastid genomes, but their genome expansions occurred independently. We show that self-splicing introns remain highly mobile in the red algal classes Rhodellophyceae and Porphyridiophyceae.

As we have improved the taxon sampling of unicellular red algae, we reconstructed the phylogeny of red algae and red complex plastids. Previous plastid phylogenomic analyses often recover red complex plastids as a monophyletic group, which supports a single red algal origin for them. With our improved taxon sampling, we confirm this result. Still, the placement of Ochrophyta is sensitive to the sequence evolution model used, sometimes resulting in Ochrophyta not grouping with the remaining red complex plastids. We find that the plastids of cryptophytes and haptophytes are closely related in all our analyses. This suggests that a relatively recent plastid transfer occurred between these two groups. Our results refute several scenarios that contrast this, such as those that propose a recent plastid transfer between haptophytes and ochrophytes. Finally, we are calibrating the plastid tree of life, which will help us estimate the age of potential endosymbiotic events that resulted in the current diversity of red plastid-bearing lineages. With the use of different sets of fossil calibrations, we aim to have confident estimates of the ages of endosymbiotic events. This way, we will be able to not only determine which endosymbiotic events occurred, but also the geological era in which they took place.

Titre : L'évolution des génomes organellaires des algues rouges et l'origine des plastes complexes rouges

Mots clés : algues, endosymbiose, plastes, phylogénomique, eucaryotes, Proteorhodophytina

Résumé : Les algues sont des eucaryotes photosynthétiques très diversifiés que l'on retrouve dans plusieurs groupes dans l'arbre des eucaryotes. Elles sont responsables de la majeure partie de la production primaire sur Terre, mais leur histoire évolutive reste obscure. Les glaucophytes, les algues rouges et les plantes vertes (algues vertes et plantes terrestres) possèdent des plastes dérivés de cyanobactéries, probablement obtenus à la suite d'une seule endosymbiose ayant eu lieu avant leur dernier ancêtre commun. En revanche, de nombreuses autres lignées eucaryotes ont acquis des plastes dits "complexes", non pas à partir d'une cyanobactérie, mais par un événement endosymbiotique dit "secondaire", avec une algue verte ou rouge. Dans le cas des plastes complexes verts (trouvés chez les euglènes et les chlorarachniophytes), à la fois l'hôte et l'endosymbionte ont été identifiés. Au contraire, les ancêtres des diverses lignées porteuses de plastes complexes rouges ne sont pas clairement identifiés. Quatre lignées divergentes abritent des plastes complexes rouges : Myzozoa, Ochrophyta, Cryptophyta et Haptophyta. Elles comprennent de nombreux groupes écologiquement importants tels que les diatomées, les laminaires et les coccolithophores.

L'identité de l'ancêtre de leurs plastes est possiblement restée obscure en raison du faible échantillonnage taxonomique disponible pour les algues rouges unicellulaires, qui sont les plus susceptibles d'être apparentées à l'ancêtre algal rouge de ces plastes complexes. Le superphylum Proteorhodophytina, récemment défini, contient de nombreuses algues unicellulaires et filamenteuses qui n'ont pas encore été incluses dans les analyses phylogénomiques des plastes. De plus, leurs génomes plastidiques sont uniques parmi les algues rouges car ils présentent une évolution très dynamiques dûe à la prolifération d'introns auto-épissés. Nous avons séquencé et analysé les génomes organellaires de 25 algues rouges unicellulaires et filamenteuses. Nous avons détecté de multiples événements de prolifération d'introns entraînant une expansion des génomes plastidiques et mitochondriaux. Les lignées sœurs *Rhodella* et *Corynoplastis* contiennent les plus grands génomes plastidiques, mais leurs expansions génomiques se sont produites indépendamment. Nous avons montré que les introns auto-épissés restent très mobiles dans les classes Rhodellophyceae et Porphyridiophyceae.

Avec cet échantillonnage taxonomique amélioré, nous avons reconstruit la phylogénie des algues rouges et des plastes complexes rouges. Les analyses phylogénomiques précédentes ont souvent retrouvé les plastes complexes rouges comme formant un groupe monophylétique, ce qui soutient une seule origine algale rouge de ces plastes complexes. Grâce à notre échantillonnage amélioré, nous confirmons ce résultat. Cependant, le positionnement des Ochrophyta est sensible au modèle d'évolution des séquences utilisé, ce qui fait que parfois les Ochrophyta ne se regroupent pas avec les autres plastes complexes rouges. Nous avons observé que les plastes des cryptophytes et des haptophytes sont étroitement liés dans toutes nos analyses, suggérant qu'un transfert de plastes relativement récent a eu lieu entre ces deux groupes. Nos résultats réfutent plusieurs scénarios évolutifs, comme ceux qui proposent un transfert récent de plastes entre les haptophytes et les ochrophytes. Enfin, nous appliquons des techniques de datation moléculaire sur la phylogénie des plastes, ce qui nous aidera à estimer l'âge des événements endosymbiotiques potentiels qui ont abouti à la diversité actuelle des lignées porteuses de plastes rouges. Grâce à l'utilisation de différents ensembles de calibrations fossiles, nous souhaitons obtenir des estimations fiables de l'âge des événements endosymbiotiques. Ainsi, nous serons en mesure de déterminer non seulement quels événements endosymbiotiques ont eu lieu, mais aussi l'ère géologique dans laquelle ils se sont déroulés.

Acknowledgements

Starting a PhD in a place where I knew no one was not easy, but I had a great time thanks to the many amazing people I met in France and those that supported me since way before.

First of all, thanks to my three supervisors for having me in the lab working on this fascinating subject and helping me become a better scientist. Thank you, **David**, for your knowledge of all things plastid-related and so much more and for your patience with my questions, and for reading the earliest drafts. No one has been clearer with pointing out my mistakes than **Puri**, which I mean in a good way! Thanks for always putting in the effort to improve my understanding of science and to ensure I always remain critical. I would never have started this PhD without the support from **Laura** during my MSc internship back in Sweden. Thank you so much for believing in me from the start, and helping me mature as a scientist through these last years. I'm also happy we were able to hang out outside of work, and I hope we can do that much more!

I am the luckiest person for having worked with all the great people in the lab! Beginning with the people I shared the PhD struggle with. **Thomas**, we started at the same time and are both doctors now! It has been great working on different aspects of plastid origin side-by-side. You made the lab much brighter with your silly jokes and your push to always make all of the lab more connected. I was also lucky enough to start in the presence of the "original four". Thank you, **Jodie**, for supporting me when I felt the worst and showing me the beautiful world under Paris. We've had many fun times (and good food); let's keep that going. Excited to see your book arrive in my mailbox soon! Thanks to **Luis** for guiding me in the world of protistology (even though you aren't really working on protists), it would now be weird to be at a protist conference without you there. It was very fun to play guitar and bass together, it motivated me a lot to play more. **Guillaume R**, you have been the funniest and silliest at many times, but also the most honest. Thanks for being a friend through this, and I hope you will soon not live so far away. And when you come back, you can teach me how to do bioinformatics properly! Also, the last one was a number 3. Thank you, **Gwendoline**, for bringing some calm into our office and helping me with a lot of the silly french administration. Sometimes I still refer to your famous bible in times of need. Of course, many PhDs joined during this adventure too. **Jazmin**, thank you for being very supportive when times were difficult. I am very happy we became good friends so quickly. I will miss the cute humming you do, it always brought a bit of charm into the office. **Brittany**, thanks for being willing to discuss phylogenomics and programming-related things and for being a great friend, especially considering the early days of the first lockdown. **Romain**, we've only known each other for a short while, but you are such a genuine and funny person. You already seem like a bioinformatics pro, so I am sure you'll do great!

Guifré, thanks for being the calmest force around! I guess now it helps that you can swim again nearly every day. Hope to visit you many times to join you for a plunge into the sea (or to play Mario Kart!). **Naoji**, you are the best microscope expert, actually better at microscopy than speaking Japanese. **Kristina**, there is no person more genuine and direct than you! It's cool to have a true protistologist around, even though you wrongly consider algae to be lame. They have much prettier colors, like the colors you find in Christmas sweaters! **Jolien**, het is altijd goed om Nederlands met iemand te kunnen spreken op het lab. Je bent een geweldige wetenschapper, en daarnaast steek je zoveel tijd in andere dingen. Ik heb geen idee hoe je het doet!

Lukas, my fellow red algae expert, your perspective is always insightful. And I wish we spend more time playing board games; perhaps soon we can get to it again! **Guillaume L**, happy to have your programming opinions on many things, and you have a great view on open science. And happy to see someone more into bikes than the average Dutch person! **Ana**, we barely talked in the first year or two, but now you are one of my best friends! It's a shame I can no longer drink cheap wine anymore, but it's worth it for all the fun and delicious dinners we've had! Thanks to the very supportive **Electra** and **Carolina**, who both given me the positive energy needed to get to my defense. And thanks to **Sergio** for your ideas on many scientific topics. Also, I want to thank the many students that have come and gone, but always improved the atmosphere in the lab. Especially **Iris**, with your help on my project but also with fun things outside of the lab. Excited to see how well you will do with your PhD!

Of course, I'd also like to thank everyone else in the DEEM team for the learning moments, and the friendliness and great atmosphere. Thank you, **Maria**, for always remaining positive when lab results looked negative. If I could be just half as cheerful and hard-working as you, I know I will do great in science. Thanks to **Miguel**, **Julien**, **Ludwig**, **Philippe**, and **Paola** for making this time at the lab as lovely as it was.

Also, thanks to many people outside the DEEM team, especially **Mohamed**, **Seïf**, **Thibault**, **Samuel**, **Tom**, and **Pauline**. And **Jacqui**, for discussing both science and weird food (including kimchi, which I unfortunately still don't like). I'm very grateful for the scientific guidance from **Richard Dorrell** and **Line Le Gall** as part of my thesis committee.

I'd like to thank everyone else who has made my time in France so fun, especially the following people! **Anjana**, I am very happy we met through French course. We went from making the silliest jokes to discussing existentialism with some Aperol by our side; what could be better? **Ameya**, you have been so kind and welcoming since we met, and you have been a great companion on adventures (role-playing or otherwise), but please note that sharks are not toxic. So happy to have played silly D&D games with the great physicists **Ben**, the always-engaging **Nat**, hilarious **Rayna**. And just fun evenings with **Nikita**, **Florian** and **Alvaro**. And, of course, my housemate **Gabriel**, your determination to work on all kinds of projects is something I could learn much from, and you even corrected a grammar mistake on the very first page of this thesis (missed by everyone else!). Some of my closest friends have not exactly been close geographically. **Rosemarie**, thanks for all the fun games we played, going back many years. We need to visit each other soon and play games in the same room! **Christian**, I've known you for even longer, and it's always been natural chatting with you about anything. Sometimes we'd not talk for months, yet each time it seemed like we had played games just a day ago.

Ik zou natuurlijk niet eens bestaan zonder mijn pa (**David**) and ma (**Christine**), ik hou zo veel van jullie, en ben enorm dankbaar voor jullie steun mijn hele leven, waar ik ook heen ging en wat ik ook ging doen, jullie waren altijd dichterbij dan het lijkt! En natuurlijk ben ik dankbaar voor mijn broer en zussen. **Gwenda**, **Ivan** en **Aileen**, ik ben niet bepaald de meest spraakzame broer, maar ben echt heel blij met jullie steun en de leuke gesprekken waneer we elkaar zien! **Elodie** en **Nathan**, jullie zijn de allerliefste en bregen altijd een glimlack op mijn gezicht! **Gerda** en **Gerard**, ook altijd dankbaar voor jullie steun, zelfs als jullie veel meer aan het hoofd hebben. Bedankt :)

Last but definitely not least, **Feriel**. The last years of the PhD are the hardest, but that is why I am so lucky to have had your support through all of this. You've always been thoughtful and kind, funny and so bright. Clearly I'm not great with putting these kind of things in words, so I will just say: I love you soooo much. None of this would have gone as smoothly without you.

Contents

1 - Introduction

Energy is necessary for life to exist, and continue existing. The largest source of energy on Earth is sunlight, and a small part of solar energy is captured by living organisms to sustain themselves and thrive. The main process—but not the only one—that captures sunlight is photosynthesis, which several bacterial and eukaryotic groups use to synthesize carbohydrate molecules and create most of the organic carbon supporting heterotrophic life (Thornton 2012). Plants, algae, and cyanobacteria perform oxygenic photosynthesis as they produce oxygen during photosynthesis. Oxygenic photosynthesis (hitherto photosynthesis) started with cyanobacteria, at least before the Great Oxygenation Event, which occurred 2.4-2.0 Ga (Giga-annum, 10^9 years ago), where oxygen started to accumulate in the at the time mostly anoxic atmosphere (Rasmussen et al. 2008; Schirrmeister et al. 2016).

Nowadays, not only cyanobacteria are photosynthesizing, but also many eukaryotes consisting of land plants and a multitude of algal lineages. Their phylogenetic, ecological, and morphological diversity is wide, including flowering plants, coralline algae, and kelp. But how did this large diversity gain the capability to perform oxygenic photosynthesis? Photosynthesizing cells of these organisms contain a membrane-bound compartment, the plastid (or chloroplast), where photosynthesis takes place. Although at first considered a part of the cell arising like other cellular structures (e.g., the Golgi apparatus), eventually researchers noted similarities between plastids and cyanobacteria (Taylor 1970). In 1885, Schimper mentioned in a footnote: "If it can be conclusively confirmed that plastids do not arise *de novo* in egg cells, the relationship between plastids and the organisms within which they are contained would be somewhat reminiscent of a symbiosis. Green plants may in fact owe their origin to the unification of a colorless organism with one uniformly tinged with chlorophyll" (Martin and Kowallik 1999). However, the first to propose a more detailed hypothesis was Constantin Mereschkowsky (Mereschkowsky 1905; Martin and Kowallik 1999). In summary, his argument was that no plastid ("chromatophore") is ever observed to differentiate from the protoplasm, i.e., that it cannot be built from the building blocks available to a cell, from scratch. Instead, plastids are passed from generation to generation and are highly similar to free-living cyanobacteria to which they are related. This theory remained much ignored until the revival and further detailing by Lynn Margulis (before, Lynn Sagan) (Sagan 1967). All photosynthesis is in essence performed by cyanobacteria or cyanobacterial-derived compartments that are (reduced) intracellular endosymbionts (cells living in cells).

Intracellular endosymbiosis is a relatively common occurrence. Endosymbionts are found as partners to the host cell in insects (e.g., *Buchnera* in aphids), corals, and many single-celled eukaryotes, although parasitic and other endosymbionts are also prevalent (Cavalier-Smith and Lee 1985; Wernegreen 2004). Some endosymbionts can no longer survive outside the host cell, and moreover the host may become dependent on endosymbiont-produced metabolites. In such cases, we speak of obligate (in contrast to facultative) endosymbiosis, and such cases are relatively common as well (Theissen and Martin 2006). As obligate endosymbionts remain in the host cell, many functions needed to be a free-living cell can be lost. However, contrasting to the plastid, obligate endosymbionts typically

still encode everything they need for their secluded life in their genomes. They do not need much more from the host other than the exchange of metabolites. Plastids instead are organelles, just as mitochondria are (Cavalier-Smith and Lee 1985). Organelles are controlled by the host, including the transport of most proteins and the division of the organelle. Due to this, few genes (and sometimes no genome at all) reside anymore within an organelle. Typically, control of division by the host, transfer of genes to the host and transport machinery for protein import are considered hallmarks of organelles, separating them from obligate endosymbionts (Cavalier-Smith and Lee 1985). At this point, the only recognized organelles are plastids and mitochondria (including mitochondrion-related organelles such as mitosomes; Shiflett and Johnson 2010), clearly contrasting with the widespread occurrence of (obligate) endosymbiosis.

As plastids have become organelles, most functions besides photosynthesis were lost—but not all. Often the plastid is responsible for the biosynthesis of aromatic amino acids (phenylalanine, tryptophan and tyrosine) (Gould, Waller, et al. 2008). The synthesis of heme, Fe-S clusters and isopentenyl diphosphate also often takes place in the plastid (Lichtenthaler et al. 1997; Gould, Waller, et al. 2008). Due to these other functions, some plastid-bearing lineages have lost photosynthesis but kept the plastid around, as is the case in apicomplexans such as the malaria-causing pathogen *Plasmodium*. In some cases, all genes needed are encoded in the nuclear genome of the host resulting in a genome-lacking plastid and, in rare cases, the plastid has been lost completely (Smith 2018). Thus although plastid acquisition is mainly related to a lifestyle change from heterotrophy to auto- or mixotrophy, a change can again occur to a heterotrophic-only lifestyle, with or without the plastid (genome).

All plastids, photosynthetic or not, have a cyanobacterial ancestry. And in many eukaryotic lineages (including land plants), a cyanobacterium was indeed the endosymbiont, resulting in what is known as a primary plastid. But other eukaryotic groups took a different path, instead gaining their plastid through endosymbiosis with a plastid-bearing eukaryote (Archibald 2009; Burki 2017). Specifically, many groups have a plastid that is red algal-derived, yet these groups are not closely related to each other, but to plastid-lacking lineages (fig. 1.1). Many ecologically relevant lineages contain such a red algal-derived plastid, such as diatoms, coccolithophores and kelps. The focus of this thesis is the origin of these red algal-derived plastids. First, the primary plastid-containing lineages (including red algae) will be introduced in section 1.1, followed by a section (1.2) on the eukaryote-derived plastids (including red algal-derived plastids). The introduction will end with an overview of the hypotheses about the origin of red complex plastids (1.3), and the relevant evidence to gauge the plausibility of these hypotheses (1.4).

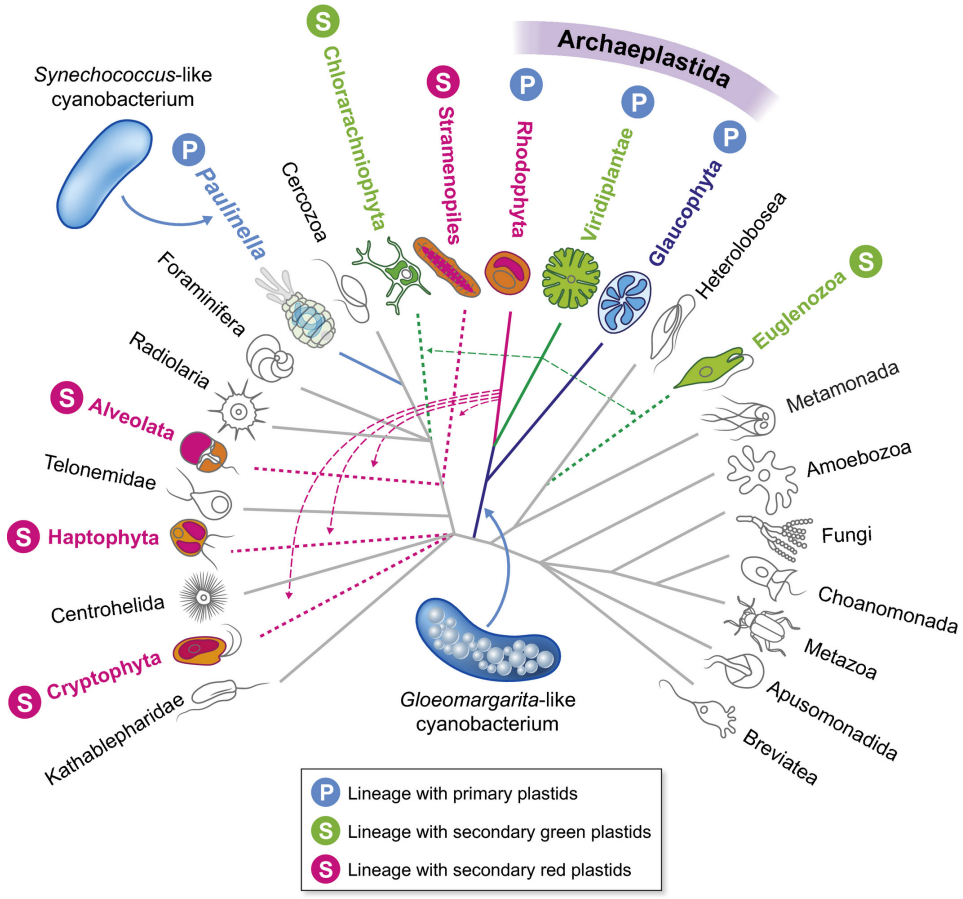


Figure 1.1: The distribution of photosynthesis in global eukaryotic phylogeny. Coloured solid branches correspond to photosynthetic lineages endowed with primary plastids and coloured dashed branches to lineages with secondary plastids (green and red colours indicate the type of secondary endosymbiont, green or red algae, respectively). Blue arrows show the two known primary endosymbioses (in Archaeplastida and *Paulinella*) and green and red arrows indicate secondary endosymbioses involving green and red algal endosymbionts. Grey branches correspond to nonphotosynthetic eukaryotic phyla. The tree has been largely modified from Adl et al. (2012). Image and caption from Ponce-Toledo et al. (2019), with authorization from the authors.

1.1 . Primary plastids

The evolutionary history of plastids has been difficult to reconstruct, partly due to their placement in many different branches in the tree of life (Burki 2017). One thing that is clear is that a subset of plastids are directly derived from a cyanobacterial endosymbiont (primary plastids), whereas others are derived from an already plastid-bearing eukaryote (secondary plastids). Furthermore, higher-level endosymbioses are possible, resulting in tertiary and quaternary plastids. It is important to note that different hypotheses on the origin of algal lineages will infer certain plastids to potentially be of different levels. For example, under some hypotheses, the plastids of ochrophytes (see 1.2.2) are secondary, but in others tertiary (e.g., Stiller, Schreiber, et al. 2014). Because of this, I will call plastids of secondary or higher-level origin *complex plastids* to avoid implicating a hypothesis when it is not relevant. As for primary plastids, there is a broad consensus on which lineages bear these plastids: green algae and land plants, red algae, and glaucophyte algae (together, Archaeplastida) (Palmer 2003; Burki 2017). These three groups will be discussed in their separate subsections, along with their likely single origin. But before, it is essential to note that multiple lineages bear potential primary plastids outside of these three main lineages.

The most likely cyanobacterial endosymbiont that may represent an independent plastid origin is the chromatophore found in some species of the testate amoeba genus *Paulinella*, with *P. chromatophora* being described by Lauterborn in 1895 (Kepner 1905). Photosynthetic *Paulinella* spp. contain two blue-green compartments within their cells (Kepner 1905; Johnson et al. 1988). These "chromatophores" are closely related to cyanobacteria of the genera *Synechococcus* and *Prochlorococcus* (Marin et al. 2005), contrary to the *Gloeomargarita*-like plastids of Archaeplastida (Ponce-Toledo, Deschamps, et al. 2017). This is a more recent acquisition, as other *Paulinella* species contain no such cyanobacterial compartment (Johnson et al. 1988). There has been discussion on whether the chromatophore can be considered a true photosynthetic organelle, a plastid (Bhattacharya and Archibald 2006; Theissen and Martin 2006). At the time of this discussion, it was known that division of the chromatophore is controlled by the host. But it is now also clear that the chromatophore genome is reduced (~30% of the average size of free-living *Synechococcus* spp.), and that some symbiont genes reside in the host nucleus and are transported across membranes into the symbiont (Nowack et al. 2008; Mackiewicz, Bodył, et al. 2012). This would make it a genuine plastid, and the only definite case outside of the Archaeplastida and archaeplastidal-derived complex plastids.

Another curious case are the cyanobacterial endosymbionts of the diatom species in the genus *Rhopalodia* and related species of the family Rhopalodiaceae (Adler et al. 2014; Schvarcz et al. 2022). Here, the cyanobacterial endosymbiont does not perform photosynthesis—the diatom already has a red algal-derived plastid—but instead performs nitrogen fixation, allowing *Rhopalodia* to thrive in nitrogen-poor conditions (Adler et al. 2014). The genome is reduced but to a lesser extent to that of *Paulinella*, being roughly 60% the size of free-living *Cyanothece* spp. (Nakayama et al. 2014). Less well-identified cyanobacterial endosymbionts are found in *Auranticordis quadriverberis* and *Hermesinum adriaticum*, both part of the phylum Cercozoa (Hargraves 2002; Chantangsi et al. 2008; Mackiewicz, Bodył, et al. 2012). More cyanobacterial endosymbionts appear in many

parts of the tree of eukaryotes, including in diatoms, corals, land plants and fungi; cyanobacterial endosymbionts may be undergoing organellogenesis more than the likely single ancient origin of Archaeplastida implies (Mackiewicz, Bodył, et al. 2012).

1.1.1 . Glaucophytes

Of the three groups of Archaeplastida, the most taxon-poor are the glaucophytes, which are ranked at the phylum level of Glaucophyta (Skuja 1948, 1954)[1], which is synonymous with Glaucocystophyta (Kies and Kremer 1986). Their name is based on their typical blue-green color (Greek: *glaukos*) and being a plant *sensu-lato* (*phyton*). Species available in culture collections only represent four genera: *Cyanophora*, *Cyanoptyche*, *Glaucocystis*, and *Gloeochaete* (Price, Steiner, et al. 2017; Figueroa-Martinez et al. 2019). More species, including new genera, have been described, but their association remains uncertain (Figueroa-Martinez et al. 2019). All representatives are unicellular and photosynthetic and can be flagellated and motile, or spherical vegetative cells that may form colonies (Jackson, Clayden, et al. 2015). The first description of any glaucophyte is of *Glaucocystis bullosa* (*Palmella bullosa*) by Friedrich Traugott Kützing in 1836 (Kies and Kremer 1986). However, the most studied representative—sometimes considered the model organism of Glaucophyta—is *Cyanophora paradoxa*, which is the only glaucophyte with a sequenced and published nuclear genome (Price, Chan, et al. 2012; Price, Goodenough, et al. 2019).

The photosynthetic compartment of glaucophytes was initially considered a "cyanelle", a cyanobacterial endosymbiont, that was even classified as its own species (*Cyanocyta korschikoffiana*) (Hall and Claus 1963; Schenk 1994). Some features of the glaucophyte plastid indeed make it seem more cyanobacterial, mainly the presence of a peptidoglycan layer between its two membranes (Schenk 1994; Reyes-Prieto, Russell, et al. 2018). On the other hand, evidence for integration on the level of an organelle started to pile up, including the lack of respiration in the cyanelles (Floener and Bothe 1982). The glaucophyte plastid genome size was shown to be small (Herdman and Stanier 1977), with most genes encoded by the host (Bayer and Schenk 1986). Due to these organellar features, the term "cyanelle" was dropped in favor of "muroplast", referring to any photosynthetic organelle with a peptidoglycan ("muropeptide") cell wall (Schenk 1994).

1.1.2 . Green algae and land plants

The green algae and land plants are with ease the most well-known photosynthetic organisms, and are grouped together in the Viridiplantae — but also called Chloroplastida, as not all representatives are "true plants" (Adl, Simpson, et al. 2005). The green lineage is highly diverse in terms of morphology, and is widespread in aquatic and—especially in comparison to other photosynthesizers—terrestrial ecosystems. The group is typically divided into two main groups: Chlorophyta and Streptophyta (fig. 1.2), although a subset of chlorophyte algae show up as sister to Streptophyta and all other Chlorophyta, making them a putative third group (Prasinodermophyta; Li et al. 2020). For photosynthesis, both chlorophyll *a* and *b* are used; the latter is missing in other archaeplastidal lineages (Jeffrey, Wright, and Zapata 2011).

[1]Glaucophyta classification by Skuja in 1948 was regarded as *nomen nudum*; the adequate description was in 1954. I was unable to retrieve these two original works and instead refer to Kies and Kremer (1986).

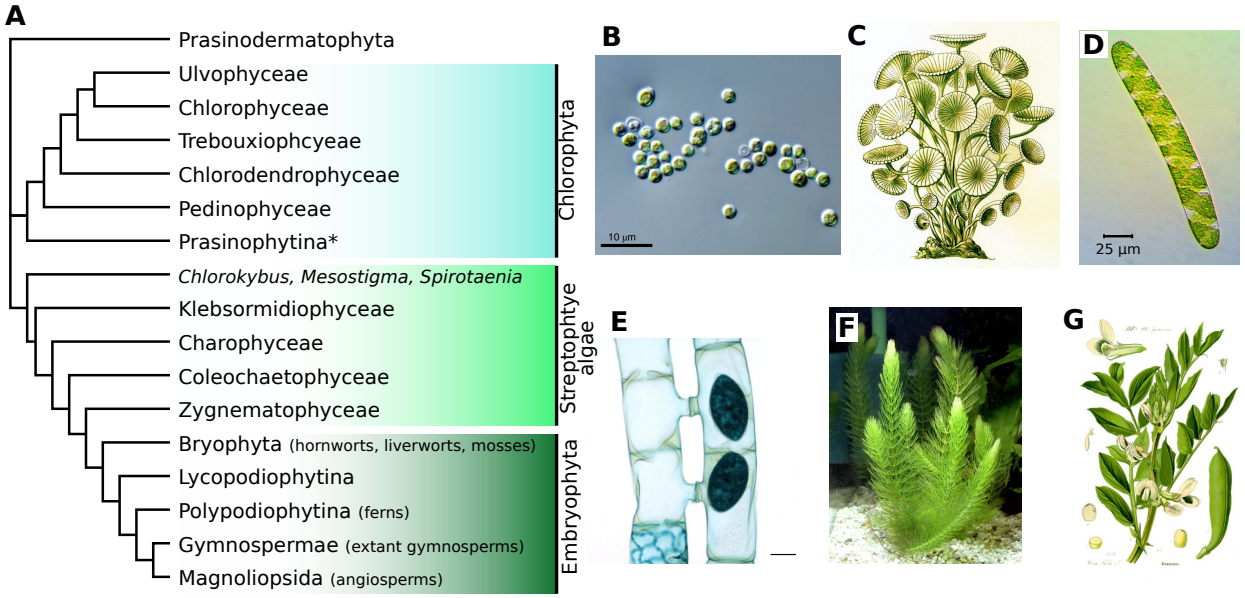


Figure 1.2: The phylogeny of green algae and land plants with images of several representatives. (A) Cladogram of Viridiplantae with a basal polytomy reflecting uncertainty in their relationships. (B) Light micrograph of *Prasinoderma coloniale* (Prasinodermatophyta) (Li et al. 2020). (C) Illustration of *Acetabularia* (Ulvophyceae) by Ernst Haeckel (1904). (D) Microscopy image of *Spirotaenia condensata* by Oliver S. (E) Photomicrograph of *Spirogyra* (Zygnematophyceae) conjugation after the two strands have combined to form a zygote (scale = 0.1 mm), taken by Jon Houseman. (F) *Ceratophyllum* sp. (hornworts, Bryophyta). (G) Illustration of *Vicia faba* (fava bean, Magnoliopsida) by Otto Wilhelm Thomé (1885). B and D were cropped for use in this thesis, and the scale bar was edited and moved for readability in D. Images C, F and G are in the public domain, remaining images are under a creative commons license (see Supplementary material). *Prasinophytina is likely polyphyletic.

Chlorophyta contains only algae, including lineages that have terrestrial niches such as those forming lichen symbioses (Lewis 2007). The subphylum Chlorophytina consists of species-rich classes Ulvophyceae, Chlorophyceae, and Trebouxiophyceae, which together account for ~7000 of the ~7300 described chlorophyte species (Guiry and Guiry 2022). Ulvophyceae includes the siphonous algae, unicellular but macroscopic organisms, formed by a single tubular cell that can branch in various patterns (Verbruggen et al. 2009). This tubular body plan allows these algae to be among the largest uninucleate cells (e.g., *Acetabularia*; fig. 1.2C) and multinucleate cells (e.g., *Caulerpa*) and due to the size and calcification of some species, they also are among the few green algal lineages with a rich fossil record (Smith 2001; Verbruggen et al. 2009). Siphonous algae are also a good example of the difficulty in resolving the phylogenetic relationships among Ulvophyceae, as they used to be considered a monophyletic group but in some analyses show up as two distinct clades (Jackson, Knoll, et al. 2018; Gulbrandsen et al. 2021).

Other green algal lineages are part of the Streptophyta and do not form a monophyletic clade but instead encompass multiple independent lineages 1.2. Three unicellular genera, *Mesostigma*, *Chlorokybus* and *Spirotaenia* form a clade sister to all other streptophytes (Lemieux, Otis, et al. 2007; Wickett et al. 2014). Filamentous species are found in the phyla Klebsormidiophyceae (unbranched filaments) and Coleochaetophyceae (branched), with no unicellular representatives known (Hall and Delwiche 2007). Similarly, the stoneworts (Charophyceae) are all macroscopic and contain root-like rhizoids, as well as whorls and branches (Hall and Delwiche 2007). This may give a false sense of continuous growing complexity, but the Conjugatophyceae (Zygnematophyceae) unambiguously show this is not the case. They are the clade typically found as sister to land plants (Lemieux, Otis, et al. 2016), and include a diverse set of both unicellular and multicellular species with more described species than the other streptophyte algae together (but still fewer than the land plants) (Gontcharov 2008).

The land plants emerged from within the streptophyte algae between 1 000 and 500 million years ago (Morris et al. 2018b; Su et al. 2021). Two clades make up the land plants: Bryophyta and Tracheophyta. Bryophytes consist of mosses, liverworts, and hornworts and their relationships have been difficult to resolve among them and the Tracheophyta. However, as methods and taxon sampling of these lineages improved—especially of the hornworts—the monophyly of the Bryophyta as a sister group to Tracheophyta was increasingly supported (Su et al. 2021). Ferns, lycophytes, gymnosperms, and angiosperms make up the rest of the land plants, together known as the Tracheophyta. Land plants are widely distributed among terrestrial environments and are responsible for a large part of all primary production (Geider et al. 2001). In contrast to the scarcity of green algae in the ocean at current times (in comparison to algae with red algal-derived plastids), land plants have made the green lineage the most common photosynthetic lineage on land (Falkowski et al. 2004).

1.1.3 . Red algae

The red algae (Rhodophyta) are a diverse group of marine, freshwater, and a few terrestrial algae. Although they come in many colors, they are known for the pigment phycoerythrin, which gives many of them their red color. Phycoerythrin is also present in cryptophytes and cyanobacteria but absent in both glaucophytes, and green algae and land plants (Jeffrey, Wright, and Zapata 2011). Besides their pigmentation, they are recognizable by their lack of flagella and centrioles in all life stages—a unique characteristic among eukaryotes (Yoon, Zuccarello, et al. 2010). Other characteristics, such as pit connections and pit plugs are also unique to red algae, but are limited to a selection of multicellular representatives (Cronquist 1960). Instead of cellulose cell walls, they are usually covered by a mucilaginous sheath, with building blocks including agaran and carrageenan (Knutsen et al. 1994).

Most of the described red algal species, and especially the multicellular species, are assigned to the Florideophyceae (Guiry and Guiry 2022). This class includes economically important genera such as *Chondrus*, and *Gelidium*, from which carrageenan and agaran are extracted (Wikfors and Ohno 2001). The florideophytes are one of the few eukaryotic lineages displaying so-called complex multicellularity, with multiple cell types and intercellular communication (Knoll 2011). This multicellularity displays itself as seaweeds but also as coral-like structures as seen in the coralline algae (fig. 1.3C). The other class in the Eurhodophytina is Bangiophyceae (*sensu stricto*), with "simple" multicellular representatives (fig. 1.3B; Yoon, Muller, et al. 2006). Some of the most well-known red algae are species of *Porphyra* and *Pyropia* which are better known as the dried edible seaweed nori.

Contrasting to this is the subphylum Cyanidiophytina, which encompasses just one class, the Cyanidiophyceae. All representatives are unicellular and colored green due to the pigment phycocyanin (fig. 1.3F). Cultured representatives are all extremophiles, being acidophilic, thermophilic, or both. However, representatives of the genus *Cyanidium* have also been recorded in caves with moderate conditions in Chile, France, Israel and Italy (Ciniglia, Yoon, et al. 2004; Reeb and Bhattacharya 2010). The first valid description of a representative was *Pleurococcus sulphurarius* by Galdieri (1899), which is now known as *Galdieria sulphuraria* (Merola et al. 1981). Currently the class contains four genera: *Cyanidium* and *Galdieria* (order Cyanidiales), and *Cyanidiococcus* and *Cyanidioschyzon* (order Cyanidioschyzonales) (Liu, Chiang, et al. 2020). Although they were proposed as a class within the red algae in 1981 (Merola et al. 1981), they were often instead placed in Porphyridiales or Bangiophyceae, typically with other unicellular red algae (Saunders and Hommersand 2004). However, phylogenomic analyses have now confidently placed them as sister to all other red algae (Müller et al. 2001; Yoon, Hackett, Pinto, et al. 2002; Saunders and Hommersand 2004; Yoon, Muller, et al. 2006).

Finally, there are four classes of mesophilic algae that are either unicellular or filamentous, including both marine and freshwater species. Two groups host only mesophilic unicellular representatives, the Porphyridiophyceae (including *Porphyridium*) and Rhodellophyceae (fig. 1.3G and H; Yoon, Muller, et al. 2006). In contrast, the Compsopogonophyceae (fig. 1.3E) only have filamentous representatives with some of them containing simple pit connections (Saunders and Hommersand 2004). Specifically, *Rhodochaete*—hosted in its own order Rhodochaetales—contains pit connections as do

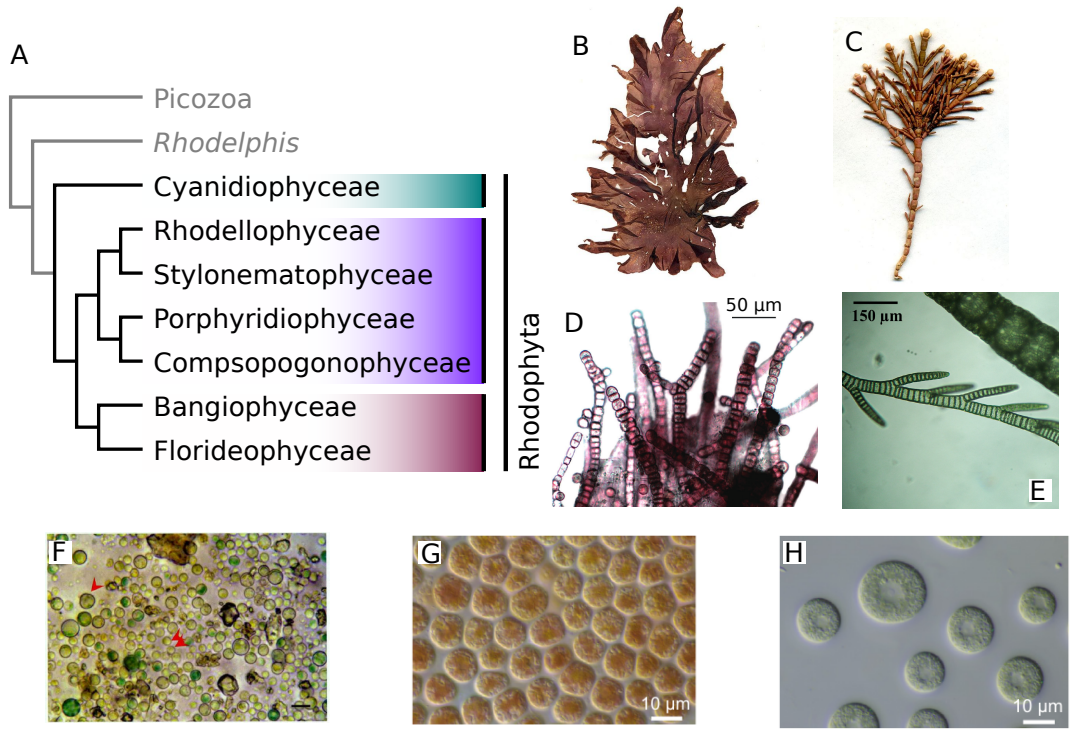


Figure 1.3: The phylogeny of red algae with images of several representatives. (A) A cladogram displaying the relationship among red algal classes of Cyanidiophytina (cyan) Proteorhodophytina (purple) and Eurhodophytina (red) and two heterotrophic groups sister to red algae (Picozoa and *Rhodelphis*). (B) Picture of herbarium sheet of *Porphyra umbilicalis* (Bangiophyceae). (C) Picture of herbarium sheet of *Corallina officinalis* (Florideophyceae). (D) Micrograph of *Viator vitreocola* (Stylonematophyceae) with "erect filaments becoming bi to multiseriate and occasionally branched" (Hansen et al. 2019). (E) Micrograph of *Compsopogon caeruleus* (Compsopogonophyceae) with "main branch and typical young uniseriate branches" (Meichtry Zaburlin et al. 2019). (F) Light microscopy image of the Cyanidiophyceae collected from Landmannalaugar (Iceland). "Single arrowhead indicates tentatively identified *Galdieria maxima*; double arrowheads indicate possible *G. sulphararia* with smaller cell size. Scale bar = 10 pm." (Ciniglia, Yang, et al. 2014) (G) "Palmelloid stage of *Flintiella sanguinaria* UTEX LB2060 (Porphyridiophyceae)" (Muñoz-Gómez, Mejía-Franco, et al. 2017). (H) "*Bulboplastis apyrenoidosa* NIES-2742 (Rhodellophyceae) as single unicells" (Muñoz-Gómez, Mejía-Franco, et al. 2017). Pictures B and C were taken by Gabriele Kothe-Heinrich. Images B-E were available under a creative commons license that allows sharing of the image and permission from images F-H was received for reuse within this thesis (see Supplementary material).

certain representatives of Compsopogonales (Zuccarello et al. 2000). The fourth class is Stylonematophyceae, which contains both unicellular (e.g., *Rhodosorus*) and filamentous representatives (e.g., *Viator*; fig. 1.3D) (Yoon, Muller, et al. 2006; Aboal et al. 2018; Hansen et al. 2019). This class shows multiple gains or losses of multicellularity, for example, *Chroothece* and *Chroodactylon* are closely related sister lineages, but only *Chroodactylon* is filamentous (Aboal et al. 2018).

The red algal tree of life

In this thesis, the main question concerns the origin of algal lineages with a red algal-derived plastid. To understand which type(s) of red algae this complex plastid has been acquired from, it is important to detail further the relationships among red algae. Modern taxonomy of red algae starts, of course, with Carl Linnaeus, who described some red algal species in the 18th century (such as Irish moss, *Corallina officinalis*) (Moestrup 2001). However, the relationships between algae and other eukaryotes—let alone between algal groups—were unclear; Linnaeus categorized algae (including lichens) in the Cryptogamia, together with mosses, ferns, and fungi. The understanding of relationships among algal groups started with Lamouroux (1813), who separated the different groups of algae based on their pigmentation (Lamouroux 1813; Saunders and Hommersand 2004).

Most important for the origin of red secondary plastids are the relationships between unicellular red algae, which are more likely candidates as a symbiont of unicellular eukaryotes. Initially, all unicellular red algae were classified in the Bangiophyceae (*sensu lato*), along with multicellular species, including the filamentous and bladed representatives of Bangiales (Gabrielson, Garbary, and Hommersand 1986). Gabrielson and colleagues (1985) were the first to show, based on morphological characters, that the Bangiophyceae *sensu lato* were not a valid group, as it is paraphyletic. This was further confirmed with phylogenetic analysis using rbcL and 18S rRNA sequences (Freshwater et al. 1994; Ragan et al. 1994). Now, Bangiophyceae (*sensu stricto*) refers only to the representatives of Bangiales.

Molecular analyses also revealed the monophyly of Compsopogonophyceae (Rhodochaetales, Erythropeltales, and Compsopogonaceae plus Boldiaceae) and the lack of monophyly of "Porphyridiales", consisting of the currently established classes Rhodellophyceae, Porphyridiophyceae, and Stylonematophyceae (Yoon, Zuccarello, et al. 2010). When the four classes Porphyridiophyceae, Rhodellophyceae, Compsopogonophyceae and Stylonematophyceae were defined; their relationships among each other remained unknown, which was attributed to rapid diversification in the Rhodophytina (i.e., Rhodophyta excluding cyanidiophytes) (Yoon, Muller, et al. 2006). Their relationships were only shown later by Muñoz-Gómez, Mejía-Franco, et al. (2017), who recovered them as a monophyletic group: the subphylum Proteorhodophytina. A short branch at the base of Proteorhodophytina does indeed reflect a likely rapid diversification of the four classes. Whether nuclear and mitochondrial data support the monophyly of Proteorhodophytina remains an open question.

1.1.4 . The origin of Archaeplastida

One key question that remains open in the evolution of algae is the origin of the primary plastid in Archaeplastida. Traditionally, many cyanobacterial endosymbioses were assumed for the many different plastid-bearing lineages (Sagan 1967; Cavalier-Smith 1982). One of the first attempts to unite the origin of plastids of Archaeplastida was by Cavalier-Smith (1982), although the plastids of Euglenozoa (see 1.2.1) and Dinozoa (part of Myzozoa, see1.2.2) were included as well. However, separate acquisitions have been suggested, including for red algae (Stiller and Hall 1997; Palmer and Delwiche 1998). Plastid phylogenies have typically found a monophyletic relationship among the Archaeplastida (e.g., Ponce-Toledo, Deschamps, et al. 2017), which may suggest a single origin. However, the monophyly of Archaeplastida has been difficult to recover with mitochondrial and nuclear data (Palmer 2003; Mackiewicz and Gagat 2014), shedding doubt on a single endosymbiotic event at the origin of Archaeplastida. Although this does not directly affect the origin of red complex plastids—the monophyly of red algae is not disputed, there may be interesting parallels to be drawn between the origin of complex red plastids and archaeplastidal plastids.

The first question is whether a single cyanobacterium is probable to lay at the origin of archaeplastidal plastids. As stated above, phylogenetic analysis of plastids strongly supports this as they are largely found as monophyletic, but multiple endosymbioses with closely related cyanobacteria can not be ruled out (Mackiewicz and Gagat 2014). However, other evidence suggests a single cyanobacterial origin of the archaeplastidal plastids. All three lineages have two plastid membranes with a homologous membrane protein transport machinery: the **t**ranslocon at the **i**nner and **o**uter envelope membrane of the **c**hloroplast (Tic/Toc supercomplex) (Price, Steiner, et al. 2017). The atpA gene cluster—consisting of genes for ATPaseA and ATPaseB, but also unrelated genes—is found across all Archaeplastida plastid genomes as well as all lineages with complex plastids (excluding Myzozoa) (Kowallik 1994; Stoebe and Kowallik 1999). Moreover, there is a large overlap between plastid-encoded proteins, with a minimum of ~251 protein-coding genes in the common ancestor of the archaeplastidal plastid (Figueroa-Martinez et al. 2019). However, similar pressures may have result in the retention of a similar gene set. Overall, it seems more probable that there was a single cyanobacterial endosymbiont at the origin of the archaeplastidal plastid.

Even if a single cyanobacterial endosymbiont is at the origin, serial endosymbiosis may play a role in explaining the origin of plastids in one or more of the archaeplastidal lineages. To address such scenarios, we ideally need to have a good idea of the relationships among the three lineages AND other eukaryotes, but unfortunately no consensus exists as every possible relationship among the three lineages has been found using different data sets from different genetic compartments (Mackiewicz and Gagat 2014). A recent study using plastid data found Glaucophyta diverging first (Ponce-Toledo, Deschamps, et al. 2017), another recent study using nuclear data found Rhodophyta diverging first instead (Irisarri et al. 2022). More substantial has been the recent placement of Picozoa and *Rhodelphis* as sister lineages to Rhodophyta (Gawryluk et al. 2019; Schön et al. 2021). Although only a single plastid loss (Picozoa) and photosynthesis loss (*Rhodelphis*) may be necessary, it may instead support a secondary origin of the plastid in red algae and *Rhodelphis*. However, the consensus remains that the archaeplastidal plastid is explained by a single endosymbiotic event.

1.2 . Complex plastids

As shortly mentioned in section 1.1, many algal groups have gained their plastid not directly from a cyanobacterium, but instead have acquired it from plastid-bearing eukaryotes. A large diversity exists with such plastids, spread among the tree of life (fig. 1.1). All known complex plastids are of green or red algal ancestry, with no complex plastids known of glaucophyte origin (Archibald 2009; Burki 2017). Although this does not necessarily mean the plastid was acquired directly from such a red or green alga, but may instead be acquired from a complex plastid-bearing lineage (Cavalier-Smith, Allsopp, et al. 1994; Archibald 2009). This is definitely true for specific dinoflagellate lineages (see section 1.2.2;Waller and Kořený 2017), but may be true for other—especially red—complex plastids (Cavalier-Smith, Allsopp, et al. 1994). In the next two subsections, the green and red complex plastids will be introduced.

1.2.1 . Green complex plastids

Three groups are known to harbor green complex plastids, these are chlorarachniophytes (Rhizaria), euglenophytes (Excavata) and several dinoflagellates with green algal-derived plastids (*Lepidodinium viride*, *L. chlorophorum* and two undescribed dinoflagellates; Sarai et al. 2020). Dinoflagellates typically have a red-algal derived plastid, but the pigmentation of *Lepidodinium* and the observation of both chlorophyll *a* and *b* shows that *Lepidodinium* has a plastid of green algal origin (Watanabe, Takeda, et al. 1987; Watanabe, Suda, et al. 1990). Due to its clear placement in dinoflagellates, its green plastid-acquisition has always been considered separate to that of chlorarachniophytes and euglenophytes. Phylogenomic analysis of the plastid genome of *L. chlorophorum* has shown that the plastid is most closely related to the green alga *Pedinomonas* (Trebouxiophyceae, Chlorophyta) (Kamikawa, Tanifuji, et al. 2015; Jackson, Knoll, et al. 2018). The two dinoflagellates with no formal description (only known as MGD and TGD) also contain pedinophyte green algal endosymbionts (Sarai et al. 2020; Matsuo, Morita, et al. 2022). Moreover, these two lineages also contain a relict genome of the endosymbiont (nucleomorph), and endosymbiotic gene transfer (EGT) seems to still be occurring as some genes are present in both the host nucleus and nucleomorph genomes (Sarai et al. 2020). It is currently unclear whether *Lepidodinium* spp. also contain a nucleomorph. A nucleomorph-like structure is present—although not as obvious as in MGD and TGD (Watanabe, Takeda, et al. 1987; Sarai et al. 2020), but a nucleomorph genome or transcriptome has not yet been sequenced, or not yet assembled from DNA and RNA sequences reported of *L. chlorophorum* (Kamikawa, Tanifuji, et al. 2015; Matsuo and Inagaki 2018).

Chlorarachniophytes and euglenophytes also both contain green complex plastids, each representing a monophyletic clade with multiple described species: Euglenophyceae contains ~1000 and Chlorarachniophyceae ~15 (Cavalier-Smith 2016; Guiry and Guiry 2022). Euglenozoa and Rhizaria (including Chlorarachniophyceae) have been considered closely related, leading to Cavalier-Smith to propose a single endosymbiotic event in their common ancestor (Cavalier-Smith 1999). Although their plastids are different—chlorarachniophytes have one more membrane (4) and a nucleomorph—Cavalier-Smith suggested that these differences occurred due to differentiation after the endosymbiotic event. However, it is now known that both belong to different parts of the eukaryotic tree of life (fig. 1.1),

and moreover that the plastids are of completely different green algal origin (Jackson, Knoll, et al. 2018; Sibbald and Archibald 2020). Euglenophytes acquired their plastid from a Pyramimonadales-like ancestor and chlorarachniophytes from Bryopsidales (Jackson, Knoll, et al. 2018). Thus there is now a consensus that the green plastids have been acquired separately in these two groups.

1.2.2 . Red complex plastids

The diversity of species that harbor red complex plastids is huge in comparison to those with green complex plastids. These species are part of four clades: Myzozoa, Ochrophyta (~22 000 described species), Cryptophyta (~200), Haptophyta (~1500) (Guiry and Guiry 2022). Together these lineages are referred to as MOCHa or chromalveolates here (Burki 2017), but they are also commonly known as CASH (A = Alveolates, including Myzozoa; S = Stramenopiles, including Ochrophyta) (Baurain et al. 2010). The exact scenario that resulted in this diversity of red plastid-bearing lineages remains unclear (see 1.3), but the ecological effect it has had, especially in marine environments, is crystal clear (Falkowski et al. 2004). Before discussing the scenarios explaining their origin, below I will outline the four lineages and their diverse evolutionary paths.

Ochrophytes

Of the four groups with red complex plastids, Ochrophyta (also known as Heterokontophyta, Ochrista; part of Stramenopiles) is the most morphologically diverse and species-rich. It includes unicellular algae like diatoms, but also seaweeds displaying complex multicellularity such as kelp (Dorrell and Bowler 2017)—multicellularity does not appear in any of the other red complex plastid-bearing groups. The number of species listed on AlgaeBase is over 22 000 but true number is likely much higher (Guiry 2012; Guiry and Guiry 2022). Their monophyly has long been recognized, being classified as Ochrista in 1986 by Cavalier-Smith, subsequently renamed as Ochrophyta to follow the International Code of Botanical Nomenclature (Cavalier-Smith and Chao 1996).

Diatoms account for the largest number of species in Ochrophyta, with more than 17 000 described but estimates of number of species reaching more than ten times as high (Guiry and Guiry 2022). Freshwater and marine species exist, though most described species are marine. Diatoms are surrounded by a shell of silica, called a frustule, which comes in many shapes and sizes (e.g., fig 1.4F and G). The frustule has been useful in tracking diatoms through geological time (starting at ~200 Mya); their fossil record is probably one of the most complete of all major unicellular eukaryotic groups (Sims et al. 2006). Classification of diatoms was largely based on the morphology of the frustule (Harwood and Nikolaev 1995). In total four groups were considered, the radial and bipolar centrics, and the raphid and araphid pennates. Phylogenomics eventually has shown that at least some of these morphological groups do not constitute monophyletic groups (Medlin, Williams, et al. 1993). The raphid pennates, as well as all pennates, appear as monophyletic by both phylogenomic and morphological evidence (fig. 1.4; Medlin 2016). However, the araphid pennates typically appear as paraphyletic, with multiple grades recovered (Kooistra, Gersonde, et al. 2007). The bipolar centrics have been proposed to be monophyletic with Thalassiosirales (a radial centric group), forming the Mediophyceae, and the remaining radial centrics being placed in the Coscinodiscophyceae (Medlin

and Kaczmarska 2004). However, other studies find a grade of radial centrics at the base of the diatoms making it still an open question how the morphological groups are related to each other (Yu et al. 2018). Finally, a lineage of mostly silica-covered (but some naked: fig. 1.4E) algae has been recovered to be the sister to all diatoms, the Bolidiophyceae (Guillou et al. 1999), which includes the genera *Triparma*, *Pentalamina*, and *Tetraparma* (Yamada, Sato, et al. 2020).

Two more algal lineages are grouped together with the diatoms and bolidophytes within the Diatomista, these are Pelagophyceae and Dictyochophyceae (fig. 1.4; Derelle et al. 2016). Representatives of Pelagophyceae have just a single flagellum and may be united morphologically by a "perforated theca" (Moestrup 2021; Wetherbee, Bringloe, et al. 2021). Dictyochophytes (silicoflagellates) are flagellate and amoeboid algae found in both marine and freshwater ecosystems and include a few cases of loss of photosynthesis (Yoon, Andersen, et al. 2009). Similar to diatoms, they make use of silica for their cell body.

The remaining algal groups are currently placed within the Chrysista (Cavalier-Smith, 1986[2]) and consist of several classes (Derelle et al. 2016). One lineage consists of solely unicellular algae known as the PESC clade (or SII clade; Yang, Boo, et al. 2012), which includes multiple described classes (fig. 1.4; Dorrell and Bowler 2017). Pinguiophyceae are named after their high fatty acid content—*pingue* meaning fat or grease in Latin (Kawachi, Inouye, et al. 2002). Eustigmatophyceae includes only unicellular algae once considered part of the Xantophyceae (see next paragraph) and contains ~100 described species (Hibberd and Leedale 1971; Guiry and Guiry 2022). Synchromophyceae is based on a new genus and species *Synchroma grande*, but was shown to also include the earlier described *Chlamydomyxa* (Horn et al. 2007; Patil et al. 2009). Moreover, they are potentially closely related to plastid-lacking genera *Leukarachnion* and *Picophagus*, suggesting putative plastid (genome) losses within Ochrophyta (Schmidt et al. 2012). Chrysophyceae has a much longer history, with first species descriptions dating back to 18th century and likely includes algae sometimes placed in Synurophyceae (such as *Mallomonas* and *Synura*) (Nicholls and Wujek 2015). Further study of the PESC clade is needed to understand their plastid evolution—including putative plastid and photosynthesis loss—and how the members relate to each other and other ochrophyte lineages (Dorrell and Bowler 2017; Franco et al. 2021).

The other group of the Chrysista is often referred to as the SI-clade and consists primarily of the Raphidophyceae, Xanthophyceae and Phaeophyceae which are typically found monophyletic (fig. 1.4) although some understudied groups are part of this group as well (Dorrell and Bowler 2017). The Phaeophyceae are better known as the brown algae, and are a large group of multicellular algae including kelps such as *Macrocystis* (fig. 1.4B). Xantophyceae is a highly diverse group of mostly freshwater species, including flagellate, amoeboid, filamentous and siphonous species (Maistro et al. 2009). Raphidophyceae are all flagellated unicellular species found in both marine and freshwater ecosystems, only a few genera are known but overall they are widespread and can cause damaging blooms (Yamaguchi et al. 2010). Multiple other groups are present and some have been described recently, reflecting especially how understudied the unicellular representatives are. One fascinating example is the "golden paradox", *Chrysoparadoxa australica*, which is a marine species placed in its

[2]I was unable to access the original work, of which the correction citation would be: Cavalier-Smith, T. (1986). The kingdom Chromista: origin and systematics. *Progress in phycological research, 4*, 309-347.

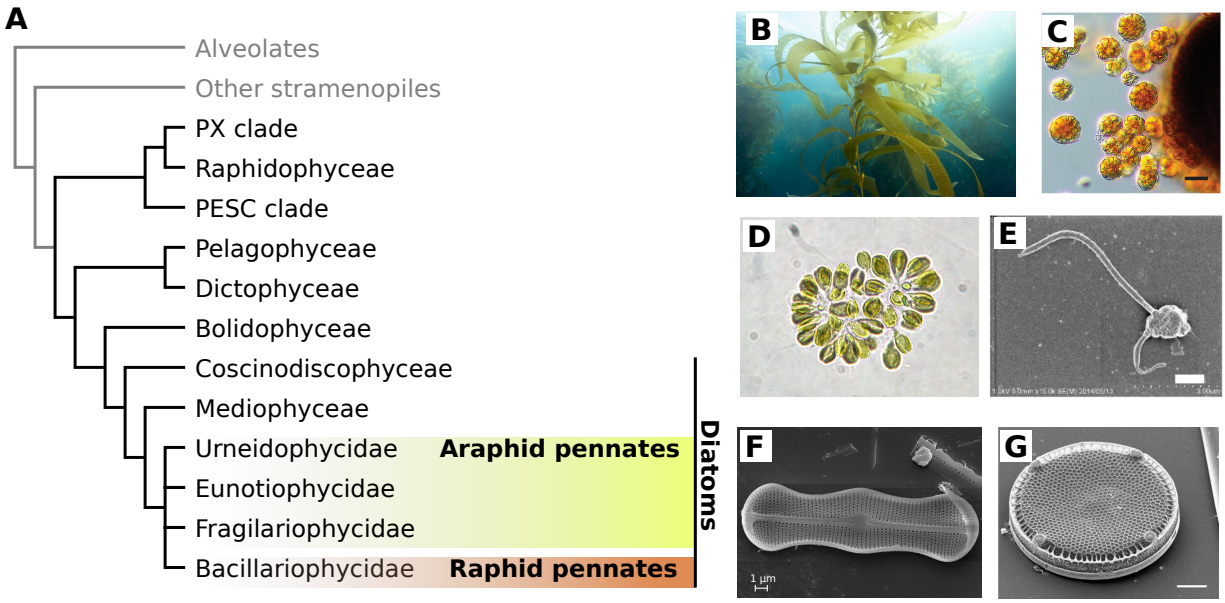


Figure 1.4: (A) The phylogeny of ochrophytes shown with their close related sister lineages (where other Stramenopiles refer to a paraphyletic grouping). (B) *Macrocystis pyrifera* (giant kelp, Phaeophyceae, PX-clade). (C) Vegatative colonies of *Gazia saundersii* (Pelagophyceae), showing daughter colonies released from a larger mother colony (Wetherbee, Bringloe, et al. 2021). Scale bar equals 20 µm. (D) Micrograph of *Synura uvella* (PESC clade) by Kristian Peters. (E) Scanning electron microscopy (SEM) image of *Triparma eleuthera* (Bolidophyceae) from Kuwata et al. 2018. (F) SEM image of *Achnanthes trinodi* by Nocila Angeli / MUSE Science Museum. (G) SEM image of *Eupodiscus radiatus* (Bradbury 2004). B is in the public domain, rights were received for C, and D-G are under a creative commons licenses (see Supplementary material).

own class (Chrysoparadoxophyceae) and is—to my knowledge—the only convincing case known of plastid membrane loss (Wetherbee, Jackson, et al. 2019; see 1.4.4).

Finally, the placement of Ochrophyta within the tree of eukaryotes is relatively well established in comparison to Haptophyta and Cryptophyta. Their relatedness to the oomycetes (pseudofungi) has been well supported as is their overall placement within the Stramenopiles (Cavalier-Smith 1998; Derelle et al. 2016). Moreover, the plastid-lacking *Actinophrys sol* has recently been shown to be the closest relative to ochrophytes and may provide more context about their origin (Azuma et al. 2022). The placement of Stramenopiles within the well supported SAR clade (also known as Harosa) places them clearly closer to Alveolata, and Rhizaria—which contain no red algal-derived plastids (Cavalier-Smith 2010; Burki, Roger, et al. 2020). Some lineages of alveolates contain red complex plastid-bearing lineages, but with a much richer history of potential separate plastid acquisitions, replacements, losses and reductions (section 1.2.2).

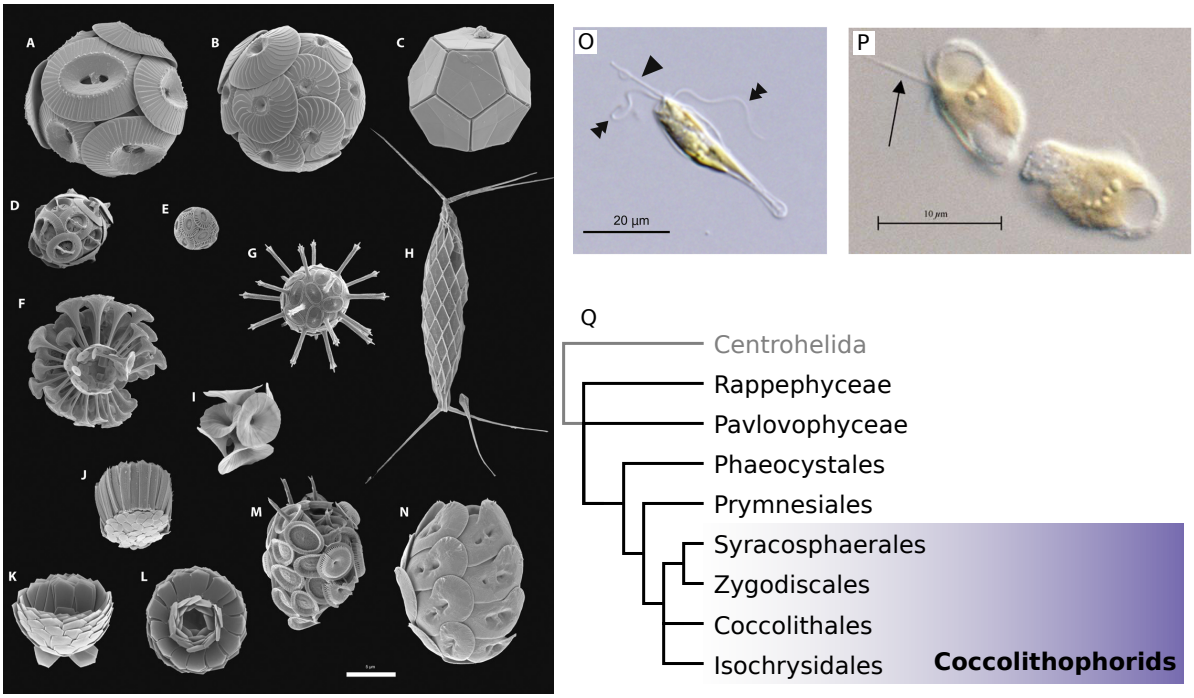


Figure 1.5: The phylogeny of haptophytes and morphology of several representatives. The diversity of coccolithophorids is shown on the left, taken from Monteiro et al. 2016 including the relevant caption below. (A to N) Scanning electron micrographs of cells collected by seawater filtration from the open ocean. Species illustrated: (A) *Coccolithus pelagicus*, (B) *Calcidiscus leptoporus*, (C) *Braarudosphaera bigelowii*, (D) *Gephyrocapsa oceanica*, (E) *E. huxleyi*, (F) *Discosphaera tubifera*, (G) *Rhabdosphaera clavigera*, (H) *Calciosolenia murrayi*, (I) U*mbellosphaera irregularis*, (J) *Gladiolithus flabellatus*, (K and L) *Florisphaera profunda*, (M) *Syracosphaera pulchra*, and (N) *Helicosphaera carteri*. Scale bar, 5 µm. (O) Micrograph of *Pavlomulina ranunculiformis* (NIES-3900; Rappephyceae) bearing a haptonema (arrowhead) between two flagella (double arrowheads) (Kawachi, Nakayama, et al. 2021). (P) Micrograph of *Prymnesium parvum* (UTEX 2797) with its haptonema (arrow) (Manning and La Claire 2010) (Q) Cladogram displaying the diversity of Haptophyta (black) and its sister lineage (gray). Images A-P are under creative commons licenses (see Supplementary material).

Haptophytes

The Haptophytes are a group of unicellular algae mostly known for the calcified scale-bearing coccolithophores such as *Emiliania huxleyi*, but also include other unicellular flagellated algae. Other haptophytes typically carry organic (non-mineralized) scales (Tsuji and Yoshida 2017). The main feature of haptophytes is the haptonema, an organelle located near the flagella that may resemble a flagellum but is distinct from it. The haptonema length varies between species and life cycle phase, and may be only residual (Jordan 2009). The total group includes roughly 1500 described species on Algaebase as of June 2022 (Guiry and Guiry 2022). Haptophytes are especially abundant on the oceans, where they occupy 30-50% of the chlorophyll *a* biomass, but are also present in freshwater ecosystems (Liu, Probert, et al. 2009; Shalchian-Tabrizi, Reier-Røberg, et al. 2011; Simon et al. 2013). Below, I will give an outline of the main haptophyte groups (fig. 1.5), but note that haptophyte diversity is probably highly understudied (Gran-Stadniczeñko et al. 2017).

Coccolithophores bear calcareous scales (coccoliths) that are highly diverse in morphology (fig.

1.5). Typically, multiple coccoliths cover the cell although in certain species only one coccolith is present, or the coccolithosphere surrounds more than one cell (Jordan 2009). Moreover, some highly modified coccoliths can in some species be connected resulting in appendages sticking out from the coccolithosphere (Young et al. 2009). Multiple functions have been ascribed to the coccoliths, including reflecting light to either increase or decrease the amount of light entering the cell, defense from predators and aiding in capturing food (Jordan 2009). Coccolithophores are part of the family Prymnesiophyceae (or Coccolithophyceae), which includes a few coccolith-lacking species. The calcereous species are typically classified into four different orders: Syracosphaerales, Zygodiscales, Coccolithales and Isochrysidales (Hagino and Young 2015). The relationships between these groups remain to be resolved, and some lineages are known with no clear placement within these groups. One example of a unplaced group is Braarudosphaeraceae, which typically contains coccoliths made of five segments, which are arranged as a regular dodecahedron in *Braarudosphaera* (fig. 1.5C). Finally, few plastid genomes are available of coccolithophores, with just a single representative outside of the Isochrysidales (Paudel et al. 2021). This may be part of the reason why the relationships among them remain unresolved, as only phylogenies with one or a few genes have been made, for example using 18S rRNA gene sequences (Edvardsen et al. 2016).

Besides the coccolithophores, Prymnesiophyceae also includes two non-calcifying orders, Phaeocystales and Prymnesiales. Prymnesiales consists of multiple genera, the most widely spread and well studied being *Chrysochromulina* and *Prymnesium*. At least two Prymnesiales species have mineralized scales (*Hyaolithus neolepsis* and *Prymnesium parvum*), bearing siliceous instead of calcareous scales (Tsuji and Yoshida 2017). Phaeocystales consists of a single genus, *Phaeocystis* (Eikrem et al. 2017; Guiry and Guiry 2022). However, species of this genus are highly ecologically relevant, as they are cosmopolitan and some commonly form blooms which may cause harmful effects to fish and other animals (Eikrem et al. 2017).

The two other groups composing Haptophyta are Rappephyceae (i.e., rappemonads) and Pavlovphyceae. Rappemonads have been known for some time from metabarcoding studies, being named after the first author of the study reporting the first 16S rDNA plastid sequence of this group (Rappé et al. 1998; Kim, Harrison, et al. 2011). From those metabarcoding studies, the rappemonads appeared as sister to Haptophyta (Kim, Harrison, et al. 2011). Recently, a photosynthetic species closely related to environmental rappemonads sequences—*Pavlomulina ranunculiformis* (fig. 1.5O)—was cultured and its plastid genome sequence was used to reconstruct the phylogeny of haptophytes, suggesting instead a sister relationship to the Prymnesiophyceae (Kawachi, Nakayama, et al. 2021). Pavlovphyceae contains just four genera and is sister to all other haptophytes (Green 1976; Fujiwara et al. 2001; Tsuji and Yoshida 2017). The placement of haptophytes within the tree of eukaryotes has remained uncertain, except for their probable sister relation to Centrohelida, a understudied group with most representatives also bearing organic or siliceous scales (Burki, Kaplan, et al. 2016; Burki, Roger, et al. 2020).

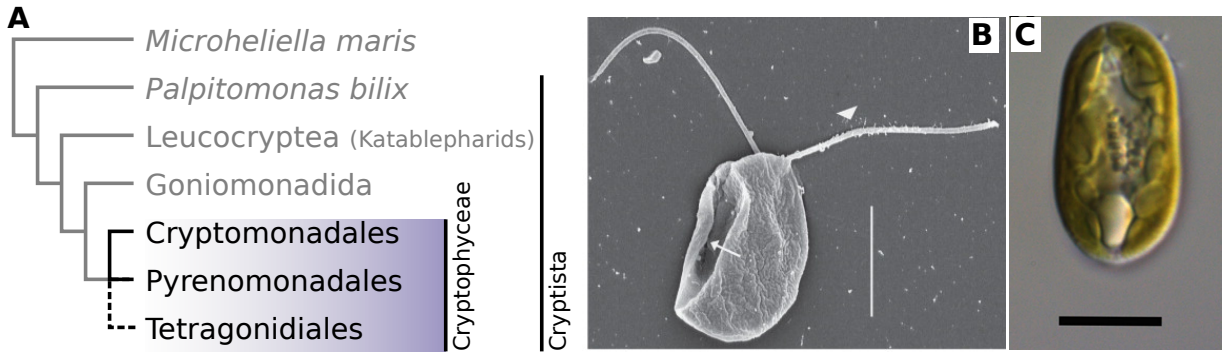


Figure 1.6: (A) The phylogeny of Cryptophyceae and closely related groups of the Pancryptista. (B) Scanning electron microscopy image of *Goniomonas avonlea* (Goniomonadida), showing the furrow (arrow) and hairs on the dorsal flagellum (arrowhead) (Kim and Archibald 2013). Scale bar equals 5 µm. (C) Micrograph of *Cryptomonas uralensis* (Cryptomonadales), ventral view, scale bar equals 10 µm (Martynenko et al. 2020). Rights were obtained to reuse B and C from their original source (see Supplementary material).

Cryptophytes

The cryptophytes are unicellular algae, of which only a few genera are known and constitute a single family: Cryptophyceae. The term cryptophytes is also used to refer to Cryptophyta as a whole, which include plastid-lacking goniomonads (Goniomonadea, or Cyathomonadacea) (fig. 1.6; Adl, Bass, et al. 2019). For clarity I will use the term cryptomonads to refer to Cryptophyta as a whole, and cryptophytes to refer to the plastid-bearing lineages (i.e., Cryptophyceae). Cryptomonads have an asymmetric cell shape, with two flagella of unequal length which are near an invagination in the cell (Archibald 2020). The cells are covered in what is known as the periplast, which is a proteinaceous layer that in some species sits both on the inside and outside of the plasma membrane (Tanifuji and Onodera 2017). Moreover, ejectosomes appear underneath the inner periplast component (Archibald 2020). Cryptomonads are known from marine, brackish and freshwater habitats (Novarino and Lucas 1993).

Cryptophytes contain one or two plastids per cell, which have chlorophyll *a* and *c*, but the most exciting feature is the nucleomorph: a reduced nucleus, from the red algal endosymbiont, within the plastid of cryptophytes (Tanifuji and Onodera 2017). Nucleomorphs are not present in any ochrophytes and haptophytes, and are otherwise known only from the green complex plastids of chlorarachniophytes (see 1.2.1) and some lineage-specific plastid acquisitions in dinoflagellates (see 1.2.2). Moreover, several cryptophtye lineages of the genus *Cryptomonas* have lost photosynthesis but typically kept their plastid (Hoef-Emden and Melkonian 2003). The other representatives of Cryptophyta are the Goniomonadea, which are plastid-lacking lineages that share many of the morphological features of cryptophytes (including ejectosomes and periplast plates) (Tanifuji and Onodera 2017).

Katablepharids were hypothesized to be sister of the cryptophytes based on morphological similarities, including the presence of the ejectosome (Okamoto and Inouye 2005). In 2005, the katablepharids (Leucocryptea) were shown to be sister to the cryptomonads using molecular data (Okamoto and Inouye 2005). Furthermore, several species have been confidently placed as sister lineages to

cryptomonads plus katablepharids. *Palpitomonas bilix* was shown to be sister and differs mainly from the lack of ejectosomes (Yabuki et al. 2014). The protist *Microheliela maris* was found to be sister to the Cryptista (Cavalier-Smith, Chao, and Lewis 2015), and the name Pancryptista was suggested to refer to *Microheliela* + Cryptista (Yazaki et al. 2022). The sister group of Pancryptista is unclear; they have been suggested to be sister to Haptista (Hacrobia), but in other analysis are recovered as sister to the Archaeplastida (see more in 1.4.5)

Myzozoans

Myzozoa (before Miozoa) is a broad group of unicellular eukaryotes, united by phylogenetic analyses (Cavalier-Smith and Chao 2004). They are part of Alveolata, with the other included group being the ciliates, a diverse group with some having symbiotic relationships although never with a true photosynthetic organelle (Palmer, Soltis, et al. 2004). Representatives of Myzozoa are highly diverse in the different lifestyles they display, including mixotrophs, mutualistic symbionts and parasites. Although many representatives contain plastids, they are typically heterotrophic or mixotrophic.

Dinoflagellates are fascinating unicellular eukaryotes, with many species harboring a three membrane-bound plastid, but also many not being photosynthetic anymore (Waller and Kořený 2017). Perkinsozoa (including *Perkinsus*) are sister to all dinoflagellates, considered as part of the larger Dinozoa group, and contains non-photosynthetic plastids similar to *Oxyrrhis*, which is sister to all other dinoflagellates (Saldarriaga et al. 2003; Waller and Kořený 2017). Dinoflagellates not only exhibit colorless plastids, but also at least one plastid loss in the parasite *Hematodinium* (Syndiniales) (Gornik et al. 2015). The ancestral dinoflagellate plastid is the peridinin plastid (as found in the "peridinin dinoflagellates"; fig. 1.7D), with a genome consisting of many minicircles carrying one or a few genes (and sometimes none) (Koumandou et al. 2004). But within the peridinin dinoflagellates, the plastid has been replaced multiple times by plastids of different origins. *Lepidodinium* and a few undescribed dinoflagellates have plastids of green algal origin (see also 1.2.1; Sarai et al. 2020). Moreover, plastids have also been acquired from haptophytes and diatoms (resulting in so-called "dinotoms"). Haptophyte plastids are found in certain members of Kareniaceae (fig. 1.7E), where the plastid membrane transport machinery of the haptophyte is used, even though a haptophyte nucleomorph is absent (Waller and Kořený 2017). Diatom-derived plastids are found in a few representatives of Peridiniales (fig. 1.7F), and in contrast to the haptophyte-derived plastids, the host nucleus and mitochondria are still present (Tomas and Cox 1973). This makes them seem more like obligate endosymbionts, although interestingly the symbionts remain in cultures for decades (Waller and Kořený 2017). Finally, it is important to note that kleptoplastidy (the theft of plastids from algae), is common in many dinoflagellate lineages (such as *Dinophysis*; fig. 1.7G) which may be a way towards plastid replacement (Bodył 2018). All of the above clearly makes the dinoflagellates an exciting group to study to understand how endosymbiosis, organellogenesis and plastid replacement, gain and loss can occur (Waller and Kořený 2017).

The chrompodellids consist of photosynthetic and heterotrophic unicellular eukaryotes. Two described species that belong to the chrompodellids are photosynthetic: *Chromera velia* and *Vitrella brassicaformis* (fig. 1.7C), with four-membrane bound plastids (Moore et al. 2008; Oborník et al.

2012). Perhaps more interesting is that they are not close relatives to each other (Janouškovec, Tikhonenkov, et al. 2015; Füssy and Oborník 2017). Importantly, non-photosynthetic colpodellids seem to still have a plastid compartment, as plastid-specific biosynthesis pathways are detected in the transcriptomes of several species (Janouškovec, Tikhonenkov, et al. 2015). This suggests that the plastid was present in the common ancestor of chrompodellids, and more likely in a further ancestor as they are sister to the plastid-bearing apicomplexans (see below).

The apicomplexans (fig. 1.7B) contain no photosynthetic members, but almost all of them bear a non-photosynthetic plastid (apicoplast) with its own, reduced genome (van Dooren and Hapuarachchi 2017). Similarly to the chromerids, this plastid is surrounded by four membranes. They are mostly known for several human parasites, including the malaria-causing pathogen *Plasmodium*. The apicoplast remains, in part due to the likely need of isoprenoid and heme biosynthesis during at least one of its life stages, where the needed metabolites are not acquired their host (Janouškovec, Tikhonenkov, et al. 2015). But, one case is known of complete plastid loss in *Cryptosporidium* (Abrahamsen et al. 2004). Plastid evolution is clearly nowhere as dynamic as in Myzozoa, and it may indicate how plastids may have evolved in other lineages.

1.3 . Hypotheses on the origins of red complex plastids

Since the first idea that some algal groups have plastids derived from red algae, different hypotheses have aimed to explain their origin. At this time, two main hypotheses have gained much attention, these are the chromalveolate hypothesis and the serial endosymbiosis hypothesis (of the latter are multiple variants). In this section, I will explain how these different hypotheses came to be, on what assumptions they are based, and the historical context is given in regards to the available evidence at the time. In the next section (1.4), I will discuss what we know about red complex plastids and their hosts, and how this information relates to the chromalveolate and serial endosymbiosis hypothesis.

1.3.1 . The chromalveolate hypothesis

Multiple hypotheses have been proposed to explain the diversity of plastid-bearing lineages as a whole, with sometimes involving many (more than 20) separate acquisitions (Cavalier-Smith 1999). However, many of these lineages have shown to be closely related, making it more parsimonious to introduce a single acquisition for multiple plastid-bearing lineages. Cavalier-Smith published multiple proposals on the origin of many plastid-bearing lineages, including one with just two endosymbiotic events, with one of these events uniting all the archaeplastidal lineages—although together with the Dinozoa and Euglenozoa (Cavalier-Smith 1982). The Dinozoa, in this proposal, were the plastid donor to the already united chromists (cryptophytes, haptophytes, stramenopiles; Cavalier-Smith 1981). Later, he proposed a hypothesis that would unite the red-complex plastid bearing lineages into a single clade, with a single red algal endosymbiosis in their common ancestor: the chromalveolate hypothesis (Cavalier-Smith 1999). This hypothesis has stand the test of time, and although under heavy scrutiny (e.g., Baurain et al. 2010) it remains a dominant, and elegant solution to the origin of red complex plastids (Dorrell and Smith 2011). At the time of the proposal, it was already clear that

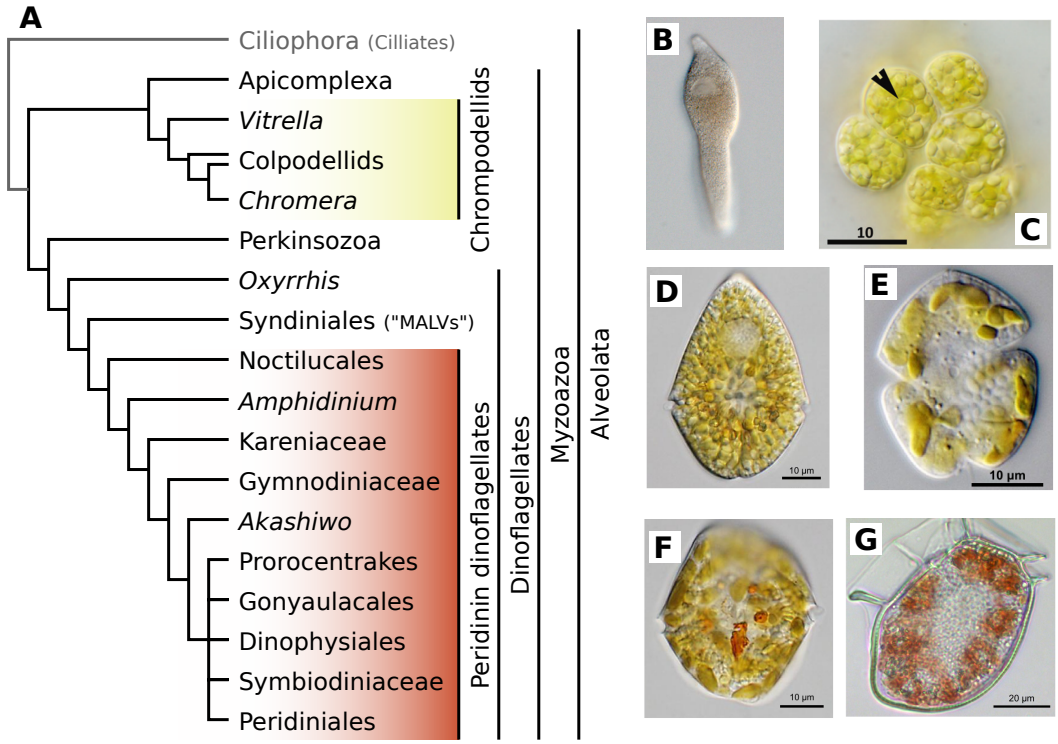


Figure 1.7: (A) Cladogram displaying relationships among Myzozoa and their sister group. MALVs = "Marine alveolates", several of which have been assigned to Syndiniales. (B) Micrograph of *Lankesteria cystodytae* (Apicomplexa) in its trophzoite stage by Sonja I. Rueckert. (C) Micrograph of *Vitrella brassicaformis*, showing a colony of small-sized young vegetative cells with a centrally located pyrenoid (arrow) (Oborník et al. 2012). (D) Micrograph of *Gymnodinium* cf. *placidum* containing perdinin plastids. (E) Micrograph of *Karenia mikimotoi* (Kareniaceae) containing haptophyte-derived plastids. (F) Micrograph of the dinotom *Kryptoperidinium foliaceum* (Perdiniales) containing diatom-derived plastids (G) Micrograph of *Dinophysis acuta* containing kleptoplasts of cryptophyte-origin. D-F from Waller and Kořený 2017, micrographs are by Gert Hansen (D-F) and Susanne Busch (G). B is under a creative commons license and rights were retrieved to reuse C-G in this thesis (see Supplementary material).

multiple lineages would most likely have lost the plastid (ciliates and heterotrophic stramenopiles). If the plastid was not yet fully integrated at this point, such plastid losses should be possible (Cavalier-Smith, Chao, and Lewis 2015).

Since the original proposal of the chromalveolate hypothesis, the eukaryotic tree of life has seen major updates. Stramenopiles, alveolates and rhizaria were placed in a single group (SAR, or Harosa), on one hand uniting the stramenopiles and alveolates as sisters but also placing them with the red complex plastid-lacking Rhizaria (Cavalier-Smith 2010; Burki, Roger, et al. 2020). Similarly, haptophytes have now been placed confidently as sister to the heterotrophic centrohelids, and cryptophytes to several heterotrophic groups of the Cryptista (Burki, Roger, et al. 2020). Thus many losses must have occurred after this endosymbiotic event (e.g., Burki 2017), and likely this would suggest that the plastid was not well integrated from the red algal acquisition up to the diversification of these lineages. Moreover, Cryptista as a whole has been proposed to share a common ancestor with Archaeplastida, which would make the chromalveolate hypothesis impossible (unless red algae invented a time machine) (Burki, Kaplan, et al. 2016). Several biosynthetic pathways (e.g., isoprenoid, heme, araomatic amino acids) have plastid and cytoplasmic versions, and typically only one of them is found in plastid-bearing lineages (Gould, Waller, et al. 2008). That both pathways would be kept for such a long time seems unlikely, especially as different myzozoan lineages carry different versions of the tetrapyrrole (required for heme biosynthesis) and fatty acid biosynthetic pathways (Waller, Gornik, et al. 2016). The once highly parsimonious chromalveolate hypothesis became less attractive, and alternative hypotheses started to come up, especially those involving endosymbioses between chromalveolate lineages.

1.3.2 . Serial endosymbiosis

As phylogenetic analysis of plastid data recover a monophyletic red complex plastid group, and all but dinoflagellates use a similar protein transport machinery (SELMA), a single red algal origin of red complex plastids remained the most parsimonious (but see also Burki, Shalchian-Tabrizi, et al. 2007). Moreover, there are more definite plastid gains than there are definite plastid losses (see 1.4.3; Burki 2017) Several authors have instead favored scenarios where the red complex plastid was acquired once (in cryptophytes), and then horizontally transferred between different chromalveolate lineages. As the myzozoan plastid origin is rather complicated (Waller and Kořený 2017), I will mainly discuss the proposed scenarios in regard to what they say about the origin of chromist algae (cryptophytes, haptophytes and ochrophytes).

A first—to my knowledge—detailed scenario of plastid transfer between chromalveolates was published by Sanchez-Puerta and Delwiche in 2008. As Cryptista and Haptista were often found closely related, they suggested a single endosymbiotic event in their common ancestor, followed by plastid losses in the heterotrophic sister lineages of cryptophytes and haptophytes. In this scenario, ochrophytes gained the plastid either after the divergence of the two groups from the haptophyte lineage (A), or before their divergence (B). In case A, the *rpl36* gene would have been lost separately in cryptophytes and haptophytes (Rice and Palmer 2006). On the other hand, in scenario B the nucleomorph would have to have been lost twice in ochrophytes and haptophytes, although one could imagine that kleptoplastidy resulted in a nucleomorph-lacking ochrophyte as seen in the

nucleomorph-lacking cryptophyte-derived kleptoplastid in *Dinophysis* spp. (Lucas and Vesk 1990; Hackett, Maranda, et al. 2003; Gagat et al. 2014). A similar hypothesis to scenario B was proposed by Dorrell and Smith 2011, with the main difference that a green algal endosymbiosis was added and the base of the Chromalveolata to account for the green algal signal found in many chromalveolates (more details about mixed algal signals in section 1.4.3).

Bodył et al. 2009 proposed a hypothesis where the acquisition of the plastid started in cryptophytes, and was then first transferred to ochrophytes, and later (also from cryptophytes) to haptophytes. Similarly to scenario B from Sanchez-Puerta and Delwiche, the loss of *rpl36* occurs only once at the expense of introducing two putative nucleomorph losses in ochrophytes and haptophytes. An advantage of this hypothesis is that it corresponds well to the putative placement of cryptophytes distantly from haptophytes, potentially sister to or nested in Archaeplastida (Burki, Kaplan, et al. 2016).

Stiller, Schreiber, et al. 2014 aimed to use a statistical approach to detect endosymbiotic gene transfer (EGT) between different lineages, specifically between red complex plastid-bearing lineages. They collected 50 nuclear genomes of different eukaryotes representing the different eukaryotic supergroups. Nuclear encoded proteins of a representative of Ochrophyta (*Phaeodactylum tricornutum*), Cryptophyta (*Guillardia theta*) and Haptophyta (*Emiliania huxleyi*) were aligned using blast against the whole database (minus its own genome). In summary, they detected stronger signal between certain groups than would be expected from phylogenetic relationships and database size. Specifically, EGT was detected between *E. huxleyi* and stramenopiles, and between *G. theta* and stramenopiles (when using *P. tricornutum*, the biggest outliers were cryptophytes and haptophytes, but the signal was not significant). They interpret this as a plastid transfer from cryptophytes to ochrophytes, and from ochrophytes to haptophytes. Others have interpreted similar results as a more recent endosymbiosis, unrelated to acquisition of the currently present plastid (1.4.3; Dorrell, Gile, et al. 2017). Moreover, this hypothesis implies two independent gene replacements of *rpl36* in cryptophytes and haptophytes (Rice and Palmer 2006). In summary, there are multiple serial endosymbiosis hypotheses with differences in assumptions on the likelihood of nucleomorph loss (or not acquired via kleptoplastidy) and gene replacement of *rpl36*. These discrepancies between them and the chromalveolate hypothesis will be discussed in more detail in the next section.

1.4 . Traces of endosymbiosis

1.4.1 . Chlorophyll *c*

Photosynthetic life uses a combination of different pigments for the absorption of light. In the case of red complex plastids, the most interesting pigment is chlorophyll *c*, which is unique to these lineages (excluding non-photosynthetic lineages such as apicomplexans, but also chromerids and eustigmatophytes; Oborník et al. 2012; Larkum 2020). This may suggest that a single origin is the most parsimonious, but there are multiple types of chlorophyll *c* (Jeffrey, Wright, and Zapata 2011; Larkum 2020). Here, I will give a short overview of these types of chlorophyll *c* and their presence in the different red complex lineages, and I will discuss the potential biosynthetic pathway of chlorophyll

c.

In total nine different chlorophyll c pigments have been characterized, the main types being c_1, c_2 and c_3 (Jeffrey and Wright 2005). Some chlorophyll *c* pigments have only been found in haptophytes and vary between different haptophyte clades (Zapata et al. 2004). Unlike chlorophyll *a* and *b*, the biosynthetic pathway to chlorophyll *c* is not well understood (Larkum 2020). Multiple suggestions have been done for possible biosynthetic precursors, but the specific enzymatic steps remain unclear. A pathway has been shown to work starting from chlorophyll *a* but it is unclear what exact pathway is actually used in algae (Xu et al. 2016). One of the most likely precursors seems to be MgDVP (3,8-divinyl-protochlorophyllide), from which only a few biosynthetic steps may be needed (Qiu et al. 2019).

Intriguingly, MgDVP is present in most algae, including in red algae and almost all red complex plastid-bearing lineages (Jeffrey, Wright, and Zapata 2011). The distribution of chlorophyll *c* is the most straightforward in cryptophytes and peridinin dinoflagellates, both having only chlorophyll c_2 and MgDVP (Jeffrey, Wright, and Zapata 2011). As stated before, haptophytes contain many different types, with the only types nearly always present being MgDVP and chlorophyll c_2 (Zapata et al. 2004). Similarly, ochrophytes have different types of chlorophyll *c*, but not a single one is present in all main lineages (Jeffrey, Wright, and Zapata 2011). As the biosynthetic pathway for chlorophyll *c* remains unclear, and the distribution among red complex plastid-bearing algae is complex, it is difficult to infer the value of this character in relation to the origin of red complex plastids. However, the simple biosynthetic pathway and patchy distribution of chlorophyll *c* types allow for the consideration of horizontal transfer (via EGT or HGT), or separate origins (Larkum 2020).

1.4.2 . Evidence from single proteins

Besides potential genes related to chlorophyll *c* biosynthesis, there are a few other genes of which the evolutionary history might support one hypothesis over another. Specifically, multiple genes involved in the Calvin cycle play such a role, which has been especially extensively discussed in case of the plastid-targeted d-glyceraldehyde-3-phosphate dehydrogenase (GAPDH) and fructose-1,6-bisphosphate aldolase (FBA) (Sanchez-Puerta and Delwiche 2008). There are cytosolic and plastid-targeted versions of these genes in Archaeplastida and red complex plastids. The plastid-targeted version of GAPDH of archaeplastidal lineages is derived from cyanobacteria, but in chromalveolates its most likely origin is through a duplication of the cytosolic version (Fast et al. 2001; Harper and Keeling 2003). A type of plastid-targeted FBA is used in chromalveolates, completely different from those in red algae and other archaeplastidal lineages, similarly to GAPDH this is likely caused by a duplication of the cytosolic FBA in chromalveolates (Patron, Rogers, et al. 2004). Similar evolutionary histories seem have occurred for phosphoribulokinase, sedoheptulose-1,7-bisphosphatase and fructose-1,6-bisphosphatase (Teich et al. 2007). Of course, in the serial endosymbiosis framework, these can be accepted as endosymbiotic gene transfers after a duplication in one of the lineages (Bodył 2005), or as more recent horizontal gene transfers, which seems to be the case for GAPDH (Takishita et al. 2009; Dorrell and Smith 2011).

Within the plastid genomes of cryptophytes and haptophytes, a gene replacement of the *rpl36*

gene occurred, which still resides within the same genetic context, located between *secY* and *rps13* (Rice and Palmer 2006). It is unclear why this gene was replaced, but this is unlikely explained by two separate events, suggesting the extant cryptophytes and haptophytes are monophyletic from a plastid perspective. For the chromalveolate hypothesis, this would suggest that these two lineages share a more recent common ancestor than either does with ochrophytes. For the serial endosymbiosis hypothesis, this would suggest that the ochrophyte plastid was gained relatively early, before the *rpl36* gene replacement.

1.4.3 . Plastid gain and loss

One of the main differences between the chromalveolate and serial endosymbiosis hypotheses is their assumption on how easy it is to gain or lose a plastid. Luckily, there are examples of both that are highly interesting in illuminating how either can happen, i.e., in which kind of scenario will a plastid be integrated or rather lost. To start, there are more cases of indisputable plastid gains. Two examples are the green complex plastids in euglenophytes and chlorarachniophytes, which are certain independent cases, specifically because we know that the respective green algal endosymbionts are different (see 1.2.1; Jackson, Knoll, et al. 2018; Sibbald and Archibald 2020). Besides this, the chromalveolate hypothesis would still pose at least a single separate event, ignoring any primary plastid origins (at least two cases: Archaeplastida and *Paulinella*). But when it comes to plastid gains, the dinoflagellates take the crown (Waller and Kořený 2017). Likely three independent green plastid acquisitions occurred in dinoflagellates, and a highly integrated haptophyte-derived plastid is present in members of Kareniaceae (Waller and Kořený 2017). To this we may be able to add the diatom- and cryptophyte-derived plastids, although they may not be fully integrated, thus it is unsure if these are true plastids (Park et al. 2006; Yamada, Bolton, et al. 2019). However, I would argue that dinoflagellate plastid replacements are not completely analogous to other plastid gains, as the machinery to maintain a plastid was already present ancestrally in these dinoflagellates. In combination with the mixotrophic lifestyle of dinoflagellates and a highly reduced genome in the ancestral peridinin plastid, it may be less surprising that this group has so many plastid replacements (Waller and Kořený 2017). But, even if disregarding the dinoflagellates, we still end up with a minimum of three complex plastid origins.

Then, what can we say about loss of complex plastids? First of all, loss of photosynthesis is highly common, and occurs in nearly every algal lineage, including cryptophytes, ochrophytes and myzozoans (Kamikawa, Yubuki, et al. 2015; Waller and Kořený 2017; Suzuki et al. 2022)—although I am not aware of any losses in haptophytes. Many of these cases are due to a lifestyle change to parasitism, but not all. Colpodellids have likely reverted back to an obligate heterotrophic lifestyle twice (as *Chromera* and *Vitrella* are photosynthetic), as have multiple diatoms of the genus *Nitzschia* (Janouškovec, Tikhonenkov, et al. 2015; Kamikawa, Yubuki, et al. 2015). However, highly probable cases of plastid loss are rare, with most lineages that gave up on their photosynthetic lifestyle still carrying the plastid around. There are two credible cases of plastid loss that unsurprisingly bring us back to the Myzozoa. The first case is the apicomplexan group *Cryptosporidium*, with no evidence for any plastid genome or nuclear encoded proteins that are targeted to the plastid (Abrahamsen et al. 2004). More recently the dinoflagellate *Hematodinium* sp. was shown to have lost its plas-

tid (Gornik et al. 2015). Why were these parasites able to lose their plastids, unlike most other non-photosynthetic myzozoans? Cytosolic biosynthetic pathways once present were replaced with plastid-associated versions when the endosymbiont was integrated (Gould, Waller, et al. 2008). In these two cases, a way was found to gain the necessary metabolites without using the plastid pathways. Both *Hematodinium* and *Cryptosporidium* use the cytosolic fatty acid biosynthesis pathway, and both lack any isoprenoid biosynthesis pathway, depending instead on their hosts (Gornik et al. 2015; Waller, Gornik, et al. 2016). One of the differences between these two parasites is that *Hematodinium* has kept the cytosolic tetrapyrrole biosynthetic pathway (related to heme biosynthesis), where *Cryptosporidium* can not produce tetrapyrrole at all and depends instead on metabolites of its host (Waller, Gornik, et al. 2016) *Cryptosporidium* also depends on its host for lysine biosynthesis, but *Hematodinium* has instead found a way to use the plastid-specific lysine biosynthesis pathway (via diaminopimelate) without the plastid, having moved the related to its nucleus and none of these genes contain sequences related to plastid-targeting (Gornik et al. 2015; Waller, Gornik, et al. 2016). The rarity of plastid loss is thus mainly related to the dependency on plastid-related metabolisms, some of which have really only been altered to work outside of the plastid once, in *Hematodinium*. This makes it difficult to imagine plastid loss in a non-parasitic lineage.

The chromalveolate hypothesis implies that such losses occurred in many free-living lineages such as ciliates and the heterotrophic stramenopiles that share a common ancestor with ochrophytes. For each of these cases, there are arguably two options: (1) the plastid was not fully integrated and at the least there was no dependency on plastid-type pathways or (2) plastid genes were transferred (similar to *Hematodinium*) and can still be found in these lineages. Potential signal of red algal-related genes was found for both oomycetes and ciliates, but these were later shown to be not convincing as they were either background noise (i.e., found by chance) or as easily explained by horizontal gene transfer (Tyler et al. 2006; Reyes-Prieto, Moustafa, et al. 2008; Elias and Archibald 2009; Stiller, Huang, et al. 2009; Archibald 2012). The katablepharid *Roombia*—closely related to cryptophytes (see 1.2.2)—was also shown to lack significant red algal signal (Burki, Okamoto, et al. 2012). The option of not having a fully-integrated plastid yet seems rather unlikely, as there is little reason to keep multiple pathways around for the same function for long (Waller, Gornik, et al. 2016). Overall, there is no convincing evidence that ciliates, oomycetes or other sister lineages to red complex plastid-bearing contained such plastids as well.

While red signal might be scarce and ambiguous in ciliates and oomycetes, it has been shown to be relatively strong in lineages with green complex plastids (Ponce-Toledo, Moreira, et al. 2018; Ponce-Toledo et al. 2019). Similarly, green signal was detected in red complex plastid-bearing lineages. Studies showed such signal (e.g., Moustafa et al. 2009; Dorrell and Smith 2011), although not without controversy in relation to poor resolution and low taxon sampling (Deschamps and Moreira 2012). However such signal remains consistent even when using more strict methods (Dorrell, Gile, et al. 2017). The red signal in green complex plastid-bearing lineages (euglenophytes and chlorarachniophytes) may be higher than green signal in red complex plastid-bearing lineages (Ponce-Toledo, Moreira, et al. 2018). Due to this signal, it has been proposed that the acquisition of green complex plastids is facilitated by the previous acquisition of red plastids (Ponce-Toledo et al. 2019). A clear

example of this may be the green complex plastids in multiple dinoflagellate lineages (Waller and Kořený 2017; Sarai et al. 2020). However, such signal may also be explained to some extend by the acquisition of genes from prey in the common ancestor of the different complex plastid bearing lineages, akin to the 'shopping bag' and 'you are what you eat' models (Doolittle 1998; Larkum et al. 2007; Bodył et al. 2009). More complex scenarios with multiple endosymbiosis events can not be excluded, and if true this would be incongruent with the assumption of the chromalveolate hypotheses that plastid acquisitions are rare.

1.4.4 . Plastid membranes and transport machinery

One of the first types of evidence that suggested that some plastids were secondary instead of primary was the number of membranes, but it was widely dismissed as speculative (Palmer and Delwiche 1998). Both green and red complex plastids come with either three membranes (euglenophytes and dinoflagellates) or four membranes (chlorarachniophytes, cryptophytes, haptophytes, ochrophytes, chrompodellids+apicomplexans), in contrast to the two membrane-bound primary plastids. The two innermost membranes are thus likely directly related to the two primary plastid membranes. Typically, these two membranes are thought to be related to the double-membrane surrounding cyanobacteria. In summary, the chromalveolate hypothesis suggests that the red complex plastid had four membranes (after integration), with a membrane loss occurring in the dinoflagellates (Cavalier-Smith 1999). In contrast, in the serial endosymbiosis hypothesis membranes were either lost after any tertiary or higher-level endosymbiosis (Gould, Maier, et al. 2015), or depending on the mode of plastid acquisition (for example, kleptoplastidy) a different number of membranes were present without membrane loss having to have occurred (Bodył 2018). To discuss these scenarios, we first need to discuss the transport machinery associated with the membranes of three and four-membrane bound complex plastids.

In four-membrane bound red complex plastids, protein transport over the "second membrane" (counting starting from outside the plastid) is likely done using the Endoplasmic Reticulum Associated Protein Degradation (ERAD) complex of the original red alga (Hempel et al. 2009). In these lineages, multiple copies can be found of subunits of the ERAD complex, those originating from the host and from the red algal symbiont. This complex in red complex plastid-bearing lineages is known as the Symbiont-specific ERAD-like machinery (SELMA). In the four-membrane bound plastids of cryptophytes, haptophytes and ochrophytes, the outermost membrane is fused with the endoplasmic reticulum (ER), allowing transfer over this fused membrane using Sec61 (Bolte et al. 2009; Hempel et al. 2009). In the plastids of myzozoans (excluding dinoflagellates), an extra step (likely vesicle trafficking) is needed for protein transport across the outermost membrane (Gould, Maier, et al. 2015). However, while all four-membrane bound red plastids contain a version of SELMA, the exact components used may differ and their evolutionary history does not suggest a straight-forward, single origin (Ponce-Toledo 2018). Moreover, in dinoflagellates, only three membranes are present, and SELMA is absent. This may suggest that the "second membrane" was lost, with no need any more for SELMA (Gould, Maier, et al. 2015), although one could wonder why this membrane is not lost in other red complex plastid-bearing lineages. Another hypothesis is that the plastid was replaced or independently gained in dinoflagellates (Waller and Kořený 2017).

Membrane loss is thus needed in dinoflagellates if we follow the chromalveolate hypothesis, but it may also be needed in the serial endosymbiosis hypothesis depending on how plastids were acquired (see below paragraph). Moreover, membrane loss likely occurred early on in the chromalveolate hypothesis before the integration of the plastid (Gould, Maier, et al. 2015). Intriguingly, definite membrane loss is exceedingly rare in established plastids. Even though genome loss, and sometimes plastid loss is observed in myzozoans (see 1.4.3), there is no evidence of membrane loss. There is a only one case of membrane loss known, identified in the ochrophyte known as the "golden paradox" *Chrysoparadoxa australica*, where two of the four membranes were lost (Wetherbee, Jackson, et al. 2019). The authors suggest that the outermost membrane has been kept as it is continuous with the ER membrane, and the innermost membrane. The proteins however still have the typical heterokont signal and transit peptides, keeping it unclear which membranes were lost (Wetherbee, Jackson, et al. 2019). Further study on the exact mode of protein transport would be fruitful, but at the least this example shows that membrane loss is not impossible.

In case of the serial endosymbiosis hypothesis, a lack of membranes may be explained differently: kleptoplastidy (Bodył 2018). In cases where only some of the cellular contents are absorbed (like in myzocytosis), expected membranes such as the plasma membrane will be absent. Such a hypothesis for some complex plastids has been suggested for the origin of the dinoflagellate peridinin plastid, for example by Whatley and Whatley (Whatley and Whatley 1981). It is an elegant way to explain higher-level plastids without membrane loss, but is there evidence for this happening in other protists? Some dinoflagellates harbor plastids of haptophyte origin (*Karenia*, *Karlodinium* and *Takayama*; Kareniaceae). The nucleus of the haptophyte is absent, and SELMA subunits (and TIC/TOC) of haptophyte origin are encoded in the nucleus, suggesting the haptophyte plastid is only surrounded by four membranes (Waller and Kořený 2017). Moreover, a closely related kleptoplastic dinoflagellate has been described that takes up haptophyte plastids temporarily, suggesting this as a likely plastid origin in the permanent haptophyte-plastids in other members of Kareniaceae (Hehenberger et al. 2019). However, it remains unclear how many membranes surround the kleptoplastid in this lineage. A best example of membrane loss through kleptoplastidy are Dinophysis sp., where a cryptophyte plastid that is within the cryptophyte four-membrane bound, is instead present as a two-membrane bound (and nucleomorph-lacking) compartment. Myzocytosis- and kleptoplastidy-like behavior are present in lineages related to ochrophytes (e.g., *Pirsonia*) and haptophytes (multiple centrohelids), further supporting this possibility (Bodył 2018).

As seen above, membrane loss may not have to be implied for the gain of red complex plastids as assumed by the serial endosymbiosis hypothesis. Moreover, it is clear that the transport machinery from the endosymbiont (SELMA) may be reused by the host. Thus membrane loss may not be a strong argument to reject the serial endosymbiosis hypothesis.

1.4.5 . Phylogenetics and phylogenomics

Phylogenetic analysis on the origin of haptophyte, cryptophyte and ochrophyte plastids have quickly confirmed their affinity with red algae, although certainty around their monophyly in such trees came with multi-gene trees (Yoon, Hackett, Pinto, et al. 2002). Since then, this monophyly is *almost* always recovered (e.g., Baurain et al. 2010; Janouškovec, Horák, et al. 2010; Dorrell, Gile, et al. 2017; Muñoz-Gómez, Mejía-Franco, et al. 2017), something that both the chromalveolate hypothesis and the serial endosymbiosis hypothesis would imply. The relationships among the three groups differ between these analyses, but to know how that supports the chromalveolate or serial endosymbiosis hypothesis, we first need to look to the host lineage phylogeny.

In contrast to the relatively stable plastid phylogenies, host phylogenies have been difficult to reconstruct (fig. 1.8). Here I will focus on the relevant changes to the tree of eukaryotes over the last two decades. Originally, classifications have followed the grouping of the red-complex plastid bearing lineages (and related lineages) into the Chromalveolata, although its lack of support was recognized (Adl, Simpson, et al. 2005). Cryptophytes and haptophytes were shown to be related to plastid-lacking groups, positioning the two in Cryptista and Haptista respectively (Okamoto and Inouye 2005; Cavalier-Smith, Chao, and Lewis 2015), but there was also support for the unification of these new groups into the Hacrobia (Hackett, Yoon, et al. 2007; Patron, Inagaki, et al. 2007; Okamoto, Chantangsi, et al. 2009). However, this grouping has not stand the test of time. Multiple studies have instead recovered the Cryptista as sister (or within) the Archaeplastida (e.g., Burki, Okamoto, et al. 2012; Burki, Kaplan, et al. 2016; Cenci et al. 2018; Schön et al. 2021; Yazaki et al. 2022), if true, this would make the chromalveolate hypothesis impossible. Stramenopiles and alveolates were also shown to be sister lineages (Hackett, Yoon, et al. 2007), but at the same time studies have shown clear support for the monophyly of these two lineages with Rhizaria, being placed into SAR/Harosa (Burki, Shalchian-Tabrizi, et al. 2007; Cavalier-Smith 2010). This grouping has stood the test of time, and recent studies suggest that Telonemia, a plastid-lacking group, are sister to SAR (Shalchian-Tabrizi, Eikrem, et al. 2006; Strassert, Jamy, et al. 2019). Because of this, potentially unlikely plastid losses (see 1.4.3 for an overview) would be implied by the chromalveolate hypothesis. The discovery of new lineages, and increase in accuracy of phylogenomic methods have kept the tree of eukaryotes updating (Burki, Roger, et al. 2020). For the specific topic of the origin of red complex plastids, confident placement of Cryptista and Haptista is necessary to test hypotheses regarding the relationships between red complex plastids, but we have not reached a consensus yet.

If we follow the chromalveolate hypothesis, it seems clear that ochrophytes and myzozoans should be recovered as monophyletic in plastid phylogenies, to the exclusion of haptophytes and cryptophytes. However, myzozoan plastids have some difficulties: dinoflagellates and apicomplexans both have fast evolving genes, and gene losses (especially in apicomplexans) and replacements (especially in dinoflagellates). The two sequenced chromerids harbor more typical plastid genomes, and they have been used to place them in plastid phylogenies. They seem to either place sister to ochrophytes or within them, although their branches are long and placement and support highly depend on the lineages included and method used (Janouškovec, Horák, et al. 2010; Ševčíková et al. 2015; Füssy and Oborník 2017). As relationships among the other lineages remain unclear, it is

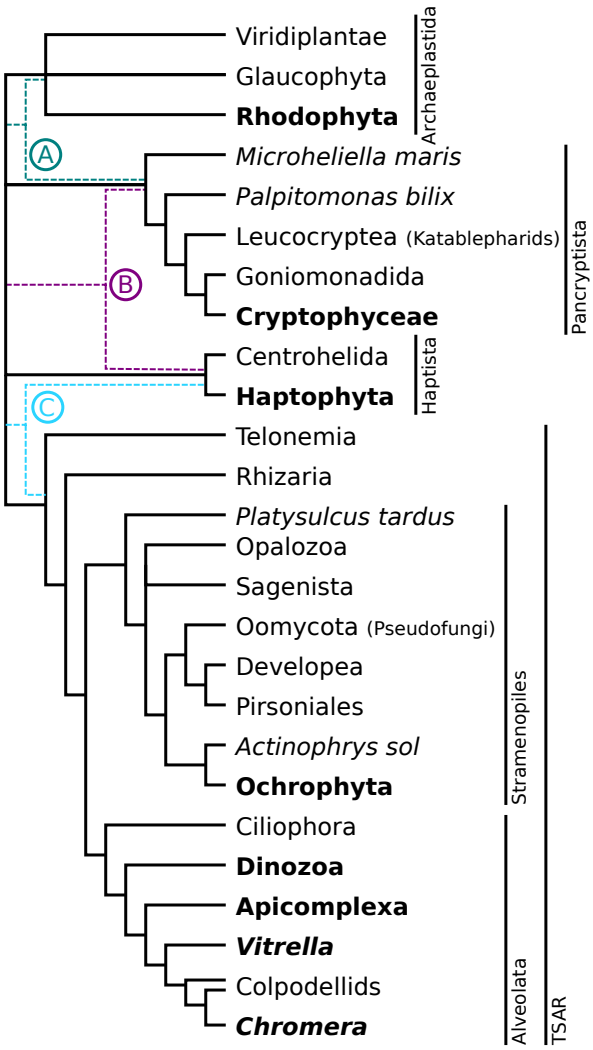


Figure 1.8: The phylogeny of Archaeplastida, Cryptophyta, Haptophyta, Alveolata, Stramenopiles and related lineages. Black lines show the well supported relationships among the shown lineages, with in certain places polytomies where the relationships remain uncertain. Dashes lines in the tree show potential relationships, with two noted that are highly important for different proposed hypothesis: (A) Pancryptista monophyletic with Archaeplastida (e.g., Yazaki et al. 2022), (B) Haptista monophyletic with Pancryptista (Hacrobia), (C) Haptista monophyletic with TSAR (e.g., Strassert, Irisarri, et al. 2021).

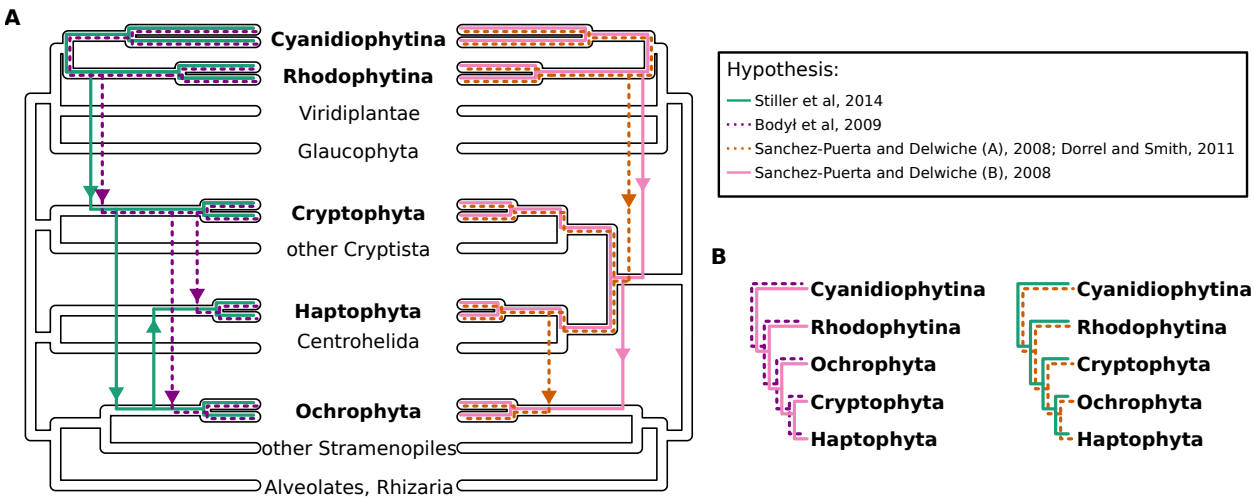


Figure 1.9: Different hypotheses suggest different plastid phylogenies. (A) Schematic tree of eukaryotes with no relationships resolved between SAR, Cryptista and Haptista (left) and one assuming the monophyly of Cryptista and Haptista (Hacrobia). Possible serial endosymbioses are shown in different colors on top of these trees (Bodył et al. 2009; Dorrell and Smith 2011; Stiller, Schreiber, et al. 2014; Sánchez-Baracaldo et al. 2017). Two hypotheses are shown for Sanchez-Puerta and Delwiche, directly referring the given alternative A and B in the original paper. (B) Plastid phylogenies expected to be found assuming the four hypotheses given, colors are the same as in A.

difficult to test the chromalveolate hypothesis specifically on relationships found within plastid trees. In contrast, several serial endosymbiosis hypotheses have been specific on between which lineages plastids have been transferred (fig. 1.9; Sanchez-Puerta and Delwiche 2008; Bodył et al. 2009; Dorrell and Smith 2011; Stiller, Schreiber, et al. 2014). Although different relationships have been found, with either a "cryptophyte-first" (cryptophytes diverging before the other groups in plastid phylogenies) or "ochrophyte-first" scenario occurring (fig. 1.9), no study has yet looked into depth at these relationships. Moreover, the monophyly of complex plastids in such phylogenies could be an artifact due to low taxon sampling of unicellular and filamentous red algae that may be more closely related to the alga (or algae) that became the red complex plastids. With a large plastid data set, and the inclusion of several plastid genomes of representatives of Proteorhodophytina, Muñoz-Gómez and Slamovits (2018) recovered red complex plastids (excluding Myzozoa) only as monophyletic with Bayesian inference. Using maximum likelihood with different models, ochrophytes were found as sister a clade of Rhodophytina, Cryptophyta and Haptophyta (Muñoz-Gómez, Mejía-Franco, et al. 2017). Thus, the question whether red complex plastids are monophyletic in plastid trees remains open, and the specific relationships remain unclear.

1.4.6 . The fossil record and molecular dating

One element that can help us understand the origin of not only algae with red complex plastids, but also other major events in the (eukaryotic) tree of life, is the fossil record. If we had good estimates for the origin and diversification of red algae and red complex plastids, we would be able to potentially exclude certain scenarios. For example, if the diversification of haptophytes occurred well after the diversification of ochrophytes, it would make it impossible that ochrophytes acquired their plastid from stem-lineage haptophytes. In this section I will address the record of (mainly photosynthetic) eukaryotes, and how this fossil record is used to date the tree of (photosynthetic) eukaryotes.

Different types of data can inform us about the origin of eukaryotic (or prokaryotic) life. Fossils, which are any preserved remains of once-living beings, may constitute clear evidence for a certain lineage existing at a specific time (e.g., Betts et al. 2018). This highly depends on the preservation of derived traits that are unique to a specific monophyletic group. For example fossils of pennate diatoms that include a raphe are clear evidence of the existence of raphid pennate diatoms, a well-known monophyletic group (Bacillariophycidae; Medlin and Kaczmarska 2004). The age of fossil can then be estimated using isotopic dating, knowledge of the rock or sediment layer the fossil was embedded in (stratigraphy), or presence of other fossils with a known stratigraphic zone (Parham et al. 2012). Besides preserved remains, biomarkers—molecules thought to only have a single biological origin—can sometimes also be used for tree calibration although there can often be issues around the exact origin or contamination (Knoll 2014). One example is alkenones, which are produced by coccolithophores and as far as is known by no others sources (organic or inorganic) (Volkman et al. 1980; Hagino and Young 2015). Maximum ages for some lineages can also be estimated (assuming they have a relatively complete fossil record), where the age is established as "*older than all the oldest possible records [of that lineage], extending back to encompass a time when the ecologic, biogeographic, geologic, and taphonomic conditions for the existence of the lineage are met, but no records are known*" (Parham et al. 2012).

Fossils of algae are typically from well fossilizing lineages such as the calcareous or siliceous coccolithophores, dinoflagellates, diatoms and coralline algae. However, these groups represent recent branches within the algal tree of life, and may not be informative to date deeper nodes. When it comes to red plastids, the earliest potential fossil information comes from red algae, although these fossils are controversial (like most Precambrian fossils; i.e., older than ~541 Ma; Gradstein 2012). The earliest fossil attributed to red algae is the ~2000 million year old *Eosphaera*, which one of the possible interpretations is that of a *Porphyridium*-like alga by Tappan (1976), although not even a eukaryotic affinity is certain (Brasier et al. 2015). One fossil often regarded as the earliest confident crown-eukaryote is *Bangiomorpha pubescens*, which is considered a multicellular alga related to the bangiophyte red algae (Bangiophyceae) (Butterfield et al. 1990; Butterfield 2000). The fossil was described in detail by Butterfield (2000), going as far as attributing it to the family Bangiaceae. Multiple characters show similarity with modern Bangiaceae including *Bangia* and genera sister to all other Bangiaceae: *Dione* and *Minerva* (Nelson et al. 2005). Specifically, the fourfold radial symmetrical arrangement of the cells has not been seen in any algae outside of the Bangiaceae

(Butterfield 2000). However, Cavalier-Smith (2002) considered that there is a definite possibility of a filamentous cyanobacterium evolving such a relatively simple morphology. Moreover, there is a large gap of red algal fossils afterwards, with the next fossil being the stem-corallinalean algae *Thallophyca* and *Paramecia* from the Doushantuo formation at ~600 Ma (Xiao et al. 2004). Besides the simple morphology, there are also no unambiguous traces of organelles or pit plugs (Butterfield 2000; Carlisle et al. 2021) keeping the (red) algal affinity of *Bangiomorpha* in the realm of uncertainty. Finally, molecular clock analyses are highly affected by the usage of this specific fossil, making ages older than when excluding it (Berney and Pawlowski 2006; Eme et al. 2014). This may suggest the taxonomic affinity is incorrect, but can also reflect a difficulty of finding clear early-eukaryotic fossils. Older (1.6 Ga), potentially red algal, fossils have been described since (Bengtson et al. 2017), but have similar issues including a lack of unambiguous pit plugs and organelles (Carlisle et al. 2021).

The fossil record directly informs about the origin of the lineage to which it is affiliated. However, using molecular clock analyses, phylogenies can be combined with this information to also estimate the age of other lineages (or rather, nodes) in the tree (for details, see 3.3). Found ages have been used to estimate the origin of eukaryotic groups based on nuclear data (e.g., Berney and Pawlowski 2006; Roger and Hug 2006; Parfrey et al. 2011; Eme et al. 2014), and have been used to test the probability of the serial endosymbiosis hypothesis (Strassert, Irisarri, et al. 2021). Such types of analysis have rarely been done before using plastid data (e.g., Sánchez-Baracaldo et al. 2017), but only a few studies include red complex plastids and discuss their origin (e.g., Yoon, Hackett, Ciniglia, et al. 2004). Yet an update remains to be done since the sequencing of many red (especially red complex) plastid genomes, and it would be fruitful to here also test the effect of the inclusion of *Bangiomorpha* and determine how the estimated plastid ages relate to that of dated host phylogenies.

2 - Objectives

As we have seen in the introduction, hypothesis on the origin of red complex plastids are at least partly supported by the fact that red complex plastids form a monophyletic group in plastid trees. However, a wide sampling of unicellular (and filamentous) red algal taxa is lacking in these plastid phylogenies. The number of species described in the subphyla including unicellular lineages (Cyanidiophytina and Proteorhodophytina) was relatively low, but there were several species described and available in culture collections of which the plastid genome has not been sequenced. The goals of this thesis were to (1) generate and explore plastid (and mitochondrial) genomes of unicellular and filamentous species in culture collections. This was followed by (2) the evaluation of hypotheses on the origin of red complex plastids with phylogenetic trees improved taxon sampling of red algae. Finally, this data and the fossil record of algae was used to (3) estimate the origin of red complex plastids and other important events that occurred in the algal tree of life.

1. Generate and explore organellar genomes of unicellular and filamentous red algae.

At the time, only ten plastid genomes of the Proteorhodophytina were sequenced. With many representatives available in culture collections, we obtained cultures to extract and sequence genomic DNA, and then assembled and annotated plastid genomes. We also searched for other genomic sequences that can be recovered from sequencing at relatively low depth, unlikely to be enough for nuclear genomes but potentially enough to assemble mitochondrial genomes. As it has been shown before that several plastid genomes of Proteorhodophytina representatives have gone through genome expansions, we analyzed the assembled plastid genomes to study these events in more detail. Mitochondrial genomes of Proteorhodophytina were understudied, as there was only a single mitochondrial genome available. Thus, one of our objectives was to determine if mitochondrial genomes of Proteorhodophytina were expanded as well, and to determine whether there is any relation between mitochondrial and plastid genome expansions.

2. Reconstruct the phylogeny of red plastids and evaluate hypotheses on the origin of red complex plastids.

Both the chromalveolate hypothesis and the serial endosymbiosis hypothesis assume a single red algal endosymbiosis at the origin of red complex plastids. Plastid phylogenies have continuously supported this, but the lack of sampling of unicellular red algae may give rise to such a result erroneously. Furthermore, even if they appear monophyletic, specific relationships among the red complex plastid lineages may support or reject specific scenarios. For example, if a stem-lineage haptophyte plastid was transferred to the common ancestor of ochrophytes (and not at all to cryptophytes), we would expect these two groups to be monophyletic in plastid phylogenies. Our objective was to construct the phylogeny of red algal and red complex plastids, and to determine the relationships among the

plastids of MOCHa lineages, using maximum likelihood and Bayesian inference, as well as topology testing and analysis of sequence compositional biases.

3. Use molecular dating to time the origin of red complex plastids in cryptophytes, haptophytes and ochrophytes.

One or more red plastid acquisitions occurred that resulted in the current diversity of red complex plastids. Although the case for Myzozoa is complex, it is clear that at least the cryptophytes, haptophytes and ochrophytes are monophyletic groups for which we could imagine a maximum of three plastid acquisition events: one for each lineage. Plastid phylogenies can inform us about which scenarios are possible (objective 2). Using the fossil record and molecular dating, we aimed to also estimate how many millions of years ago this events occurred. We aimed to gather calibrations points based on fossil records and other geological information to be able to calibrate the plastid tree of life. One obstacle however was the uncertainty around the phylogenetic affiliation and age of many—especially Precambrian—fossils. Thus, we made use of different calibration sets and models on evolutionary rates to establish estimated dates for the origin of red complex plastids in cryptophytes, haptophytes and ochrophytes.

3 - Materials and methods

This section gives an overview of the methods used to generate annotated organellar genomes and to perform phylogenetic analysis, and provide more detail on the use of the fossil record for molecular dating. An overview of the methods used per chapter are also given as figures (figs. 3.1 and 3.2). Details on the methods used for specific projects can be found in the method sections of Chapter 4 and 5.

3.1 . From culture to organellar genomes

DNA extraction and sequencing

To improve the resolution of the tree of red algae and red complex plastids, we collected cultures of unicellular and filamentous species that were otherwise understudied. Our cultures consisted mainly of representatives of Proteorhodophytina plus a representative of Bangiophyceae, and several representatives of Cyanidiophyceae. There was little organellar genome sequence data available for Proteorhodophytina species at the start of this thesis: there were ten species with plastid genomes and just one with a mitochondrial genome sequenced (Nan et al. 2017). However, there were more cultured representatives available in culture collections, and in total, we acquired organellar data for 25 species using short- and long-read sequencing. For short-read sequencing, the sequenced DNA fragments can (of course) be short, which allowed us to use relatively aggressive methods to extract the DNA, such as bead-beating. For our Proteorhodophytina cultures, we extracted the DNA using the DNeasy PowerBiofilm kit, which includes bead-beating. We also added a heat-shock step to disrupt the cells, using liquid nitrogen and a hot water bath, although we did not test whether this improved DNA yield. Long-read sequencing (with Nanopore) needs longer DNA fragments and very clean DNA, which can be challenging to obtain with red algae due to the thick mucilage (polysaccharide-rich) layer they produce (Evans et al. 1974). We were able to generate plastid genome data with Nanopore sequences for two species—including *Rhodella violacea*, which has a large, repeat-rich plastid genome. For further details on the DNA extraction methods, see the method sections in the manuscript of Chapter 4.

Organellar genome assembly

Once the genomic reads were ready to be analyzed, quality checks were done using fastqc (Andrews 2016), followed by trimming using Trimmomatic or bbmap (Bolger et al. 2014; Bushnell 2014). The following steps differ from eukaryotic or prokaryotic nuclear genome assembly. The main difference for organellar genome assembly is the assembler used. Although assemblers may work that are for general purpose (e.g., SPAdes) or metagenome-specific (e.g., metaspades), the assembly may be faster and of higher quality with organelle-specific assemblers such as NOVOPlasty and GetOrganelle (Dierckxsens et al. 2017; Nurk et al. 2017; Jin et al. 2020; Prjibelski et al. 2020). However, we did assemble all sequence data sets with SPAdes, as an entire assembly allowed us

to detect contamination and evaluate the presence of other genomic data (mitochondrial, nuclear). To do so, we used Bandage (Wick, Schultz, et al. 2015) to visualize the assemblies made with SPAdes-based assemblers (including GetOrganelle).

In this thesis, plastid genomes were mainly assembled using NOVOPlasty, which is designed explicitly for plastid genomes but also has options to assemble mitochondrial genomes (Dierckxsens et al. 2017). However, as we could not use Nanopore reads with this tool, we did not use it to assemble repeat-rich organellar genomes due to short sequences not being able to cross large repeats. In those cases, long-read (Nanopore) sequences are a must. We assembled long-reads with Unicycler (Wick, Judd, et al. 2017), which is designed for prokaryotic genomes, but works well for organellar genomes, as they are, in essence, reduced prokaryotic genomes.

Organellar genome annotation

Typically, plastid genome annotation is relatively straightforward. The genome follows the bacterial genetic code, and similar to bacteria the encoded genes are rarely interrupted by introns. A benefit over bacterial genomes is that there are fewer genes, with at most ~200 protein-coding genes in a red algal plastid genome, which are the most gene-rich (excluding recently acquired plastids, such as the chromatophore of *Paulinella*; Nowack et al. 2008). Moreover, genes are widely conserved between distant species. Multiple tools are available that are made specifically to annotate plastid genomes, such as GeSeq, CpGAVAS and DOGMA (Wyman et al. 2004; Liu, Shi, et al. 2012; Tillich et al. 2017). Even if these tools make mistakes in predicting gene sequences, they are relatively easy to identify in an annotation editor. As the possible start and stop codons are well defined, viewers that show open reading frames make it easy to spot mistakes and fix erroneous predictions. Here, we made use of GeSeq (Tillich et al. 2017) and the annotation editor Artemis (Carver et al. 2012) to annotate the plastid genomes of two florideophycean algae (*Mesophyllum expansum* and *Lithophyllum* sp.) and the coccolithophore *Chrysotila dentata* (see Chapter 5).

However, the plastid genomes of red algae belonging to the subphylum Proteorhodophytina are often intron-rich (Muñoz-Gómez, Mejía-Franco, et al. 2017). In some cases, introns have proliferated, resulting in expanded plastid genomes (Muñoz-Gómez, Mejía-Franco, et al. 2017). Thus, instead of using typical plastid annotation pipelines, we used Exonerate, an intron-aware tool that aligns proteins to genomic contigs (Slater and Birney 2005). We used a custom python script to merge results with variations in the allowed intron size to create a draft annotation (in .gff format; see Chapter 4 for details). We used the same pipeline for the mitochondrial genomes assembled in this study, as these also often contain introns in protein-coding genes. Finally, rRNA and tRNA genes were predicted using tRNAscan-SE for the Proteorhodophytina organellar genomes (Chan and Lowe 2019).

After the organellar genomes were assembled and annotated, we used comparative genomics (see Chapter 4) and phylogenomics to study them in further detail.

3.2 . Phylogenomics

There are multiple ways to study the evolutionary relationships between lineages. One way is to use morphological characters, which can especially be helpful for macroscopic species. However, even for macroscopic lineages, only a limited number of morphological characters can be used. For microbial organisms, only a few distinctive morphological characters may set distantly related species apart. DNA, RNA, and protein sequences can easily contain hundreds of nucleotides or amino acids per gene. For example, the plastid genome of *Cyanidium caldarium* contains an *rbcL* gene of 1464 nucleotides, which encodes for a 488 amino acid-long protein (Glöckner et al. 2000; The UniProt Consortium 2015). Using phylogenetics, phylogenies can be reconstructed using just one or a few genes. However, a single gene may not contain enough phylogenetic signal to reconstruct ancient evolutionary events accurately. Instead, multiple genes per taxon can be studied within a single analysis, which we refer to as phylogenomics. Phylogenomics can be done either by concatenating the sequences of multiple genes or by combining single-gene trees using a supertree approach.

This thesis focuses on relatively deep relationships among algae: between classes and superphyla of red algae, and between all archaeplastidal and complex plastids. As this covers much evolutionary history, a high amount of phylogenetic signal is needed. Protein sequences of plastid-encoded genes were gathered for either red algal lineages, or of representatives of all main algal taxa (for details, see Chapters 4 and 5). To detect orthologous proteins—meaning they come from a common ancestral sequence—we used PSI-BLAST to score homology and orthAgogue with mcl for clustering and orthology prediction (Altschul et al. 1997; van Dongen 2000; Ekseth et al. 2014). However, adjustments were made in both cases as no orthology predictor is perfect. To improve the orthologous groups made, we used the gene names and domains already known, which is feasible as plastid genes are well conserved and well studied. In such cases, it may be more productive to group based on orthologous groups defined in databases such as eggNOG (Huerta-Cepas et al. 2019). We refined the orthologous groups using preliminary alignments and single gene trees to detect non-orthologous genes. Paralogous genes are rarely present in plastid genomes, but in some cases, two copies of a gene are present when they are encoded in the two inverted repeats present in many plastid genomes (e.g., Lemieux, Turmel, et al. 1985).

Protein alignments were done in this study using MAFFT (Katoh, Kuma, et al. 2005; Katoh and Standley 2013), using the G-INS-i algorithm, which uses global pairwise alignment information, which is likely more suitable for well-conserved plastid genes than the local version of this method (L-INS-i). For chapter 4, we mainly made manual changes to remove erroneous sequences, for example, when a stop codon was not identified and the sequence continued for longer than expected. For chapter 5 we instead made use of the program PREQUAL, which can identify non-homologous characters and mask them (Whelan et al. 2018). In that case, we manually compared the alignments to verify that the masking was done correctly. At this point, some sequences should still be removed, mainly parts of the alignment that show an unexpected amount of variability that is unlikely to have a biological explanation. Such sites were trimmed using the program BMGE (Criscuolo and Gribaldo 2010). Finally, single-gene trees were made to verify that there was no evidence of horizontal gene transfer, as in such cases the gene tree does not follow the species tree.

After making orthologous groups, we concatenated the alignments resulting in a single, large alignment. We used both maximum likelihood (ML) with IQ-TREE and Bayesian inference (BI) with PhyloBayes to reconstruct the phylogenies (Lartillot, Lepage, et al. 2009; Lartillot, Rodrigue, et al. 2013; Nguyen et al. 2015; Minh, Schmidt, et al. 2020). In IQ-TREE, we focused mainly on the substitution matrices defined by LG and cpREV as the former is highly accurate overall, and the latter is based on plastid data (Adachi et al. 2000; Le and Gascuel 2008). We performed model tests using IQ-TREE and specifically added complex models that were not tested by default. We used the C10-C60 empirical mixture models in IQ-TREE, which are variants of the CAT model in PhyloBayes (Lartillot and Philippe 2004). Generally, a single substitution matrix is used for all positions in the alignment, whereas in reality, different sites in a protein have specific constraints (Scornavacca et al. 2020). Within PhyloBayes, we used CAT+GTR(+gamma), which infers exchange rates based on the given alignment instead of empirical matrices such as LG and cpREV (Lartillot, Lepage, et al. 2009). Support of nodes is given by bootstrap values in the maximum likelihood framework. Standard bootstrap is based on randomly selecting characters from the original alignment to create a similar alignment of the same length and rerunning the phylogeny. However, as this is computationally intensive, we used approximate bootstrap values such as ultrafast bootstrap (Minh, Nguyen, et al. 2013; Hoang et al. 2018). When using Bayesian inference, support is instead given as posterior probability based on the visited optimal trees during the tree search. It is important to note that these supports are interpreted differently. For example, standard bootstrap is often more conservative than ultrafast bootstrap (Hoang et al. 2018) or Bayesian probability.

3.3 . Molecular clock analyses

> *It is absurd to think that our current relatively simple stochastic substitution models and models of the rate of evolution over the tree of life adequately describe the process of molecular evolution over billions of years of evolution.*
>
> Andrew Roger & Laura Hug (2006)

As mentioned in the introduction, many eukaryotic fossils have uncertainties surrounding their age and phylogenetic affinity (i.e., to which branch in the tree they belong). This section overviews the methodological aspect of including fossils and their uncertainties and other methodological aspects of molecular clock analyses.

Once a well-resolved phylogenetic tree has been reconstructed, nodes in the tree can be constrained based on the fossil record of included lineages and other information that may indicate the presence or absence of a lineage in geological time. Methods exist to perform tree search and node calibration at the same time. However, these are computationally expensive, especially when already dealing with difficult-to-resolve phylogenies such as the algal tree of life (Sauquet 2013), so here the focus is only on methods that calibrate an already reconstructed tree. Any node can be given a minimum and maximum constraint. A node can be minimally constrained by a fossil or biomarker of

a lineage that is a descendant of that node, typically displaying a character unique to that lineage. For example, to set a minimum possible age for the node defining the split between raphid and a group of araphid pennates, one can use the earliest diatom fossil with a raphe: a character unique to raphid diatoms (Medlin, Williams, et al. 1993). Ideally, this age is used for the closest possible node. As the earliest fossil of raphid pennate diatoms is of the genus *Lyrella* (originally as "*Navicula*" by Witt in 1886), if there was genomic data of this genus in the tree, one could constrain the node defined by the split between *Lyrella* and a sister branch (Sims et al. 2006). This fossil can also constrain any parental nodes, but it would be superfluous as they are older by definition.

Maximum calibrations define the oldest age a node can have, i.e., it argues that the speciation event it represents must have happened before a specific time. Typically, if there are sediments representing an environment where the conditions are met for the descendants to establish, but there is no evidence for it, the age of these sediments could serve as a maximum constraint (Parham et al. 2012). However, often little is known about the taphonomic (how organisms decay and fossilize) or ecological conditions. Furthermore, it is difficult to exclude that either the fossil record has not been extensively studied or that the early years of a lineage were constrained to a specific location. In case of diatoms, the oldest often accepted (but still controversial; Falkowski et al. 2004; Girard et al. 2020) fossils are of the genus *Pyxidicula*, from the early Jurassic (174.1 - 201.3 Ma) (Kooistra and Medlin 1996; Sims et al. 2006; Kooistra, Gersonde, et al. 2007; Gradstein 2012). As no diatom fossils are found from older sediments, one could argue that the common ancestor of diatoms existed in the Triassic. A potential maximum calibration could be set at ~250 Ma, at the border between the Triassic and Permian (Gradstein 2012). Some have made this assumption (e.g., Medlin 2015), whereas others have instead used the more relaxed maximum constraint of the start of the Cambrian, ~550 Ma ago (e.g., Sims et al. 2006; Strassert, Irisarri, et al. 2021). In light of uncertainty, it is difficult to favor one constraint over the other.

Would it then be best to favor only the oldest likely maximum age and to use very few maximum calibration points? No. Being too conservative with maximum calibrations can result in much older age estimates, some being improbable or even impossible (Marjanović 2021). Older maxima may give more space for molecular data and evolutionary rate models to give estimates (Su et al. 2021). However, as even the most complex models are unable to model billions of years of evolutionary history, (Roger and Hug 2006), it becomes clear that high confidence can be placed in neither the models nor the fossil record. In light of this uncertainty, it is best to try multiple models (relaxed molecular clock models that model evolutionary rates) and sets of calibrations—especially when dealing with deep evolution—to allow this uncertainty to be visible. Moreover, calibrations do not have to be given as hard limits. Instead, a distribution allowing for the posterior ages to go outside of these limits can be used. This way, posterior ages can go outside the given limits, which can be used as an indication that the age or phylogenetic affinity of the calibration is incorrect (e.g. Eme et al. 2014).

Finally, as there is arbitrariness in how the fossil record is interpreted when calibrating a phylogenetic tree, there will be differences between studies, even if otherwise using comparable methods. The chosen calibrations must be clearly reported, and the argumentation for their age and phyloge-

netic affinity (Parham et al. 2012). This way, studies can be more easily compared, and differences in the use of calibrations can be readily understood.

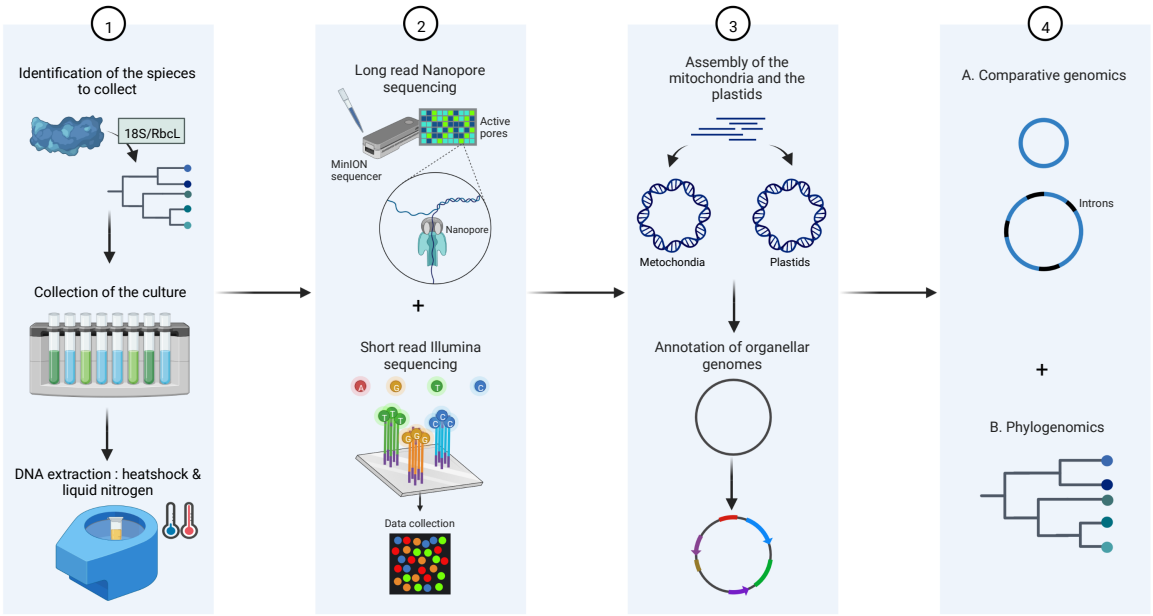


Figure 3.1: Overview of methods used in Chapter 4.

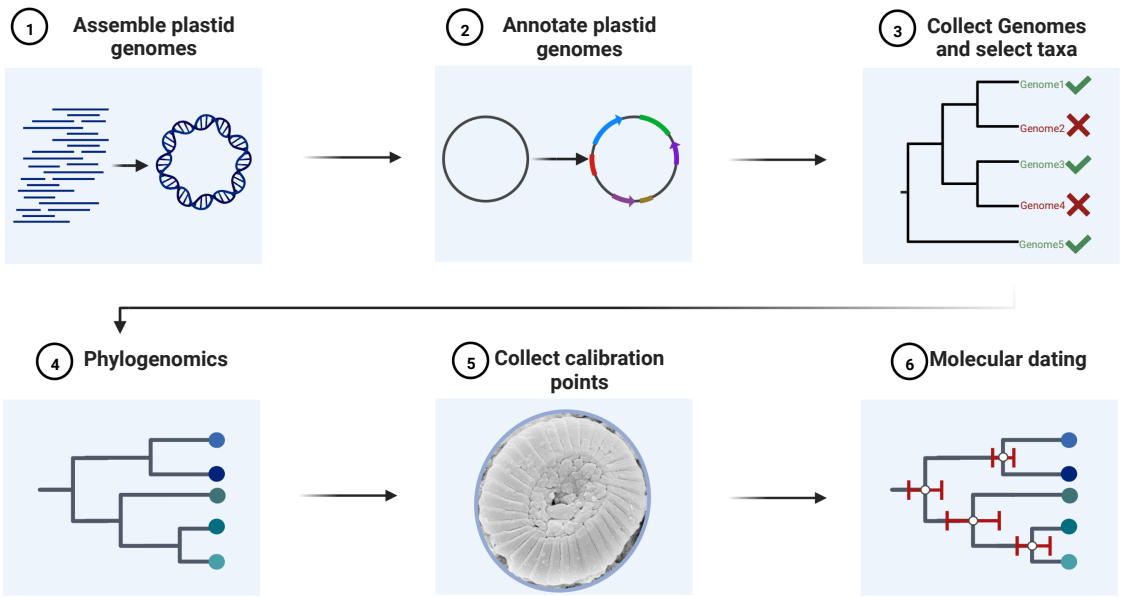


Figure 3.2: Overview of methods used in Chapter 5.

4 - Independent Size Expansions in Red Algal Plastid and Mitochondrial Genomes

4.1 . Context

Introduction and objectives

Red algae are a diverse group of photosynthetic eukaryotes, consisting of unicellular, filamentous and complex multicellular algae. Red algae are sometimes considered to have retained more ancestral features than other algae and to be slow evolving, partly due to their gene-rich plastid genomes (Janouškovec, Liu, et al. 2013). Plastid genomes have been widely studied for subset of the multicellular species of Bangiophyceae and Florideophyceae, and the extremophilic unicellular species of Cyanidiophyceae, but other red algal classes (Stylonematophyceae, Compsopogonophyceae, Rhodellophyceae, Porphyridiophyceae) have remained mostly overlooked (Janouškovec, Liu, et al. 2013; Cao et al. 2018). In contrast, the few plastid genomes studied in the remaining four classes - together constituting the subphylum Proteorhodophytina - are highly diverse in size, and vary in gene synteny and group II intron content (Muñoz-Gómez, Mejía-Franco, et al. 2017; Preuss et al. 2021). Group II introns can copy and paste their sequences when complete (Jacquier and Dujon 1985), which likely explains their proliferation in these large genomes. Moreover, the mitochondrial genome is especially understudied in Proteorhodophytina, with *Compsopogon caeruleus* as the only sequenced representative (Nan et al. 2017).

To fill this gap in our understanding of red algal organellar genomes, we sequenced 25 plastid and mitochondrial genomes of unicellular and filamentous species, mainly representatives of Proteorhodophytina. We used this data to study (1) the relationships between red algal classes based

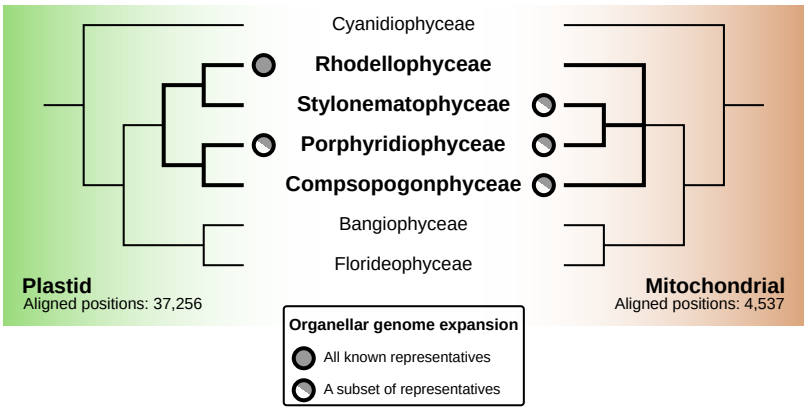


Figure 4.1: Schematic phylogeny of red algae, highlighting the relationships between the four classes of Proteorhodophytina (**bold**), based on plastid data (left, green) and mitochondrial data (right, brown). Presence of plastid and mitochondrial genome expansion are shown besides the classes, where genome expansion is defined as a plastid genome $>$ 250 kb and a mitochondrial genome $>$ 60 kb.

on plastid and mitochondrial data, with a focus on the monophyly of Proteorhodophytina, (2) the genome expansion events in the plastid genome and (3) the mitochondrial genome diversity.

Results and discussion

Muñoz-Gómez, Mejía-Franco, et al. (2017) found the four classes of mesophilic simple red algae to be monophyletic, placing them in the subphylum Proteorhodophytina. We revisited the red algal phylogeny using both a plastid and mitochondrial data set. We found strong support for the monophyly of the Proteorhodophytina using both data sets, although the internal relationship between the four classes differ between the two data sets (Fig. 1). However, topology testing showed that the plastid-derived phylogeny (Fig. 1, left) can not be rejected by the mitochondrial data. As the deep branches of the Proteorhodophytina are very short, and mitochondrial data set is smaller and contains highly diverging lineages, it is more likely that the plastid topology (Fig. 1, left) reflects the true relationships.

Group II introns are present in all plastid genomes of the Proteorhodophytina, but the genomes typically remain relatively gene-dense, similar to red algal plastid genomes in the Florideophyceae. But in certain cases these introns proliferated, resulting in genome expansion, the largest being found in *Corynoplastis japonica*, with 1.1 Mb (Muñoz-Gómez, Mejía-Franco, et al. 2017). We find that although these large genomes are restricted to two unicellular classes - the Rhodellophyceae and Porphyridiophyceae, their intron content suggests that expansion events are often independent. This is even clearer for the mitochondrial genomes, where species with large mitochondrial genomes (~100kb) are often sister to species with small (30-40kb) ones. Stylonematophyceae contain the most divergent mitochondrial genomes, having an elevated GC-content with fewer introns than expected. Although mitochondrial and plastid genome expansions are mainly caused by group II intron invasion, we find no recent transfer of these introns between different organellar genomes.

4.2 . Manuscript

Independent Size Expansions in Red Algal Plastid and Mitochondrial Genomes
(Genome Biol. Evol. 14(4), evac037)

Fabian van Beveren, Laura Eme, Purificación López-García, Maria Ciobanu, and David Moreira*

Ecologie Systématique Evolution, Centre National de la Recherche Scientifique—CNRS, Université Paris-Saclay, AgroParisTech, Orsay, France

*Corresponding author: E-mail: david.moreira@universite-paris-saclay.fr.

Accepted: 07 March 2022
Published: 15 March 2022

GBE

Independent Size Expansions and Intron Proliferation in Red Algal Plastid and Mitochondrial Genomes

Fabian van Beveren, Laura Eme, Purificación López-García, Maria Ciobanu, and David Moreira *

Ecologie Systématique Evolution, Centre National de la Recherche Scientifique—CNRS, Université Paris-Saclay, AgroParisTech, Orsay, France

*Corresponding author: E-mail: david.moreira@universite-paris-saclay.fr.

Accepted: 07 March 2022

Abstract

Proliferation of selfish genetic elements has led to significant genome size expansion in plastid and mitochondrial genomes of various eukaryotic lineages. Within the red algae, such expansion events are only known in the plastid genomes of the Proteorhodophytina, a highly diverse group of mesophilic microalgae. By contrast, they have never been described in the much understudied red algal mitochondrial genomes. Therefore, it remains unclear how widespread such organellar genome expansion events are in this eukaryotic phylum. Here, we describe new mitochondrial and plastid genomes from 25 red algal species, thereby substantially expanding the amount of organellar sequence data available, especially for Proteorhodophytina, and show that genome expansions are common in this group. We confirm that large plastid genomes are limited to the classes Rhodellophyceae and Porphyridiophyceae, which, in part, are caused by lineage-specific expansion events. Independently expanded mitochondrial genomes—up to three times larger than typical red algal mitogenomes—occur across Proteorhodophytina classes and a large shift toward high GC content occurred in the Stylonematophyceae. Although intron proliferation is the main cause of plastid and mitochondrial genome expansion in red algae, we do not observe recent intron transfer between different organelles. Phylogenomic analyses of mitochondrial and plastid genes from our expanded taxon sampling yielded well-resolved phylogenies of red algae with strong support for the monophyly of Proteorhodophytina. Our work shows that organellar genomes followed different evolutionary dynamics across red algal lineages.

Key words: Proteorhodophytina, Rhodophyta, phylogenomics, group II introns, genome expansion.

Significance

Red algal plastids exhibit large genome size variation, especially in the recently described subphylum Proteorhodophytina. By contrast, their mitochondrial genomes remain unexplored. Here, we sequenced, assembled, and analyzed new plastid and mitochondrial genomes of 25 species of red microalgae. We uncovered new cases of genome size expansion in plastids and mitochondria, in both cases associated with group II intron proliferation. The phylogenomic analysis supported that these expansions occurred several times independently.

Introduction

Red algae (Rhodophyta) are an ancient and diverse group of photosynthetic eukaryotes, consisting of more than 7,000 described species (Guiry and Guiry 2021), which range from unicellular species to complex multicellular seaweeds such as nori (*Pyropia yezoensis*) and Irish moss (*Chondrus crispus*). They are characterized by their red plastid pigments and the lack of both flagella and centrioles (Woelkerling 1990). Together with glaucophytes and green algae and plants, they form the supergroup Archaeplastida (Adl et al. 2005). The monophyly of red algae is well supported and the study of their phylogeny has elucidated

many of their intragroup relationships. The subphylum Eurhodophytina comprises the Bangiophyceae and Florideophyceae, the classes containing seaweeds, whereas the unicellular and extremophilic Cyanidiophyceae solely form the subphylum Cyanidiophytina, sister to the rest of red algae (Qiu et al. 2016). The placement of the remaining four classes (Porphyridiophyceae, Compsopogonophyceae, Rhodellophyceae, and Stylonematophyceae) was long unclear before a phylogenetic study including a representative taxon sampling showed that they were monophyletic and proposed the subphylum Proteorhodophytina to host them (Muñoz-Gómez et al. 2017).

Contemporary plastids of the three phyla of Archaeplastida evolved from an endosymbiotic cyanobacterium through significant metabolic and genomic reduction (Schwartz and Dayhoff 1978; Martin et al. 1998). Among them, the plastid genomes (plastomes) of red algae are usually thought to retain several primitive features and to evolve slowly (Butterfield 2000; Glöckner et al. 2000; Janouškovec et al. 2013). For example, plastome alignments within the classes Bangiophyceae and Florideophyceae show little change in synteny (Janouškovec et al. 2013; Cao et al. 2018), although Cyanidiaceae and inter-class comparisons show less conservation (Janouškovec et al. 2013; Muñoz-Gómez et al. 2017). Genome size shows little variation as well, typically ranging from 150 to 200 kb. By contrast, intron invasion has resulted in massive plastome expansion in the Proteorhodophytina: the Rhodellophyceae include the largest red algal plastid genome (1.1 Mb in *Corynoplastis japonica*), more than five times larger than the typical florideophyte plastomes (Muñoz-Gómez et al. 2017).

Self-splicing introns are found not only in plastomes of red algae, but also in the organellar genomes of land plants, green algae, euglenoids, and other protists (Plant and Gray 1988; Copertino and Hallick 1993; Lambowitz and Zimmerly 2011; Brouard et al. 2016). As observed in the Proteorhodophytina, in some of these organisms, such as certain green algae and the euglenoid *Eutreptiella pomquetensis*, the proliferation of self-splicing introns has led to plastid genome expansion (Brouard et al. 2016; Dabbagh et al. 2017). These introns do not only spread within one genome, but they can also travel between organellar genomes, intra- and interspecifically, and even between prokaryotes and eukaryotic organelles (Burger et al. 1999; Sheveleva and Hallick 2004; Pombert et al. 2005; Khan and Archibald 2008). Intron proliferation happens due to the copy–paste ability of the spliced intron, which is aided by an intron-encoded protein (IEP), a maturase often containing a reverse-transcriptase domain (Jacquier and Dujon 1985). However, this copy–paste ability is usually lost in organellar introns, which can have degenerate IEPs like in the plastome of the red alga *Porphyridium purpureum* (Perrineau et al. 2015).

In contrast to their plastomes, genome expansion has never been reported for mitochondrial genomes (mitogenomes) of red algae, which remain much less studied. Although mitogenomes of Florideophyceae species exhibit well-conserved gene content and synteny (Yang et al. 2015), they show more variability in Bangiophyceae and Cyanidiophyceae (Yang et al. 2015; Liu et al. 2020a). Nevertheless, size remains stable within the currently known diversity of red algal mitogenomes, ranging from 21 kb (*Galdieria sulphuraria*) to 43 kb (*Bangia fuscopurpurea*). However, there is only one mitochondrial genome sequenced for the Proteorhodophytina, from the species *Compsopogon caeruleus* (Nan et al. 2017), which hinders the description of the overall mitogenome diversity in all red algal classes.

To fill this gap and better characterize the diversity of organellar genomes in red algae, we have sequenced the plastomes and mitogenomes of 25 species with a particular focus on the poorly known Proteorhodophytina, for which only the *C. caeruleus* mitogenome and few plastomes were available at the start of this work. These new data allowed us to propose a well-resolved phylogenetic framework for Rhodophyta and to unveil divergent evolutionary patterns in these organellar genomes, including multiple genome size expansions.

Results

To explore the evolutionary relationships and organellar genome evolution in red algae, we sequenced DNA of 25 red algal species: five Cyanidiophyceae (including two uncultured ones from an environmental sample), seven Stylonematophyceae, three Porphyridiophyceae, five Compsopogonophyceae, four Rhodellophyceae, and one Bangiophyceae (supplementary table S1, Supplementary Material online; Materials and Methods). We assembled their organellar genome sequences and predicted gene features that we used for genome comparison and phylogenetic analyses.

Plastid and Mitochondrial Phylogenies Support the Monophyly of Proteorhodophytina

To build a phylogenetic framework for subsequent comparative genomic studies, we first carried out a phylogenomic analysis based on the plastid genomes from this study and those publicly available, including representatives from all known red algal classes. Our plastid data set contained 69 taxa and 189 proteins (supplementary fig. S1, Supplementary Material online), which, after trimming, consisted of 37,256 conserved amino acid positions. We made a corresponding mitochondrial data set, with 10 fewer taxa (from non-Eurhodophytina lineages lacking available mitochondrial genome sequences). This data set consisted of 59 taxa and 21 proteins (4,537 conserved amino acid positions after trimming; supplementary fig. S2,

Supplementary Material online). We used Bayesian inference with the CAT + GTR model and maximum likelihood approaches with data set-specific substitution models (see Materials and Methods) to elucidate the evolutionary relationships among red algae with both data sets. These phylogenies were rooted at the base of the Cyanidiophyceae since this class is known to be sister to all other red algae (Ciniglia et al. 2004; Yoon et al. 2006).

Both the plastid and mitochondrial phylogenies recovered the seven known classes of red algae with full support as well as the monophyly of Eurhodophytina (figs. 1 and 2; supplementary fig. S3, Supplementary Material online). The monophyly of Proteorhodophytina was also recovered with strong support for both data sets (plastid: 1/84/100, mitochondrial: 0.99/96/98; support: Bayesian posterior probability/ultrafast bootstrap/nonparametric bootstrap).

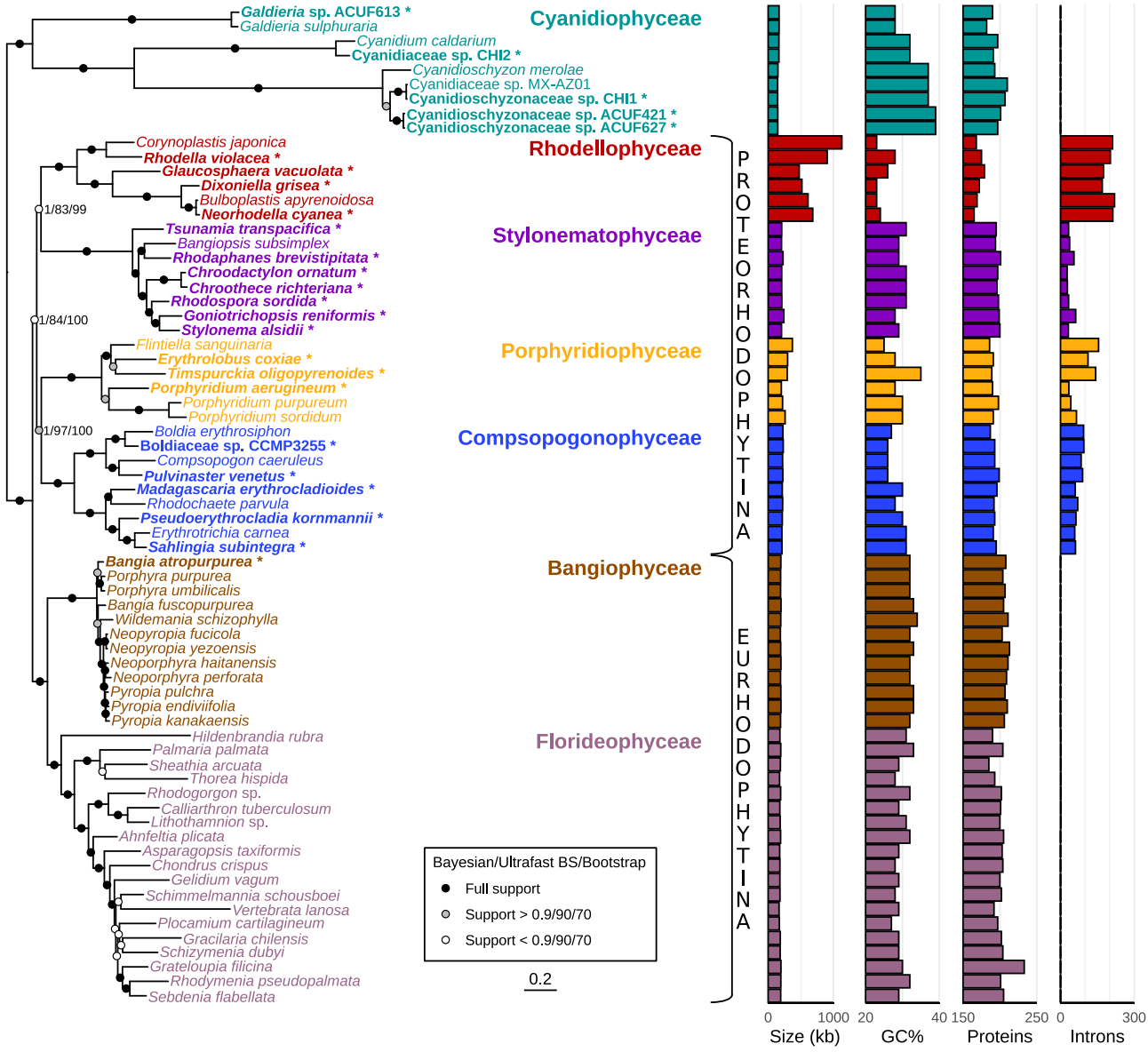


Fig. 1.—Phylogenetic tree of red algae based on plastid markers and statistics of plastid genome features. Species names bolded with an asterisk indicate sequences from this study; the colors indicate the different red algal classes. The tree was constructed from a concatenation of 189 proteins (37,256 positions, 69 taxa) using Bayesian inference (CAT + GTR model) and maximum likelihood (cpREV + C60 + F + R7 with 1,000 ultrafast bootstrap replicates; cpREV + C60 + F + R7 + PMSF with 100 nonparametric bootstrap replicates). Support values for the deep relationships among Proteorhodophytina classes are shown with labels (Bayesian probability/ultrafast bootstrap/nonparametric bootstrap). Genome statistics are indicated next to each taxon: genome size (kb), GC content (%), the number of protein-coding genes (notice that the scale ranges between 150 and 250 to emphasize the differences), and the number of introns in protein-coding genes.

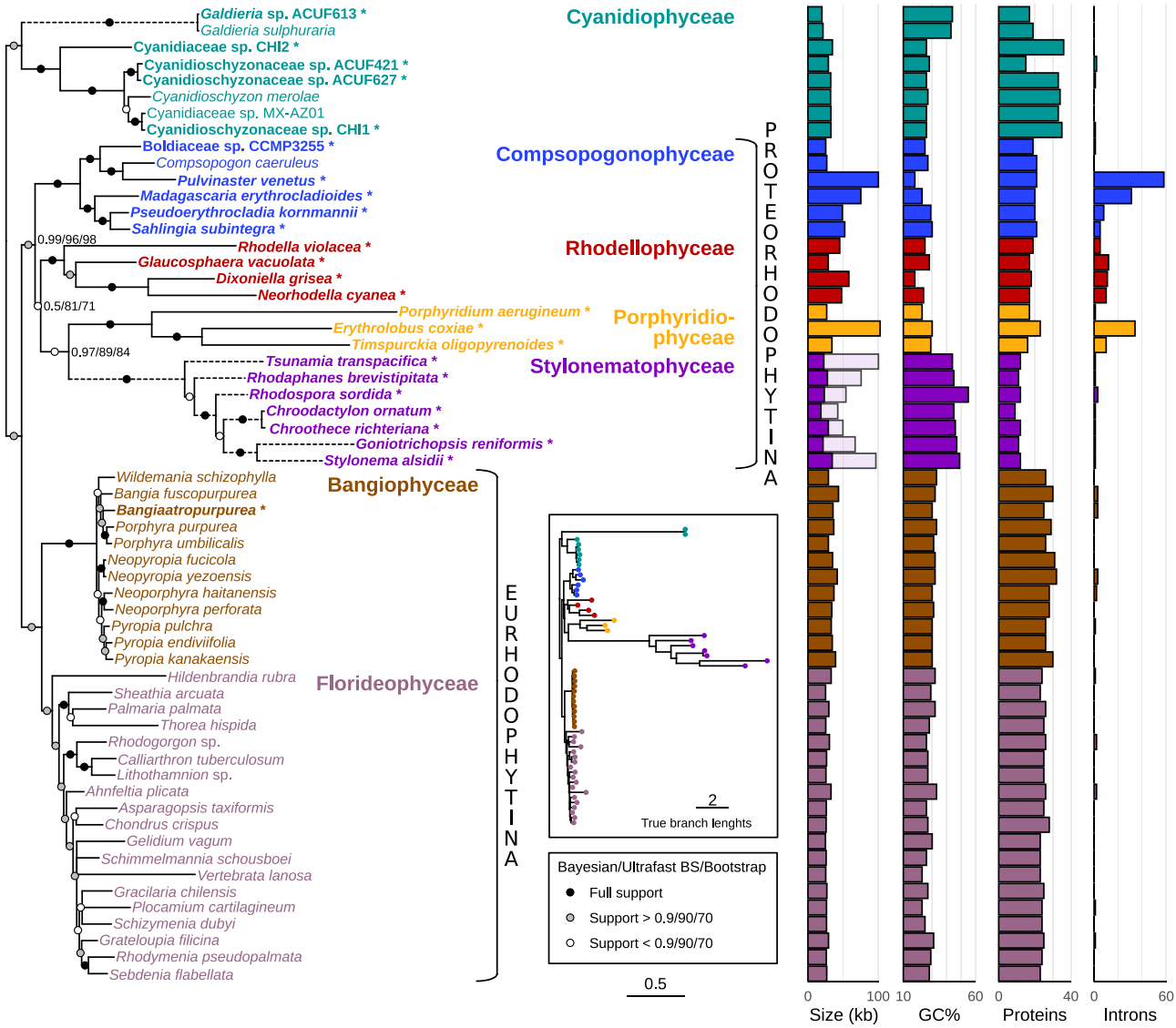


Fig. 2.—Phylogenetic tree of red algae based on mitochondrial markers and statistics of mitochondrial genome features. Species names with an asterisk indicate sequences from this study; the colors indicate the different red algal classes. The tree was constructed from a concatenation of 21 proteins (4,537 positions, 59 taxa) with Bayesian inference (CAT + GTR) and maximum likelihood (mtZOA + C60 + F + R9 with ultrafast bootstrap; mtZOA + C60 + F + R9 + PMSF with nonparametric bootstrap). Branch support is shown with circles on the branches, actual values are shown for bipartitions important for understanding the relationships among the Proteorhodophytina (Bayesian probability/ultrafast bootstrap/nonparametric bootstrap). All branches for the *Galdieria* clade (*G. phlegrea* + *G. sulphuraria*) and the Stylonematophyceae have been shortened to one-fourth of their actual length for readability (indicated by dashed lines), the true branch lengths are shown in the inset beside the phylogeny. Genome statistics are shown next to each taxon: genome size (kb), GC content (%), the number of protein-coding genes, and the number of introns in protein-coding genes. When mitochondrial genomes were fragmented, a minimal size is shown considering only the contigs with protein-coding genes, and the total size (lighter bar) corresponds to the final assembly including contigs without protein-coding genes.

Additionally, the plastid data set recovered the monophyly of Rhodellophyceae + Stylonematophyceae (1/83/99) as well as the monophyly of Compsopogonophyceae + Porphyridiophyceae (1/97/100) (fig. 1). By contrast, the mitochondrial data set recovered a well-supported clade of Stylonematophyceae + Porphyridiophyceae (0.97/89/84), but no clear support for any other relationship among the Proteorhodophytina classes (fig. 2). We carried out an approximately unbiased (AU) test (Shimodaira 2002) to compare these two different tree topologies (supplementary fig. S4, Supplementary Material online). Although the topology recovered by the plastid data set (fig. 1) was not

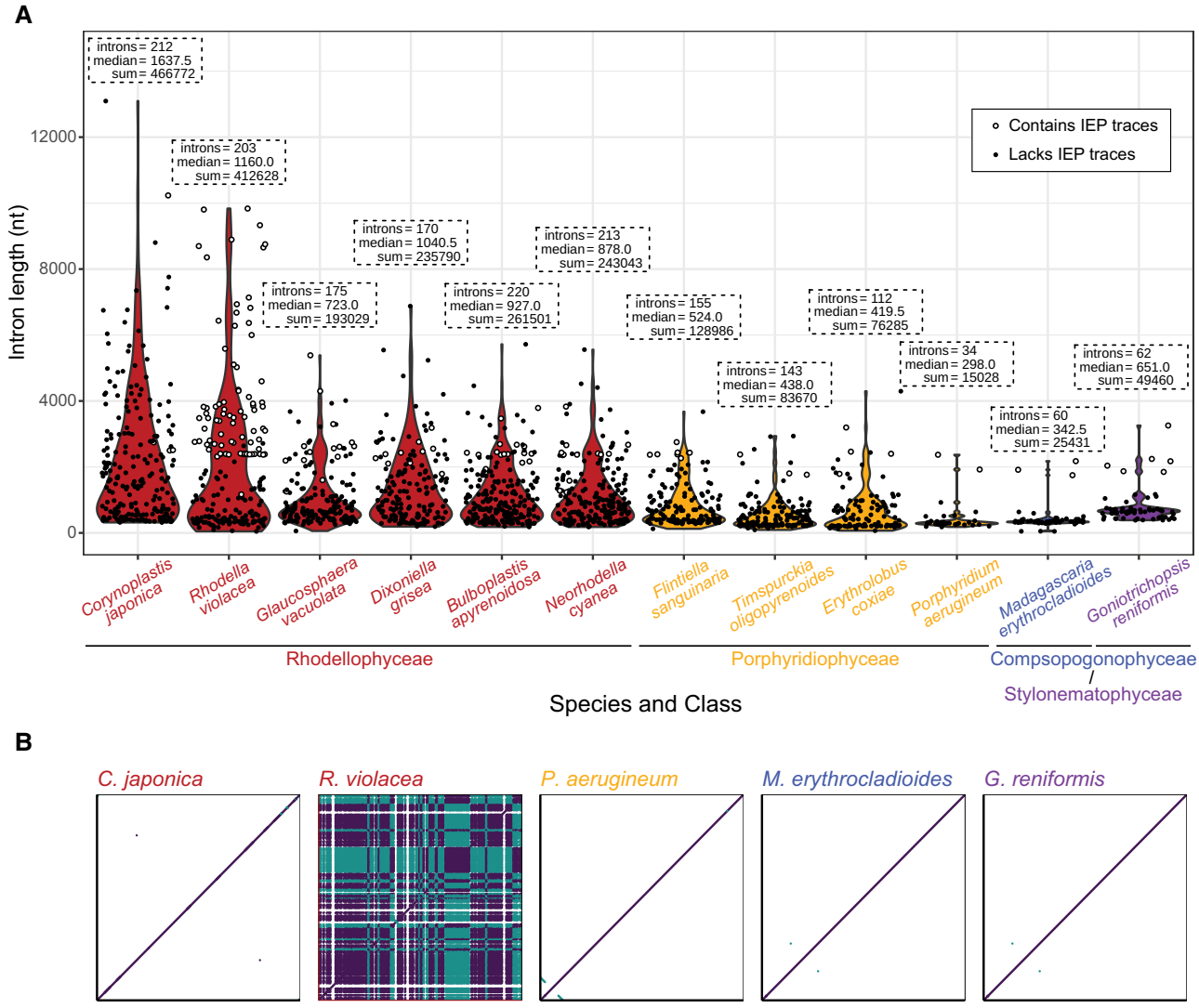


Fig. 3.—Introns and repeated regions in a selection of red algal plastid genomes. (*A*) Violin plots showing the intron size distribution of selected genomes. The total number of introns, the median length of introns, and the total size of all introns are shown in dashed boxes. White dots indicate that the intron has traces of an intron-encoded protein based on BLAST searches of group II intron-encoded reverse transcriptase sequences. (*B*) Genome self-alignments showing the repetitive content of each genome. Alignments in the forward and reverse directions are colored in purple and cyan, respectively.

rejected by the mitochondrial data set (*P*-value = 0.4), the mitochondrial topology (fig. 2) was strongly rejected by the plastid data set (*P*-value = 0.00238).

Large Plastid Genome Sizes Explained by Relatively Recent Group-II Intron Proliferation

Plastid genomes available so far for the Proteorhodophytina are intron-rich (Tajima et al. 2014; Muñoz-Gómez et al. 2017; Preuss et al. 2021). Our study confirmed this trend; our enriched taxon sampling for this group showed that all currently sequenced Proteorhodophytina plastid genomes had at least 27 introns (fig. 1). Typically, these plastomes contained mostly small (200–600 bp) introns and a few larger ones (1,500–2,500 bp) that coded for an IEP (fig. 3*A*). The only exceptions were the highly intron-rich plastomes of Rhodellophyceae and a clade within the Porphyridiophyceae (containing the species *Flintiella sanguinaria*, *Erythrolobus coxiae*, and *Timspurckia oligopyrenoides*), which had introns ranging from small (~100 bp) to huge (up to 13 kb in *Corynoplastis japonica*), with bigger introns often lacking IEPs.

The large number and size of introns in plastomes of Rhodellophyceae and a subset of Porphyridiophyceae explained the large genome sizes in these two clades, ranging from 290 kb up to 1.1 Mb (fig. 1; Muñoz-Gómez et al. 2017). Within the two classes, there was a large variation in genome size. In Porphyridiophyceae, the plastome of

F. sanguinaria (370 kb) was significantly larger than the plastomes of its sister lineages *E. coxiae* (298 kb) and *T. oligopyrenoides* (292 kb), which was at least partly explained by an increase in intron content (fig. 3*A*). Many differently sized plastomes were present in the Rhodellophyceae, all larger than any of the other red algal plastomes, yet they encoded fewer proteins (fig. 1; supplementary fig. S1, Supplementary Material online). Instead, they contained more and typically larger introns than those found in the expanded plastomes of Porphyridiophyceae (fig. 3*A*). The two largest red algal plastomes were of the sister lineages *C. japonica* and *Rhodella violacea*, which may suggest that most of the genome expansion occurred in their common ancestor. However, when aligning each genome to itself to find repetitive sequences, there were few highly identical sequences present in *C. japonica*, whereas that of *R. violacea* had many (fig. 3*B*). These highly identical regions corresponded to introns, many of which contained IEPs, although most were degenerate (fig. 3*A*; supplementary fig. S5, Supplementary Material online). A possible explanation for this observation is that the plastome of *R. violacea* underwent genome expansion by intron invasion independently of, and more recently than, that of *C. japonica*. An alternative could be that the expansion occurred in the common ancestor of these two species, but that the rate of subsequent intron degeneration was much higher in *C. japonica* than in *R. violacea.*

The analysis of the position of introns can help to distinguish between these two possibilities. If the intron expansion occurred in a common ancestor of *R. violacea* and *C. japonica*, the introns should mostly reside at the same locations, whereas if there were independent intron proliferation events, the introns should often be found at different locations. For all our Proteorhodophytina plastomes, we determined the host gene and the position of each intron and compared it with the other species (fig. 4). Only 42% of the *R. violacea* introns had a corresponding intron at the same position in *C. japonica* (40% for the *C. japonica* versus *R. violacea* comparison). This result, together with the easily detectable IEPs in the plastome of *R. violacea* (supplementary fig. S5, Supplementary Material online) and the close phylogenetic affinity of these IEPs (supplementary fig. S6, Supplementary Material online) supported a single, relatively recent, period of intron proliferation in *R. violacea*. In general, the species of Rhodellophyceae and Porphyridiophyceae had less introns in common with their sister species than the species of Stylonematophyceae and Compsopogonophyceae. Intron proliferation in Rhodellophyceae and Porphyridiophyceae was highly correlated with plastome size (fig. 5*A*).

Genome rearrangements are known to have occurred in Proteorhodophytina (Muñoz-Gómez et al. 2017). To detect additional putative recent rearrangements within each class, we aligned the plastomes of their representatives (supplementary fig. S7, Supplementary Material online). Among plastomes of Stylonematophyceae, only two rearrangements (both inversions) were detected, whereas multiple ones were visible in the plastomes of Compsopogonophyceae, especially between Compsopogonales species (represented here by *Pulvinaster venetus*, *C. caeruleus*, *Boldia erythrosiphon*, and Boldiaceae sp. CCMP3255) and the remaining representatives of Compsopogonophyceae. Finally, plastome rearrangements were especially common in representatives of Rhodellophyceae and Porphyridiophyceae, where many occurred among sister lineages. These results suggested different genome rearrangement dynamics in the different red algal classes, potentially caused by the high mobility of introns, especially in the intron-rich plastomes of Rhodellophyceae and Porphyridiophyceae.

Proteorhodophytina Contain the Largest Mitochondrial Genomes Among Red Algae

All known mitochondrial genomes of red algae are small, ranging from 21 up to 43 kb. This includes *C. caeruleus* (29 kb), the only mitogenome characterized in the Proteorhodophytina (Nan et al. 2017). Like for the plastid genomes, our considerably richer taxon sampling allowed us to reveal a large size variation in red algal mitochondrial genomes (fig. 2), with sizes over twice those previously reported. The largest ones were found in *E. coxiae* (Porphyridiophyceae, 102 kb) and *P. venetus* (Compospogonophyceae, 100 kb). The Stylonematophyceae also included highly expanded mitogenomes (e.g., ~100 kb in *Tsunamia transpacifica*), but we only obtained fragmented assemblies for these species, making it difficult to determine their precise size. Overall, although expanded plastid genomes were limited to two specific clades (see above), expanded mitogenomes were found sporadically within different lineages of the Proteorhodophytina (fig. 2).

We examined whether these mitogenome size increases were related to the number of introns. Typically, red algae have up to five mitochondrial introns, including introns in rRNA genes (fig. 2; Yang et al. 2015). We found that the two largest mitogenomes had many more introns interrupting protein-coding genes (58 in *P. venetus* and 34 in *E. coxiae*) and their presence was correlated with mitogenome size in these species (fig. 5*D*). However, these intron-rich mitogenomes did not appear as highly rearranged as the intron-rich plastomes of Porphyridiophyceae and Rhodellophyceae representatives (supplementary figs. S7 and S8, Supplementary Material online). Since introns are also common in plastomes, it could be hypothesized that group II introns have been horizontally transferred between the plastid and mitochondrial organellar genomes in different Proteorhodophytina classes. However, a phylogeny of

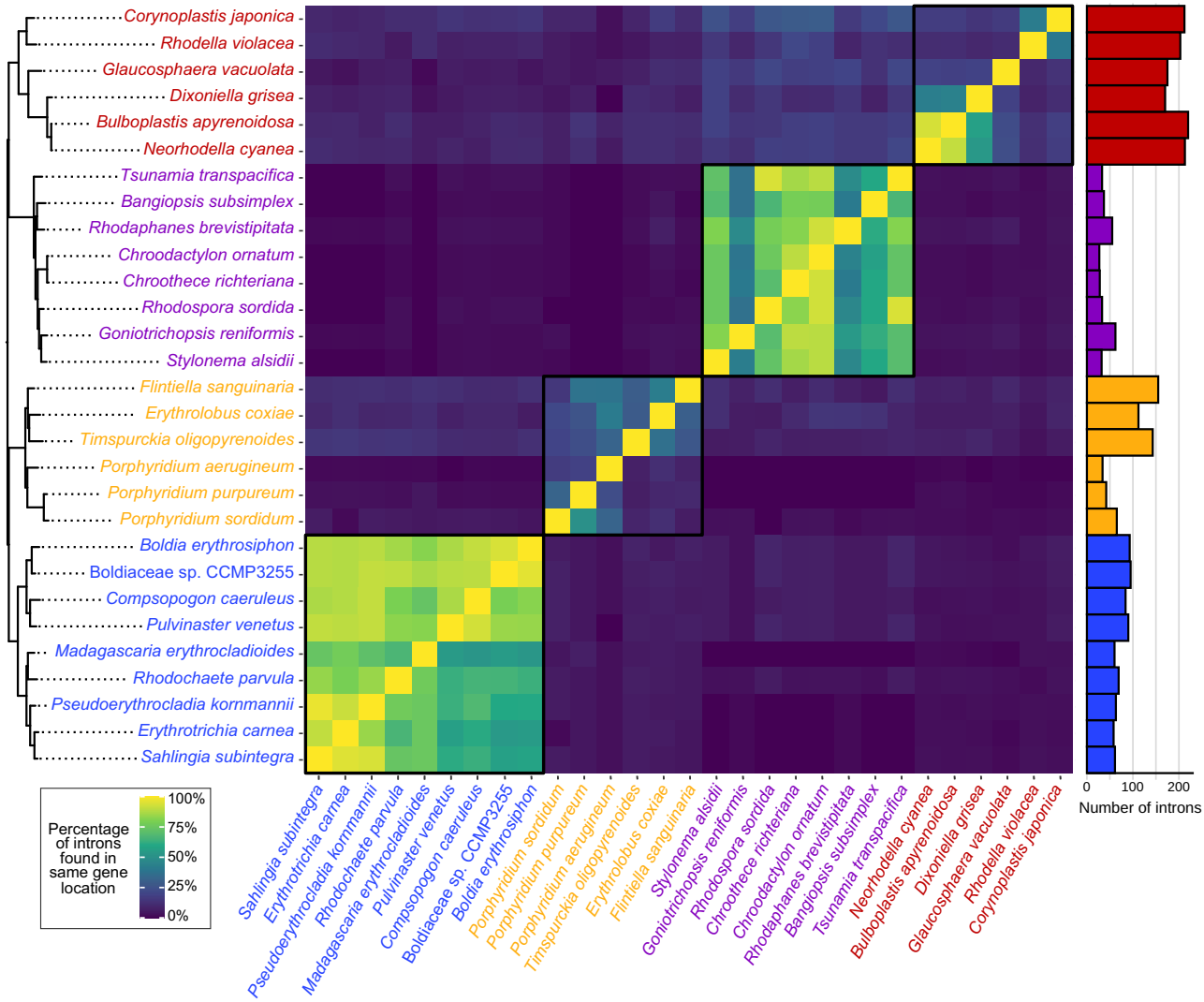


Fig. 4.—Percentage of introns present in the same position in red algal plastomes. The position was considered the same if it was in the range of ±15 nt in the same gene. Barplots on the right show the total number of introns in each plastome (as in fig. 1).

all red algal intron-encoded proteins did not show any clear recent horizontal transfer of group II introns between the two types of organellar genomes (supplementary fig. S7, Supplementary Material online).

The most divergent mitogenomes were found in the Stylonematophyceae. They contained the smallest number of encoded genes and exhibited the highest GC content (figs. 2 and 5). These mitogenomes also contained only one intron, residing in the *cox3* gene, with the exception of *R. sordida*, which had two more introns. It is possible that the low number of genes and introns identified in these mitogenomes was due to the difficulty to assemble and annotate these more divergent genomes. However, they all appear to lack the same genes, which suggest that our observations are not due to assembly or annotation errors (supplementary fig. S2, Supplementary Material online).

Discussion

The resolution of the red algal phylogeny, especially the relationships among the various lineages of unicellular and filamentous mesophilic species, has remained controversial for several decades (Gabrielson et al. 1985; Saunders and Hommersand 2004; Yoon et al. 2006). Based on the phylogenomic analysis of six plastid genomes from representatives of different classes, Muñoz-Gómez et al. (2017) observed that these lineages form a monophyletic group, supporting the erection of the subphylum Proteorhodophytina. This result was subsequently confirmed by a phylogenetic analysis of plastid data including additional species of Compsogonophyceae and Stylonematophyceae (Preuss et al. 2021). In this study, we significantly improved the taxon sampling and showed that the monophyly of Proteorhodophytina is strongly

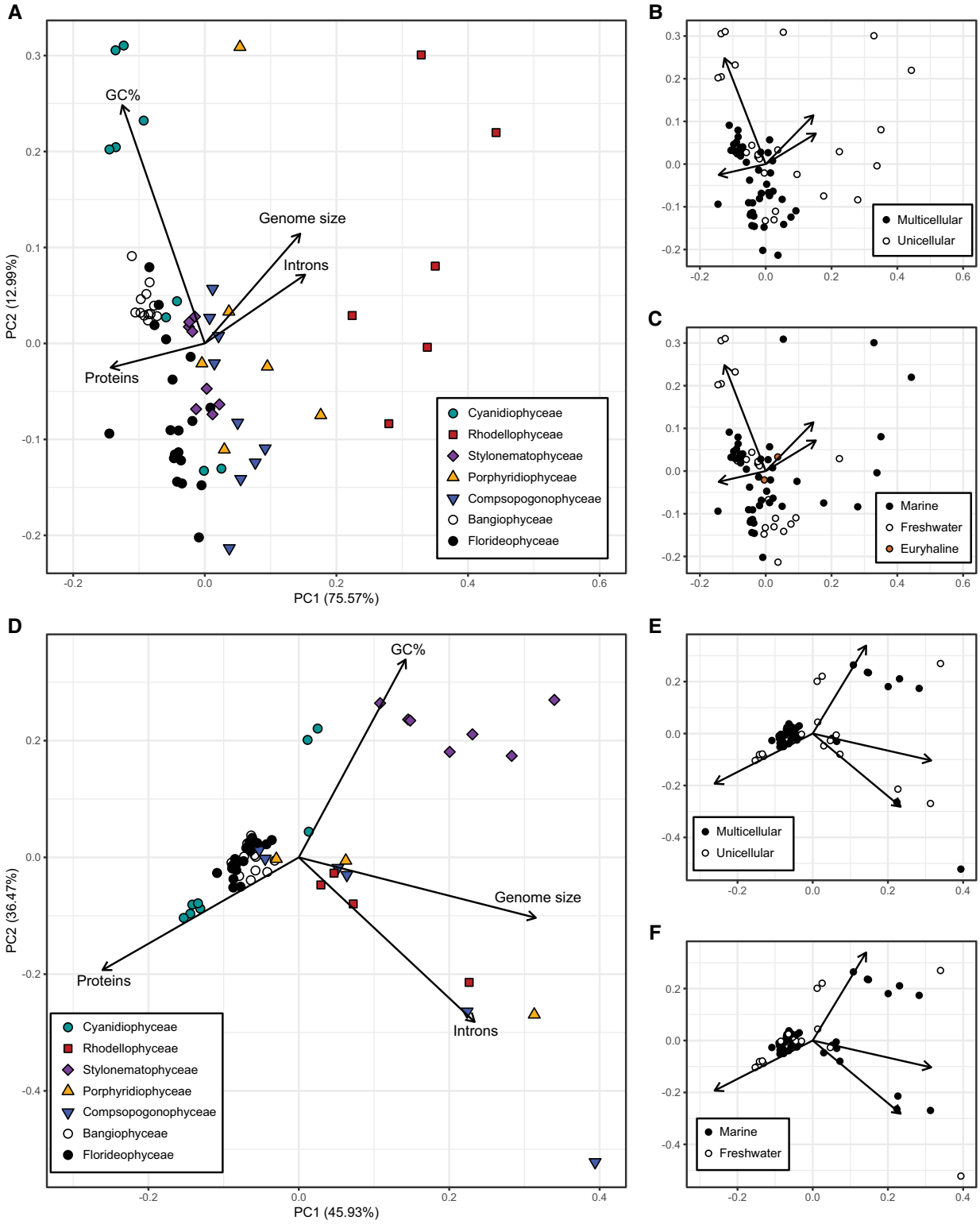


Fig. 5.—PCA of red algal organellar genomes based on genome size, GC content, the number of encoded proteins, and the number of introns. (*A*–*C*) PCA for plastid genomes showing (*A*) the red algal classes to which they belong, (*B*) the uni- or multicellular character of each species, and (*C*) the ecosystem type they live in. Euryhaline is used for species known to grow in both freshwater and marine conditions (*Porphyridium purpureum* and *P. sordidum*). (*D*–*F*) PCA for mitochondrial genomes showing (*D*) red algal classes, (*E*) uni- or multicellular species, and (*F*) the ecosystem they live in (mitochondrial genome sequences were not available for any euryhaline species).

supported not only based on plastid data, but also on mitochondrial data.

Within Proteorhodophytina, our plastid data set supported the same relationships among classes as found in the previous studies (fig. 1; Muñoz-Gómez et al. 2017; Preuss et al. 2021). The mitochondrial phylogeny showed different relationships among the four Proteorhodophytina classes (fig. 2), but branch support was typically lower and the plastid-derived relationships were not rejected by the mitochondrial data set (supplementary fig. S4, Supplementary Material online). Likely, a combination of (1) the high divergence of the mitochondrial sequences of *Galdieria* sp. ACUF613, *G. sulphuraria*, and Stylonematophyceae (fig. 5*D*); (2) the substantially smaller size of the mitochondrial data set (4,537 amino acids versus 37,256 for the plastid data set); and (3) the relatively short branches at the base of the Proteorhodophytina resulted in a different, less resolved mitochondrial phylogeny. Altogether this suggests that the plastid-encoded protein data set is more reliable to reconstruct the phylogeny of red algae. Future analyses based on nuclear genomes, still very poorly represented for Proteorhodophytina, will provide an additional test for the robustness of the plastid phylogeny.

Plastid genomes of Cyanidiophyceae, Bangiophyceae, and Florideophyceae are consistent in size and gene density (Janouškovec et al. 2013; Cao et al. 2018; Liu et al. 2020a). By contrast, those of Proteorhodophytina are intron-rich and can reach much larger sizes (fig. 1; Muñoz-Gómez et al. 2017; Preuss et al. 2021). Our larger taxon sampling of Proteorhodophytina allowed us to observe that expanded genomes were restricted to Rhodellophyceae and a subclade within Porphyridiophyceae. As both Rhodellophyceae and Porphyridiophyceae have only unicellular representatives, plastome expansion is limited to unicellular species but includes both freshwater and marine species (fig. 5*A*–*C*). A previous comparison of several *P. purpureum* strains revealed that some introns are mobile (Perrineau et al. 2015). Our results supported that this appears to be a general feature in Porphyridiophyceae and Rhodellophyceae (fig. 4), causing plastome expansion in these two red algal classes (fig. 1).

Genome rearrangements are common in the expanded plastomes (supplementary fig. S7, Supplementary Material online). Such rearrangements are often related to the presence of repetitive sequences, for example, in the plastomes of the green alga *Volvox carteri* and many land plants (Smith and Lee 2009; Wicke et al. 2011). Introns have been shown to be directly responsible for genome rearrangements in the bacterial endosymbiont *Wolbachia* (Leclercq et al. 2011). High mobility of group II introns in the plastomes of Rhodellophyceae and Porphyridiophyceae may be the cause of the observed rearrangements, although we cannot exclude that differences in plastome DNA repair mechanisms may have also played a role (Robart and Zimmerly 2005; Smith 2020). Additional research on the structure of introns and genome comparisons of closely related species will help to better understand the relationship between high intron content and the frequency of genome rearrangement as well as the mechanisms responsible for intron proliferation.

Although expanded genomes in red algae were restricted to the Proteorhodophytina, there are many other algae and land plants known to have large plastid genomes. This is the case of many species of Chlamydomonadales, among which *Haematococcus lacustris* displays the largest known plastome (1.35 Mb) (Bauman et al. 2018; Smith 2018). In contrast with the Proteorhodophytina, these genomes are often expanded due to the proliferation of palindromic repeats, which are also found in mitogenomes of *Volvox* spp. (Aono et al. 2002; Smith and Lee 2009; Zhang et al. 2019).

There is a large variation in mitochondrial genome size among eukaryotes, being famously small and gene-dense in animals and often large in plants—such as the largest known mitogenome of the Siberian larch (*Larix sibirica*) of ~11.7 Mb, which is rich in mobile genetic elements, including group II introns (Putintseva et al. 2020). The mitogenome of the fungus *Morchella crassipes* (~500 kb) is the largest known outside plants and is also rich in group II (but also group I) introns (Liu et al. 2020b). Mitogenome expansion occurs in red algae to a less extreme extent but has resulted in the largest red algal mitochondrial genome (~132 kb), recently described in a strain of *P. purpureum* (Kim et al. 2022). In this study, we show multiple lineage-specific genome expansions, including a more than 3-fold genome size increase in *P. venetus* in comparison to *C. caeruleus* (fig. 2). In contrast to plastome expansions, mitogenome expansions occur not only in unicellular, but also in multicellular (filamentous) species, including both marine and freshwater species (figs. 5*D*–*F*).

Organellar introns are known to be mobile, and putative transfers have been reported between red and brown algal mitochondria (Bhattacharya et al. 2001) and from cyanobacteria to red algal mitochondria (Burger et al. 1999). Moreover, intraspecific transfer of intronic elements between organelles was suggested to have occurred in green algae (Pombert et al. 2005; Zhang et al. 2019). One hypothesis is that genome expansion of plastid and mitochondria both result from a general invasion of these organelles by group II introns and/or from intron transfer between them. However, our data do not support these scenarios in red algae since (1) there is only one red algal species (*E. coxiae*) where both organellar genomes are substantially expanded, and (2) there is no phylogenetic evidence of recent intron transfer between plastid and mitochondrial genomes in the same species based on phylogenies from intron-encoded proteins (supplementary fig. S6,

Supplementary Material online). Although intron transfer between organellar genomes appears not to be responsible for the genome expansions described here, it is clear that selfish genetic elements, in particular group II introns, are largely responsible for organellar genome expansion across the eukaryotic domain, including red algae.

Materials and Methods

Species Selection and Culturing

Red algal species were selected based on their phylogenetic position in 18S rRNA and RbcL phylogenies and availability in culture collections. Selected cultures were obtained from the Sammlung von Algenkulturen der Universität Göttingen (SAG; Germany), the Culture Collection of Algae and Protozoa (CCAP; Scotland), the Algal Collection University Federico II (ACUF, Italy), and the National Center for Marine Algae and Microbiota (NCMA; Maine, USA). In addition, we also processed a natural sample rich in cyanidiophyte algae that was collected in the El Chichon volcano (Mexico), which contained two distinct *Galderia* species. A complete overview of the acquired cultures and method of DNA extraction can be found in supplementary table S1, Supplementary Material online. All cultures were grown in the laboratory using the Provasoli culture medium for several weeks at 21 °C with a 12 h/12 h light/dark cycle (West and McBride 1999).

DNA Extraction and Sequencing

Cells were first disrupted by heat-shock, freezing them in liquid nitrogen (90 s), and then thawing them at 65 °C (90 s) three times. Then, different DNA extraction methods were tested to maximize the DNA yield and quality (supplementary table S1, Supplementary Material online). For all species, the DNeasy PowerBiofilm kit (Qiagen) was used to extract DNA for short-read (Illumina) sequencing, following the manufacturer protocol. Because of a low DNA yield, whole genome DNA amplification was used for *Galdieria phlegrea* and *G. maxima* using the Phi 29 isothermal amplification method with the kit EquiPhi29™ DNA Polymerase (Thermofisher) as described by the manufacturer. To obtain high molecular weight DNA for long-read (Nanopore) sequencing, we used the DNeasy PowerBiofilm kit (Qiagen) for *R. violacea* and a CTAB-based method for *Sahlinga subintegra.* Briefly, 50 ml of culture was pelleted for 5 min at 500 × g. To eliminate a maximum of polysaccharides and bacteria attached to the red algal cells, the pellet was washed with a mix of culture medium and a nonionic surfactant (Pluronic) at a final concentration of 0.05% by performing three cycles of vortexing, sonication at room temperature (two cycles of 1 min, 37 Hz), and centrifugation (500 g, 5 min). The cleaned pellet was then lysed with 500 µl of Carlston buffer containing 100 mM Tris-Cl, pH 9.5, 2% CTAB, 1.4 M NaCl, 1% PEG 8000, and 20 mM EDTA, preheated to 65 °C and 300 µl of chloroform, and incubated for 30 min at 65 °C while shaking. The aqueous phase containing the DNA was then further purified with one volume of chloroform/isoamyl alcohol. The DNA was precipitated by incubation with 0.8 M sodium citrate and 1.2 M NaCl and 100% isopropanol at room temperature for 15 min and centrifugation 15 min at 10,000 × g at room temperature. The DNA pellet was washed twice with cold 70% ethanol and dried at room temperature. The DNA was resuspended in 50 µl of preheated (50 °C) 10 mM Tris-HCl pH 8 buffer and quantified with the Qubit™ dsDNA HS Assay Kit (Thermofisher) following the manufacturer protocol. DNA fragment sizes were visualized on a 0.7% agarose gel stained with GelRed® Nucleic Acid Gel Stain (Biotium). The presence of carryover contaminants was assessed with a Nanodrop spectrophotometer (Thermofisher).

2 × 150 bp or 2 × 100 bp paired-end Illumina reads were obtained by Eurofins Genomics (Konstanz, Germany) and CNAG-CRG (Barcelona, Spain) using the Illumina HiSeq 2500 and NovaSeq 6000 technologies (supplementary table S1, Supplementary Material online). Samples prepared for Nanopore sequencing were sequenced on a MinION Mk 1B device on R9.4.1 flow cells. Prior to Nanopore library preparation, the high molecular weight DNA was size-selected with the Short Read Eliminator XS kit (Circulomics) as described by the manufacturer. Nanopore libraries were constructed with the SQK-LSK 109 kit following the Genomic DNA ligation protocol proposed by the manufacturer, with minor modifications: The ligation time was extended to 30 min, all the AmPure beads purification steps were extended to 10 min of incubation with the magnetic beads and the elution from the beads was performed at 37 °C min for several hours. Real time base-calling was performed on a MinIT (MinKNOW v3.6.3) using Guppy (v3.2.9).

Organellar Genome Assembly

Illumina reads were trimmed with trimmomatic (v0.38; ILLUMINACLIP: adapters.fa: 2: 30: 10 LEADING: 30 TRAILING: 30 SLIDINGWINDOW: 4: 30 MINLEN: 36; Bolger et al. 2014) using the extended list of adapters from BBMap (v38.41; Bushnell 2014). The quality of trimmed Illumina reads was checked with FastQC (v0.11.5; Andrews 2016). Genomes sequenced with both Nanopore and Illumina were assembled with Unicycler (v0.4.9b; using bold mode for *R. violacea* and conservative mode for *S. subintegra*; Wick et al. 2017). The plastid genomes of *Neorhodella cyanea* and *Rhodophanes brevistipitata* were assembled with GetOrganelle (v1.5.1c; with variable k-mers and -R 200; Jin et al. 2020) due to the presence of many short repeats and a single large repeated

region, respectively. All other plastid genomes were assembled using NOVOPlasty (v3.4; Dierckxsens et al. 2017), with the *rbcL* sequence in the assembly as the seed sequence and variable k-mer sizes. No single assembler worked optimally for all mitogenomes and NOVOPlasty, GetOrganelle, and SPAdes (v3.11.0; with or without meta; Nurk et al. 2017; Prjibelski et al. 2020) were used (see supplementary table S1, Supplementary Material online). Genome assemblies obtained with GetOrganelle and SPAdes were manually inspected using Bandage (Wick et al. 2015). When an organellar genome assembly remained fragmented, only the contigs with protein-coding genes of plastid or mitochondrial origin were selected. For the mitogenomes of the Stylonematophyceae, this resulted in a large discrepancy between the assembled genome size and the cumulative size of selected contigs; thus, both were used to estimate genome size and other genome statistics (fig. 2).

Organellar reads were gathered by mapping the reads to the assembled organellar contigs, using BBmap with option paired only = t for short reads, and Minimap2 (v2.17-r941; Li 2018) with option -ax map-ont followed by *samtools fastq* (v1.9-52-g651bf14; Li et al. 2009) with option-F 4 for long reads. All organellar reads and unfragmented organellar genomes are available on SRA (BioProject PRJNA744153) and GenBank, respectively (supplementary table S1, Supplementary Material online). Fragmented organellar genomes are provided as fasta files, along with the predicted protein sequences (see fragmented organellar genome data in the Figshare repository).

To identify plastome and mitogenome rearrangements within non-Eurhodophytina classes, we aligned the organellar genomes of representatives for each class (supplementary figs. S7 and S8, Supplementary Material online). As starting points of circular genomes are arbitrary, the genomes were aligned with Mauve (snapshot 2015-02-13; Darling et al. 2004) to identify a common syntenic region to use as the starting point. The adjusted plastomes were then aligned using Mauve with default settings for each of the classes.

Genome Annotation

tRNAs were predicted using TRNAscan-SE with parameter -O (v2.0.3; Chan and Lowe 2019) and rRNAs were predicted using rnammer with parameters -S bac -m lsu, ssu, tsu (v1.2; Lagesen et al. 2007). Protein-coding sequences were predicted using protein sequences of closely related species with published organellar genomes as reference. Exonerate (v2.3; Slater and Birney 2005) was used to align known proteins to the genome assemblies as it can infer introns, using the applicable genetic code and a variation of allowed intron sizes (other options: –model protein2dna: bestfit -E -n 1 -s 60 –percent 10). Nucleotides for splice sites were assumed unknown, as manual inspection of the protein sequences did not show consistent patterns (data not shown). An in-house python script (cds_from_exonerate.py, available in Figshare) was used to create a valid coding sequence based on the Exonerate data by finding a correct start codon near the start of the protein alignment and a stop codon near the end. Manual curation of the genome annotation was done using BLAST (Johnson et al. 2008) and the Artemis genome browser (Carver et al. 2012). The published plastid genome sequence of *C. caeruleus* was reannotated using the same method, as intron predictions were inconsistent with that of other Proteorhodophytina plastomes. The number of predicted introns in plastid-encoded proteins of *C. caeruleus* changed from 20 to 85.

Genome Statistics

To compare the organellar genomes, genome length and GC content were determined using *stats.sh* from the BBTools suite (Bushnell 2014). For the fragmented mitogenomes of the Stylonematophyceae, it was run both for the complete mitochondrial assembly ("maximum size") and the protein-coding contigs ("minimum size"). The number of encoded proteins was estimated by counting the number of genes represented by CDS features in the .gff files. For each gene, the number of introns was assumed to be equal to the number of CDS features minus one. A principal component analysis (PCA) was made based on these statistics for both the plastid and mitochondrial genomes using the R stats function *prcomp* (R v4.0.4; options scale. = T; R Core Team 2021). PCA plots were made using the ggfortify function *autoplot* (v0.4.11; Tang et al. 2016) in ggplot2 (v3.3.3; Wickham 2016).

Phylogenetic Analyses

Published sequence data of organellar genomes from other red algal species were retrieved from GenBank (supplementary table S1, Supplementary Material online; Benson et al. 2013). In total, 43 plastid genomes were collected comprising all species of Cyanidiophyceae and Proteorhodophytina with available plastid genomes as well as a selection of Florideophyceae and Bangiophyceae. Including our own data, this resulted in a data set of 69 taxa (plastid dataset in Figshare). Among these species, ten of the Cyanidiophyceae and Proteorhodophytina representatives had no mitochondrial genome sequence available, which resulted in a mitochondrial data set containing 59 taxa (mitochondrial dataset in Figshare).

For each data set, we used an all-against-all BLAST search using psi-blast (BLAST 2.6.0+; options: -evalue 10 -outfmt 6 -max_target_seqs 200 -seg yes -soft_masking true -use_sw_tback -word_size 2 -matrix BLOSUM45;

Altschul et al. 1997), and draft orthologous groups were created using orthAgogue (v1.0.3; default options; Ekseth et al. 2014) and running mcl (v1: 14-137; van Dongen 2000) on the "all.abc" output file. An in-house script was used to deal with the presence of paralogous sequences (mcl_to_og.py, available in Figshare). These were either merged if the separated genes were considered two portions of a unique gene separated due to misannotation, or in the case of duplicated sequences specific to a genome, the sequence most similar to other sequences in the orthologous group was retained.

For the plastid and mitochondrial data sets, the multiple sequence alignment (MSA) for each protein was constructed with MAFFT G-INS-i or MAFFT L-INS-i, respectively (v7.310; Katoh et al. 2005). The MSAs were manually refined using single-protein trees made with IQ-TREE (v1.6.11; using -m cpREV + C60 + F + G for plastid MSAs, and -mset LG -mrate G, I + G, R -mfreq FU, F for mitochondrial MSAs; Nguyen et al. 2015) and inspecting the alignments with Aliview (v1.24; Larsson 2014). Final MSAs were trimmed using BMGE (default options; Criscuolo and Gribaldo 2010) and concatenated, resulting in a plastid data set of 37,256 amino acids and a mitochondrial data set of 4,537 amino acids.

The concatenated plastid data set was analyzed by maximum likelihood (ML) with IQ-TREE using the model cpREV + C60 + F + R7 and ultrafast bootstrap (Nguyen et al. 2015; Hoang et al. 2018). The resulting ML tree was used as the guide tree for rapid approximation of posterior mean site frequency (PMSF) under the same model and the generation of 100 nonparametric bootstrap replicates (Wang et al. 2018). The mitochondrial data set was similarly analyzed with the models mtZOA + C60 + F + R9 and mtZOA + PMSF(C60)+F + R9. The Bayesian inference phylogenetic analysis of the plastid and mitochondrial data sets was performed using Phylobayes-MPI (v1.8c; Lartillot et al. 2013) with the CAT-GTR model (Lartillot and Philippe 2004) and four chains run to 10,000 generations (maxdiff remained 1 for both data sets due to the lack of resolution of several branches within the Florideophyceae). Trace files were inspected with graphylo (supplementary figs. S9 and S10, Supplementary Material online; https://github.com/wrf/graphphylo) and Tracer (Rambaut et al. 2018) to determine a burn in of 200 generations for both mitochondrial and plastid phylogenies. Phylogenetic trees were plotted with the R package ggtree (v2.4.1; Yu 2020). The above phylogenies, along with the full alignments and single gene alignments, are available in the Figshare repository.

To compare the trees based on the plastid and mitochondrial data sets, we reconstructed trees for both the plastid and mitochondrial taxon sampling and constrained two possible relationships within the Proteorhodophytina: (1) a clade of Stylonematophyceae + Porphyridiophyceae, with remaining relationships unresolved, and (2) a clade of Stylonematophyceae + Rhodellophyceae, plus a clade of Compsopogonophyceae + Porphyridiophyceae (supplementary fig. S4, Supplementary Material online). All other relationships were kept the same as in the original trees. For both data sets, we tested the three trees from the original analysis (Bayesian and maximum likelihood with and without PMSF) and 100 nonparametric bootstrap trees of the related maximum likelihood analysis with PMSF. The AU-test (Shimodaira 2002) was used, as implemented in IQ-TREE, with the options -n 0 -zb 10,000 -au -zw (Nguyen et al. 2015). The models used were as before: mtZOA + C60 + F + R9 for the mitochondrial data set cpREV + C60 + F + R7 for the plastid data set.

A single-gene phylogeny using the protein RbcL was made to include more available plastid data for comparison (RbcL data set in Figshare). We gathered RbcL sequences of the 69 representatives in the plastid data set and all RbcL sequences of red algal origin not part of the Eurhodophytina from Uniprot (The UniProt Consortium 2015), resulting in a total of 404 sequences. Sequences were aligned with MAFFT L-INS-i (Katoh et al. 2005) and phylogenetic inference was done using IQ-TREE (Nguyen et al. 2015) with options -m cpREV + C40 + F + R4 -bb 1,000 -alrt 1,000 (supplementary fig. S11, Supplementary Material online).

Analysis of Introns and Intron-Encoded Proteins

First, for each intron interrupting a protein-coding gene in plastid genomes, we determined to which orthologous group it was associated (based on ortholog grouping, see "Phylogenetic Analyses" above). Second, the intron location within the gene was determined by considering the first nucleotide of the start codon as position 0 and disregarding the size of previous introns within that gene sequence (i.e., as if the introns were spliced out). To estimate the number of plastome introns that a species A shared with another species B, we calculated the fraction of introns of species A that are at the same location ($\pm$15 nt) in the same gene in species B. These results were visualized with a heatmap made with ggplot2 (Wickham 2016).

To detect whether plastid introns have traces of IEPs, they were aligned to a set of 144 proteins with the reverse-transcriptase domain of group II introns from the conserved domain database (CDD; cd01651; Lu et al. 2020) using blastx (*E*-value threshold of 1×10^{-10}; Altschul et al. 1997). Size distribution of the introns of different species was plotted using violin plots and raw data points with ggplot2 (Wickham 2016).

To construct a phylogeny of all IEPs encoded in red algal group II introns, we used the cd01651 protein sequences to search all our plastid and mitochondrial genomes and all red algal organellar genomes in GenBank using tblastn

with the option -max_intron_length 15,000 (Altschul et al. 1997). Regions with blast hits were considered to have traces of an IEP if there was more than one blast hit and an *E*-value $< 1 \times 10^{-5}$. These parameters were chosen based on manually checking the results obtained with the plastid genome of *E. coxiae*. Sequences were aligned using MAFFT L-INS-i (Katoh et al. 2005) and nonhomologous sequences were removed. The final alignment consisted of 546 IEP sequences and 2,084 amino acid positions and was trimmed with BMGE (to 462 amino acids) with the options -w 1 -h 1 -g 0.7 (Criscuolo and Gribaldo 2010). The IEP phylogeny was reconstructed with IQ-TREE (-m LG + C20 + G + F -bb 1,000) and with the R package ggtree (supplementary fig. S6, Supplementary Material online; intron data set in Figshare; Yu 2020).

To detect the presence of repetitive regions, including multiple copies of intron-encoded proteins and other repeats, we aligned each genome with itself using NUCmer (–maxmatch –nosimplify; MUMmer, 4.0.0beta2; Marçais et al. 2018). Files with coordinates were created using show-coords (-c -l -r -T) and plots were created using ggplot2.

Supplementary Material

Supplementary data are available at *Genome Biology and Evolution* online (http://www.gbe.oxfordjournals.org/).

Acknowledgments

We thank the UNICELL single-cell genomics platform (https://www.deemteam.fr/en/unicell) for help in DNA preparation and Nanopore sequencing, Iris Rizos for preliminary analysis on intron size distribution in red algal plastids, and Line Le Gall for critical reading of the manuscript. We are grateful for the comments and suggestions of two anonymous reviewers that helped to improve this article. This work was funded by the European Research Council Advanced Grants ProtistWorld (No. 322669, P.L.-G.) and Plast-Evol (No. 787904, D.M.) and the ERC Starting grant MacroEpik (No. 803151, L.E.).

Data Availability

The Nanopore and Illumina sequencing reads were deposited in the NCBI SRA under BioProject PRJNA744153. Complete organellar genomes were submitted to GenBank, accessions can be found in supplementary table S1, Supplementary Material online. Sequencing reads and organellar genomes are also available at Figshare (https://doi.org/10.6084/m9.figshare.17111693), along with fragmented organellar genome and protein sequences, phylogenomic data sets, and custom python scripts.

Literature Cited

Adl SM, et al. 2005. The new higher level classification of eukaryotes with emphasis on the taxonomy of protists. J Eukaryot Microbiol. 52(5):399–451.

Altschul SF, et al. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 25(17):3389–3402.

Andrews S. 2016. FastQC: A quality control tool for high throughput sequence data. Available from: https://www.bioinformatics.babraham.ac.uk/projects/fastqc/.

Aono N, Shimizu T, Inoue T, Shiraishi H. 2002. Palindromic repetitive elements in the mitochondrial genome of *Volvox*[1]. FEBS Lett. 521(1):95–99.

R Core Team. 2021. R: a language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. Available from: https://www.R-project.org/.

Bauman N, et al. 2018. Next-generation sequencing of *Haematococcus lacustris* reveals an extremely large 1.35-megabase chloroplast genome. Genome Announc. 6(12): e00181-18.

Benson DA, et al. 2013. GenBank. Nucleic Acids Res. 41(D1):D36–D42.

Bhattacharya D, Cannone JJ, Gutell RR. 2001. Group I intron lateral transfer between red and brown algal ribosomal RNA. Curr Genet. 40(1):82–90.

Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30(15):2114–2120.

Brouard J-S, Turmel M, Otis C, Lemieux C. 2016. Proliferation of group II introns in the chloroplast genome of the green alga *Oedocladium carolinianum* (Chlorophyceae). PeerJ. 4:e2627.

Burger G, Saint-Louis D, Gray MW, Lang BF. 1999. Complete sequence of the mitochondrial DNA of the red alga *Porphyra purpurea*. Cyanobacterial introns and shared ancestry of red and green algae. Plant Cell 11(9):1675–1694.

Bushnell B. 2014. BBMap: a fast, accurate, splice-aware aligner. Berkeley (CA): Lawrence Berkeley National Lab (LBNL). Available from: https://www.osti.gov/biblio/1241166-bbmap-fast-accurate-splice-aware-aligner.

Butterfield NJ. 2000. *Bangiomorpha pubescens* n. gen., n. sp.: implications for the evolution of sex, multicellularity, and the Mesoproterozoic/Neoproterozoic radiation of eukaryotes. Paleobiology 26(3):386–404.

Cao M, Bi G, Mao Y, Li G, Kong F. 2018. The first plastid genome of a filamentous taxon '*Bangia*' sp. OUCPT-01 in the Bangiales. Sci Rep. 8:10688.

Carver T, Harris SR, Berriman M, Parkhill J, McQuillan JA. 2012. Artemis: an integrated platform for visualization and analysis of high-throughput sequence-based experimental data. Bioinformatics 28(4):464–469.

Chan PP, Lowe TM. 2019. tRNAscan-SE: searching for tRNA genes in genomic sequences. In: Kollmar M, editor. Gene prediction. Methods Mol Biol. Vol. 1962. New York (NY): Humana. pp. 1–14.

Ciniglia C, Yoon HS, Pollio A, Pinto G, Bhattacharya D. 2004. Hidden biodiversity of the extremophilic Cyanidiales red algae. Mol Ecol. 13(7):1827–1838.

Copertino DW, Hallick RB. 1993. Group II and group III introns of twintrons: potential relationships with nuclear pre-mRNA introns. Trends Biochem Sci. 18(12):467–471.

Criscuolo A, Gribaldo S. 2010. BMGE (Block Mapping and Gathering with Entropy): a new software for selection of phylogenetic informative regions from multiple sequence alignments. BMC Evol Biol. 10(1):210.

Dabbagh N, Bennett MS, Triemer RE, Preisfeld A. 2017. Chloroplast genome expansion by intron multiplication in the basal psychrophilic euglenoid *Eutreptiella pomquetensis*. PeerJ. 5:e3725.

Darling ACE, Mau B, Blattner FR, Perna NT. 2004. Mauve: multiple alignment of conserved genomic sequence with rearrangements. Genome Res. 14(7):1394–1403.
Dierckxsens N, Mardulyn P, Smits G. 2017. NOVOPlasty: de novo assembly of organelle genomes from whole genome data. Nucleic Acids Res. 45(4):e18.
Ekseth OK, Kuiper M, Mironov V. 2014. orthAgogue: an agile tool for the rapid prediction of orthology relations. Bioinformatics 30(5): 734–736.
Gabrielson PW, Garbary DJ, Scagel RF. 1985. The nature of the ancestral red alga: inferences from a cladistic analysis. Biosystems 18(3): 335–346.
Glöckner G, Rosenthal A, Valentin K. 2000. The structure and gene repertoire of an ancient red algal plastid genome. J Mol Evol. 51(4):382–390.
Guiry MD, Guiry GM. 2021. Algaebase. Galway: AlgaeBase World-wide electronic publication, National University of Ireland. Available from: https://www.algaebase.org (Last accessed August 25, 2021).
Hoang DT, Chernomor O, von Haeseler A, Minh BQ, Vinh LS. 2018. UFBoot2: improving the ultrafast bootstrap approximation. Mol Biol Evol. 35(2):518–522.
Jacquier A, Dujon B. 1985. An intron-encoded protein is active in a gene conversion process that spreads an intron into a mitochondrial gene. Cell 41(2):383–394.
Janouškovec J, et al. 2013. Evolution of red algal plastid genomes: ancient architectures, introns, horizontal gene transfer, and taxonomic utility of plastid markers. PLoS One 8(3):e59001.
Jin J-J, et al. 2020. GetOrganelle: a fast and versatile toolkit for accurate de novo assembly of organelle genomes. Genome Biol. 21(2): 241.
Johnson M, et al. 2008. NCBI BLAST: a better web interface. Nucleic Acids Res. 36(suppl_2):W5-9.
Katoh K, Kuma K, Miyata T, Toh H. 2005. Improvement in the accuracy of multiple sequence alignment program MAFFT. Genome Inform. 16(1):22–33.
Khan H, Archibald JM. 2008. Lateral transfer of introns in the cryptophyte plastid genome. Nucleic Acids Res. 36(9):3043–3053.
Kim D, et al. 2022. Group II intron and repeat-rich red algal mitochondrial genomes demonstrate the dynamic recent history of autocatalytic RNAs. BMC Biol. 20(1):2.
Lagesen K, et al. 2007. RNAmmer: consistent and rapid annotation of ribosomal RNA genes. Nucleic Acids Res. 35(9):3100–3108.
Lambowitz AM, Zimmerly S. 2011. Group II Introns: mobile ribozymes that invade DNA. Cold Spring Harb. Perspect Biol. 3(8):a003616.
Larsson A. 2014. AliView: a fast and lightweight alignment viewer and editor for large datasets. Bioinformatics 30(22): 3276–3278.
Lartillot N, Philippe H. 2004. A Bayesian mixture model for across-site heterogeneities in the amino-acid replacement process. Mol Biol Evol. 21(6):1095–1109.
Lartillot N, Rodrigue N, Stubbs D, Richer J. 2013. PhyloBayes MPI: phylogenetic reconstruction with infinite mixtures of profiles in a parallel environment. Syst Biol. 62(4):611–615.
Leclercq S, Giraud I, Cordaux R. 2011. Remarkable abundance and evolution of mobile group ii introns in *wolbachia* bacterial endosymbionts. Mol Biol Evol. 28(1):685–697.
Li H, et al. 2009. The Sequence Alignment/Map format and SAMtools. Bioinformatics 25(16):2078–2079.
Li H. 2018. Minimap2: pairwise alignment for nucleotide sequences. Bioinformatics 34(18):3094–3100.
Liu W, et al. 2020b. Subchromosome-scale nuclear and complete mitochondrial genome characteristics of *Morchella crassipes*. Int J Mol Sci. 21(2):483.
Liu SL, Chiang Y-R, Yoon HS, Fu H-Y. 2020a. Comparative genome analysis reveals *Cyanidiococcus* gen. nov., a new extremophilic red algal genus sister to *Cyanidioschyzon* (Cyanidioschyzonaceae, Rhodophyta). J Phycol. 56(6):1428–1442.
Lu S, et al. 2020. CDD/SPARCLE: the conserved domain database in 2020. Nucleic Acids Res. 48(D1):D265–D268.
Marçais G, et al. 2018. MUMmer4: a fast and versatile genome alignment system. PLOS Comput Biol. 14(1):e1005944.
Martin W, et al. 1998. Gene transfer to the nucleus and the evolution of chloroplasts. Nature 393:162–165.
Muñoz-Gómez SA, et al. 2017. The new red algal subphylum Proteorhodophytina comprises the largest and most divergent plastid genomes known. Curr Biol. 27(11):1677–1684.e4.
Nan F, et al. 2017. Origin and evolutionary history of freshwater Rhodophyta: further insights based on phylogenomic evidence. Sci Rep. 7:2934.
Nguyen L-T, Schmidt HA, von Haeseler A, Minh BQ. 2015. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol Biol Evol. 32(1):268–274.
Nurk S, Meleshko D, Korobeynikov A, Pevzner PA. 2017. metaSPAdes: a new versatile metagenomic assembler. Genome Res. 27(5): 824–834.
Perrineau M-M, Price DC, Mohr G, Bhattacharya D. 2015. Recent mobility of plastid encoded group II introns and twintrons in five strains of the unicellular red alga *Porphyridium*. PeerJ. 3:e1017.
Plant AL, Gray JC. 1988. Introns in chloroplast protein-coding genes of land plants. Photosynth Res. 16(1):23–39.
Pombert J-F, Otis C, Lemieux C, Turmel M. 2005. The chloroplast genome sequence of the green alga *Pseudendoclonium akinetum* (Ulvophyceae) reveals unusual structural features and new insights into the branching order of chlorophyte lineages. Mol Biol Evol. 22(9):1903–1918.
Preuss M, Verbruggen H, West JA, Zuccarello GC. 2021. Divergence times and plastid phylogenomics within the intron-rich order Erythropeltales (Compsopogonophyceae, Rhodophyta). J Phycol. 57(3):1035–1044.
Prjibelski A, Antipov D, Meleshko D, Lapidus A, Korobeynikov A. 2020. Using SPAdes de novo assembler. Curr Protoc Bioinformatics 70(1): e102.
Putintseva YA, et al. 2020. Siberian larch (*Larix sibirica* Ledeb.) mitochondrial genome assembled using both short and long nucleotide sequence reads is currently the largest known mitogenome. BMC Genomics 21(1):654.
Qiu H, Yoon HS, Bhattacharya D. 2016. Red algal phylogenomics provides a robust framework for inferring evolution of key metabolic pathways. PLoS Curr. 8, doi:10.1371/currents.tol.7b037376e6d84a1be34af756a4d90846.
Rambaut A, Drummond AJ, Xie D, Baele G, Suchard MA. 2018. Posterior summarization in Bayesian phylogenetics using Tracer 1.7. Syst Biol. 67(5):901–904.
Robart AR, Zimmerly S. 2005. Group II intron retroelements: function and diversity. Cytogenet Genome Res. 110(1-4):589–597.
Saunders GW, Hommersand MH. 2004. Assessing red algal supraordinal diversity and taxonomy in the context of contemporary systematic data. Am J Bot. 91(10):1494–1507.
Schwartz RM, Dayhoff MO. 1978. Origins of prokaryotes, eukaryotes, mitochondria, and chloroplasts. Science 199(4327):395–403.
Sheveleva EV, Hallick RB. 2004. Recent horizontal intron transfer to a chloroplast genome. Nucleic Acids Res. 32(2):803–810.
Shimodaira H. 2002. An approximately unbiased test of phylogenetic tree selection. Syst Biol. 51(3):492–508.
Slater GSC, Birney E. 2005. Automated generation of heuristics for biological sequence comparison. BMC Bioinformatics 6(1):31.

Smith DR. 2018. *Haematococcus lacustris*: the makings of a giant-sized chloroplast genome. AoB PLANTS 10(5):ply058.
Smith DR. 2020. Can green algal plastid genome size be explained by DNA repair mechanisms? Genome Biol Evol. 12(2):3797–3802.
Smith DR, Lee RW. 2009. The mitochondrial and plastid genomes of *Volvox carteri*: bloated molecules rich in repetitive DNA. BMC Genomics 10:132.
Tajima N, et al. 2014. Analysis of the complete plastid genome of the unicellular red alga *Porphyridium purpureum*. J Plant Res. 127(3): 389–397.
Tang Y, Horikoshi M, Li W. 2016. ggfortify: unified interface to visualize statistical results of popular R packages. R J. 8(2):474.
The UniProt Consortium. 2015. UniProt: a hub for protein information. Nucleic Acids Res. 43(D1):D204–D212.
van Dongen SM. 2000. Graph clustering by flow simulation [Doctoral dissertation]. [Utrecht (Netherlands)]: University Utrecht.
Wang H-C, Minh BQ, Susko E, Roger AJ. 2018. Modeling site heterogeneity with posterior mean site frequency profiles accelerates accurate phylogenomic estimation. Syst Biol. 67(2): 216–235.
West JA, McBride DL1999. Long-term and diurnal carpospore discharge patterns in the Ceramiaceae Rhodomelaceae and Delesseriaceae (Rhodophyta). Hydrobiol. 398-399:101–114.
Wick RR, Judd LM, Gorrie CL, Holt KE. 2017. Unicycler: resolving bacterial genome assemblies from short and long sequencing reads. PLOS Comput Biol. 13(6):e1005595.
Wick RR, Schultz MB, Zobel J, Holt KE. 2015. Bandage: interactive visualization of de novo genome assemblies. Bioinformatics 31(20): 3350–3352.
Wicke S, Schneeweiss GM, dePamphilis CW, Müller KF, Quandt D. 2011. The evolution of the plastid chromosome in land plants: gene content, gene order, gene function. Plant Mol Biol. 76(3): 273–297.
Wickham H. 2016. ggplot2: elegant graphics for data analysis. New York: Springer-Verlag. Available from: https://ggplot2.tidyverse.org.
Woelkerling WMJ. 1990. An Introduction. In: Cole KM, Sheath RG, editors. Biology of the Red Algae. New York (NY): Cambridge University Press. pp. 1–6.
Yang EC, et al. 2015. Highly conserved mitochondrial genomes among multicellular red algae of the Florideophyceae. Genome Biol Evol. 7(8):2394–2406.
Yoon HS, Muller KM, Sheath RG, Ott FD, Bhattacharya D. 2006. Defining the major lineages of red algae (Rhodophyta). J Phycol. 42(2):482–492.
Yu G. 2020. Using ggtree to visualize data on tree-like structures. Curr Protoc Bioinformatics 69(1):e96.
Zhang X, et al. 2019. The mitochondrial and chloroplast genomes of the green alga *Haematococcus* are made up of nearly identical repetitive sequences. Curr Biol. 29(15):R736–R737.

Associate editor: Dr. Daniel Sloan

5 - Evaluating serial endosymbiosis hypotheses in the light of red plastid phylogenomics

5.1 . Context

Introduction and objectives

Lineages such as land plants, green algae and red algae have gained their plastid form an endosymbiosis between their common ancestor and a cyanobacterium. Many other eukaryotic lineages instead harbor a complex plastid, derived from a green or red alga. The host and the plastid donor are well identified for green complex plastids (Jackson, Knoll, et al. 2018). In contrast, many lineages such as diatoms and coccolithophores harbor plastids of red algal origin but how exactly they acquired them remains unclear (Archibald 2015). The chromalveolate hypothesis proposes that the four different lineages containing complex red plastids (Myzozoa, Ochrophyta, Cryptophyta and Haptophyta; together MOCHa) form a monophyletic group, and that their common ancestor acquired the red algal plastid (Cavalier-Smith 1999). However, it has become clear that many plastid-lacking lineages are also descendants of this common ancestor, implying many plastid losses in these groups (Burki 2017). Some studies even suggest that cryptophytes are more closely related to red algae than to other MOCHa lineages, which would make the chromalveolate hypothesis impossible as red algae would be descendants from the common ancestor of MOCHa (Burki, Okamoto, et al. 2012; Burki, Kaplan, et al. 2016; Cenci et al. 2018; Schön et al. 2021; Yazaki et al. 2022). Other hypotheses, known as serial endosymbiosis hypotheses, have instead been proposed that involve plastid transfers between MOCHa lineages (Sanchez-Puerta and Delwiche 2008; Bodył et al. 2009; Dorrell and Smith 2011; Stiller, Schreiber, et al. 2014).

Serial endosymbiosis hypotheses mainly differ in which endosymbiotic events occurred and in which order. For example, Bodył et al. (2009) propose that an ancestor of cryptophytes first acquired its plastid from red algae, which was then first transferred to ochrophytes, and later again from cryptophytes to haptophytes. Further transfers from haptophytes and ochrophytes to different myzozoan lineage are proposed to have occurred as well. These kinds of scenarios should result in different plastid phylogenies, with the relationships among MOCHa lineages varying between these scenarios. However, these hypotheses have not yet been tested thoroughly. Furthermore, it has been shown based on eukaryotic phylogenies, calibrated with known fossils, that serial endosymbiosis hypotheses are possible as lineages that underwent endosymbiosis overlap in geological time (Strassert, Irisarri, et al. 2021). Arguably, estimate ages for plastid transfer events found in plastid phylogenies should line up with the estimated ages of their hosts, but this has not been tested yet.

Results and discussion

We created two phylogenomic data sets, one consisting of plastid encoded proteins of red algal red complex-plastid bearing lineages, and another with the same lineages plus green algae and glaucophytes. The taxon sampling of red algae, especially of Proteorhodophytina, is largely improved due to our (Chapter 4) and previous studies (Muñoz-Gómez, Mejía-Franco, et al. 2017). Using Bayesian inference (BI), the red complex plastids form a monophyletic group within the red algae, as previously reported (Yoon, Hackett, Pinto, et al. 2002; Muñoz-Gómez, Mejía-Franco, et al. 2017). However, using maximum likelihood (ML), ochrophytes place instead as sister to all other red plastids (including red algae). We show that this placement is however sensitive to some extent to fast evolving site-removal. As there are amino acid preferences among all red plastid-bearing groups, we decided to recode the data which however still gave no consistent result. From this we argue that a single red algal origin still fits best, as the probably better fitting model in BI analyses (CAT+GTR+G4) recover the red complex plastids as monophyletic. Of course, both the chromalveolate hypothesis and serial endosymbiosis hypotheses assume this.

We also find that haptophyte and cryptophyte plastids are closely related, which is well supported in BI and ML frameworks, and remains as well after the removal of fast-evolving sites and amino acid recoding to reduce bias. Serial endosymbiosis hypotheses that propose a plastid transfer from the stem lineage of haptophytes to the stem lineage of ochrophytes (Sanchez-Puerta and Delwiche 2008; Stiller, Schreiber, et al. 2014) must in light of this evidence be refuted. Instead, our data favor the hypothesis proposed by Bodył et al., although many uncertainties remain.

Finally, we have made five sets of calibrations to calibrate the plastid tree of life. These sets vary in their use of controversial fossils such as *Bangiomorpha* and *Proterocladus* (Butterfield 2000; Parfrey et al. 2011; Eme et al. 2014; Tang et al. 2020), as well as several maximum constraints that are up to debate (Hedges et al. 2018; Morris et al. 2018a,b). These analyses are still running, but will help us estimate when certain plastid transfers occurred.

5.2 . Manuscript

Evaluating serial endosymbiosis hypotheses in the light of red plastid phylogenomics
(In prep.)

Fabian van Beveren[1], Laura Eme[1], Purificación López-García[1], Guifré Torruella[1], Line Le Gall[2], and David Moreira[1]

[1]Ecologie Systématique Evolution, Centre National de la Recherche Scientifique—CNRS, Université Paris-Saclay, AgroParisTech, Gif-sur-Yvette, France
[2]Institut de Systématique, Évolution, Biodiversité (ISYEB), Muséum National d'Histoire Naturelle, CNRS, Sorbonne Université, EPHE, Université des Antilles, Paris, France

Evaluating serial endosymbiosis hypotheses in the light of red plastid phylogenomics.
(in prep.)

Fabian van Beveren[1], Laura Eme[1], Purificación López-García[1], Guifré Torruella[1], Line Le Gall[2], and David Moreira[1]

[1] Ecologie Systématique Evolution, Centre National de la Recherche Scientifique—CNRS, Université Paris-Saclay, AgroParisTech, Gif-sur-Yvette, France
[2] Institut de Systématique, Évolution, Biodiversité (ISYEB), Muséum National d'Histoire Naturelle, CNRS, Sorbonne Université, EPHE, Université des Antilles, Paris, France

Abstract

The origin of the red algal-derived plastids found in several eukaryotic lineages remains one of the major open enigmas in eukaryotic evolution. The original idea of a single origin of all these groups (the 'chromalveolate hypothesis') has been progressively abandoned in favor of serial endosymbiosis scenarios, which suggest that secondary red plastids have been transferred between these different plastid-bearing lineages. These scenarios imply specific phylogenetic relationships among these lineages in plastid phylogenies. Here, we test these hypothetical relationships using plastid sequence data from a diverse set of red algae, cryptophytes, haptophytes, and ochrophytes. We found a tight relationship between haptophyte and cryptophyte plastids, which contradicts several serial endosymbiosis scenarios. Moreover, we show that ochrophytes have an unstable position in plastid phylogenies, which is compatible with different scenarios, including a separate plastid acquisition in this lineage. Finally, we used a comprehensive set of fossil calibration points to estimate when plastids were transferred between lineages. These molecular dating analyses are still ongoing, and will reveal in which geological period plastids were transferred. We propose that two hypotheses best explain the known diversity of red algal plastids in light of our results.

Introduction

Plants and algae photosynthesize using a cyanobacterial-derived organelle known as the plastid. Cyanobacteria have been converted into a primary plastid at least two times, likely once in the common ancestor of Archaeplastida (glaucophytes, red algae, and green algae plus land plants) and more recently in a common ancestor of photosynthetic *Paulinella* species (Marin et al. 2005; Archibald 2009). Many other diverse algal lineages did not acquire their plastids directly from cyanobacteria but possess complex plastids of green or red algal origin (Burki 2017; Sibbald and Archibald 2020). Two independent cases—besides green dinoflagellates (Sarai et al. 2020)—are known for the origin of green complex plastids found in the euglenophytes and chlorarachniophytes, where both the host and symbiont lineages have been broadly identified (Jackson et al. 2018; Sibbald and Archibald 2020). By contrast, red complex plastids are present in four diverse eukaryotic groups (Myzozoa, Ochrophyta, Cryptophyta, and Haptophyta; together called MOCHa) and include ecologically significant lineages such as dinoflagellates, diatoms, and kelps (Archibald 2012; Burki 2017). Multiple hypotheses have aimed to explain the origin of these diverse lineages; a consensus has not yet been reached.

The chromalveolate hypothesis proposed that a single endosymbiotic event between a red alga and a common ancestor of MOCHa occurred (Cavalier-Smith 1999). However, that common ancestor would have had many red plastid-lacking descendants, such as oomycetes, ciliates, centrohelids, and rhizarians (Burki et al. 2020; Sibbald and Archibald 2020). In fact, this hypothesis implies that at least ten lineages have lost their plastid independently (Burki 2017). Moreover, while plastid phylogenies easily recover the monophyly of chromalveolates, it is much more challenging to recover it with nuclear or mitochondrial data, even when excluding plastid-lacking sister lineages (Baurain et al. 2010).

To deal with these discrepancies, several authors have proposed different serial endosymbiosis scenarios where the red complex plastid was acquired directly from a red alga only once and then transferred several times between the different MOCHa lineages (Sanchez-Puerta and Delwiche 2008; Bodył et al. 2009; Dorrell and Smith 2011; Petersen et al. 2014; Stiller et al. 2014). Flavors of this hypothesis differ in the endosymbiotic events proposed and their order. For example, Stiller et al. (2014) proposed that the plastid was transferred from cryptophytes to ochrophytes, followed by a transfer from ochrophytes to haptophytes. This scenario implies that ochrophytes and haptophytes should be sister lineages in plastid phylogenies. Nevertheless, such assumptions have not been extensively tested in a plastid phylogenomic framework.

Molecular dating based on nuclear genome data has also been used to estimate the origin of different eukaryotic groups, including the MOCHa lineages. Specifically, Strassert et al. (2021) have shown that several serial endosymbiosis hypotheses fit the dated tree. In a plastid phylogenetic framework, such an analysis should give similar ranges for transfer times, but this remains to be tested. With a further increase of available plastid genomes of unicellular red algal lineages (Muñoz-Gómez et al. 2017; Preuss et al. 2021; van Beveren et al. 2022), we here revisit the phylogeny of red algae and red complex plastids, and their diversification through geological time using molecular dating.

Results and Discussion

We assembled and annotated two new plastid genomes of the species *Mesophyllum expansum* MEX45 and *Lithophyllum* sp. LIT44, belonging to the red algal order Corallinales. We also annotated the plastid genome of the coccolithophore *Chrysotila dentata* MZ819921.1 (Paudel et al. 2021). To reconstruct the phylogeny of red complex plastids, two data sets were made (see Materials and Methods for details). A 'red-only' data set encompassed 140 plastid-encoded proteins of red algae and chromalveolates, consisting of 87 taxa (37 red algae, 40 ochrophytes, 6 haptophytes and 4 cryptophytes). Myzozoa were not included due to their fast-evolving plastid gene sequences, and the lack of many genes in Dinozoa and Apicomplexa (Zhang et al. 2000). A second data set, 'all plastids', consisted of 120 plastid-encoded proteins for the same taxa with additional green lineages (including euglenophytes and chlorarachniophytes) and glaucophyte sequences.

Multiple red algal origins of red complex plastids are unlikely

We reconstructed phylogenetic trees for both data sets using Bayesian inference (BI) with PhyloBayes-MPI (Lartillot et al. 2013), maximum likelihood (ML) with IQ-TREE (Minh et al. 2020), and a coalescent-based supertree analysis with ASTRAL-III (Zhang et al. 2018). BI recovered similar topologies of red algae and red plastids with both data sets (fig. 1, supplementary fig. S1). Specifically, the monophyly of red complex plastids was well supported by both the 'red-only' (0.94 posterior probability; pp) and 'all plastids' (1.0 pp) data sets. Similarly, ASTRAL recovered red complex plastids as monophyletic, but with lower support (0.62 pp) and only with the 'red-only' data set (supplementary figs. S2 and S3).

All the ML phylogenies, with both data sets and the different models, recovered a middle to highly supported placement (up to 97 ultrafast bootstrap support) of Ochrophyta as sister to Rhodophytina, Haptophyta and Cryptophyta when placing the root between Cyanidiophytina and all other red plastids (supplementary figs. S4-S6). When including the outgroup ('all plastids' data set), Ochrophyta was placed as sister to all other red plastids, including Cyanidiophytina (supplementary fig. S7-S9). A similar result was found by Muñoz-Gómez and colleagues (2017) using ML tree reconstruction, where Ochrophyta was recovered as sister to Rhodophytina, Haptophyta and Cryptophyta with several models. This may suggest a separate plastid origin in ochrophytes from a different red algal group, but this result might also be explained by a potential long branch attraction (LBA) artifact.

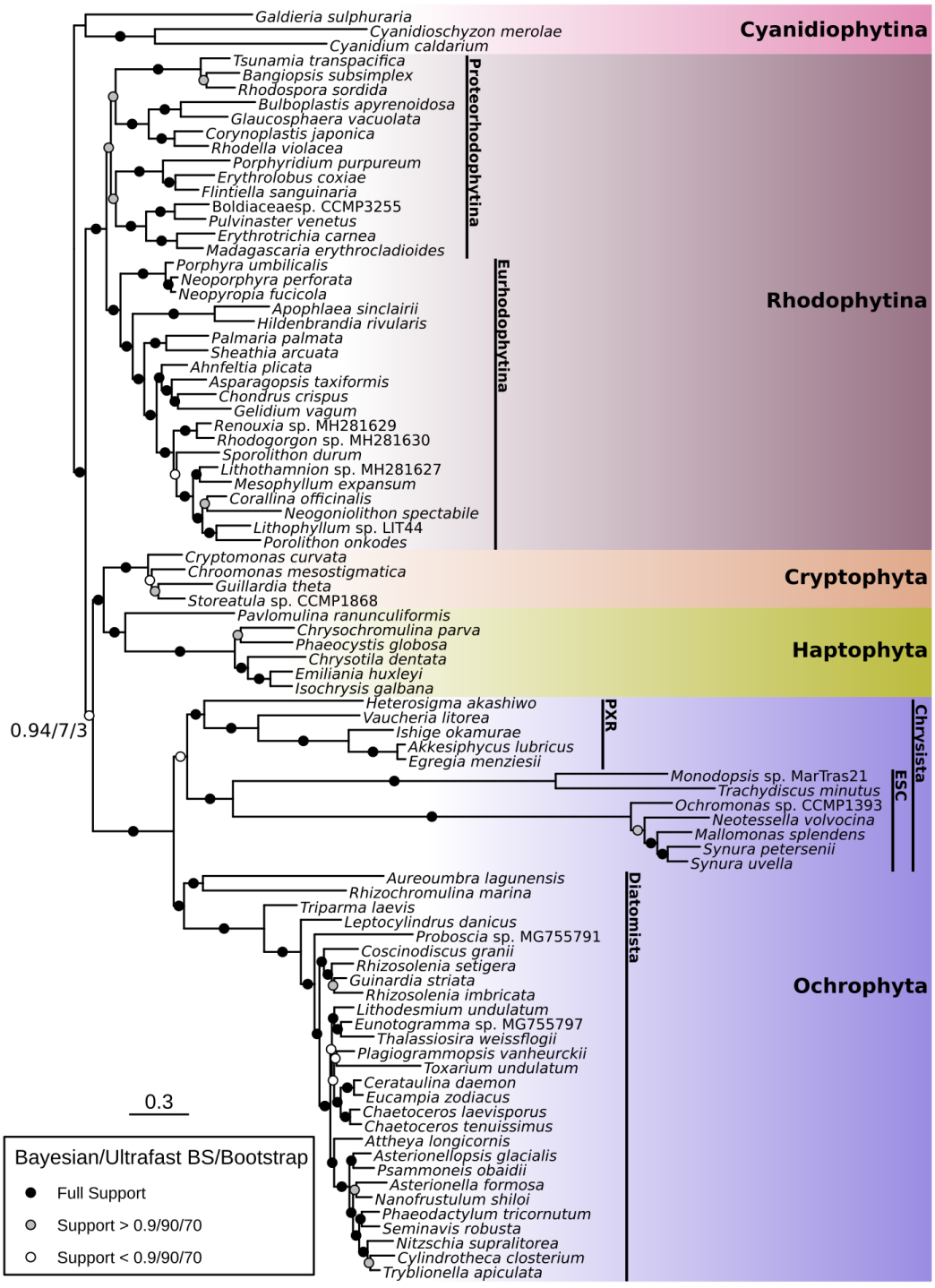


Figure 1: Phylogenetic reconstruction of red algae and lineages with red complex plastids based on plastid genome data. The tree was reconstructed with a trimmed alignment of 140 plastid-encoded protein sequences (27,658 amino acid positions), using both Bayesian inference (CAT+GTR+G4) and maximum likelihood (cpREV+C60+F+R8 with ultrafast bootstrap support and cpREV+C60+R8+PMSF with nonparametric bootstrap support). The tree topology shown is based on Bayesian inference, and was rooted between Cyanidiophytina and all other taxa. Support values are shown with colored circles (see legend), with the support for the monophyly of red complex plastids verbose, reflecting the lack of monophyly in maximum likelihood analyses. PXR = Phaeophyceae, Xanthophyceae and Raphidophyceae; ESC = Eustigmatophyceae, Synurophyceae, Chrysophyceae. The scale bar indicates the average number of substitutions per site.

Within the ochrophytes, the placement of the ESC clade (Eustigmatophyceae, Chrysophyceae and Synurophyceae) differed between BI and ML trees. The BI phylogeny (fig. 1; supplementary fig. S1) was in agreement with previous analyses, placing the ESC clade together with the PXR clade (Phaeophyceae, Xanthophyceae and Raphidophyceae) within the Chrysista (Cavalier-Smith 1986; Derelle et al. 2016). Using ML, the ESC clade was instead placed as sister to all ochrophytes (supplementary figs. S4-S9). As the root-to-tip distances are longer for the ESC clade than the other ochrophytes, the placement may be affected by LBA, and potentially result in a misplacement of Ochrophyta overall. To test this possibility in the ML framework, we removed the long-branching representatives of the ESC clade from our data set, but this still resulted in topologies with the same placement of ochrophytes (supplementary fig. S10).

The stem of Ochrophyta is longer in comparison to Cryptophyta and Haptophyta (fig.1; supplementary fig. S1) so that the possible erroneous placement could be due to fast evolution on this branch. To test this hypothesis, we removed fast evolving sites from the red-only data set in steps of 10% and calculated ultrafast bootstrap support for several groups on the resulting ML phylogenies (fig. 2). As a reference, we included the supports for Ochrophyta and Rhodophytina, which both remained at 100% even when only 30 or 40% of the slowest evolving sites remained. Only at one step was the monophyly of complex plastids moderately supported (91.3%), when half of the fastest evolving sites were removed. However, when 60% of the sites were removed, the support lowered back to just 7.6%. This may indicate that the lack of support for the monophyly of red complex plastids is not only related to the erroneous placement of the fast-evolving ESC lineages.

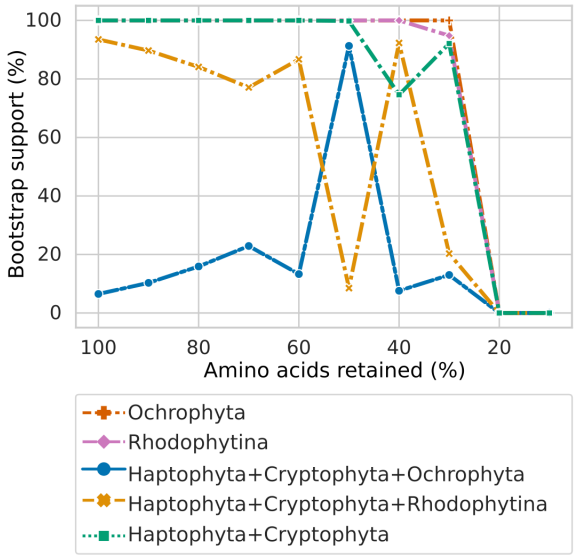


Figure 2: Change of support for several groups as fast-evolving sites were removed in steps of 10%. The complete alignment consists of 27,658 amino acids. The y-axis shows ultrafast bootstrap support based on 1000 replicates.

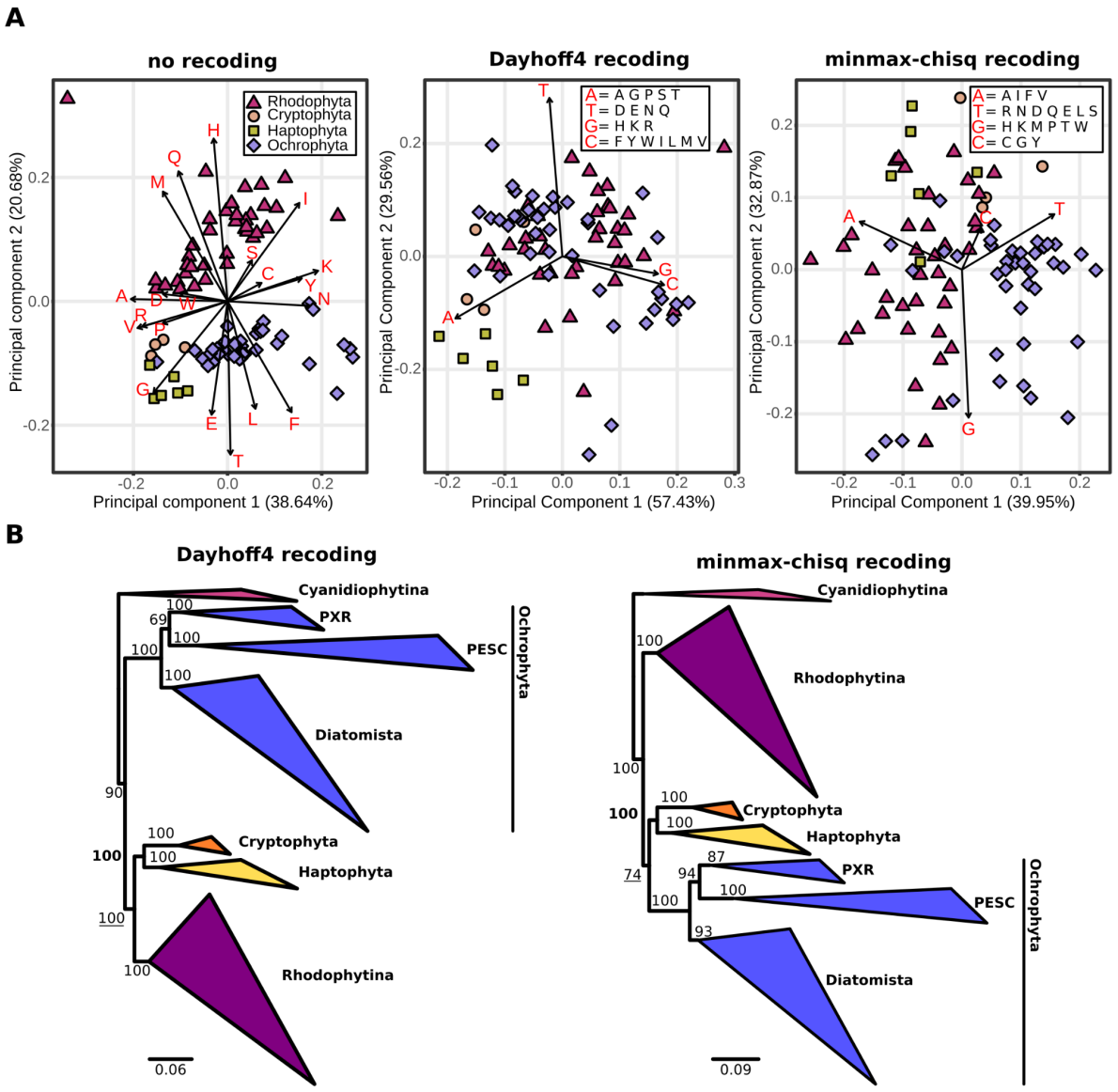


Figure 3: (A) Principal component analysis showing compositional biases in the untreated 'red-only' data set and the same data set recoded with the Dayhoff4 and minmax-chisq approaches (see main text). Left panel shows the loadings (arrows) for each amino acid and in the other two panels for the recoded categories (see legend). Note that Cysteine is absent in the four Dayhoff categories and is instead recoded to a question mark in the alignment. (B) Phylogenetic trees based on the Dayhoff (left) and minmax-chisq (right) recoded data sets, both rooted between Cyanidiophytina and all other taxa. Main groups of interest are grouped together (for the complete trees, see supplementary fig. S12 and S13). Taxon sampling is equal to that of fig. 1. Ultrafast bootstrap support is given for the shown branches, including the support for the monophyly of Cryptophyta and Haptophyta (bold) and their monophyly with either Eurhodophytina (left, underlined) or Ochrophyta (right, underlined).

Another possibility to explain the unstable position of Ochrophyta might be related to different amino acid compositional preferences between plastid genomes, resulting in potentially erroneous grouping of lineages with similar amino acid composition (Foster and Hickey 1999). A principal component analysis (PCA) of the amino acid composition of the inferred proteomes of the red-plastid lineages showed that all the different groups, not only ochrophytes, exhibit different amino acid preferences (fig. 3A). As this observation might be biased due to the absence of certain genes in some lineages, we did the same analysis using only the positions where three or less taxa have a gap, which produced a similar result (supplementary fig. S11). To alleviate the observed compositional bias, we recoded the red-only alignment using the Dayhoff (four categories) and minmax-chisq (four categories) approaches, the latter calculating an optimal recoding scheme to improve composition homogeneity (Susko and Roger 2007). Both recoded data sets produced less distinct groups in a PCA analysis (fig. 3A). After Dayhoff recoding, the corresponding ML phylogenetic tree still placed Ochrophyta as sister to Rhodophytina, Haptophyta and Cryptophyta (full support; fig. 3B; supplementary fig. S12). By contrast, the minmax-chisq recoding produced a tree that recovered the monophyly of complex plastids although with low support (74% ultrafast bootstrap, fig. 3B; supplementary fig S13).

Within the ML framework, different topologies can be tested using the approximately unbiased (AU) test (Shimodaira 2002). We made use of this to test if the monophyly of all complex red plastids was significantly worse than the original ML topology with ochrophytes branching separately (fig. 4). This alternative topology was not rejected by the AU-test using the ‘red-plastid’ data set including and excluding the fast evolving ESC clade. As the fast-evolving site removal and amino acid recoding also do not clearly reject the monophyly of red complex plastids, we consider this the most likely topology. Although these results do not necessarily rule out the possibility of several independent origins of red complex plastids from closely related red algae, the plastid phylogeny is more parsimoniously explained by a single red algal origin. Thus, our results are congruent with both serial endosymbioses hypotheses (Sanchez-Puerta and Delwiche 2008; Bodył et al. 2009; Stiller et al. 2014) and the chromalveolate hypothesis (Cavalier-Smith 1999).

The strongly supported monophyly of Haptophyta and Cryptophyta rejects some serial endosymbiosis scenarios

Serial endosymbiosis hypotheses and the chromalveolate hypothesis also evoke specific relationships between the red complex plastids. The chromalveolate hypothesis implies that the phylogeny of red complex plastids should follow the phylogeny of their hosts (Baurain et al. 2010). However, the relationships among these groups remain to be resolved (Burki et al. 2020) so we focused instead on the serial endosymbiosis hypotheses, as different published versions imply different relationships among the plastid lineages (fig. 4A). We compare four scenarios, two of which are from Sanchez-Puerta and Delwiche (2008), the other two from Bodył et al. (2009) and Stiller et al. (2014). Note that these scenarios assume that endosymbiotic events did not involve crown-lineages of the three chromist groups, so that the corresponding plastid phylogenies should result in each chromist group being monophyletic, as is the case here and in previous plastid phylogeny studies (fig. 1; Yoon et al. 2002; Dorrell et al. 2017). Under these conditions, only three plastid phylogenies are possible (fig. 4B), although the scenario where haptophytes are sister to the other lineages (haptophyte-first) has not been proposed to our knowledge. Using the ‘red-only’ data set with and without the ESC clade, our analyses do not reject the ochrophyte-first scenario as

previously mentioned, but do reject both the cryptophyte-first and haptophyte-first scenarios (fig 4B). Therefore, our topology test rejects two of the included scenarios of the serial endosymbiosis hypothesis (Sanchez-Puerta and Delwiche 2008; Stiller et al. 2014) that imply that cryptophytes would be sister to the remaining two lineages (fig. 4B). This is further strengthened by previous studies recovering ochrophytes as sister to haptophophytes and cryptophytes (Dorrell et al. 2017; Muñoz-Gómez et al. 2017). Furthermore, our previous results show that the monophyly of Haptophyta and Cryptophyta, to the exclusion of Ochrophyta, is highly supported in all our phylogenetic analyses (fig. 1). This result is resilient to the removal of fast-evolving sites (fig. 2) and to both recoding strategies (fig. 3B). Our results clearly favor hypothesis A from Sanchez-Puerta and Delwiche (2008), as well as the hypothesis from Bodył et al (2009), which both include an early tertiary plastid acquisition in Ochrophyta.

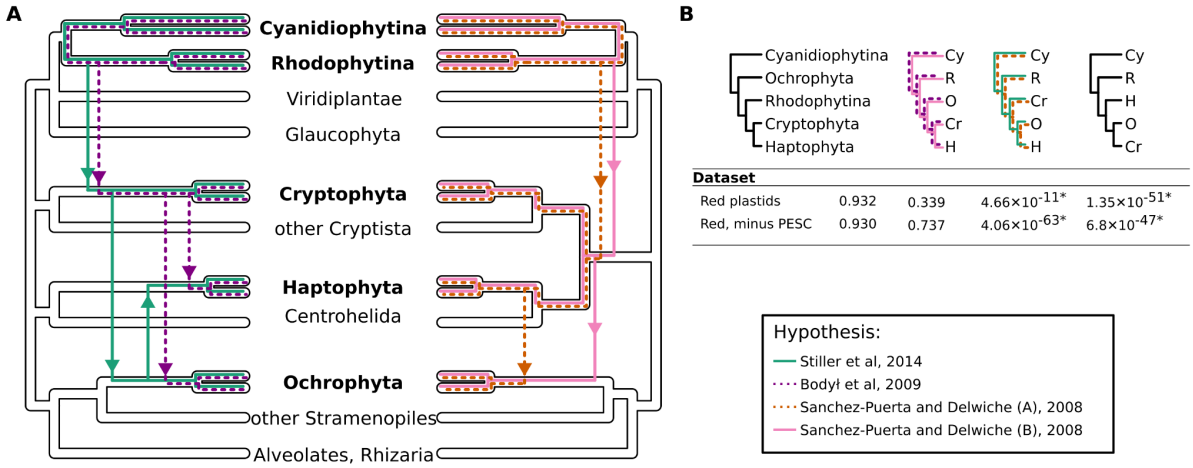


Figure 4: Topology test of the plastid phylogenetic relationships implied by several serial endosymbiosis hypotheses. (A) Four hypotheses are shown, two of which make no relevant assumption on the relationships between Haptophyta and Cryptophyta (left tree), or assume that they are more closely related to each other than to Ochrophyta (right tree). (B) Four possible plastid phylogenies (all rooted here between Cyanidiophytina and all other taxa), two of which are implied by a serial endosymbiosis hypothesis (colored). The table shows the results of the approximate unbiased (AU) test based on the 'red-only' data set and a data set excluding the fast evolving ochrophyte group ESC (see main text). Asterisks denote topologies rejected by the AU-test.

Table 1: Fossil calibrations used for molecular dating analysis. Nodes are listed as either the origin of one group or a split of multiple groups (separated by a dash). All ages are listed in millions of years ago (Ma). For details on the calibrations see supplementary information S1. Five calibration sets were made based on these data: Default: Default minimum and maximum calibrations. Default+B: Default calibrations + *Bangiomorpha.* Default+BP: Default calibrations, plus *Bangiomorpha* and *Proterocladus*. Extended: Default calibrations, plus extra maximum calibrations. Extended+BP: Extended set, plus *Bangiomorpha* and *Proterocladus*.

Default minimum calibrations	**Age (Ma)**		**Default maximum calibrations**	**Age (Ma)**
Corallinophycidae – Ahnfeltiophycidae, Rhodymeniophycidae	550.4		Root	3000.0
Sporolithales – Corallinales, Hapaldiales	133.3		Corallinophycidae	635.8
Corallinales	63.8		Dasycladales	635.8
Phaeocystales – Isochrysidales, Coccolithales	209.5		Bryopsidales	635.8
Coccolithales – Isochrysidales	180.4		Embryophyta	1042.0
Rhizosoleniales – Coscinodiscophycidae	90.0		Angiosperms	247.2
Eunotogramma – *Thalassiosira*	65.9			
Hemiaulales – Chaetocerotales	145.0			
Urneidophycidae	73.0		**Extra maximum calibrations**	**Age (Ma)**
Bacillariophycidae - Fragilariophycidae	56.0		Coccolithales – Isochrysidales	252.2
Synura – *Mallomonas*	46.4		Bacillariophyta	252.2
Dasycladales – Oltmannsiellopsidales	439.2		Bacillariophyceae	165.0
Caulerpa – *Halimeda*, *Pseudocodium*, *Rhipilia*	271.8		Bacillariophycidae	75.0
Halimeda – *Pseudocodium*, *Rhipilia*	209.5		Embryophyta[1]	515.5
Embryophyta – Zygnematophyceae	466.2		Spermatophytina	365.6
Embryophyta	426.9		Pinales	321.4
Jungermanniopsida - Marchantiopsida	405.0			
Marchantiales - Lunulariales	226.4			
Bryophyta	330.7		**Extra minimum: *Bangiomorpha***	**Age (Ma)**
Lycopodiophytina - Polypodiophytina, Spermatophytina	420.7		Rhodophyta	1030.0
Isoetales, Selaginellales – Lycopodiales	392.1			
Spermatophytina - Polypodiophytina	385.5			
Ophioglossidae, Marattiidae - Polypodiidae	318.7		**Extra minimum: *Proterocladus***	**Age (Ma)**
Pinopsida – Gnetopsida	308.1		Ulvophyceae[2] – Trebouxiophyceae	940.4
Eudicots – Monocots	125.9			
Calycanthus – *Chimonanthus*	85.8			
Annonaceae, Magnoliaceae – Myristica	112.6			
Canellales – Piperales	125.9			
Ranunculales – Fabales, Apiales	119.6			

[1]Replaces the default maximum calibration for Embryophyta of 1042.0 Ma.

[2]Including the green complex plastids of chlorarachniophytes, which are ulvophyte-derived (Jackson et al. 2018).

Dating the plastid tree of life

Not only the topology of phylogenies can be used to test scenarios on the origin of red complex plastids. One option recently explored is the use of the fossil record and molecular dating to test whether certain scenarios are possible, i.e. whether it is likely that the proposed host and endosymbiont have lived around the same time (Strassert et al. 2021). With our updated red plastid tree of life, it can be useful to revisit this question using plastid data and see whether there is any discrepancy between these data sets. To do this, we have gathered information on the fossil record of plant and algal lineages that can be used to constrain our phylogeny (supplementary information S1). Several calibration points that we have included here were not used before for molecular dating of the algal tree of life, partly made possible by the inclusion of plastid data of the coralline algae *Mesophylum expansum* and *Lithophyllum* sp., and the coccolithophore *Chrysotila dentata* (Paudel et al. 2021). Our collected calibrations include several fossils of coralline algae such as *Karpathia sphaerocellulosa* (Aguirre et al. 2007; Rösler et al. 2017), as well as fossils of siphonous algae such as *Caulerpa* sp. and *Palaeocymopolia silurica* (Gustavson and Delevoryas 1992; Mastik and Tinn 2015).

The selection of fossils and other data in the fossil record for constraining nodes is to some extent subjective, especially when establishing maximum ages for certain lineages. To increase the robustness of our analysis we have varied the use of certain calibration points (table 1; supplementary fig. S14) and described the reasoning for the selection of our constraints (supplementary information S1). Specifically we have varied the use of two ancient putative algal fossil species: *Bangiomorpha pubescens* (Butterfield et al. 1990; Butterfield 2000) and *Proterocladus antiquus* (Tang et al. 2020). *Bangiomorpha*, presumably belonging to the Rhodophyta, has often been regarded as the oldest *bona fide* crown-group eukaryotic fossil and is currently dated at ~1050 Ma (Butterfield 2000; Gibson et al. 2018). Nevertheless, its use in several molecular clock analyses has at times suggested that its age is too old or the phylogenetic affinity erroneous, although an older estimated age was used at the time of ~1200 Ma (Butterfield 2000; Berney and Pawlowski 2006; Eme et al. 2014). There are even older (~1600 Ma) potentially red algal fossils that are however also contested (Bengtson et al. 2017; Carlisle et al. 2021). Fossils grouped into the fossil genus *Proterocladus*, potentially a green alga, were first described in ~700 Ma fossil formations (Butterfield et al. 1994), but older fossils of ~1000 Ma have been more recently described and are classified as a single species of siphonous algae (*Proterocladus antiquus*), belonging to the Ulvophyceae (Tang et al. 2020). This finding is important as it is one of the few putative algal fossils outside of red algae that is significantly older than the many algal fossils from the Doushantuo formation (~600 Ma) (Zhang et al. 1998). Finally, another controversial calibration concerns the maximum age of land plants (Embryophyta), which in previous studies has been either set to ~515 Ma or ~1050 Ma based on the earliest occurrence of cryptopores (Morris et al. 2018) or of Precambrian sediments where land plants are expected to be found if they had originated, but no fossils are found, respectively (Clarke et al. 2011; Strother et al. 2011; Su et al. 2021). These ages have shown to give highly divergent results (Su et al. 2021), thus we varied the use of this calibration. In total, five sets were made with changes related to the above listed calibrations and several more maximum calibrations. Among them, one set of calibrations was a subset of those used by Strassert et al (2021) that are applicable to the algal groups present in our tree for a more straightforward comparison with their nuclear data-based molecular clock analysis (supplementary fig. S15).

First, we tested our data set for the presence of autocorrelation between rates using CorrTest (Tao et al. 2019), which indicated that it was likely the case (score: 0.991; p-value < 0.001). Still, for comparison we used one autocorrelated and two uncorrelated models of evolutionary rates available in PhyloBayes (Thorne et al. 1998; Drummond et al. 2006; Lepage et al. 2007). Currently, these analyses are ongoing, and after one month they have not yet gone outside of the burn-in phase as the log-likelihood is still steadily improving (Figure 5). We are looking into decreasing the length of the alignment, either by selecting clock-like genes or selecting genes with the highest taxon coverage. We can also further decrease the overall taxon sampling of the alignment if necessary.

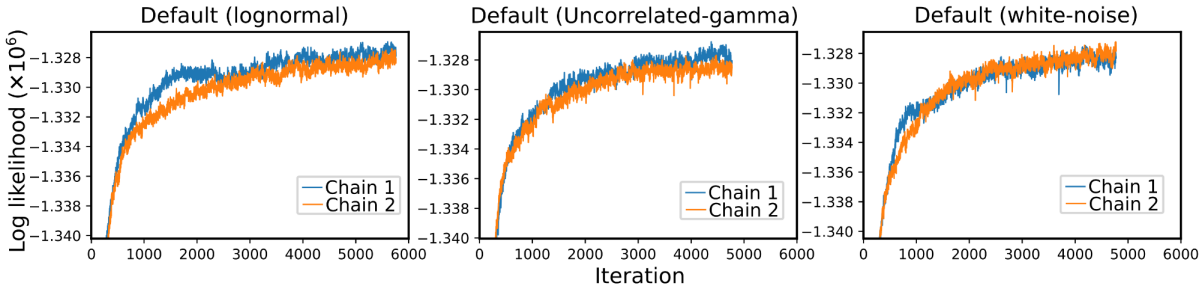


Figure 5: Change in log likelihood during molecular dating of the all plastids data set using Phylobayes. The two chains running for each data set show that the burn-in phase has not been passed after ~5000 generations. The default calibration set has been used to produce these figures, but similar results are obtained with the other calibration sets.

Once these analyses are finished, we aim to compare the results of the different calibration sets to evaluate the effect of changing some of the calibrations. We will also be able to look into the potential effects of different relaxed molecular clock models. We hope to be able to compare these results with that of previous analyses of plastid (e.g. Yoon et al. 2002) and nuclear (e.g. Strassert et al. 2021) data. At this stage, it is difficult to say what the results might look like, although it is most probable that the presence or absence of *Proterocladus* and *Bangiomorpha* will have the most significant effect on the estimated ages. Such results could be interpreted as potential errors with these fossils, although they may instead suggest a general lack of sampling of Precambrian algal fossils. As for the ages estimated for the origin of red complex plastids, comparisons can be made between potential time of transfers given by our plastid tree, and the time of diversification inferred for host trees.

Methods

Algae collection, DNA extraction and sequencing, and plastid genome assembly

Samples of the coralline red algae *Mesophyllum expansum* and *Lithophyllum* sp. were collected in May of 2016 from the mediterranean sea by Sausset-les-Pins and Marseille, France. A sample of each alga was submerged in a Petri dish with 50% HCl for 5 seconds to remove part of the calcareous material and rinsed and cut in pieces. DNA was then extracted using the DNAPowerSoil kit (Qiagen) following the manufacturer protocol. Extracted DNA was sent for sequencing with Illumina HiSeq 2500, by CNAG-CRG (Barcelona, Spain). Read quality was verified using fastqc v0.11.9 (Andrews 2016). Adapter sequences were trimmed using *bbduk.sh* (options: ktrim=r k=23 mink=11 hdist=1 tpe tbo) of the BBTools suite v38.93 (Bushnell 2014), and trimmed reads were verified using fastqc. To

assemble the plastid genome, trimmed reads of both species were aligned against the *rbcL* nucleotide sequence of *Sporolithon durum* (KT266785.1) using *bbmap.sh* of the BBTools suite (options: reads=100000 nodisk). The trimmed reads were assembled into plastid genomes with NOVOPlasty v4.2, using the mapped reads as a seed (Dierckxsens et al. 2017). The complete assembly pipeline can be found in the Fighsare repository (analysis/01_assembly_and_annotation/).

Plastid annotation of *Mesophyllum expansum*, *Lithophyllum* sp., and *Chrysotila dentata*
In this study, we have included the unannotated plastid genome of the haptophyte *Chrysotila dentata* (MZ819921.1 on GenBank, retrieved February 10, 2022). Although considered closely related to *Phaeocystis antarctica* (Order Phaeocystales) by the original authors based on whole plastome phylogeny (Paudel et al. 2021), this genus is considered as part of the order Coccolithales (Andersen et al. 2014). We annotated the genome of *C. dentata*, *M. expansum,* and *Lithophyllum* sp. using the online web tool GeSeq (Tillich et al. 2017), accessed on October 14, 2021, for the two red algae, and on February 14, 2022 for *C. dentata*. Genome annotations were manually verified and adjusted within the annotation viewer Artemis 18.0.2 (Carver et al. 2012). The annotated protein sequences RbcL and TufA of *C. dentata* were used to confirm its position in the tree of haptophytes, using available RbcL and TufA sequences of haptophytes in Uniprot (The UniProt Consortium 2015), which showed its affinity to other *Chrysotila* spp. within the clade of Coccolithales (supplementary fig. S16 and S17). The complete annotation pipeline can be found in the Fighsare repository (analysis/01_assembly_and_annotation/).

Genome collection
Plastid genome sequences were gathered from GenBank (Benson et al. 2013) to represent all major lineages of photosynthetic eukaryotes (last accessed February 21, 2022). We also included the then not yet public 25 plastid genomes of unicellular and filamentous red algae determined by van Beveren et al. (2022). Plastid-encoded gene sequences with no complete genome available for six siphonous green algae from the order Dasycladales were also collected (*Acetabularia acetabulum*, *A. peniculus*, *Neomeris* sp. HV02668, *Dasycladus* sp. HV04045, *Halicoryne* sp. HV04044, *Batophora oerstedii* and *B. occidentalis*). To further expand the taxon sampling of Dasycladales, we gathered plastid protein sequences by aligning known Dasycladales plastid-encoded protein sequences to proteomes provided by Jon Bråte (*Acetabularia crenulata*, *Bornetella oligospora*, *Neomeris dumetosa,* and *Parvocaulis polyphysoides*). In total, we gathered plastid-encoded protein sequences from 440 taxa including *C. dentata*, *M. expansum,* and *Lithophyllum* sp. Taxonomic information for all algae was added following AlgaeBase (last accessed August 30, 2021; Guiry and Guiry 2022), and for land plants following the Integrated Taxonomic Information System (ITIS; https://www.itis.gov/ ; last accessed February 21, 2022) with some exceptions (supplementary Table 1) . All code to gather data from Genbank, and to gather plastid-encoded protein sequences from proteomes of Dasycladales is available in the Figshare repository (data/ and analysis/00_dascalydales/).

To alleviate the computational burden due to the size of the data set, we removed plastid genomes based on their similarity to other included plastid genomes. With this purpose, we reconstructed a preliminary tree using RbcL and PsaA protein sequences (excluding Dasycladales and a set of Angiosperms) and excluded 139 plastid genomes (supplementary fig. S18; analysis/02_prelim_tree and analysis/03_protein_dataset in Figshare). The resulting data set contained 301 taxa.

Phylogenomic data set preparation

Protein sequences of all selected 301 taxa were gathered and a blast database was made using *makeblastdb* v2.12.0+ (Altschul et al. 1990). To create orthologous groups, PSI-blast (Altschul et al. 1997) was run with options -evalue 10 -outfmt 6 -max_target_seqs 200 -seg yes -soft_masking true -use_sw_tback -word_size 2 -matrix BLOSUM45. Homology relations were estimated based on the blast results using orthAgogue with option -u (Ekseth et al. 2014). We used mcl (van Dongen 2000) with multiple inflation values and chose the value 1.5 to continue as it seemed to have a good balance between number of paralogs included and amount of true orthologs likely missing. Genes erroneously split in multiple orthologous groups were merged. The groups were further improved by making single gene alignments with MAFFT G-INS-i v7.487 (Katoh et al. 2005; Katoh and Standley 2013) and single gene trees with IQ-TREE 2.2.0.3 with the options -fast -alrt 1000 -m cpREV+F+G (Nguyen et al. 2015; Minh et al. 2020). All changes to be made to the orthologous groups were listed in a table and a custom python script was used to apply the changes (analysis/03_protein_dataset in Figshare). We kept 154 orthologous groups, each including at least 60 taxa. Then, we ran PREQUAL v1.02 (Whelan et al. 2018) with default settings to remove errors in the alignments, such as those caused by sequencing errors, which were manually verified but unaltered. Alignments were trimmed using BMGE v1.12 with the BLOSUM30 model (Criscuolo and Gribaldo 2010). All code related to the creation of orthologous groups can be found in the Figshare repository (analysis/03_protein_dataset).

We created a preliminary tree using the 20 orthologous groups with the highest taxon sampling for our 301 selected taxa. After testing multiple combinations of taxa and genes included, we kept two data sets: one with only red algae and red complex plastids (“red only”, 87 taxa, 140 genes, 27,658 characters) and another including the same red lineages, plus green lineages and glaucophytes, but with fewer genes included (“all plastids”, 176 taxa, 120 genes, 22,209 characters). All code related to the subsampling and creation of the final data sets can be found in the Figshare repository (analysis/04_final_subsampling).

Phylogenomic analysis

To reconstruct phylogenies using maximum likelihood (ML), we made use of IQ-TREE (Nguyen et al. 2015). A modeltest was performed on the ‘red-only’ data set to compare the fit of cpREV+C60+F+R8 versus LG+C60+F+R8 (Le and Gascuel 2008), with the best model being cpREV+C60+F+R8 (Adachi et al. 2000). A similar test was done for the ‘all plastids’ data set, where we compared the same models but with R9, the best model being cpREV+C60+F+R9. Both the cpREV and LG based models were used on both data sets with IQ-TREE including 1000 ultrafast bootstrap replicates and 1000 SH-aLRT replicates (Guindon et al. 2010; Hoang et al. 2018). The output tree for each analysis with cpREV was then used as a guide tree to reconstruct the phylogenies with the same models plus PMSF with 100 standard nonparametric bootstrap replicates (Wang et al. 2018).

The phylogenies for both data sets were also reconstructed using bayesian inference (BI) with PhyloBayes-MPI 1.8c (Lartillot et al. 2013), using the CAT+GTR+G4 model. For each data set, three chains were run up to 10,000 generations (supplementary figs. S19-S24). Convergence was not reached after 10,000 generations, but very few topological differences remained. The burn-in was estimated by inspecting the trace files with Tracer 1.7.1 (Rambaut et al. 2018) and by plotting the trace values using a Python script (see /analysis/ in Figshare.), a burn-in of 1000 was chosen for both data sets. Bayesian posteriors

for the branches were calculated using bpcomp and the effective size of variables with tracecomp (Lartillot et al. 2013).

To test whether the different placement of ochrophytes in ML analyses are due to the long branches of the ESC group, we reran the ML analysis on the red-only data set with the same alignment and models, but excluding the seven ESC lineages included in the original data sets. IQ-TREE was run to reconstruct the phylogeny using the cpREV+C60+F+R8 model.

To assess which relationships between red algae and red complex plastids are statistically supported we reconstructed constrained trees with IQ-TREE of four different topologies: ochrophyte-outside (Ochrophyta, (Rhodophytina,(Cryptophyta, Haptophyta))); ochrophyte-first (R, (O, (C, H))); haptophyte-first (R, (H, (C, O))); cryptophyte-first (R, (C, (H, O))). The resulting tree topologies were tested within IQ-TREE on the red-only data set using the options -n 0 -zb 10000 -zw -au. The same constrained trees and topology test were done with the red-only data set excluding the ESC lineages.

We reconstructed the phylogenies of both data sets also using a concatenation-free approach, the multi-species coalescent method ASTRAL (Zhang et al. 2018). First we made single gene trees using IQ-TREE including a modeltest (-mset LG,cpREV -mfreq FU,F,FO -mrate E,I,G,I+G) to which we also added the following mixture models: LG+C20+F+G, LG+C40+F+G, LG+C60+F+G, cpREV+C20+F+G, cpREV+C40+F+G, and cpREV+C60+F+G. Tree reconstruction included ultrafast bootstrap replicates, which were used to make a collapsed set of trees, where low support (<10) nodes collapsed as this can improve accuracy slightly (Zhang et al. 2018). ASTRAL was run with the original ML trees, the ML trees with low-supported nodes collapsed, and a run with the ML trees and the 1000 ultrafast bootstrap trees to include bootstrap values. All related code can be found in the Figshare repository (analysis/07_supertree).

When constructing the phylogeny of red plastids with the model cpREV+C60+F+R8+PMSF, we used the --rate parameter to get the evolutionary rates of each position. We used these data to remove the fastest evolving sites in steps of 10%. Phylogenies for each group were reconstructed using the cpREV+C60+F+R8 model with ultrafast bootstrap, and the bootstrap support for several groups (fig. 2) was determined and plotted (analysis/10_fast_evolving_site_remoal on Figshare).

To recode the red plastid alignment, we made use of Dayhoff recoding in four categories. We also created a separate four category recoding using minmax-chisq (Susko and Roger 2007). The original data set was trimmed again using BMGE (-h 1 -w 1 -g 0.04) to see if the compositional bias was due to positions not present in some of the lineages. The principal component analysis was done using prcomp (with scale. = T) in R v4.0.4 (R Core Team 2021), and visualized using autoplot (using the packages ggfortify and ggplot2; Tang et al. 2016; Wickham 2016) of the original and recoded alignments. To reconstruct phylogenies of the recoded data, we used IQ-TREE with the model GTR+G4 with ultrafast bootstrap and SH-aLRT with 1000 replicates each (supplementary figs. S12 and S13; analysis/09_AA_composition_alignments on Figshare).

Molecular dating

To estimate divergence times of different algal groups, we made use of the phylogeny reconstructed using PhyloBayes for the ‘all plastids’ data set. To calibrate this tree, we gathered calibration points from the literature. An overview of all calibrations used in this analysis, and the fossil or other data related to them, their phylogenetic affiliation, and age

are listed in the Supplementary info, and largely follow the guidelines from Parham et al. (Parham et al. 2012). As certain calibrations are uncertain, we made different sets. A conservative set was created with uncontroversial calibrations, a second set included the fossil *Bangiomorpha* and a third one included *Bangiomorpha* and *Proterocladus* (Butterfield 2000; Tang et al. 2020) as both fossils are followed by a large gap of no red or green algal fossils. A fourth set equals the conservative set plus extra maximum calibration, plus a change of the maximum of 1,042 Ma on land plants to 515.5 Ma, following Morris et al (Morris et al. 2018). The fifth set equals the fourth, plus the *Bangiomorpha* and *Proterocladus* calibrations. Finally, we also used the dates from Strassert et al (2021) that were applicable to the lineages we include to directly compare our results.

We calibrated the phylogeny based on the all-plastids data set and the topology retrieved based on this data set. We tested whether evolutionary rates are autocorrelated in this phylogeny using CorrTest as provided in the MegaX v10.2 suite using the cpREV+F+I+G4 model, the default maximum relative rate ratio of 20, and using all sites (Adachi et al. 2000; Kumar et al. 2018; Tao et al. 2019). The calibrations were used to date the tree using PhyloBayes 4.1, as it includes the sophisticated substitution model CAT+GTR+G4, as well as the possibility to use both autocorrelated and uncorrelated relaxed molecular clock models (Lartillot et al. 2009). Although fossils used as minimum calibrations with certain ages and phylogenetic affinities are best used as hard bounds (Parham et al. 2012; Marjanović 2021), this is likely rare for our set of calibrations as especially phylogenetic affinity can be uncertain. Thus, we used PhyloBayes with soft bounds for all our analyses, with a birth-death prior of divergence times. We used all calibration sets with one autocorrelated (log-normal) and two uncorrelated models (white-noise and uncorrelated gamma multipliers) (Thorne et al. 1998; Drummond et al. 2006; Lepage et al. 2007). Convergence was checked with bpcomp and tracecomp, and by plotting the variables in the trace file, including the log-likelihood. All code related to the molecular clock analysis can be found in analysis/11_dating in Figshare.

References

Adachi J, Waddell PJ, Martin W, Hasegawa M. 2000. Plastid Genome Phylogeny and a Model of Amino Acid Substitution for Proteins Encoded by Chloroplast DNA. *J. Mol. Evol.* 50:348–358.

Aguirre J, Baceta JI, Braga JC. 2007. Recovery of marine primary producers after the Cretaceous–Tertiary mass extinction: Paleocene calcareous red algae from the Iberian Peninsula. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* 249:393–411.

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. *J. Mol. Biol.* 215:403–410.

Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 25:3389–3402.

Andrews S. 2016. FastQC: A quality control tool for high throughput sequence data. Available from: https://www.bioinformatics.babraham.ac.uk/projects/fastqc/

Archibald JM. 2009. The Puzzle of Plastid Evolution. *Curr. Biol.* 19:R81–R88.

Archibald JM. 2012. Chapter Three - The Evolution of Algae by Secondary and Tertiary Endosymbiosis. In: Piganeau G, editor. Advances in Botanical Research. Vol. 64. Genomic Insights into the Biology of Algae. Academic Press. p. 87–118. Available from: http://www.sciencedirect.com/science/article/pii/B9780123914996000037

Baurain D, Brinkmann H, Petersen J, Rodríguez-Ezpeleta N, Stechmann A, Demoulin V, Roger AJ, Burger G, Lang BF, Philippe H. 2010. Phylogenomic Evidence for Separate Acquisition of Plastids in Cryptophytes, Haptophytes, and Stramenopiles. *Mol. Biol. Evol.* 27:1698–1709.

Bengtson S, Sallstedt T, Belivanova V, Whitehouse M. 2017. Three-dimensional preservation of cellular and subcellular structures suggests 1.6 billion-year-old crown-group red algae. *PLOS Biol.* 15:e2000735.

Benson DA, Cavanaugh M, Clark K, Karsch-Mizrachi I, Lipman DJ, Ostell J, Sayers EW. 2013. GenBank. *Nucleic Acids Res.* 41:D36-42.

Berney C, Pawlowski J. 2006. A molecular time-scale for eukaryote evolution recalibrated with the continuous microfossil record. *Proc. R. Soc. B Biol. Sci.* 273:1867–1872.

van Beveren F, Eme L, López-García P, Ciobanu M, Moreira D. 2022. Independent Size Expansions and Intron Proliferation in Red Algal Plastid and Mitochondrial Genomes. *Genome Biol. Evol.* 14:evac037.

Bodył A, Stiller JW, Mackiewicz P. 2009. Chromalveolate plastids: direct descent or multiple endosymbioses? *Trends Ecol. Evol.* 24:119–121.

Burki F. 2017. Chapter One - The Convoluted Evolution of Eukaryotes With Complex Plastids. In: Hirakawa Y, editor. Advances in Botanical Research. Vol. 84. Secondary Endosymbioses. Academic Press. p. 1–30. Available from: http://www.sciencedirect.com/science/article/pii/S0065229617300484

Burki F, Roger AJ, Brown MW, Simpson AGB. 2020. The New Tree of Eukaryotes. *Trends Ecol. Evol.* 35:43–55.

Bushnell B. 2014. BBMap: A Fast, Accurate, Splice-Aware Aligner. Lawrence Berkeley National Lab. (LBNL), Berkeley, CA (United States) Available from: https://www.osti.gov/biblio/1241166-bbmap-fast-accurate-splice-aware-aligner

Butterfield NJ. 2000. Bangiomorpha pubescens n. gen., n. sp.: implications for the evolution of sex, multicellularity, and the Mesoproterozoic/Neoproterozoic radiation of eukaryotes. *Paleobiology* 26:386–404.

Butterfield NJ, Knoll AH, Swett K. 1990. A bangiophyte red alga from the Proterozoic of arctic Canada. *Science* 250:104–107.

Butterfield NJ, Knoll AH, Swett K. 1994. Paleobiology of the neoproterozoic svanbergfjellet formation, Spitsbergen. *Lethaia* 27:76–76.

Carlisle EM, Jobbins M, Pankhania V, Cunningham JA, Donoghue PCJ. 2021. Experimental taphonomy of organelles and the fossil record of early eukaryote evolution. *Sci. Adv.* 7:eabe9487.

Carver T, Harris SR, Berriman M, Parkhill J, McQuillan JA. 2012. Artemis: an integrated platform for visualization and analysis of high-throughput sequence-based experimental data. *Bioinformatics* 28:464–469.

Cavalier-Smith T. 1986. The kingdom Chromista: origin and systematics. *Prog. Phycol. Res.* 4:309–347.

Cavalier‐Smith T. 1999. Principles of Protein and Lipid Targeting in Secondary Symbiogenesis: Euglenoid, Dinoflagellate, and Sporozoan Plastid Origins and the Eukaryote Family Tree1,2. *J. Eukaryot. Microbiol.* 46:347–366.

Clarke JT, Warnock RCM, Donoghue PCJ. 2011. Establishing a time-scale for plant evolution. *New Phytol.* 192:266–301.

Criscuolo A, Gribaldo S. 2010. BMGE (Block Mapping and Gathering with Entropy): a new software for selection of phylogenetic informative regions from multiple sequence alignments. *BMC Evol. Biol.* 10:210.

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.* 33:2890–2898.

Dierckxsens N, Mardulyn P, Smits G. 2017. NOVOPlasty: de novo assembly of organelle genomes from whole genome data. *Nucleic Acids Res.* 45:e18–e18.

van Dongen SM. 2000. Graph clustering by flow simulation.

Dorrell RG, Gile G, McCallum G, Méheust R, Bapteste EP, Klinger CM, Brillet-Guéguen L, Freeman KD, Richter DJ, Bowler C. 2017. Chimeric origins of ochrophytes and haptophytes revealed through an ancient plastid proteome.Bhattacharya D, editor. *eLife* 6:e23717.

Dorrell RG, Smith AG. 2011. Do red and green make brown?: perspectives on plastid acquisitions within chromalveolates. *Eukaryot. Cell* 10:856–868.

Drummond AJ, Ho SYW, Phillips MJ, Rambaut A. 2006. Relaxed Phylogenetics and Dating with Confidence. *PLOS Biol.* 4:e88.

Ekseth OK, Kuiper M, Mironov V. 2014. orthAgogue: an agile tool for the rapid prediction of orthology relations. *Bioinformatics* 30:734–736.

Eme L, Sharpe SC, Brown MW, Roger AJ. 2014. On the Age of Eukaryotes: Evaluating Evidence from Fossils and Molecular Clocks. *Cold Spring Harb. Perspect. Biol.* 6:a016139.

Foster PG, Hickey DA. 1999. Compositional Bias May Affect Both DNA-Based and Protein-Based Phylogenetic Reconstructions. *J. Mol. Evol.* 48:284–290.

Gibson TM, Shih PM, Cumming VM, Fischer WW, Crockford PW, Hodgskiss MSW, Wörndle S, Creaser RA, Rainbird RH, Skulski TM, et al. 2018. Precise age of Bangiomorpha pubescens dates the origin of eukaryotic photosynthesis. *Geology* 46:135–138.

Guindon S, Dufayard J-F, Lefort V, Anisimova M, Hordijk W, Gascuel O. 2010. New Algorithms and Methods to Estimate Maximum-Likelihood Phylogenies: Assessing the Performance of PhyML 3.0. *Syst. Biol.* 59:307–321.

Guiry MD, Guiry GM. 2022. Algaebase. *AlgaeBase World-Wide Electron. Publ. Natl. Univ. Irel. Galway* [Internet]. Available from: https://www.algaebase.org

Gustavson TC, Delevoryas T. 1992. Caulerpa-like marine alga from Permian strata, Palo Duro Basin, West Texas. *J. Paleontol.* 66:160–161.

Hoang DT, Chernomor O, von Haeseler A, Minh BQ, Vinh LS. 2018. UFBoot2: Improving the Ultrafast Bootstrap Approximation. *Mol. Biol. Evol.* 35:518–522.

Jackson C, Knoll AH, Chan CX, Verbruggen H. 2018. Plastid phylogenomics with broad taxon sampling further elucidates the distinct evolutionary origins and timing of secondary green plastids. *Sci. Rep.* 8:1523.

Katoh K, Kuma K, Miyata T, Toh H. 2005. Improvement in the accuracy of multiple sequence alignment program MAFFT. *Genome Inform.* 16:22–33.

Katoh K, Standley DM. 2013. MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. *Mol. Biol. Evol.* 30:772–780.

Kumar S, Stecher G, Li M, Knyaz C, Tamura K. 2018. MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Mol. Biol. Evol.* 35:1547–1549.

Lartillot N, Lepage T, Blanquart S. 2009. PhyloBayes 3: a Bayesian software package for phylogenetic reconstruction and molecular dating. *Bioinformatics* 25:2286–2288.
Lartillot N, Rodrigue N, Stubbs D, Richer J. 2013. PhyloBayes MPI: Phylogenetic Reconstruction with Infinite Mixtures of Profiles in a Parallel Environment. *Syst. Biol.* 62:611–615.
Le SQ, Gascuel O. 2008. An Improved General Amino Acid Replacement Matrix. *Mol. Biol. Evol.* 25:1307–1320.
Lepage T, Bryant D, Philippe H, Lartillot N. 2007. A general comparison of relaxed molecular clock models. *Mol. Biol. Evol.* 24:2669–2680.
Marin B, M. Nowack EC, Melkonian M. 2005. A Plastid in the Making: Evidence for a Second Primary Endosymbiosis. *Protist* 156:425–432.
Marjanović D. 2021. The Making of Calibration Sausage Exemplified by Recalibrating the Transcriptomic Timetree of Jawed Vertebrates. *Front. Genet.* 12:535.
Mastik V, Tinn O. 2015. New dasycladalean algal species from the Kalana Lagerstätte (Silurian, Estonia). *J. Paleontol.* 89:262–268.
Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, von Haeseler A, Lanfear R. 2020. IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. *Mol. Biol. Evol.* 37:1530–1534.
Morris JL, Puttick MN, Clark JW, Edwards D, Kenrick P, Pressel S, Wellman CH, Yang Z, Schneider H, Donoghue PCJ. 2018. The timescale of early land plant evolution. *Proc. Natl. Acad. Sci.* 115:E2274–E2283.
Muñoz-Gómez SA, Mejía-Franco FG, Durnin K, Colp M, Grisdale CJ, Archibald JM, Slamovits CH. 2017. The New Red Algal Subphylum Proteorhodophytina Comprises the Largest and Most Divergent Plastid Genomes Known. *Curr. Biol.* 27:1677-1684.e4.
Nguyen L-T, Schmidt HA, von Haeseler A, Minh BQ. 2015. IQ-TREE: A Fast and Effective Stochastic Algorithm for Estimating Maximum-Likelihood Phylogenies. *Mol. Biol. Evol.* 32:268–274.
Parham JF, Donoghue PCJ, Bell CJ, Calway TD, Head JJ, Holroyd PA, Inoue JG, Irmis RB, Joyce WG, Ksepka DT, et al. 2012. Best Practices for Justifying Fossil Calibrations. *Syst. Biol.* 61:346–359.
Paudel YP, Hu Z, Khatiwada JR, Fan L, Pradhan S, Liu B, Qin W. 2021. Chloroplast genome analysis of Chrysotila dentata. *Gene* 804:145871.
Petersen J, Ludewig A-K, Michael V, Bunk B, Jarek M, Baurain D, Brinkmann H. 2014. Chromera velia, endosymbioses and the rhodoplex hypothesis—plastid evolution in cryptophytes, alveolates, stramenopiles, and haptophytes (CASH lineages). *Genome Biol. Evol.* 6:666–684.
Preuss M, Verbruggen H, West JA, Zuccarello GC. 2021. Divergence times and plastid phylogenomics within the intron-rich order Erythropeltales (Compsopogonophyceae, Rhodophyta). *J. Phycol.* [Internet] n/a. Available from: https://onlinelibrary.wiley.com/doi/abs/10.1111/jpy.13159
R Core Team. 2021. R: A Language and Environment for Statistical Computing. Available from: https://www.R-project.org/
Rambaut A, Drummond AJ, Xie D, Baele G, Suchard MA. 2018. Posterior Summarization in Bayesian Phylogenetics Using Tracer 1.7. *Syst. Biol.* 67:901–904.
Rösler A, Perfectti F, Peña V, Aguirre J, Braga JC. 2017. Timing of the evolutionary history of Corallinaceae (Corallinales, Rhodophyta). *J. Phycol.* 53:567–576.
Sanchez-Puerta MV, Delwiche CF. 2008. A HYPOTHESIS FOR PLASTID EVOLUTION IN CHROMALVEOLATES 1. *J. Phycol.* 44:1097–1107.
Sarai C, Tanifuji G, Nakayama T, Kamikawa R, Takahashi K, Yazaki E, Matsuo E, Miyashita H, Ishida K, Iwataki M, et al. 2020. Dinoflagellates with relic endosymbiont nuclei as models for elucidating organellogenesis. *Proc. Natl. Acad. Sci.* 117:5364–5375.
Shimodaira H. 2002. An Approximately Unbiased Test of Phylogenetic Tree Selection. *Syst. Biol.* 51:492–508.
Sibbald SJ, Archibald JM. 2020. Genomic insights into plastid evolution. *Genome Biol. Evol.*

[Internet]. Available from: https://academic.oup.com/gbe/advance-article/doi/10.1093/gbe/evaa096/5836826

Stiller JW, Schreiber J, Yue J, Guo H, Ding Q, Huang J. 2014. The evolution of photosynthesis in chromist algae through serial endosymbioses. *Nat. Commun.* 5:5764.

Strassert JFH, Irisarri I, Williams TA, Burki F. 2021. A molecular timescale for eukaryote evolution with implications for the origin of red algal-derived plastids. *Nat. Commun.* 12:1879.

Strother PK, Battison L, Brasier MD, Wellman CH. 2011. Earth's earliest non-marine eukaryotes. *Nature* 473:505–509.

Su D, Yang L, Shi X, Ma X, Zhou X, Hedges SB, Zhong B. 2021. Large-Scale Phylogenomic Analyses Reveal the Monophyly of Bryophytes and Neoproterozoic Origin of Land Plants. *Mol. Biol. Evol.* 38:3332–3344.

Susko E, Roger AJ. 2007. On Reduced Amino Acid Alphabets for Phylogenetic Inference. *Mol. Biol. Evol.* 24:2139–2150.

Tang Q, Pang K, Yuan X, Xiao S. 2020. A one-billion-year-old multicellular chlorophyte. *Nat. Ecol. Evol.* 4:543–549.

Tang Y, Horikoshi M, Li W. 2016. ggfortify: unified interface to visualize statistical results of popular R packages. *R J* 8:474.

Tao Q, Tamura K, U. Battistuzzi F, Kumar S. 2019. A Machine Learning Method for Detecting Autocorrelation of Evolutionary Rates in Large Phylogenies. *Mol. Biol. Evol.* 36:811–824.

The UniProt Consortium. 2015. UniProt: a hub for protein information. *Nucleic Acids Res.* 43:D204–D212.

Thorne JL, Kishino H, Painter IS. 1998. Estimating the rate of evolution of the rate of molecular evolution. *Mol. Biol. Evol.* 15:1647–1657.

Tillich M, Lehwark P, Pellizzer T, Ulbricht-Jones ES, Fischer A, Bock R, Greiner S. 2017. GeSeq – versatile and accurate annotation of organelle genomes. *Nucleic Acids Res.* 45:W6–W11.

Wang H-C, Minh BQ, Susko E, Roger AJ. 2018. Modeling Site Heterogeneity with Posterior Mean Site Frequency Profiles Accelerates Accurate Phylogenomic Estimation. *Syst. Biol.* 67:216–235.

Whelan S, Irisarri I, Burki F. 2018. PREQUAL: detecting non-homologous characters in sets of unaligned homologous sequences. *Bioinformatics* 34:3929–3930.

Wickham H. 2016. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York Available from: https://ggplot2.tidyverse.org

Yoon HS, Hackett JD, Pinto G, Bhattacharya D. 2002. The Single, Ancient Origin of Chromist Plastids. *J. Phycol.* 38:40–40.

Zhang C, Rabiee M, Sayyari E, Mirarab S. 2018. ASTRAL-III: polynomial time species tree reconstruction from partially resolved gene trees. *BMC Bioinformatics* 19:15–30.

Zhang Y, Yin L, Xiao S, Knoll AH. 1998. Permineralized Fossils from the Terminal Proterozoic Doushantuo Formation, South China. *Mem. Paleontol. Soc.* 50:1–52.

Zhang Z, Green BR, Cavalier-Smith T. 2000. Phylogeny of Ultra-Rapidly Evolving Dinoflagellate Chloroplast Genes: A Possible Common Origin for Sporozoan and Dinoflagellate Plastids. *J. Mol. Evol.* 51:26–40.

6 - Discussion

One of the main questions remaining in eukaryotic evolution is the origin of the many lineages with red algal-derived plastids. Hypotheses that have been proposed, specifically the chromalveolate and serial endosymbiosis hypotheses, are partly based on the monophyly of red complex plastids in plastid phylogenies. Yet, the sampling of potentially deep-branching, unicellular red algae has been lacking, even though many species have been described and are available in culture collections. In this thesis, the goal has been to acquire plastid (and mitochondrial) genome sequences of such red algae, especially of the poorly studied Proteorhodophytina. Using these newly acquired data, we analyzed the organellar genome evolution and expansion by group II introns, followed by phylogenomic analyses of red algae and red complex plastids.

Proteorhodophytina are a group of understudied red algae, which can however reveal much about the evolution of red algae and the origin of red complex plastids (Muñoz-Gómez, Mejía-Franco, et al. 2017). Plastid phylogenies typically place the red complex plastids between Cyanidiophytina and all other red algae (Yoon, Hackett, Pinto, et al. 2002; Dorrell, Gile, et al. 2017; Muñoz-Gómez, Mejía-Franco, et al. 2017). These other algae consist of the multicellular groups Bangiophyceae and Florideophyceae (Yoon, Muller, et al. 2006), of which many plastid genomes have been sequenced and studied (Lee et al. 2016). However, the Proteorhodophytina, which include unicellular and filamentous species are more likely to be closely related to the probably unicellular alga (or algae) that was the ancestor of all red complex plastids. It has even been suggested that the ancestor of red complex plastids may be a member of the Porphyridiales (*sensu lato*, then referring to unicellular red algae including members of Cyanidiophyceae, and Proteorhodophytina; Shalchian-Tabrizi, Skånseng, et al. 2006). Yet, sampling of Proteorhodophytina species has been very incomplete for a long time, with a first coverage of all Proteorhodophytina classes in a plastid genome phylogeny being from 2017 (Muñoz-Gómez, Mejía-Franco, et al. 2017). Intriguingly, the monophyly of complex plastids was not easily recovered with that data set, especially with maximum likelihood methods. Moreover, as found by Muñoz-Gómez, Mejía-Franco, et al. (2017), these plastid genomes are also of interest as they have undergone genome expansion and reshuffling of plastid genomic segments. We observed similar results with a much larger sampling of plastid genomes of Proteorhodophytina, but also of their mitochondrial genomes (Chapter 4). These data may have implications for the evolution of red algae overall, which will be discussed first.

6.1 . Organellar genome evolution in red algae

Using both plastid and mitochondrial-encoded proteins, we reconstructed the phylogeny of red algae (Chapter 4). Although the monophyly of Proteorhodophytina was shown before using plastid data (Muñoz-Gómez, Mejía-Franco, et al. 2017; Preuss et al. 2021), we get support for it with a larger taxon sampling with plastid data, and recover it for the first time with mitochondrial data. Moreover, the inclusion of red complex plastids in this phylogeny does not affect the monophyly of Proteorhodophytina (Chapter 5). Whereas our plastid phylogeny shows the same relationships among classes of Proteorhodophytina from previous studies, our mitochondrial genome data does not. As the mitochondrial genomes encode fewer proteins, there is likely less phylogenomic signal. Moreover, some lineages—especially stylonematophytes—have very fast evolving mitochondrial protein sequences. The advantages of the plastid data set make the phylogenetic construction based on this data more likely, but several deep, short branches in this phylogeny may make it that very few substitutions are responsible for the distinction between these lineages. All in all, both plastid and mitochondrial data support the monophyly of Proteorhodophytina and support the idea that a rapid diversification occurred after their last common ancestor (Yoon, Muller, et al. 2006).

Plastid genomes of red algae from the classes Florideophyceae and Bangiophyceae vary little is size and organization (Janouškovec, Liu, et al. 2013; Lee et al. 2016). This has been seen as evidence for slow evolution since the origin of the red algae, so that they have remained relatively ancestral-like in comparison to other archaeplastidal lineages (Glöckner et al. 2000; Janouškovec, Liu, et al. 2013). In contrast, rearrangements of the plastid genomes have been more frequent in lineages of Cyanidiophyceae and especially Proteorhodophytina (Chapter 4; Muñoz-Gómez, Mejía-Franco, et al. 2017). We have shown that these rearrangements are particularly common in the unicellular classes Porphyridiophyceae and Rhodellophyceae (see Chapter 4), where also the largest plastid genomes are found (section 4; Muñoz-Gómez, Mejía-Franco, et al. 2017). These rearrangements are largely mediated by group II introns, which are self-splicing introns able to move within and between different genomes (Lambowitz and Zimmerly 2004, 2011). This clearly indicates that plastid genomes of red algae have not been necessarily slow-evolving, although this may be true for only some groups of red algae as plastid genome organization is relatively stable within Florideophyceae and Bangiophyceae (Lee et al. 2016).

Group II intron proliferation has expanded plastid genomes in only the two strictly unicellular classes of Proteorhodophytina, but the question remains open why genome expansion only happened in these classes. No obvious link could be made considering the ecosystem they thrive in (i.e., freshwater vs marine; see chapter 4). As introns are present in all plastid genomes of Proteorhodophytina representatives, and some clearly encoding a group II intron-related maturase, it is probable that the common ancestor carried these introns in its plastid genome (Chapter 4; Muñoz-Gómez, Mejía-Franco, et al. 2017; Preuss et al. 2021). In representatives of Compsopogonophyceae and Stylonematophyceae, there are few introns left that still carry signal of the encoded maturase. Depending on the exact regions that are present in this maturase and the remaining sequence of the intron, the ability for the intron to copy and paste itself may be lost (Perrineau et al. 2015). Although we have no direct evidence on the domains present in the introns and their intron encoded

proteins, it has been shown that they are present in the unexpanded plastid genomes of *Porphyridium* spp. (Perrineau et al. 2015). Moreover, we show that introns remain highly mobile, with difference in intron positions between sister lineages. Especially in the case of *Rhodella violacea*, where several highly identical introns have proliferated in its plastid genome. We may get a better understanding of these events by comparing these plastid genomes with other expanded plastid genomes outside of red algae, but these genome expansion events are not necessarily as much related to group II intron proliferation (Aono et al. 2002; Smith and Lee 2009; Zhang et al. 2019).

Interestingly, expanded plastid genomes do not occur within red complex plastids, instead having plastids smaller than those of typical red algae, and encoding fewer genes (Kim and Archibald 2009; Janouškovec, Horák, et al. 2010). However, red complex plastids may carry introns, some of which are group II introns (Bhattacharya, Qiu, et al. 2018). As self-splicing introns have been shown to not only move around within a genome, but also between genomes of different species (Burger et al. 1999; Bhattacharya, Cannone, et al. 2001), introns in red complex plastids can have various origins. If a red algal origin of some of these introns would be established, this could however be highly interesting. Although red complex plastids typically are found as a sister group to the plastids of Proteorhodophytina plus Eurhodophytina, it could suggest that at least such introns were present in the red alga that is the common ancestor of red complex plastids. Introns could potentially play a role in the mobility of other DNA sequences (Pombert et al. 2005; Zhang et al. 2019), which could putatively aid in endosymbiotic gene transfer, especially as many more plastid genes are moved to the nucleus in red complex plastid-bearing lineages than in red algae. If some introns in red complex plastids are of red algal origin, it may indicate that they originated from the ancient red algal endosymbiont.

Introns have also been observed to move between organellar genomes of the same species (e.g., Pombert et al. 2005; Turmel et al. 2016). As we were able to assemble and annotate several mitochondrial genomes of proteorhodophytes, we were able to establish that several of these mitochondrial genomes are also expanded (Chapter 4; Kim, Lee, et al. 2022). However, no proteorhodophytes has both an expanded plastid and mitochondrial genome, making it unlikely that the expansion is related to movement of introns between organellar genomes. Moreover, our phylogenies of intron encoded maturases suggest that these have not been recently transferred between organellar genomes. This does not rule out that such events have occurred, either within a current day lineage of in a common ancestor of lineages with expanded genomes, but no clear evidence exist.

Most intriguing are the mitochondrial genomes of Stylonematophyceae, which have a much higher GC content than the mitochondrial genomes of all other red algae, except for acidophilic *Galdieria* spp. (Chapter 4). But in contrast to *Galdieria* spp., all representatives of Stylonematophyceae are mesophilic. As the mitochondrial genomes were also repeat rich, we were unable to assemble them in single circular contigs. This makes it difficult to understand the alternative evolutionary path these mitochondrial genomes took in comparison to other red algae.

6.2 . Red algal multicellularity

Finally, the placement of Proteorhodophytina as sister to Eurhodophytina may inform us about the origin of (simple) multicellularity within red algae. All representatives of Eurhodophytina are multicellular and among proteorhodophytes there are also multicellular lineages (Muñoz-Gómez, Mejía-Franco, et al. 2017). Stylonematophyceae hosts both unicellular and multicellular species, showing that gains and losses of multicellularity are quite common in red algae (Aboal et al. 2018). As these are changes from simple filaments to unicellarity, it is indeed likely to be simple as such changes happen all over the eukaryotic tree of life (Knoll 2011). Interestingly, Compsopogonophyceae only hosts multicellular species, and some display what can be considered as simple pit plugs (Pueschel and Cole 1982; Saunders and Hommersand 2004). As pit plugs and pit connections are unique to red algae, one could question the likelihood that this evolved separately in Compsopogonophyceae and Eurhodophytina. In view of this, the lack of multicellularity in other species of Proteorhodophytina may be more easily interpreted as loss of ancestral multicellularity. Moreover, the red algal fossil *Bangiomorpha*, which displays simple multicellularity, could thus very well be interpreted as similar to a common ancestor of Eurhodophytina and Proteorhodophytina, although pit plugs are not observed in this fossil (Butterfield 2000) and are controversial in older putative red algal fossils (Bengtson et al. 2017; Carlisle et al. 2021). *Bangiomorpha* was originally considered as part of Bangiaceae, and its radial symmetry is definitely similar to that of bangiophyte species like *Bangia*, *Minerva* and *Dione* (Nelson et al. 2005). Because of this, the fossil has been used to minimally constrain the divergence between Bangiophyceae and its sister class Florideophyceae (e.g., Strassert, Irisarri, et al. 2021). However, this fossil has also been used to calibrate deeper nodes in the tree of red algae, including the node representing the divergence between Cyanidiophyceae and all other red algae (Rhodophytina) (e.g., Parfrey et al. 2011), or even between red algae and other archaeplastida (e.g., Betts et al. 2018). Considering that the common ancestor of Rhodophytina was likely multicellular, using *Bangiomorpha* to date the divergence of this group and the unicellular Cyanidiophyceae seems most plausible, although it cannot be definitely excluded that the common ancestor of all red algae was multicellular.

6.3 . The origin of red algal-derived plastids

The chromalveolate hypothesis poses that a single endosymbiotic event in the common ancestor of red complex plastid-bearing lineages occurred (Cavalier-Smith 1999). However, it has become clear that this has to go hand in hand with many losses in plastid-lacking lineages that are descendants as this common ancestor (Burki 2017). Assuming that the common ancestor of chromalveolates does not have red algae as one of its descendants, this would not be impossible. It would however imply that many lineages did not become dependent on the plastid for potentially a very long time. For example, as ochrophytes contain red complex plastid, the loss in its sister (*Actinophyris sol*, but similarly also oomycetes and other sister lineages) can only have occurred after the divergence between these two groups (fig. 6.1), which may be unlikely. Can we estimate how many millions of years of evolution this entails? Strassert, Irisarri, et al. (2021) get a range for the stem-lineage of

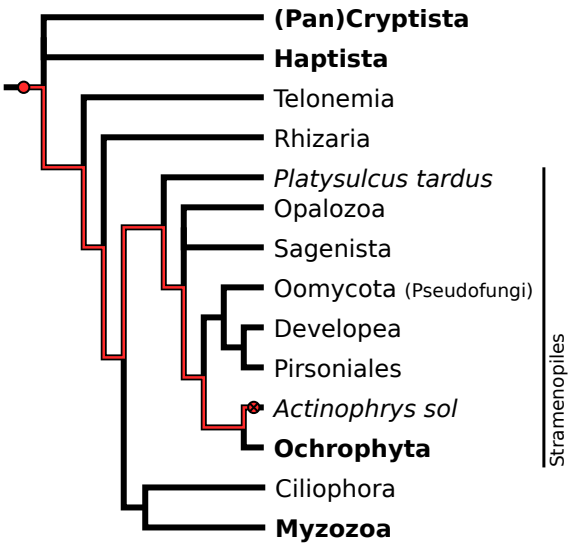


Figure 6.1: Schematic overview of the phylogeny of red complex plastid-bearing lineages and their relatives (**bold**). The plastid acquisition assumed by the chromalveolate hypothesis is shown with a red circle, with the loss of the plastid in a sister group of ochrophytes shown with a crossed red circle. The red line follows the evolutionary path in which the plastid was retained without the host having any (metabolic) dependency on the plastid.

Rhodophytina (from which the complex plastids evolve) of 1675-1281 Ma, whereas the ochrophytes diverge from their heterotrophic sister lineage roughly ~1000 Ma, leaving a putative gap of ~300 Ma where lineages evolved without plastid integration. Such an amount of time of continuous evolution without plastid integration may seem unlikely, especially considering that the integrated primary plastid of *Paulinella chromatophora* probably originated ~90 to 140 Ma (Delaye et al. 2016). Other evidence, such as chlorophyll *c*, and several genes of the Calvin cycle do not argue very strongly for or against the chromalveolate hypothesis (see sections 1.4.1 and 1.4.2), but do at least slightly favor the chromalveolate hypothesis over any serial endosymbiosis hypothesis as fewer transfers of plastids or genes are implied. The main way that the chromalveolate hypothesis may lose all credibility, is if the suggested sister relation of Cryptista and Archaeplastida rings true (Schön et al. 2021; Strassert, Irisarri, et al. 2021; Yazaki et al. 2022). If however Archaeplastida are not descendants of the common ancestor of chromalveolates, and the relationships among chromalveolates are well resolved, plastid phylogenies such as those reconstructed here (see chapter 5) may reject the chromalveolate hypothesis if plastid and host phylogenies are not congruent.

Considering the traces of endosymbiosis discussed in the introduction (section 1.4), and the plastid phylogenomic analyses from Chapter 5, it would be interesting to evaluate the different serial endosymbiosis hypotheses (Sanchez-Puerta and Delwiche 2008; Bodył et al. 2009; Stiller, Schreiber, et al. 2014). We have shown that the sister relation of Cryptophyta and Haptophyta in plastid phylogenies is strong (Chapter 5) using multiple phylogenetic reconstruction methods and after removal of fast evolving sites, or removal of amino acid biases. The probable monophyly of cryptophyte and haptophyte plastids is also supported by the *rpl36* gene replacement in their plastid genomes of these two lineages (Rice and Palmer 2006). This gene replacement is of specific interest because (1) plastid gene replacements are highly rare and (2) *rpl36* phylogenies including bacterial

representatives suggest the source is also the same, making it extremely likely to be a singular event (Rice and Palmer 2006).

One potential issue is that haptophytes and ochrophytes have a feature in common: the loss of the nucleomorph. However, this could be considered more likely to occur twice, as for example shown by the ease of nucleomorph loss in the cryptophyte-derived kleptoplastid in *Dinophysis* spp. (Lucas and Vesk 1990; Gagat et al. 2014). Moreover, nucleomorphs are absent in the green plastids of euglenophytes (Cavalier-Smith 1999) and the haptophyte-derived plastids of dinoflagellates of the family Kareniaceae (Waller and Kořený 2017). Thus, loss of the original nucleomorph may be relatively easy, and likely easier to explain than independent replacements of *rpl36*.

All in all, considering the evidence of plastid and host phylogenies, as well as of the gene replacement of *rpl36*, only one proposed serial endosymbiosis hypothesis fits well, which is the proposal from Bodył et al. One of the larger issues in my opinion is that ochrophytes acquire the plastid relatively early, but the split between them and heterotrophic stramenopiles seems to appear later than the same diversification of cryptophytes and haptophytes from their heterotrophic sisters (Strassert, Irisarri, et al. 2021), although such estimates are very uncertain.

Can we consider other scenarios that fit well but have not been proposed? To try to get to another hypothesis, there is first of all some evidence we can do little with. We do not know much about how likely it is to evolve the membrane transport machinery (SELMA) multiple times (Waller, Gornik, et al. 2016), the ability to synthesize chlorophyll *c* (Larkum 2020) and the potentially ancient duplications of Calvin-cycle related genes in red complex plastid-bearing algae (Teich et al. 2007). These three lines of evidence may make multiple independent red algal acquisitions unlikely, but it would be an interesting idea to consider. As ochrophytes seem quite distantly related to the other two lineages from a plastid perspective (Chapter 5; Dorrell, Gile, et al. 2017), one option would be to propose an independent red algal acquisition in ochrophytes. This can be proposed without changing anything about the *rpl36* gene replacement, or the number of nucleomorph losses (which remains at two losses). However, this hypothesis should be considered less parsimonious than the serial endosymbiosis hypotheses from Sanchez-Puerta and Delwiche (2008) and Bodył et al. (2009) as an independent origin of the ochrophyte plastid also involves an independent origin or acquisition of SELMA and chlorophyll *c* biosynthesis. Finally, all serial endosymbioses hypotheses suggest that the first red algal endosymbiosis occurred either in an ancestor of cryptophytes, or an ancestor of Hacrobia. The nucleomorph present in cryptophytes may be considered an argument for this, but it is important to note that all plastid transfers occurred in stem-lineages of the red-plastid bearing groups. There is no reason why, for example, ochrophytes could ancestrally have contained a nucleomorph, but was lost before the common ancestor of ochrophytes (but after a potential transfer to cryptophytes). Although a small step has been taken in this thesis in understanding the relationships among red complex plastids, many hypotheses remain congruent with the available data.

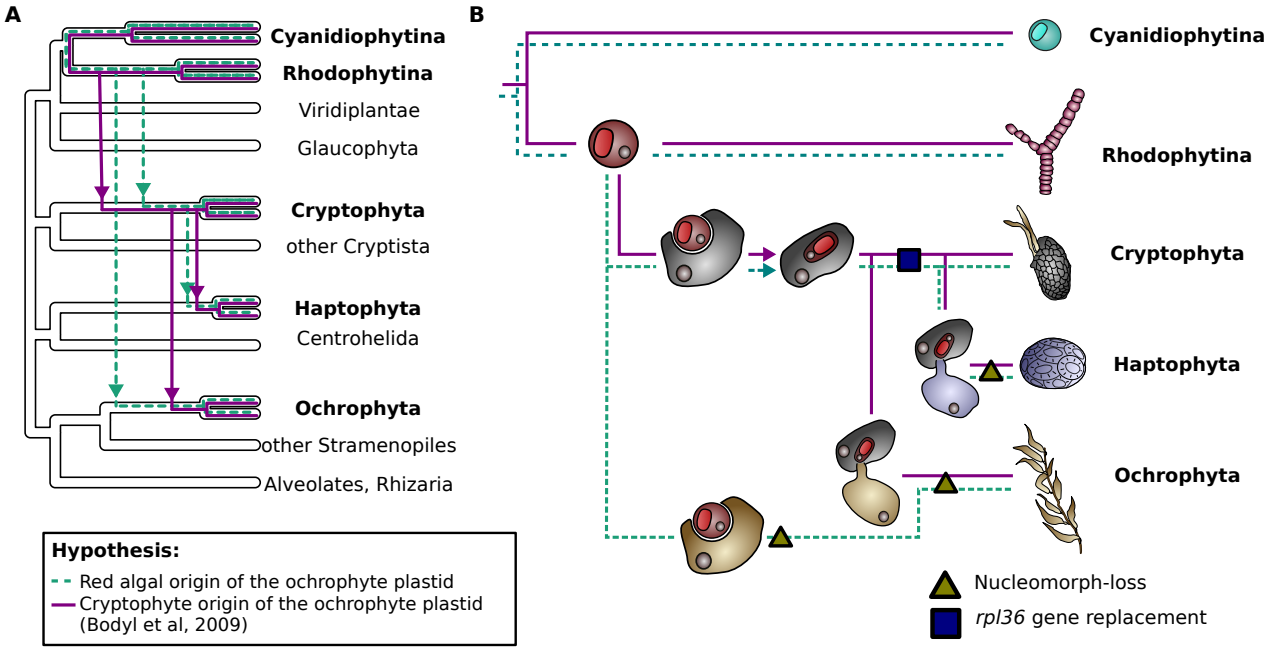


Figure 6.2: Potential hypotheses best fitting the current evidence on the origin of red complex plastids. (A) Subset of the eukaryotic tree of life including Archaeplastida, red complex plastid-bearing lineages and their relatives. Two potential hypotheses are shown, one with a single red algal origin of red complex plastids (purple; Bodył et al. 2009), and another with a separate origin of the ochrophyte plastid (green). (B) Schematic of the tree in A, with potential endosymbiotic origins of plastids of the different groups by taking up a whole cell, or using kleptoplastidy (shown as using myzocytosis). Both scenarios imply a single gene replacement of *rpl36* and two losses of the nucleomorph, in the haptophyte and ochrophyte stem lineages.

6.4 . When did the red complex plastid originate?

At the time of writing this thesis, our molecular dating analysis is still in progress and little can be said about preliminary results (Chapter 5). However, some things we can be said simply based on the tree topology and the calibrations used. The topology at the very least will suggest that the putative transfer from cryptophytes to haptophytes occurred before the plastid acquisition in ochrophytes (see also fig. 6.2). Interestingly, the estimated age for the stem group of ochrophytes is fewer millions of years ago than that of haptophytes and cryptophytes (Strassert, Irisarri, et al. 2021). Of course, these branches have a length, and the nodes at either end have a margin of error, which in the end results in an overlap of all the branches of the stem lineages of ochrophytes, haptophytes and cryptophytes. Because of this, the scenario proposed by Bodył et al. (2009) is not refuted. Even a direct transfer from the stem lineage of Rhodophytina to the stem lineage of Ochrophyta is not impossible. Considering the fact that most imaginable scenarios fit the estimated ages from Strassert, Irisarri, et al. (2021) and that estimated ages in our analysis and theirs have large confidence intervals, it is unlikely that this analysis will strongly argue for or against any scenario on the origin of complex plastids.

What can potentially be of more interest is the use of different calibration sets. I mentioned earlier that *Bangiomorpha* is—if considered red algal—best seen as a stem lineage of Rhodophytina. However, using this fossil as a soft, minimum constrain for the divergence between Rhodophytina and Cyanidiophytina has resulted in the estimated age being much younger than this fossil (Eme et al. 2014). When calibrating only the algal tree of life, few other Proterozoic (older than 540 Ma) fossils remain to be used. However, we here make use of multiple fossils of coralline algae, including stem-coralline algae from the Doushantuo formation (estimated age of 550-635 Ma) (Xiao et al. 2004). Moreover, in this analysis we include the putative green algal fossil of *Proterocladus antiquus* (Tang et al. 2020), allowing us to test the effect of including putative and ancient fossils of red and green algae.

6.5 . Perspectives

The discussion above may have already revealed some of the many outstanding questions in this field. And of course, more questions have popped up after the results from this thesis. Here, I will discuss a few topics where further research could improve our understanding of the early evolution of red algae and red complex plastids.

Will nuclear data confirm the found relationships among Proteorhodophytina?

Our phylogenomic analysis on the organellar genomes of red algae have recovered a clear monophyly of Proteorhodophytina (Chapter 4). However, the relationships among the lineages of Proteorhodophytina are not completely certain as the results of the mitochondrial phylogeny are different from those of the plastid phylogeny. Moreover, deep branches representing the evolutionary distances between these lineages are very short. The phylogenetic reconstruction based on plastid data is likely the most accurate, due to its size and a lack of fast-evolving sequences in the Proteorhodophytina, support from different sources would be ideal. Nuclear genome data, such as proteins predicted from transcriptomes, could be highly useful as many more genes are possible to use, that evolve under different constraints than the plastid and mitochondrial encoded proteins (Baurain et al. 2010; Cavalier-Smith, Chao, and Lewis 2015). Morphological data may also help understand the potential close relationships between some classes of Proteorhodophytina. Identifying characters that, for example, unite the Porphyridiophyceae and Compsopogonophyceae could help to understand the evolutionary path they took.

Why do plastid genomes expand only in rhodellophytes and porphyridiophytes?

Group II introns are highly mobile in the red algal plastid genomes in Rhodellophyceae and Porphyridiophyceae. Even in strains of *Porphyridium purpureum*, where the plastid genomes have not expanded (Chapter 4; Perrineau et al. 2015). We were able to show this intron mobility by comparing the genetic locations of introns between plastid genomes of rhodellophytes and porphyridiophytes, but it is unknown why genome expansion occurred in all but the representatives of *Porphyridium*. Moreover, group II introns are present in the two other classes of Proteorhodophytina (Compsopogonophyceae, Stylonematophyceae), where introns are copied and pasted less frequently. Analysis of these group II introns and their encoded proteins may help understanding why these introns still have mobility and perhaps why they cause genome expansion in only a subset of lineages. Moreover, self-splicing introns—such as group II introns—may aid gene transfer (Pombert et al. 2005; Turmel et al. 2016), which could putatively include gene transfer from the plastid to the nucleus. Plastid genomes of rhodellophytes, although larger than other red algal plastid genomes, actually encode fewer proteins than other red algal plastid genomes. Thus, it may be interesting to see whether these genes have been moved to the nucleus and if introns were responsible for this transfer. Interestingly, the red complex plastids found in chromalveolates all encode fewer genes than red algal plastids, and one could wonder if these gene transfers are related to group II intron mobility as well (Kim and Archibald 2009). Studying the introns of red algae and red complex plastids may potentially show a link, and there are potential first hints of transfer of red algal mitochondrial introns (Kim, Lee, et al. 2022) with similar ones being found in stramenopiles and fungi. Of course, as these introns are also found in fungi, the intron transfer may have little to do with endosymbiotic gene transfer. But with more plastid and mitochondrial genomes now available of Proteorhodophytina, a detailed study may identify intron transfers between red algae and red complex plastids.

Are there red algal out there, not belonging to the currently known classes?

The basis of this thesis has been the organellar genome sequencing of several known red algal species that have been understudied, specifically belonging to Proteorhodophytina. Other red algal lineages remain to be studied, such as two bangiophyte species that are sister to the rest of Bangiophyceae: *Minerva* and *Dione* (Nelson et al. 2005). However, all described red algae seem to nicely place within the seven classes of red algae (Yoon, Muller, et al. 2006), thus one can question how likely it is that we find something more closely related to the elusive alga that was engulfed and reduced to the red complex plastid. Metabarcoding studies at times identify red algae in marine and freshwater environments, although many studies do not detect a single one. Moreover, red algae have typically a relatively large cell size (Yang, Scott, et al. 2010; Scott et al. 2011; Kushibiki et al. 2012), making them easy to spot in samples and making it more questionable why we would still be missing some. This may suggest that the red complex plastids may have outcompeted those closely related red algae. On the other hand, new eukaryotic lineages are frequently identified that may change our view on plastid evolution again (Gawryluk et al. 2019) and again (Wetherbee, Jackson, et al. 2019) and again (Kawachi, Nakayama, et al. 2021). Perhaps we are not sampling the right niches where these algae thrive, or perhaps the typical DNA extraction methods do not destroy the tough mucilage layers of many red algae, making them invisible in metabarcoding studies.

Discovery of new eukaryotic lineages: how will it change our view on plastid origins?

Besides the discovery of new red algae, finding new lineages of red complex plastid-bearing lineages could be exciting as well, even if these lineages still belong to MOCHa lineages. Discovery of the chromerids (Moore et al. 2008; Oborník et al. 2012) has been highly important in understanding the presence of non-photosynthetic plastids in their sister lineages (colpodellids and apicomplexans). Similarly, Rappephyceae are a potentially ecologically significant group of haptophytes of which only one representative has been currently described (Kawachi, Nakayama, et al. 2021). Findings of heterotrophic lineages closely related to photosynthetic groups can also better help us understand what their common ancestor may have looked like, and how the plastid was acquired in the photosynthetic group. Recent examples include *Actinophrys sol*, sister to all ochrophytes (Azuma et al. 2022). Of course, any new eukaryotic lineage may help us reconstruct the phylogeny of eukaryotes with more accuracy, with potentially significant implications to proposed scenarios on the origin of red complex plastids (e.g., Cavalier-Smith, Chao, and Lewis 2015; Yazaki et al. 2022). Finally, there is no reason to think there are no more fascinating algae to find out there, with plastid origins different from what we have seen before. Rather, it would be a surprise if we were unable to find another dinoflagellate out there with a unique plastid replacement. But even further understanding the odd plastids in dinoflagellates, or the membrane loss in the golden paradox may tell us a lot about how plastids can be acquired and changed (Waller and Kořený 2017; Wetherbee, Jackson, et al. 2019).

How much red and green signal can be found in the plastid proteome of haptophytes and cryptophytes?

Green algal signal is present in some lineages with red complex plastids, and an even stronger red signal in lineages with green complex plastids (Moustafa et al. 2009; Dorrell and Smith 2011; Deschamps and Moreira 2012; Ponce-Toledo, Moreira, et al. 2018). The high green signal in ochrophytes is specifically interesting, and could perhaps suggest that stem ochrophytes went through an endosymbiosis with green algae, or at the least a high amount of horizontal gene transfer occurred (Dorrell, Gile, et al. 2017). If ochrophytes (and haptophytes) gained their red complex plastid not by taking up a whole cell, but through something like myzocytosis, it may make sense that some of the plastid-related proteome came from elsewhere, similar to plastid replacements in some dinoflagellates (Dorrell and Howe 2015). Such scenarios have been proposed for these lineages (Bodył 2018), but not for cryptophytes. It would thus be interesting to look into the signal present in cryptophytes, and directly compare this to ochrophytes and haptophytes. If the plastid of cryptophytes mainly work on red algal genes, more so than ochrophytes and haptophytes, it could be ascribed to the different mode of plastid acquisition. Thus, such results could shed light on the origin of red complex plastids.

Which genes are responsible for chlorophyll *c* biosynthesis, and how did they spread among chromalveolates

As discussed in section 1.4.1, we still know very little about chlorophyll *c* biosynthesis, in contrast to chlorophyll *a* and *b* (Qiu et al. 2019; Larkum 2020). It is clear that chlorophyll *c* is only present in certain chromalveolates, which may make it a striking evidence for a singular red algal origin of red complex plastids, followed by the origin of chlorophyll *c* biosynthesis in the lineage that acquired this red complex plastid (Larkum 2020). However, the case of chlorophyll *c* is more complicated, as there are many types of it, and these types have a patchy distribution among chromalveolate lineages (especially ochrophytes and haptophytes; Jeffrey and Wright 2005; Jeffrey, Wright, and Zapata 2011), suggesting that horizontal transfer may play a significant role in the spread of chlorophyll *c*, or that it is relatively easy to evolve. However, which genes function in chlorophyll *c* biosynthesis remains unknown, even if a proposed pathways exist (Xu et al. 2016; Qiu et al. 2019). Identifying the genes that play a role in its biosynthesis in different lineages will help us understand why only chromalveolates use chlorophyll *c*. And comparative and phylogenetic studies of these genes may reveal how the different lineages have acquired this biosynthetic pathway.

7 - Conclusions

Proteorhodophytina is a monophyletic lineage.

Using both plastid- and mitochondrial-based phylogenies, we have shown that the subphylum Proteorhodophytina is monophyletic, further building on results from previous studies (Muñoz-Gómez, Mejía-Franco, et al. 2017; Preuss et al. 2021). This monophyly remains well supported after the inclusion of red complex plastids in the red algal phylogeny.

Mobile group II introns are largely responsible for plastid genome expansion in rhodellophytes and porphyridiophytes.

Group II introns are present in all representatives of Proteorhodophytina, but the mobility of introns is low in representatives of Compsopogonophyceae and Stylonematophyceae. The mobility of introns is however high in the plastid genomes of rhodellophytes and porphyridiophytes, resulting in many genome rearrangements.

Recent lineage-specific intron proliferation is the cause for plastid genome expansion in *Corynoplastis japonica* and *Rhodella violacea*.

The two largest red algal plastid genomes are found in *Corynoplastis japonica* and *Rhodella violacea*. Their expansion is clearly caused by group II intron proliferation, but intriguingly these events are lineage-specific, as the genomic location of introns is different, and many group II introns of *R. violacea* are in an early stage of degradation.

Mitochondrial genomes of Proteorhodophytina are highly variable in size and intron content.

Not only do many proteorhodophytes carry expanded plastid genomes, many contain expanded mitochondrial genomes. These are found in almost all classes of Proteorhodophytina, except in Rhodellophyceae.

Direct intron transfer between organellar genomes is unlikely to be the cause of genome expansion in proteorhodophytes.

While both expanded plastid and mitochondrial genomes are found in representatives of Proteorhodophytina, only one species as both organellar genomes expanded. As there is no transfer of intron-encoded proteins between these genomes, we suggest that plastid and mitochondrial genome expansion are not directly related.

If plastids were serially transferred between chromist lineages, the ochrophyte plastid was acquired before the divergence of haptophytes (or plastid acquisition in haptophytes).

Considering the number of plastid losses implied by the chromalveolate hypothesis, a serial endosymbiosis scenario may be more parsimonious to explain the presence of plastids in cryptophytes, haptophytes, ochrophytes (and myzozoans). Our phylogenomic reconstructions showed that the plastids of haptophytes and cryptophytes share a more recent common ancestor than they do with ochrophytes, which is also supported by the *rpl36* gene replacement shared by these lineages. If the haptophytes gained their plastid from cryptophytes, this must have occurred after the plastid acquisition in ochrophytes. If instead the plastid of haptophytes and cryptophytes was gained in their common ancestor, the ochrophyte plastid must have been acquired before the diversification of those two groups.

A separate acquisition of the plastid of ochrophytes cannot be excluded.

The connection between haptophyte and cryptophyte plastids is clear, but how exactly they relate to the ochrophyte plastid is unknown. Our phylogenomic analyses indicate that a separate origin of the ochrophyte plastid cannot be excluded, even if it is not the most parsimonious scenario.

French summary

Introduction

L'énergie a été nécessaire à la vie pour apparaître et pour continuer à exister. La plus grande source d'énergie sur Terre est la lumière du soleil, une petite partie de celle-ci est captée par les organismes vivants pour se maintenir et prospérer. Le principal processus - mais pas le seul - qui permet de capter la lumière solaire est la photosynthèse, que plusieurs groupes bactériens et eucaryotes utilisent pour synthétiser des molécules d'hydrates de carbone et créer ainsi la majeure partie du carbone organique nécessaire à la vie hétérotrophe (Thornton 2012). Les plantes, les algues et les cyanobactéries réalisent une photosynthèse oxygénique car elles produisent de l'oxygène pendant la photosynthèse. La photosynthèse oxygénique (que dorénavant je nommerai "photosynthèse") a commencé avec les cyanobactéries, au moins avant la Grande Oxygénation, qui s'est produite 2,4-2,0 Ga (Giga-annum, il y a 10^9 années), où l'oxygène a commencé à s'accumuler dans l'atmosphère alors essentiellement anoxique (Rasmussen et al. 2008 ; Schirrmeister et al. 2016).

De nos jours, les cyanobactéries ne sont pas les seules à réaliser la photosynthèse, de nombreux eucaryotes dont les plantes terrestres et une multitude de lignées d'algues la réalisent également. Leur diversité phylogénétique, écologique et morphologique est grande, incluant les plantes à fleurs, les algues coralliennes ou les laminaires. Mais comment cette grande diversité a-t-elle acquis la capacité à réaliser la photosynthèse ? Les algues rouges, les glaucophytes, les algues vertes et les plantes terrestres contiennent des organites directement dérivés d'un endosymbiont cyanobactérien, appelés plastes primaires. Mais de nombreux autres eucaryotes possèdent des plastes dites complexes, car dérivés d'un eucaryote déjà porteur de plastes. Tous les plastes complexes connus sont issus d'algues vertes ou rouges, et aucun n'est issu de glaucophytes (Archibald 2009 ; Burki 2017).

Le cas des plastes complexes rouges est particulièrement intéressant, car on ne sait toujours pas comment les différentes lignées ont acquis ces plastes. On trouve des plastes complexes rouges chez les diatomées, les algues brunes, les coccolithophores et même chez des organismes non photosynthétiques comme l'agent pathogène *Plasmodium*. Toutes ces lignées se retrouvent dans quatre grands groupes eucaryotes : Myzozoa (dont *Plasmodium*), Ochrophyta (dont les diatomées et les laminaires), Cryptophyta et Haptophyta (dont les coccolithophores), appelés ensemble MOCHa (figure 8.1; Burki 2017).

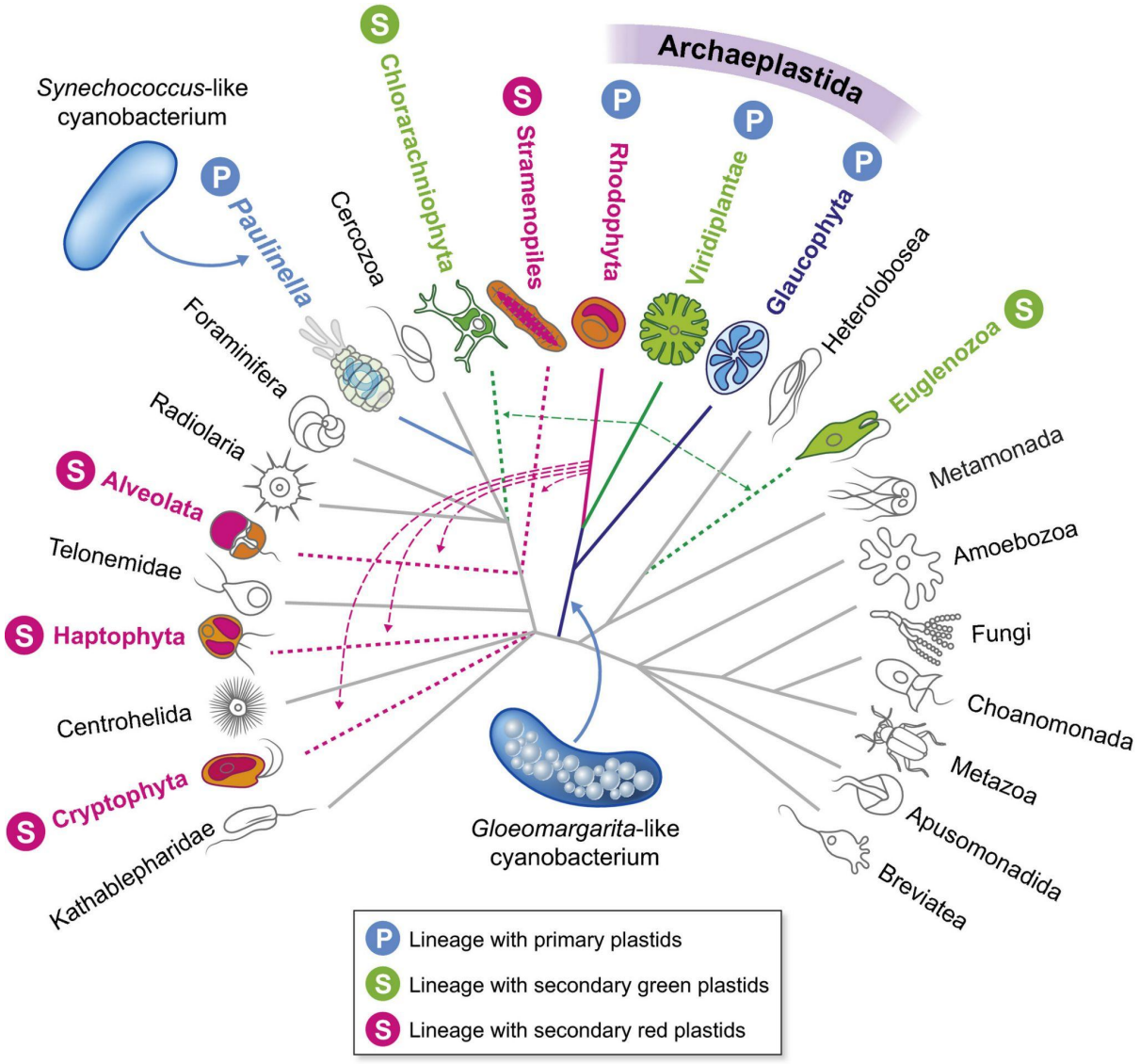


Figure 8.1 : Distribution de la photosynthèse dans la phylogénie eucaryote globale. Les branches pleines colorées correspondent aux lignées photosynthétiques dotées de plastes primaires et les branches pointillées colorées aux lignées dotées de plastes secondaires (les couleurs verte et rouge indiquent le type d'endosymbiont secondaire, respectivement les algues vertes ou rouges). Les flèches bleues montrent les deux endosymbioses primaires connues (chez Archaeplastida et *Paulinella*) et les flèches vertes et rouges indiquent les endosymbioses secondaires impliquant des endosymbiontes d'algues vertes et rouges. Les branches grises correspondent aux phyla eucaryotes non photosynthétiques. L'arbre a été largement modifié à partir de Adl et al. (2012). Image et légende de Ponce-Toledo et al. (2019), avec l'autorisation des auteurs.

De multiples hypothèses ont été proposées pour expliquer l'origine des plastes de ces diverses lignées, mais on n'est pas encore parvenu à un consensus. L'hypothèse des chromalvéolés proposait qu'un seul événement endosymbiotique entre une algue rouge et l'ancêtre commun des MOCHa se soit produit (Cavalier-Smith 1999). Cependant, cet ancêtre commun aurait eu de nombreux descendants dépourvus de plastes rouges, comme les oomycètes, les ciliés, les centrohelides et les rhizariens (Burki et al. 2020 ; Sibbald et Archibald 2020). Cette hypothèse implique donc qu'au moins dix lignées ont perdu leur plaste (Burki 2017). De plus, alors que les phylogénies plastidiques permettent de retrouver facilement la monophylie des chromalvéolés, il est beaucoup plus difficile de la retrouver avec des données nucléaires ou mitochondriales, même en excluant les lignées sœurs dépourvues de plastes (Baurain et al. 2010).

Pour faire face à ces divergences, de nombreux auteurs ont proposé différents scénarios d'endosymbiose en série où le plaste complexe rouge a été acquis une seule fois directement à partir d'une algue rouge, suivi de transferts horizontaux entre les lignées de MOCHa (Sanchez-Puerta et Delwiche 2008 ; Bodył et al. 2009 ; Dorrell et Smith 2011 ; Petersen et al. 2014 ; Stiller et al. 2014). Les versions de cette hypothèse diffèrent dans les événements endosymbiotiques proposés et leur ordre. Par exemple, Stiller et al. (2014) ont proposé que le plaste ait été transféré des cryptophytes aux ochrophytes, suivi d'un transfert des ochrophytes aux haptophytes. Ce scénario implique que les ochrophytes et les haptophytes devraient être des lignées sœurs dans les phylogénies des plastes. Néanmoins, ces implications n'ont pas été testées de manière approfondie dans un cadre phylogénomique des plastes.

Pour mieux comprendre les relations entre les plastes des lignées de MOCHa, nous avons commencé par améliorer l'échantillonnage des taxons d'algues rouges. Les algues rouges unicellulaires sont plus susceptibles d'être étroitement liées à l'ancêtre commun des plastes trouvés dans le MOCHa, car il était probablement aussi unicellulaire. Pourtant, elles restent peu étudiées, en particulier les algues rouges unicellulaires du sous-phylum Proteorhodophytina (Muñoz-Gómez, Mejía-Franco, et al. 2017). Nous avons étudié les génomes plastidiaux et mitochondriaux de ces taxa, non seulement pour comprendre l'origine des plastes complexes rouges, mais aussi l'évolution du génome dans ce groupe fascinant d'algues rouges. Ensuite, nous avons utilisé ces nouvelles données plastidiques pour reconstruire la phylogénie des algues rouges et des plastes complexes rouges, et pour déterminer quelles hypothèses d'endosymbiose en série correspondent à nos données. De plus, nous utilisons également les archives fossiles des algues et des plantes terrestres pour estimer les dates de ces événements endosymbiotiques.

Expansions de taille indépendantes et prolifération d'introns dans les génomes plastidiaux et mitochondriaux des algues rouges

Les algues rouges constituent un groupe diversifié d'eucaryotes photosynthétiques, composé d'algues unicellulaires, filamenteuses et multicellulaires complexes. Les algues rouges sont parfois considérées comme ancestrales ou à évolution lente, en partie en raison de leurs génomes plastidiaux riches en gènes (Janouškovec et al. 2013). Les génomes plastidiaux ont été largement étudiés pour un sous-ensemble d'espèces multicellulaires de Bangiophyceae et Florideophyceae, et pour les espèces unicellulaires extrêmophiles de Cyanidiophyceae, mais d'autres classes d'algues rouges (Stylonematophyceae, Compsopogonophyceae, Rhodellophyceae, Porphyridiophyceae) sont restées pour la plupart négligées (Cao et al. 2018 ; Janouškovec et al. 2013). Les quelques génomes plastidiaux étudiés dans ses quatre classes - constituant ensemble le sous-phylum Proteorhodophytina - sont de taille très diverse et varient en termes de synténie génétique et de contenu en introns du groupe II (Muñoz-Gómez, Mejía-Franco, et al. 2017 ; Preuss et al. 2021). Les introns du groupe II peuvent copier et coller leurs séquences lorsqu'ils sont complets (Jacquier et Dujon 1985), ce qui explique probablement leur contenu élevé dans ces grands génomes. De plus, le génome mitochondrial est particulièrement peu étudié chez les Proteorhodophytina, avec *Compsopogon caeruleus* comme seul représentant (Nan et al. 2017).

Pour combler cette lacune dans notre compréhension des génomes organellaires des algues rouges, nous avons séquencé 25 génomes plastidiaux et mitochondriaux d'espèces unicellulaires et filamenteuses, principalement des représentants de Proteorhodophytina. Nous avons utilisé ces données pour étudier (1) les relations entre les classes d'algues rouges basées sur les données plastidiales et mitochondriales, en mettant l'accent sur la monophylie de Proteorhodophytina, (2) les événements d'expansion du génome plastidial et (3) la diversité des génomes mitochondriaux.

Muñoz-Gómez, Mejía-Franco et al. (2017) ont trouvé que les quatre classes d'algues rouges simples mésophiles étaient monophylétiques, les plaçant dans le sous-phylum Proteorhodophytina. Nous avons réexaminé la phylogénie des algues rouges en utilisant un ensemble de données sur les plastes et les mitochondries. Nous avons trouvé un fort support pour la monophylie des Proteorhodophytina en utilisant les deux ensembles de données, bien que la relation interne entre les quatre classes des Proteorhodophytina diffère entre les deux ensembles de données (Fig. 8.2). Cependant, les tests topologiques ont montré que la phylogénie dérivée des plastes (Fig. 8.2, à gauche) ne peut pas être rejetée

par les données mitochondriales. Comme les branches profondes des Proteorhodophytina sont très courtes et que l'ensemble de données mitochondriales est plus petit et contient des lignées très divergentes, il est plus probable que la topologie plastidiale (Fig. 8.2, à gauche) reflète les véritables relations.

Les introns du groupe II sont présents dans tous les génomes plastidiaux des Proteorhodophytina, mais les génomes restent généralement denses en gènes, comme c'est le cas aussi pour les génomes plastidiaux des algues rouges de la classe Florideophyceae. Dans certains cas, ces introns ont proliféré, entraînant une expansion du génome, le plus grand étant celui de *Corynoplastis japonica* à 1,1 Mb (Muñoz-Gómez, Mejía-Franco, et al. 2017). Nous constatons que, bien que ces grands génomes soient limités à deux classes unicellulaires - les Rhodellophyceae et Porphyridiophyceae, leur contenu en introns suggère que les événements d'expansion sont souvent indépendants. Ceci est encore plus clair pour les génomes mitochondriaux, où les espèces avec de grands génomes mitochondriaux (~100 kb) sont souvent sœurs des espèces qui en ont un petit (30-40 kb). Les Stylonematophyceae contiennent les génomes mitochondriaux les plus divergents, avec un contenu en GC élevé et moins d'introns qu'attendu. Bien que l'expansion des génomes des deux organites soit principalement causée par l'invasion des introns du groupe II, nous ne trouvons pas de transfert récent de ces introns entre les génomes organellaires.

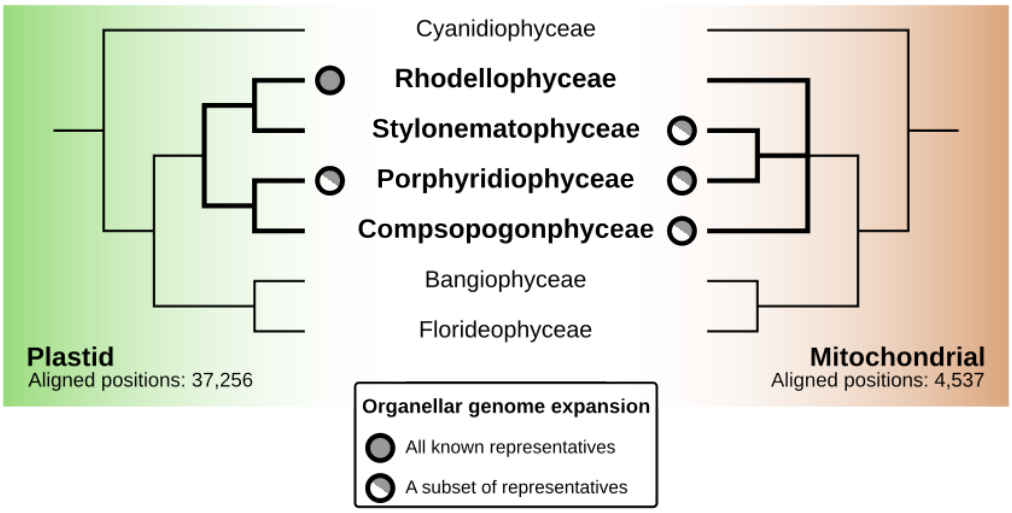


Figure 8.2 : Phylogénie schématique des algues rouges, mettant en évidence les relations entre les quatre classes de Proteorhodophytina (en gras), sur la base des données relatives aux plastes (à gauche, en vert) et aux mitochondries (à droite, en marron). La présence d'une expansion du génome plastidial et mitochondrial est indiquée à côté des classes, où l'expansion du génome est définie comme un génome plastidial > 250 kb et un génome mitochondrial > 60 kb.

Évaluation des hypothèses d'endosymbioses en série à la lumière de la phylogénomique des plastes rouges

Comme indiqué dans l'introduction, de multiples hypothèses ont été proposées pour expliquer la diversité des plastes rouges complexes. Les hypothèses d'endosymbiose en série diffèrent principalement quant aux événements endosymbiotiques qui se sont produits et à leur ordre. Par exemple, Bodył et al. (2009) proposent qu'un ancêtre des cryptophytes a d'abord acquis son plaste à partir d'algues rouges, qui a ensuite été transféré aux ochrophytes, puis à nouveau des cryptophytes aux haptophytes. D'autres transferts des haptophytes et des ochrophytes vers différentes lignées de myzozoaires auraient également eu lieu. Ces types de scénarios devraient donner lieu à différentes phylogénies des plastes, les relations entre les lignées de myzozoaires variant selon ces scénarios. Cependant, ces hypothèses n'ont pas encore été testées de manière approfondie. En outre, il a été démontré, sur la base de phylogénies eucaryotes calibrées avec des fossiles connus, que les hypothèses d'endosymbiose en série sont possibles car les lignées qui ont subi une endosymbiose se chevauchent dans les temps géologiques (Strassert et al. 2021). On peut soutenir que les âges estimés des événements de transfert de plastes trouvés dans leurs phylogénies devraient correspondre aux âges estimés de leurs hôtes, mais ceci n'a pas encore été testé.

Nous avons créé deux ensembles de données phylogénomiques, l'un constitué des protéines codées dans les plastes d'algues rouges et des lignées porteuses de plastes complexes rouges, et un autre avec les mêmes lignées plus les algues vertes et les glaucophytes. L'échantillonnage taxonomique d'algues rouges, en particulier des Proteorhodophytina, a été largement amélioré grâce à notre étude et à des études précédentes (Muñoz-Gómez, Mejía-Franco, et al. 2017). En utilisant l'inférence bayésienne (IB), les plastes complexes rouges forment un groupe monophylétique au sein des algues rouges, comme trouvé précédemment (Muñoz-Gómez, Mejía-Franco, et al. 2017 ; Yoon, Hackett, et al. 2002). Cependant, en utilisant le maximum de vraisemblance (ML), les ochrophytes se placent plutôt comme groupe sœur de toutes les autres plastes rouges (y compris les algues rouges). Nous montrons que ce placement est cependant sensible, dans une certaine mesure, à la suppression de sites à évolution rapide. Comme il existe des préférences en matière d'acides aminés parmi tous les groupes porteurs de plastes rouges, nous avons décidé de recoder les données, ce qui n'a cependant donné aucun résultat cohérent. Nous en déduisons qu'une origine algale rouge unique reste l'hypothèse la plus appropriée, car le modèle probablement le mieux ajusté dans les analyses BI (CAT+GTR+G4), retrouve les plastes complexes rouges comme monophylétiques (Fig. 8.3).

Bien entendu, l'hypothèse des chromalvéolés et celle de l'endosymbiose en série le supposent.

Nous constatons également que les plastes des haptophytes et des cryptophytes sont étroitement apparentés, ce qui est bien étayé dans les cadres IB et ML, et le reste après l'élimination des sites à évolution rapide et le recodage des acides aminés pour réduire les biais de composition. Les hypothèses d'endosymbiose en série qui proposent un transfert de plastes de la lignée souche des haptophytes à la lignée souche des ochrophytes (Sanchez-Puerta et Delwiche 2008 ; Stiller, Schreiber, et al. 2014) doivent être réfutées à la lumière de ces résultats. Au contraire, nos données favorisent l'hypothèse proposée par Bodył et al, bien que de nombreuses incertitudes subsistent.

Enfin, nous avons construit cinq ensembles de calibrations fossiles pour étalonner la phylogénie des plastes. Ces ensembles varient dans leur utilisation de fossiles controversés tels que *Bangiomorpha* et *Proterocladus* (Butterfield 2000 ; Eme et al. 2014 ; Parfrey et al. 2011 ; Tang et al. 2020), ainsi que de plusieurs contraintes maximales qui sont sujettes à débat (Hedges et al. 2018 ; Morris et al. 2018a,b). Ces analyses sont encore en cours, mais elles nous aideront à estimer l’âge de certains transferts de plastes.

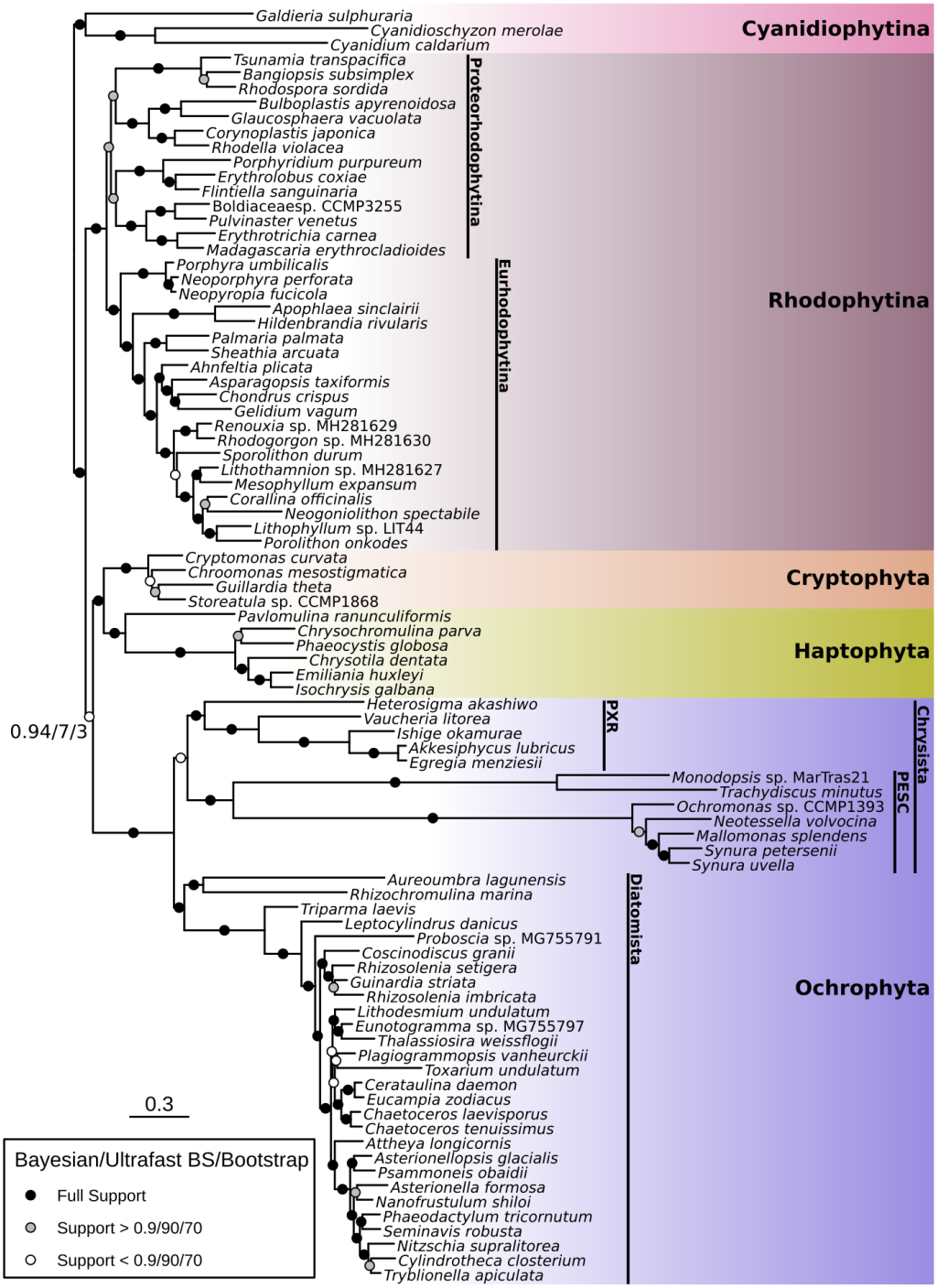


Figure 8.3: Reconstruction phylogénétique des algues rouges et des lignées avec des plastes complexes rouges basée sur les données des génomes plastidiaux. L'arbre a été reconstruit à partir d'un alignement de 140 séquences de protéines codées par les plastes (27 658 positions), en utilisant à la fois l'inférence bayésienne (CAT+GTR+G4) et le maximum de vraisemblance (cpREV+C60+F+R8 avec un support bootstrap ultra-rapide et cpREV+C60+R8+PMSF avec un support bootstrap non-paramétrique). Les valeurs de soutien sont indiquées par des cercles colorés (voir la légende), le soutien pour la monophylie des plastes complexes rouges est également explicitement écrit sur la figure, ce qui reflète l'absence de monophylie dans les analyses de maximum de vraisemblance . PXR = Phaeophyceae, Xanthophyceae et Raphidophyceae ; ESC = Eustigmatophyceae, Synurophyceae, Chrysophyceae.

Discussion et perspectives

Ce travail a amélioré notre compréhension de l'arbre de vie des algues rouges. En utilisant des phylogénies basées sur les plastes et les mitochondries, nous avons montré que le sous-phylum Proteorhodophytina est monophylétique, ce qui renforce les résultats d'études précédentes (Muñoz-Gómez, Mejía-Franco, et al. 2017 ; Preuss et al. 2021). Cette monophylie reste bien soutenue après l'inclusion des plastes complexes rouges dans la phylogénie des algues rouges (Fig. 8.3). Outre leur phylogénie, nous avons également amélioré notre compréhension de l'évolution des organelles dans le sous-phylum Proteorhodophytina. Les génomes plastidiaux de grande taille étaient déjà connus mais seulement sur la base d'un petit échantillonnage de taxons (Muñoz-Gómez, Mejía-Franco, et al. 2017), de plus à l'époque seul un génome mitochondrial avait été séquencé pour un représentant des Proteorhodophytina. Tous les génomes plastidiaux des Proteorhodophytina contiennent des introns du groupe II, mais la mobilité des introns est faible chez les représentants des Compsopogonophyceae et des Stylonematophyceae. La mobilité des introns est cependant élevée dans les génomes plastidiaux des rhodellophytes et des porphyridiophytes, ce qui entraîne de nombreux réarrangements du génome. Les deux plus grands génomes plastidiaux d'algues rouges se trouvent chez *Corynoplastis japonica* et *Rhodella violacea*. Leur expansion est clairement causée par la prolifération d'introns du groupe II, mais, chose intrigante, ces événements sont spécifiques de la lignée, car l'emplacement génomique des introns est différent, et de nombreux introns du groupe II de *R. violacea* sont à un stade précoce de dégradation. Non seulement de nombreux proteorhodophytes portent des génomes plastidiaux plus grand, mais beaucoup contiennent des génomes mitochondriaux étendus. Ceux-ci sont trouvés dans toutes les classes de Proteorhodophytina excepté dans les Rhodellophyceae. Alors que les génomes plastidiaux et mitochondriaux plus grand sont trouvés dans les représentants de Proteorhodophytina, une seule espèce a les deux génomes organellaires expansés. Comme il n'y a pas de transfert de protéines codées par les introns entre ces génomes, nous suggérons que l'expansion des génomes plastidiaux et mitochondriaux n'est pas directement liée.
Nos analyses phylogénomiques des plastes des algues rouges et des plastes complexes rouges ont des implications importantes pour les hypothèses actuelles sur l'origine des plastes complexes rouges. Nous avons constaté que les plastes des ochrophytes sont difficiles à placer dans nos arbres, leur position changeant en fonction des méthodes et des modèles d'évolution des séquences utilisés. Cela peut indiquer une origine indépendante de leur plaste, bien qu'une explication plus parcimonieuse puisse être une origine tertiaire précoce de ce plaste. Enfin, nous avons montré que les plastes des cryptophytes et des haptophytes sont très étroitement apparentés. Cette constatation est appuyée par le fait que

le gène *rpl36* a été remplacé dans les génomes des plastes des deux groupes (Rice et Palmer 2006). Si l'hypothèse des chromalvéolés est correcte, cela impliquerait que ces lignées soient plus étroitement apparentées entre elles qu'aux ochrophytes. Compte tenu du nombre de pertes de plastes impliquées par cette hypothèse, un scénario d'endosymbiose en série pourrait être plus parcimonieux pour expliquer la présence de plastes chez les cryptophytes, les haptophytes, les ochrophytes (et les myzozoaires) (Burki 2017). Si l'origine du plaste des haptophytes est un événement de symbiose tertiaire avec une cryptophyte comme endosymbiont, alors cet événement doit avoir eu lieu après l'acquisition du plaste de type cryptophyte chez les ochrophytes. Ces résultats sont incompatibles avec certains scénarios proposés (Sanchez-Puerta et Delwiche 2008 ; Stiller et al. 2014).

Bien sûr, de nombreuses questions restent ouvertes. Les relations entre les lignées de Proteorhodophytina ne sont pas complètement certaines car les résultats de la phylogénie mitochondriale sont différents de ceux de la phylogénie plastidiale. De plus, les branches profondes représentant les distances évolutives entre ces lignées sont très courtes et difficiles à résoudre. La reconstruction phylogénétique basée sur les données des plastes est probablement la plus précise, en raison de la plus grande taille du jeu de données et de l'absence de séquences à évolution rapide dans les Proteorhodophytina, mais un support supplémentaire pourrait être nécessaire. Les données du génome nucléaire, comme les protéines prédites à partir des transcriptomes, pourraient être très utiles car elles permettent d'utiliser beaucoup plus de gènes qui évoluent sous des contraintes différentes de celles des protéines codées par les plastes et les mitochondries (Baurain et al. 2010 ; Cavalier-Smith, Chao et Lewis 2015). En outre, l'échantillonnage taxonomique des algues rouges pourrait être amélioré. En particulier, la découverte d'algues rouges en dehors des sept classes connues serait étonnante, mais il est difficile de dire s'il existe des algues rouges qui ne font pas partie de ces classes connues. D'autres lignées d'algues rouges restent à étudier, comme deux espèces de bangiophytes qui sont sœurs du reste des Bangiophyceae : *Minerva* et *Dione* (Nelson et al. 2005). En général, la découverte de nouvelles lignées eucaryotes et la génération de données pour ces lignées seront également importantes pour améliorer notre compréhension de l'évolution des plastes rouges complexes. Par exemple, une lignée sœur récente d'ochrophytes a été découverte et séquencée (Azuma et al. 2022). Il reste ainsi de nombreuses possibilités d'améliorer notre compréhension de l'évolution des plastes.

Bibliography

Aboal, M. et al. (2018). "Diversity of *Chroothece* (Rhodophyta, Stylonematales) Including Two New Species". In: *European Journal of Phycology* 53.2, pp. 189–197.

Abrahamsen, M. S. et al. (2004). "Complete Genome Sequence of the Apicomplexan, *Cryptosporidium parvum*". In: *Science* 304.5669, pp. 441–445.

Adachi, J. et al. (2000). "Plastid Genome Phylogeny and a Model of Amino Acid Substitution for Proteins Encoded by Chloroplast DNA". In: *Journal of Molecular Evolution* 50.4, pp. 348–358.

Adl, S. M., D. Bass, et al. (2019). "Revisions to the Classification, Nomenclature, and Diversity of Eukaryotes". In: *Journal of Eukaryotic Microbiology* 66.1, pp. 4–119.

Adl, S. M., A. G. B. Simpson, et al. (2005). "The New Higher Level Classification of Eukaryotes with Emphasis on the Taxonomy of Protists". In: *Journal of Eukaryotic Microbiology* 52.5, pp. 399–451.

Adler, S. et al. (2014). "*Rhopalodia gibba*: The First Steps in the Birth of a Novel Organelle?" In: *Endosymbiosis*. Ed. by W. Löffelhardt. Vienna: Springer Vienna, pp. 167–179.

Altschul, S. F. et al. (1997). "Gapped BLAST and PSI-BLAST: A New Generation of Protein Database Search Programs". In: *Nucleic Acids Research* 25.17, pp. 3389–3402.

Andrews, S. (2016). *FastQC: A Quality Control Tool for High Throughput Sequence Data*. Version 0.11.9.

Aono, N. et al. (2002). "Palindromic Repetitive Elements in the Mitochondrial Genome of *Volvox*". In: *FEBS Letters* 521.1, pp. 95–99.

Archibald, J. M. (2009). "The Puzzle of Plastid Evolution". In: *Current Biology* 19.2, R81–R88.

— (2012). "Chapter Three - The Evolution of Algae by Secondary and Tertiary Endosymbiosis". In: *Advances in Botanical Research*. Ed. by G. Piganeau. Vol. 64. Genomic Insights into the Biology of Algae. Academic Press, pp. 87–118.

— (2015). "Endosymbiosis and Eukaryotic Cell Evolution". In: *Current Biology* 25.19, R911–R921.

— (2020). "Cryptomonads". In: *Current Biology* 30.19, R1114–R1116.

Azuma, T. et al. (2022). "An Enigmatic Stramenopile Sheds Light on Early Evolution in Ochrophyta Plastid Organellogenesis". In: *Molecular Biology and Evolution*, msac065.

Baurain, D. et al. (2010). "Phylogenomic Evidence for Separate Acquisition of Plastids in Cryptophytes, Haptophytes, and Stramenopiles". In: *Molecular Biology and Evolution* 27.7, pp. 1698–1709.

Bayer, M. G. and H. E. A. Schenk (1986). "Biosynthesis Of Proteins In *Cyanophora paradoxa*. I. Protein Import Into The Endocyanelle Analyzed By Micro Two- Dimensional Gel Electrophoresis". In: p. 6.

Bengtson, S. et al. (2017). "Three-Dimensional Preservation of Cellular and Subcellular Structures Suggests 1.6 Billion-Year-Old Crown-Group Red Algae". In: *PLOS Biology* 15.3, e2000735.

Berney, C. and J. Pawlowski (2006). "A Molecular Time-Scale for Eukaryote Evolution Recalibrated with the Continuous Microfossil Record". In: *Proceedings of the Royal Society B: Biological Sciences* 273.1596, pp. 1867–1872.

Betts, H. C. et al. (2018). "Integrated Genomic and Fossil Evidence Illuminates Life's Early Evolution and Eukaryote Origin". In: *Nature Ecology & Evolution* 2.10 (10), pp. 1556–1562.

Bhattacharya, D. and J. M. Archibald (2006). "Response to Theissen and Martin". In: *Current Biology* 16.24, R1017–R1018.

Bhattacharya, D., J. J. Cannone, and R. R. Gutell (2001). "Group I Intron Lateral Transfer between Red and Brown Algal Ribosomal RNA". In: *Current Genetics* 40.1, pp. 82–90.

Bhattacharya, D., H. Qiu, et al. (2018). "When Less Is More: Red Algae as Models for Studying Gene Loss and Genome Evolution in Eukaryotes". In: *Critical Reviews in Plant Sciences* 37.1, pp. 81–99.

Bodył, A. (2005). "Do Plastid-Related Characters Support the Chromalveolate Hypothesis?" In: *Journal of Phycology* 41.3, pp. 712–719.

— (2018). "Did Some Red Alga-Derived Plastids Evolve via Kleptoplastidy? A Hypothesis". In: *Biological Reviews* 93.1, pp. 201–222.

Bodył, A., J. W. Stiller, and P. Mackiewicz (2009). "Chromalveolate Plastids: Direct Descent or Multiple Endosymbioses?" In: *Trends in Ecology & Evolution* 24.3, pp. 119–121.

Bolger, A. M., M. Lohse, and B. Usadel (2014). "Trimmomatic: A Flexible Trimmer for Illumina Sequence Data". In: *Bioinformatics* 30.15, pp. 2114–2120.

Bolte, K. et al. (2009). "Protein Targeting into Secondary Plastids". In: *Journal of Eukaryotic Microbiology* 56.1, pp. 9–15.

Bradbury, J. (2004). "Nature's Nanotechnologists: Unveiling the Secrets of Diatoms". In: *PLOS Biology* 2.10, e306.

Brasier, M. D. et al. (2015). "Changing the Picture of Earth's Earliest Fossils (3.5–1.9 Ga) with New Approaches and New Discoveries". In: *Proceedings of the National Academy of Sciences* 112.16, pp. 4859–4864.

Burger, G. et al. (1999). "Complete Sequence of the Mitochondrial DNA of the Red Alga *Porphyra purpurea*. Cyanobacterial Introns and Shared Ancestry of Red and Green Algae." In: *The Plant Cell* 11.9, pp. 1675–1694.

Burki, F. (2017). "Chapter One - The Convoluted Evolution of Eukaryotes With Complex Plastids". In: *Advances in Botanical Research*. Ed. by Y. Hirakawa. Vol. 84. Secondary Endosymbioses. Academic Press, pp. 1–30.

Burki, F., M. Kaplan, et al. (2016). "Untangling the Early Diversification of Eukaryotes: A Phylogenomic Study of the Evolutionary Origins of Centrohelida, Haptophyta and Cryptista". In: *Proceedings of the Royal Society B: Biological Sciences* 283.1823, p. 20152802.

Burki, F., N. Okamoto, et al. (2012). "The Evolutionary History of Haptophytes and Cryptophytes: Phylogenomic Evidence for Separate Origins". In: *Proceedings of the Royal Society B: Biological Sciences* 279.1736, pp. 2246–2254.

Burki, F., A. J. Roger, et al. (2020). "The New Tree of Eukaryotes". In: *Trends in Ecology & Evolution* 35.1, pp. 43–55.

Burki, F., K. Shalchian-Tabrizi, et al. (2007). "Phylogenomics Reshuffles the Eukaryotic Supergroups". In: *PLOS ONE* 2.8, e790.

Bushnell, B. (2014). *BBMap: A Fast, Accurate, Splice-Aware Aligner*. LBNL-7065E. Lawrence Berkeley National Lab. (LBNL), Berkeley, CA (United States).

Butterfield, N. J. (2000). "*Bangiomorpha pubescens* n. Gen., n. Sp.: Implications for the Evolution of Sex, Multicellularity, and the Mesoproterozoic/Neoproterozoic Radiation of Eukaryotes". In: *Paleobiology* 26.3, pp. 386–404.

Butterfield, N. J., A. H. Knoll, and K. Swett (1990). "A Bangiophyte Red Alga from the Proterozoic of Arctic Canada". In: *Science* 250.4977, pp. 104–107.

Cao, M. et al. (2018). "The First Plastid Genome of a Filamentous Taxon '*Bangia*' Sp . OUCPT-01 in the Bangiales". In: *Scientific Reports* 8.1 (1), p. 10688.

Carlisle, E. M. et al. (2021). "Experimental Taphonomy of Organelles and the Fossil Record of Early Eukaryote Evolution". In: *Science Advances* 7.5, eabe9487.

Carver, T. et al. (2012). "Artemis: An Integrated Platform for Visualization and Analysis of High-Throughput Sequence-Based Experimental Data". In: *Bioinformatics* 28.4, pp. 464–469.

Cavalier-Smith, T. (1981). "Eukaryote Kingdoms: Seven or Nine?" In: *Biosystems* 14.3, pp. 461–481.

— (1982). "The Origins of Plastids". In: *Biological Journal of the Linnean Society* 17.3, pp. 289–306.

— (1998). "A Revised Six-Kingdom System of Life". In: *Biological Reviews* 73.3, pp. 203–266.

— (2002). "The Neomuran Origin of Archaebacteria, the Negibacterial Root of the Universal Tree and Bacterial Megaclassification." In: *International Journal of Systematic and Evolutionary Microbiology* 52.1, pp. 7–76.

— (2010). "Kingdoms Protozoa and Chromista and the Eozoan Root of the Eukaryotic Tree". In: *Biology Letters* 6.3, pp. 342–345.

— (2016). "Higher Classification and Phylogeny of Euglenozoa". In: *European Journal of Protistology* 56, pp. 250–276.

Cavalier-Smith, T., M. T. Allsopp, and E. E. Chao (1994). "Chimeric Conundra: Are Nucleomorphs and Chromists Monophyletic or Polyphyletic?" In: *Proceedings of the National Academy of Sciences* 91.24, pp. 11368–11372.

Cavalier-Smith, T. and E. E. Chao (1996). "18S rRNA Sequence of *Heterosigma carterae* (Raphidophyceae), and the Phylogeny of Heterokont Algae (Ochrophyta)". In: *Phycologia* 35.6, pp. 500–510.

Cavalier-Smith, T. and E. E. Chao (2004). "Protalveolate Phylogeny and Systematics and the Origins of Sporozoa and Dinoflagellates (Phylum Myzozoa Nom. Nov.)" In: *European Journal of Protistology* 40.3, pp. 185–212.

Cavalier-Smith, T., E. E. Chao, and R. Lewis (2015). "Multiple Origins of Heliozoa from Flagellate Ancestors: New Cryptist Subphylum Corbihelia, Superclass Corbistoma, and Monophyly of Haptista, Cryptista, Hacrobia and Chromista". In: *Molecular Phylogenetics and Evolution* 93, pp. 331–362.

Cavalier-Smith, T. and J. J. Lee (1985). "Protozoa as Hosts for Endosymbioses and the Conversion of Symbionts into Organelles". In: *The Journal of Protozoology* 32.3, pp. 376–379.

Cavalier-Smith, T. (1999). "Principles of Protein and Lipid Targeting in Secondary Symbiogenesis: Euglenoid, Dinoflagellate, and Sporozoan Plastid Origins and the Eukaryote Family Tree". In: *Journal of Eukaryotic Microbiology* 46.4, pp. 347–366.

Cenci, U. et al. (2018). "Nuclear Genome Sequence of the Plastid-Lacking Cryptomonad *Goniomonas avonlea* Provides Insights into the Evolution of Secondary Plastids". In: *BMC Biology* 16.1, p. 137.

Chan, P. P. and T. M. Lowe (2019). "tRNAscan-SE: Searching for tRNA Genes in Genomic Sequences". In: *Gene Prediction: Methods and Protocols*. Ed. by M. Kollmar. Methods in Molecular Biology. New York, NY: Springer, pp. 1–14.

Chantangsi, C., H. J. Esson, and B. S. Leander (2008). "Morphology and Molecular Phylogeny of a Marine Interstitial Tetraflagellate with Putative Endosymbionts: *Auranticordis quadriverberis* n. Gen. et Sp. (Cercozoa)". In: *BMC Microbiology* 8.1 (1), pp. 1–16.

Ciniglia, C., E. C. Yang, et al. (2014). "Cyanidiophyceae in Iceland: Plastid *rbcL* Gene Elucidates Origin and Dispersal of Extremophilic *Galdieria sulphuraria* and *G. maxima* (Galdieriaceae, Rhodophyta)". In: *Phycologia* 53.6, pp. 542–551.

Ciniglia, C., H. S. Yoon, et al. (2004). "Hidden Biodiversity of the Extremophilic Cyanidiales Red Algae". In: *Molecular Ecology* 13.7, pp. 1827–1838.

Criscuolo, A. and S. Gribaldo (2010). "BMGE (Block Mapping and Gathering with Entropy): A New Software for Selection of Phylogenetic Informative Regions from Multiple Sequence Alignments". In: *BMC Evolutionary Biology* 10.1, p. 210.

Cronquist, A. (1960). "The Divisions and Classes of Plants". In: *The Botanical Review* 26.4, pp. 425–482.

Delaye, L., C. Valadez-Cano, and B. Pérez-Zamorano (2016). "How Really Ancient Is *Paulinella chromatophora*?" In: *PLoS Currents* 8.

Derelle, R. et al. (2016). "A Phylogenomic Framework to Study the Diversity and Evolution of Stramenopiles (=Heterokonts)". In: *Molecular Biology and Evolution* 33.11, pp. 2890–2898.

Deschamps, P. and D. Moreira (2012). "Reevaluating the Green Contribution to Diatom Genomes". In: *Genome Biology and Evolution* 4.7, pp. 683–688.

Dierckxsens, N., P. Mardulyn, and G. Smits (2017). "NOVOPlasty: De Novo Assembly of Organelle Genomes from Whole Genome Data". In: *Nucleic Acids Research* 45.4, e18–e18.

Doolittle, W. F. (1998). "You Are What You Eat: A Gene Transfer Ratchet Could Account for Bacterial Genes in Eukaryotic Nuclear Genomes". In: *Trends in Genetics* 14.8, pp. 307–311.

Dorrell, R. G. and C. Bowler (2017). "Chapter Three - Secondary Plastids of Stramenopiles". In: *Advances in Botanical Research*. Ed. by Y. Hirakawa. Vol. 84. Secondary Endosymbioses. Academic Press, pp. 57–103.

Dorrell, R. G., G. Gile, et al. (2017). "Chimeric Origins of Ochrophytes and Haptophytes Revealed through an Ancient Plastid Proteome". In: *eLife* 6. Ed. by D. Bhattacharya, e23717.

Dorrell, R. G. and C. J. Howe (2015). "Integration of Plastids with Their Hosts: Lessons Learned from Dinoflagellates". In: *Proceedings of the National Academy of Sciences* 112.33, pp. 10247–10254.

Dorrell, R. G. and A. G. Smith (2011). "Do Red and Green Make Brown?: Perspectives on Plastid Acquisitions within Chromalveolates". In: *Eukaryotic Cell* 10.7, pp. 856–868.

Edvardsen, B., E. S. Egge, and D. Vaulot (2016). "Diversity and Distribution of Haptophytes Revealed by Environmental Sequencing and Metabarcoding–a Review". In: *Perspect. Phycol* 3, pp. 77–91.

Eikrem, W. et al. (2017). "Haptophyta". In: *Handbook of the Protists*. Ed. by J. M. Archibald et al. Cham: Springer International Publishing, pp. 1–61.

Ekseth, O. K., M. Kuiper, and V. Mironov (2014). "*orthAgogue*: An Agile Tool for the Rapid Prediction of Orthology Relations". In: *Bioinformatics* 30.5, pp. 734–736.

Elias, M. and J. M. Archibald (2009). "Sizing up the Genomic Footprint of Endosymbiosis". In: *BioEssays* 31.12, pp. 1273–1279.

Eme, L. et al. (2014). "On the Age of Eukaryotes: Evaluating Evidence from Fossils and Molecular Clocks". In: *Cold Spring Harbor Perspectives in Biology* 6.8, a016139.

Evans, L. V. et al. (1974). "Studies on the Synthesis and Composition of Extracellular Mucilage in the Unicellular Red Alga *Rhodella*". In: *Journal of Cell Science* 16.1, pp. 1–21.

Falkowski, P. G. et al. (2004). "Why Is the Land Green and the Ocean Red?" In: *Coccolithophores: From Molecular Processes to Global Impact*. Ed. by H. R. Thierstein and J. R. Young. Berlin, Heidelberg: Springer, pp. 429–453.

Fast, N. M. et al. (2001). "Nuclear-Encoded, Plastid-Targeted Genes Suggest a Single Common Origin for Apicomplexan and Dinoflagellate Plastids". In: *Molecular Biology and Evolution* 18.3, pp. 418–426.

Figueroa-Martinez, F., C. Jackson, and A. Reyes-Prieto (2019). "Plastid Genomes from Diverse Glaucophyte Genera Reveal a Largely Conserved Gene Content and Limited Architectural Diversity". In: *Genome Biology and Evolution* 11.1, pp. 174–188.

Floener, L. and H. Bothe (1982). "Metabolic Activities in *Cyanophora paradoxa* and Its Cyanelles: II. Photosynthesis and Respiration". In: *Planta* 156.1, pp. 78–83.

Franco, A. D. et al. (2021). "Lower Statistical Support with Larger Datasets: Insights from the Ochrophyta Radiation". In: *bioRxiv*, p. 2021.01.14.426536.

Freshwater, D. W. et al. (1994). "A Gene Phylogeny of the Red Algae (Rhodophyta) Based on Plastid *rbcL*." In: *Proceedings of the National Academy of Sciences* 91.15, pp. 7281–7285.

Fujiwara, S. et al. (2001). "Molecular Phylogeny Of The Haptophyta Based On The *rbcL* Gene And Sequence Variation In The Spacer Region Of The Rubisco Operon". In: *Journal of Phycology* 37.1, pp. 121–129.

Füssy, Z. and M. Oborník (2017). "Chapter Six - Chromerids and Their Plastids". In: *Advances in Botanical Research*. Ed. by Y. Hirakawa. Vol. 84. Secondary Endosymbioses. Academic Press, pp. 187–218.

Gabrielson, P. W., D. J. Garbary, and R. F. Scagel (1985). "The Nature of the Ancestral Red Alga: Inferences from a Cladistic Analysis". In: *Biosystems* 18.3, pp. 335–346.

Gabrielson, P. W., D. Garbary, and M. H. Hommersand (1986). "Systematics of Red Algae (Rhodophyta)". In: *Critical Reviews in Plant Sciences* 3.4, pp. 325–366.

Gagat, P. et al. (2014). "Tertiary Plastid Endosymbioses in Dinoflagellates". In: *Endosymbiosis*. Springer, pp. 233–290.

Gawryluk, R. M. R. et al. (2019). "Non-Photosynthetic Predators Are Sister to Red Algae". In: *Nature* 572.7768 (7768), pp. 240–243.

Geider, R. J. et al. (2001). "Primary Productivity of Planet Earth: Biological Determinants and Physical Constraints in Terrestrial and Aquatic Habitats". In: *Global Change Biology* 7.8, pp. 849–882.

Girard, V. et al. (2020). "Thai Amber: Insights into Early Diatom History?" In: *BSGF - Earth Sciences Bulletin* 191, p. 23.

Glöckner, G., A. Rosenthal, and K. Valentin (2000). "The Structure and Gene Repertoire of an Ancient Red Algal Plastid Genome". In: *Journal of Molecular Evolution* 51.4, pp. 382–390.

Gontcharov, A. A. (2008). "Phylogeny and Classification of Zygnematophyceae (Streptophyta): Current State of Affairs". In: *Fottea* 8.2, pp. 87–104.

Gornik, S. G. et al. (2015). "Endosymbiosis Undone by Stepwise Elimination of the Plastid in a Parasitic Dinoflagellate". In: *Proceedings of the National Academy of Sciences* 112.18, pp. 5767–5772.

Gould, S., U.-G. Maier, and W. Martin (2015). "Protein Import and the Origin of Red Complex Plastids". In: *Current Biology* 25.12, R515–R521.

Gould, S. B., R. F. Waller, and G. I. McFadden (2008). "Plastid Evolution". In: *Annual Review of Plant Biology* 59.1, pp. 491–517.

Gradstein, F. M., ed. (2012). *The Geologic Time Scale 2012*. 1st ed. Amsterdam ; Boston: Elsevier. 2 pp.

Gran-Stadniczeñko, S. et al. (2017). "Haptophyte Diversity and Vertical Distribution Explored by 18S and 28S Ribosomal RNA Gene Metabarcoding and Scanning Electron Microscopy". In: *Journal of Eukaryotic Microbiology* 64.4, pp. 514–532.

Green, J. C. (1976). "Notes on the Flagellar Apparatus and Taxonomy of *Pavlova mesolychnon* Van Der Veer, and on the Status of *Pavlova* Butcher and Related Genera within the Haptophyceae". In: *Journal of the Marine Biological Association of the United Kingdom* 56.3, pp. 595–602.

Guillou, L. et al. (1999). "*Bolidomonas*: A New Genus with Two Species Belonging to a New Algal Class, the Bolidophyceae (Heterokonta)". In: *Journal of Phycology* 35.2, pp. 368–381.

Guiry, M. D. (2012). "How Many Species of Algae Are There?" In: *Journal of Phycology* 48.5, pp. 1057–1063.

Guiry, M. D. and G. Guiry (2022). *Algaebase*. AlgaeBase. World-wide electronic publication, National University of Ireland, Galway. URL: `https://www.algaebase.org`.

Gulbrandsen, Ø. S. et al. (2021). "Phylogenomic Analysis Restructures the Ulvophyceae". In: *Journal of Phycology* 57.4, pp. 1223–1233.

Hackett, J. D., L. Maranda, et al. (2003). "Phylogenetic Evidence for the Cryptophyte Origin of the Plastid of *Dinophysis* (Dinophysiales, Dinophyceae)". In: *Journal of Phycology* 39.2, pp. 440–448.

Hackett, J. D., H. S. Yoon, et al. (2007). "Phylogenomic Analysis Supports the Monophyly of Cryptophytes and Haptophytes and the Association of Rhizaria with Chromalveolates". In: *Molecular Biology and Evolution* 24.8, pp. 1702–1713.

Hagino, K. and J. R. Young (2015). "Biology and Paleontology of Coccolithophores (Haptophytes)". In: *Marine Protists: Diversity and Dynamics*. Ed. by S. Ohtsuka et al. Tokyo: Springer Japan, pp. 311–330.

Hall, J. D. and C. F. Delwiche (2007). "In the Shadow of Giants: Systematics of the Charophyte Green Algae". In: *Unravelling the algae: the past, present, and future of algal systematics.* The Systematics Association Special Volume Series 75. Ed. by J. Brodie and J. Lewis, pp. 155–169.

Hall, W. T. and G. Claus (1963). "Ultrastructural Studies On The Blue-Green Algal Symbiont In *Cyanophora paradoxa* Korschikoff". In: *Journal of Cell Biology* 19.3, pp. 551–563.

Hansen, G. I. et al. (2019). "*Viator vitreocola* Gen. et Sp. Nov. (Stylonematophyceae), a New Red Alga on Drift Glass Debris in Oregon and Washington, USA". In: *ALGAE* 34.2, pp. 71–90.

Hargraves, P. (2002). "The Ebridian Flagellates *Ebria* and *Hermesinum*". In: *Plankton Biology and Ecology* 49, pp. 9–16.

Harper, J. T. and P. J. Keeling (2003). "Nucleus-Encoded, Plastid-Targeted Glyceraldehyde-3-Phosphate Dehydrogenase (GAPDH) Indicates a Single Origin for Chromalveolate Plastids". In: *Molecular Biology and Evolution* 20.10, pp. 1730–1735.

Harwood, D. M. and V. A. Nikolaev (1995). "Cretaceous Diatoms: Morphology, Taxonomy, Biostratigraphy". In: *Short Courses in Paleontology* 8, pp. 81–106.

Hedges, S. B. et al. (2018). "Accurate Timetrees Require Accurate Calibrations". In: *Proceedings of the National Academy of Sciences* 115.41, E9510–E9511.

Hehenberger, E., R. J. Gast, and P. J. Keeling (2019). "A Kleptoplastidic Dinoflagellate and the Tipping Point between Transient and Fully Integrated Plastid Endosymbiosis". In: *Proceedings of the National Academy of Sciences* 116.36, pp. 17934–17942.

Hempel, F. et al. (2009). "ERAD-Derived Preprotein Transport across the Second Outermost Plastid Membrane of Diatoms". In: *Molecular Biology and Evolution* 26.8, pp. 1781–1790.

Herdman, M. and R. Stanier (1977). "The Cyanelle: Chloroplast or Endosymbiotic Prokaryote?" In: *FEMS Microbiology Letters* 1.1, pp. 7–11.

Hibberd, D. J. and G. F. Leedale (1971). "A New Algal Class — the Eustigmatophyceae". In: *TAXON* 20.4, pp. 523–525.

Hoang, D. T. et al. (2018). "UFBoot2: Improving the Ultrafast Bootstrap Approximation". In: *Molecular Biology and Evolution* 35.2, pp. 518–522.

Hoef-Emden, K. and M. Melkonian (2003). "Revision of the Genus *Cryptomonas* (Cryptophyceae): A Combination of Molecular Phylogeny and Morphology Provides Insights into a Long-Hidden Dimorphism". In: *Protist* 154.3, pp. 371–409.

Horn, S. et al. (2007). "*Synchroma grande* Spec. Nov. (Synchromophyceae Class. Nov., Heterokontophyta): An Amoeboid Marine Alga with Unique Plastid Complexes". In: *Protist* 158.3, pp. 277–293.

Huerta-Cepas, J. et al. (2019). "eggNOG 5.0: A Hierarchical, Functionally and Phylogenetically Annotated Orthology Resource Based on 5090 Organisms and 2502 Viruses". In: *Nucleic Acids Research* 47.D1, pp. D309–D314.

Irisarri, I., J. F. H. Strassert, and F. Burki (2022). "Phylogenomic Insights into the Origin of Primary Plastids". In: *Systematic Biology* 71.1, pp. 105–120.

Jackson, C., S. Clayden, and A. Reyes-Prieto (2015). "The Glaucophyta: The Blue-Green Plants in a Nutshell". In: *Acta Societatis Botanicorum Poloniae* 84.2.

Jackson, C., A. H. Knoll, et al. (2018). "Plastid Phylogenomics with Broad Taxon Sampling Further Elucidates the Distinct Evolutionary Origins and Timing of Secondary Green Plastids". In: *Scientific Reports* 8.1 (1), p. 1523.

Jacquier, A. and B. Dujon (1985). "An Intron-Encoded Protein Is Active in a Gene Conversion Process That Spreads an Intron into a Mitochondrial Gene". In: *Cell* 41.2, pp. 383–394.

Janouškovec, J., A. Horák, et al. (2010). "A Common Red Algal Origin of the Apicomplexan, Dinoflagellate, and Heterokont Plastids". In: *Proceedings of the National Academy of Sciences* 107.24, pp. 10949–10954.

Janouškovec, J., S.-L. Liu, et al. (2013). "Evolution of Red Algal Plastid Genomes: Ancient Architectures, Introns, Horizontal Gene Transfer, and Taxonomic Utility of Plastid Markers". In: *PLOS ONE* 8.3, e59001.

Janouškovec, J., D. V. Tikhonenkov, et al. (2015). "Factors Mediating Plastid Dependency and the Origins of Parasitism in Apicomplexans and Their Close Relatives". In: *Proceedings of the National Academy of Sciences* 112.33, pp. 10200–10207.

Jeffrey, S. W. and S. W. Wright (2005). "Photosynthetic Pigments in Marine Microalgae: Insights from Cultures and the Sea". In: *Algal Cultures, Analogues of Blooms and Applications*. Ed. by D. Subba Rao. 1. Enfield, USA: Science Publishers, pp. 33–90.

Jeffrey, S. W., S. W. Wright, and M. Zapata (2011). "Microalgal Classes and Their Signature Pigments". In: *Phytoplankton Pigments*. Ed. by S. Roy et al. 1st ed. Cambridge University Press, pp. 3–77.

Jin, J.-J. et al. (2020). "GetOrganelle: A Fast and Versatile Toolkit for Accurate de Novo Assembly of Organelle Genomes". In: *Genome Biology* 21.1, p. 241.

Johnson, P. W., P. E. Hargraves, and J. M. Sieburth (1988). "Ultrastructure and Ecology of *Calycomonas ovalis* Wulff, 1919, (Chrysophyceae) and Its Redescription as a Testate Rhizopod, *Paulinella ovalis* N. Comb. (Filosea: Euglyphina)". In: *The Journal of Protozoology* 35.4, pp. 618–626.

Jordan, R. W. (2009). "Coccolithophores". In: *Encyclopedia of Microbiology (Third Edition)*. Ed. by M. Schaechter. Oxford: Academic Press, pp. 593–605.

Kamikawa, R., G. Tanifuji, et al. (2015). "Plastid Genome-Based Phylogeny Pinpointed the Origin of the Green-Colored Plastid in the Dinoflagellate *Lepidodinium chlorophorum*". In: *Genome Biology and Evolution* 7.4, pp. 1133–1140.

Kamikawa, R., N. Yubuki, et al. (2015). "Multiple Losses of Photosynthesis in *Nitzschia* (Bacillariophyceae)". In: *Phycological Research* 63.1, pp. 19–28.

Katoh, K., K.-i. Kuma, et al. (2005). "Improvement in the Accuracy of Multiple Sequence Alignment Program MAFFT". In: *Genome informatics* 16.1, pp. 22–33.

Katoh, K. and D. M. Standley (2013). "MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability". In: *Molecular Biology and Evolution* 30.4, pp. 772–780.

Kawachi, M., I. Inouye, et al. (2002). "The Pinguiophyceae Classis Nova, a New Class of Photosynthetic Stramenopiles Whose Members Produce Large Amounts of Omega-3 Fatty Acids". In: *Phycological Research* 50.1, pp. 31–47.

Kawachi, M., T. Nakayama, et al. (2021). "Rappemonads Are Haptophyte Phytoplankton". In: *Current Biology*.

Kepner, W. A. (1905). "*Paulinella chromatophora*". In: *The Biological Bulletin* 9.2, pp. 128–129.

Kies, L. and B. P. Kremer (1986). "Typification Of The Glaucocystophyta". In: *TAXON* 35.1, pp. 128–133.

Kim, D., J. Lee, et al. (2022). "Group II Intron and Repeat-Rich Red Algal Mitochondrial Genomes Demonstrate the Dynamic Recent History of Autocatalytic RNAs". In: *BMC Biology* 20.1 (1), pp. 1–16.

Kim, E. and J. M. Archibald (2009). "Diversity and Evolution of Plastids and Their Genomes". In: *The Chloroplast*. Ed. by A. S. Sandelius and H. Aronsson. Vol. 13. Plant Cell Monographs. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 1–39.

Kim, E. and J. M. Archibald (2013). "Ultrastructure and Molecular Phylogeny of the Cryptomonad *Goniomonas avonlea* Sp. Nov." In: *Protist* 164.2, pp. 160–182.

Kim, E., J. W. Harrison, et al. (2011). "Newly Identified and Diverse Plastid-Bearing Branch on the Eukaryotic Tree of Life". In: *Proceedings of the National Academy of Sciences* 108.4, pp. 1496–1500.

Knoll, A. H. (2011). "The Multiple Origins of Complex Multicellularity". In: *Annual Review of Earth and Planetary Sciences* 39, pp. 217–239.

Knoll, A. H. (2014). "Paleobiological Perspectives on Early Eukaryotic Evolution". In: *Cold Spring Harbor Perspectives in Biology* 6.1, a016121.

Knutsen, S. H. et al. (1994). "A Modified System of Nomenclature for Red Algal Galactans". In: 37.2, pp. 163–170.

Kooistra, W. H. C. F., R. Gersonde, et al. (2007). "Chapter 11 - The Origin and Evolution of the Diatoms: Their Adaptation to a Planktonic Existence". In: *Evolution of Primary Producers in the Sea*. Ed. by P. G. Falkowski and A. H. Knoll. Burlington: Academic Press, pp. 207–249.

Kooistra, W. H. C. F. and L. K. Medlin (1996). "Evolution of the Diatoms (Bacillariophyta): IV. A Reconstruction of Their Age from Small Subunit rRNA Coding Regions and the Fossil Record". In: *Molecular Phylogenetics and Evolution* 6.3, pp. 391–407.

Koumandou, V. L. et al. (2004). "Dinoflagellate Chloroplasts – Where Have All the Genes Gone?" In: *Trends in Genetics*. Pagination Error in This Issue, See Publisher's Note in Vol. 21 Issue 1 p. 36 20.5, pp. 261–267.

Kowallik, K. V. (1994). "From Endosymbionts to Chloroplasts-Evidence for a Single Prokaryotic Eukaryotic Endocytobiosis". In: *Endocytobiosis and Cell Research* 10.1-2, pp. 137–149.

Kushibiki, A. et al. (2012). "New Unicellular Red Alga, *Bulboplastis apyrenoidosa* Gen. et Sp. Nov. (Rhodellophyceae, Rhodophyta) from the Mangroves of Japan: Phylogenetic and Ultrastructural Observations". In: *Phycological Research* 60.2, pp. 114–122.

Kuwata, A. et al. (2018). "Bolidophyceae, a Sister Picoplanktonic Group of Diatoms – A Review". In: *Frontiers in Marine Science* 5.

Lambowitz, A. M. and S. Zimmerly (2004). "Mobile Group II Introns". In: *Annual Review of Genetics* 38.1, pp. 1–35.

— (2011). "Group II Introns: Mobile Ribozymes That Invade DNA". In: *Cold Spring Harbor Perspectives in Biology* 3.8, a003616.

Lamouroux, J. V. (1813). *Essai Sur Les Genres de La Famille Des Thalassiophytes Non Articulées*. Dufour.

Larkum, A. W. D. (2020). "Light-Harvesting in Cyanobacteria and Eukaryotic Algae: An Overview". In: *Photosynthesis in Algae: Biochemical and Physiological Mechanisms*. Ed. by A. W. Larkum, A. R. Grossman, and J. A. Raven. Advances in Photosynthesis and Respiration. Cham: Springer International Publishing, pp. 207–260.

Larkum, A. W. D., P. J. Lockhart, and C. J. Howe (2007). "Shopping for Plastids". In: *Trends in Plant Science* 12.5, pp. 189–195.

Lartillot, N., T. Lepage, and S. Blanquart (2009). "PhyloBayes 3: A Bayesian Software Package for Phylogenetic Reconstruction and Molecular Dating". In: *Bioinformatics* 25.17, pp. 2286–2288.

Lartillot, N. and H. Philippe (2004). "A Bayesian Mixture Model for Across-Site Heterogeneities in the Amino-Acid Replacement Process". In: *Molecular Biology and Evolution* 21.6, pp. 1095–1109.

Lartillot, N., N. Rodrigue, et al. (2013). "PhyloBayes MPI: Phylogenetic Reconstruction with Infinite Mixtures of Profiles in a Parallel Environment". In: *Systematic Biology* 62.4, pp. 611–615.

Le, S. Q. and O. Gascuel (2008). "An Improved General Amino Acid Replacement Matrix". In: *Molecular Biology and Evolution* 25.7, pp. 1307–1320.

Lee, J. et al. (2016). "Parallel Evolution of Highly Conserved Plastid Genome Architecture in Red Seaweeds and Seed Plants". In: *BMC Biology* 14.1.

Lemieux, C., C. Otis, and M. Turmel (2007). "A Clade Uniting the Green Algae *Mesostigma viride* and *Chlorokybus atmophyticus* Represents the Deepest Branch of the Streptophyta in Chloroplast Genome-Based Phylogenies". In: *BMC Biology* 5.1 (1), pp. 1–17.

— (2016). "Comparative Chloroplast Genome Analyses of Streptophyte Green Algae Uncover Major Structural Alterations in the Klebsormidiophyceae, Coleochaetophyceae and Zygnematophyceae". In: *Frontiers in Plant Science* 7, p. 697.

Lemieux, C., M. Turmel, et al. (1985). "The Large Subunit of Ribulose-1,5-Bisphosphate Carboxylase-Oxygenase Is Encoded in the Inverted Repeat Sequence of the *Chlamydomonas eugametos* Chloroplast Genome". In: *Current Genetics* 9.2, pp. 139–145.

Lewis, L. A. (2007). "Chlorophyta on Land". In: *Algae and Cyanobacteria in Extreme Environments*. Ed. by J. Seckbach. Red. by J. Seckbach. Vol. 11. Cellular Origin, Life in Extreme Habitats and Astrobiology. Dordrecht: Springer Netherlands, pp. 569–582.

Li, L. et al. (2020). "The Genome of *Prasinoderma coloniale* Unveils the Existence of a Third Phylum within Green Plants". In: *Nature Ecology & Evolution* 4.9 (9), pp. 1220–1231.

Lichtenthaler, H. K. et al. (1997). "Biosynthesis of Isoprenoids in Higher Plant Chloroplasts Proceeds via a Mevalonate-Independent Pathway". In: *FEBS Letters* 400.3, pp. 271–274.

Liu, C., L. Shi, et al. (2012). "CpGAVAS, an Integrated Web Server for the Annotation, Visualization, Analysis, and GenBank Submission of Completely Sequenced Chloroplast Genome Sequences". In: *BMC Genomics* 13.1, p. 715.

Liu, H., I. Probert, et al. (2009). "Extreme Diversity in Noncalcifying Haptophytes Explains a Major Pigment Paradox in Open Oceans". In: *Proceedings of the National Academy of Sciences of the United States of America* 106.31, pp. 12803–12808.

Liu, S.-L., Y.-R. Chiang, et al. (2020). "Comparative Genome Analysis Reveals *Cyanidiococcus* Gen. Nov., A New Extremophilic Red Algal Genus Sister to *Cyanidioschyzon* (Cyanidioschyzonaceae, Rhodophyta)". In: *Journal of Phycology* 56.6, pp. 1428–1442.

Lucas, I. A. N. and M. Vesk (1990). "The Fine Structure Of Two Photosynthetic Species Of *Dinophysis* (Dinophysiales, Dinophyceae)". In: *Journal of Phycology* 26.2, pp. 345–357.

Mackiewicz, P., A. Bodył, and P. Gagat (2012). "Protein Import into the Photosynthetic Organelles of *Paulinella chromatophora* and Its Implications for Primary Plastid Endosymbiosis". In: *Symbiosis* 58.1, pp. 99–107.

Mackiewicz, P. and P. Gagat (2014). "Monophyly of Archaeplastida Supergroup and Relationships among Its Lineages in the Light of Phylogenetic and Phylogenomic Studies. Are We Close to a Consensus?" In: *Acta Societatis Botanicorum Poloniae* 83.4 (4), pp. 263–280.

Maistro, S. et al. (2009). "Phylogeny and Taxonomy of Xanthophyceae (Stramenopiles, Chromalveolata)". In: *Protist* 160.3, pp. 412–426.

Manning, S. R. and J. W. La Claire (2010). "Prymnesins: Toxic Metabolites of the Golden Alga, *Prymnesium parvum* Carter (Haptophyta)". In: *Marine Drugs* 8.3 (3), pp. 678–704.

Marin, B., E. C. M. Nowack, and M. Melkonian (2005). "A Plastid in the Making: Evidence for a Second Primary Endosymbiosis". In: *Protist* 156.4, pp. 425–432.

Marjanović, D. (2021). "The Making of Calibration Sausage Exemplified by Recalibrating the Transcriptomic Timetree of Jawed Vertebrates". In: *Frontiers in Genetics* 12, p. 535.

Martin, W. and K. V. Kowallik (1999). "Annotated English Translation of Mereschkowsky's 1905 Paper 'Über Natur Und Ursprung Der Chromatophoren Im Pflanzenreiche'". In: *European Journal of Phycology* 34.3, pp. 287–295.

Martynenko, N. A. et al. (2020). "A New Species of *Cryptomonas* (Cryptophyceae) from the Western Urals (Russia)". In: *European Journal of Taxonomy* 649 (649).

Matsuo, E. and Y. Inagaki (2018). "Patterns in Evolutionary Origins of Heme, Chlorophyll *a* and Isopentenyl Diphosphate Biosynthetic Pathways Suggest Non-Photosynthetic Periods Prior to Plastid Replacements in Dinoflagellates". In: *PeerJ* 6, e5345.

Matsuo, E., K. Morita, et al. (2022). "Comparative Plastid Genomics of Green-Colored Dinoflagellates Unveils Parallel Genome Compaction and RNA Editing". In: *Frontiers in Plant Science* 13.

Medlin, L. K. (2015). "A Timescale for Diatom Evolution Based on Four Molecular Markers: Reassessment of Ghost Lineages and Major Steps Defining Diatom Evolution." In: *Vie et milieu-life and enVironment*.

— (2016). "Opinion: Can Coalescent Models Explain Deep Divergences in the Diatoms and Argue for the Acceptance of Paraphyletic Taxa at All Taxonomic Hierarchies?" In: *Nova Hedwigia*, pp. 107–128.

Medlin, L. K. and I. Kaczmarska (2004). "Evolution of the Diatoms: V. Morphological and Cytological Support for the Major Clades and a Taxonomic Revision". In: *Phycologia* 43.3, pp. 245–270.

Medlin, L. K., D. M. Williams, and P. A. Sims (1993). "The Evolution of the Diatoms (Bacillariophyta). I. Origin of the Group and Assessment of the Monophyly of Its Major Divisions". In: *European Journal of Phycology* 28.4, pp. 261–275.

Meichtry Zaburlin, N. R. et al. (2019). "First Record of the Red Alga *Compsopogon caeruleus* (Balbis Ex C. Agardh) Montagne 1846 in the High Paraná River, Argentina-Paraguay". In:

Mereschkowsky, C. (1905). "Uber Natur Und Ursprung Der Chromatophoren Im Pflanzenreiche". In: *Biologisches Centralblatt* 25, pp. 593–604.

Merola, A. et al. (1981). "Revision of *Cyanidium caldarium*. Three Species of Acidophilic Algae". In: *Giornale botanico italiano* 115.4-5, pp. 189–195.

Minh, B. Q., M. A. T. Nguyen, and A. von Haeseler (2013). "Ultrafast Approximation for Phylogenetic Bootstrap". In: *Molecular Biology and Evolution* 30.5, pp. 1188–1195.

Minh, B. Q., H. A. Schmidt, et al. (2020). "IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era". In: *Molecular Biology and Evolution* 37.5, pp. 1530–1534.

Moestrup, Ø. (2001). "Algal Taxonomy: Historical Overview". In: *e LS*.

— (2021). "The Strange Pelagophyceae: Now Also Defined Ultrastructurally?" In: *Journal of Phycology* 57.2, pp. 393–395.

Monteiro, F. M. et al. (2016). "Why Marine Phytoplankton Calcify". In: *Science Advances* 2.7, e1501822.

Moore, R. B. et al. (2008). "A Photosynthetic Alveolate Closely Related to Apicomplexan Parasites". In: *Nature* 451.7181 (7181), pp. 959–963.

Morris, J. L. et al. (2018a). "Reply to Hedges et al.: Accurate Timetrees Do Indeed Require Accurate Calibrations". In: *Proceedings of the National Academy of Sciences* 115.41, E9512–E9513.

— (2018b). "The Timescale of Early Land Plant Evolution". In: *Proceedings of the National Academy of Sciences* 115.10, E2274–E2283.

Moustafa, A. et al. (2009). "Genomic Footprints of a Cryptic Plastid Endosymbiosis in Diatoms". In: *Science* 324.5935, pp. 1724–1726.

Müller, K. M. et al. (2001). "Ribosomal DNA Phylogeny of the Bangiophycidae (Rhodophyta) and the Origin of Secondary Plastids". In: *American Journal of Botany* 88.8, pp. 1390–1400.

Muñoz-Gómez, S. A., F. G. Mejía-Franco, et al. (2017). "The New Red Algal Subphylum Proteorhodophytina Comprises the Largest and Most Divergent Plastid Genomes Known". In: *Current Biology* 27.11, 1677–1684.e4.

Muñoz-Gómez, S. A. and C. H. Slamovits (2018). "Chapter Three - Plastid Genomes in the Myzozoa". In: *Advances in Botanical Research*. Ed. by S.-M. Chaw and R. K. Jansen. Vol. 85. Plastid Genome Evolution. Academic Press, pp. 55–94.

Nakayama, T. et al. (2014). "Complete Genome of a Nonphotosynthetic Cyanobacterium in a Diatom Reveals Recent Adaptations to an Intracellular Lifestyle". In: *Proceedings of the National Academy of Sciences* 111.31, pp. 11407–11412.

Nan, F. et al. (2017). "Origin and Evolutionary History of Freshwater Rhodophyta: Further Insights Based on Phylogenomic Evidence". In: *Scientific Reports* 7.

Nelson, W. A., T. J. Farr, and J. E. S. Broom (2005). "*Dione* and *Minerva*, Two New Genera from New Zealand Circumscribed for Basal Taxa in the Bangiales (Rhodophyta)". In: *Phycologia* 44.2, pp. 139–145.

Nguyen, L.-T. et al. (2015). "IQ-TREE: A Fast and Effective Stochastic Algorithm for Estimating Maximum-Likelihood Phylogenies". In: *Molecular Biology and Evolution* 32.1, pp. 268–274.

Nicholls, K. H. and D. E. Wujek (2015). "Chapter 12 - Chrysophyceae and Phaeothamniophyceae". In: *Freshwater Algae of North America (Second Edition)*. Ed. by J. D. Wehr, R. G. Sheath, and J. P. Kociolek. Aquatic Ecology. Boston: Academic Press, pp. 537–586.

Novarino, G. and I. A. N. Lucas (1993). "Some Proposals for a New Classification System of the Cryptophyceae". In: *Botanical Journal of the Linnean Society* 111.1, pp. 3–21.

Nowack, E. C. M., M. Melkonian, and G. Glöckner (2008). "Chromatophore Genome Sequence of *Paulinella* Sheds Light on Acquisition of Photosynthesis by Eukaryotes". In: *Current Biology* 18.6, pp. 410–418.

Nurk, S. et al. (2017). "metaSPAdes: A New Versatile Metagenomic Assembler". In: *Genome Research* 27.5, pp. 824–834.

Oborník, M. et al. (2012). "Morphology, Ultrastructure and Life Cycle of *Vitrella brassicaformis* n. Sp., n. Gen., a Novel Chromerid from the Great Barrier Reef". In: *Protist* 163.2, pp. 306–323.

Okamoto, N., C. Chantangsi, et al. (2009). "Molecular Phylogeny and Description of the Novel Katablepharid *Roombia truncata* Gen. et Sp. Nov., and Establishment of the Hacrobia Taxon Nov". In: *PLoS ONE* 4.9, e7080.

Okamoto, N. and I. Inouye (2005). "The Katablepharids Are a Distant Sister Group of the Cryptophyta: A Proposal for Katablepharidophyta Divisio Nova/Kathablepharida Phylum Novum Based on SSU rDNA and Beta-Tubulin Phylogeny". In: *Protist* 156.2, pp. 163–179.

Palmer, J. D. (2003). "The Symbiotic Birth and Spread of Plastids: How Many Times and Whodunit?" In: *Journal of Phycology* 39.1, pp. 4–12.

Palmer, J. D. and C. F. Delwiche (1998). "The Origin and Evolution of Plastids and Their Genomes". In: *Molecular Systematics of Plants II: DNA Sequencing*. Ed. by D. E. Soltis, P. S. Soltis, and J. J. Doyle. Boston, MA: Springer US, pp. 375–409.

Palmer, J. D., D. E. Soltis, and M. W. Chase (2004). "The Plant Tree of Life: An Overview and Some Points of View". In: *American Journal of Botany* 91.10, pp. 1437–1445.

Parfrey, L. W. et al. (2011). "Estimating the Timing of Early Eukaryotic Diversification with Multigene Molecular Clocks". In: *Proceedings of the National Academy of Sciences* 108.33, pp. 13624–13629.

Parham, J. F. et al. (2012). "Best Practices for Justifying Fossil Calibrations". In: *Systematic Biology* 61.2, pp. 346–359.

Park, M. G. et al. (2006). "First Successful Culture of the Marine Dinoflagellate *Dinophysis acuminata*". In: *Aquatic Microbial Ecology* 45.2, pp. 101–106.

Patil, V. et al. (2009). "Revisiting the Phylogenetic Position of Synchroma Grande". In: *Journal of Eukaryotic Microbiology* 56.4, pp. 394–396.

Patron, N. J., Y. Inagaki, and P. J. Keeling (2007). "Multiple Gene Phylogenies Support the Monophyly of Cryptomonad and Haptophyte Host Lineages". In: *Current Biology* 17.10, pp. 887–891.

Patron, N. J., M. B. Rogers, and P. J. Keeling (2004). "Gene Replacement of Fructose-1,6-Bisphosphate Aldolase Supports the Hypothesis of a Single Photosynthetic Ancestor of Chromalveolates". In: *Eukaryotic Cell* 3.5, pp. 1169–1175.

Paudel, Y. P. et al. (2021). "Chloroplast Genome Analysis of *Chrysotila dentata*". In: *Gene* 804, p. 145871.

Perrineau, M.-M. et al. (2015). "Recent Mobility of Plastid Encoded Group II Introns and Twintrons in Five Strains of the Unicellular Red Alga *Porphyridium*". In: *PeerJ* 3, e1017.

Pombert, J.-F. et al. (2005). "The Chloroplast Genome Sequence of the Green Alga *Pseudendoclonium akinetum* (Ulvophyceae) Reveals Unusual Structural Features and New Insights into the Branching Order of Chlorophyte Lineages". In: *Molecular Biology and Evolution* 22.9, pp. 1903–1918.

Ponce-Toledo, R. I. (2018). "Origins and Early Evolution of Photosynthetic Eukaryotes". PhD thesis. Université Paris-Saclay.

Ponce-Toledo, R. I., P. Deschamps, et al. (2017). "An Early-Branching Freshwater Cyanobacterium at the Origin of Plastids". In: *Current Biology* 27.3, pp. 386–391.

Ponce-Toledo, R. I., D. Moreira, et al. (2018). "Secondary Plastids of Euglenids and Chlorarachniophytes Function with a Mix of Genes of Red and Green Algal Ancestry". In: *Molecular Biology and Evolution* 35.9, pp. 2198–2204.

Ponce-Toledo, R. I., P. López-García, and D. Moreira (2019). "Horizontal and Endosymbiotic Gene Transfer in Early Plastid Evolution". In: *New Phytologist* 0 (ja).

Preuss, M. et al. (2021). "Divergence Times and Plastid Phylogenomics within the Intron-Rich Order Erythropeltales (Compsopogonophyceae, Rhodophyta)". In: *Journal of Phycology* 57.3, pp. 1035–1044.

Price, D. C., C. X. Chan, et al. (2012). "*Cyanophora paradoxa* Genome Elucidates Origin of Photosynthesis in Algae and Plants". In: *Science* 335.6070, pp. 843–847.

Price, D. C., U. W. Goodenough, et al. (2019). "Analysis of an Improved *Cyanophora paradoxa* Genome Assembly". In: *DNA Research* 26.4, pp. 287–299.

Price, D. C., J. M. Steiner, et al. (2017). "Glaucophyta". In: *Handbook of the Protists*. Ed. by J. M. Archibald, A. G. Simpson, and C. H. Slamovits. Cham: Springer International Publishing, pp. 23–87.

Prjibelski, A. et al. (2020). "Using SPAdes De Novo Assembler". In: *Current Protocols in Bioinformatics* 70.1, e102.

Pueschel, C. M. and K. M. Cole (1982). "Rhodophycean Pit Plugs: An Ultrastructural Survey with Taxonomic Implications". In: *American journal of botany* 69.5, pp. 703–720.

Qiu, N. W. et al. (2019). "Advances in the Members and Biosynthesis of Chlorophyll Family". In: *Photosynthetica* 57.4, pp. 974–984.

Ragan, M. A. et al. (1994). "A Molecular Phylogeny of the Marine Red Algae (Rhodophyta) Based on the Nuclear Small-Subunit rRNA Gene." In: *Proceedings of the National Academy of Sciences* 91.15, pp. 7276–7280.

Rappé, M. S. et al. (1998). "Phylogenetic Diversity of Ultraplankton Plastid Small-Subunit rRNA Genes Recovered in Environmental Nucleic Acid Samples from the Pacific and Atlantic Coasts of the United States". In: *Applied and Environmental Microbiology* 64.1, pp. 294–303.

Rasmussen, B. et al. (2008). "Reassessing the First Appearance of Eukaryotes and Cyanobacteria". In: *Nature* 455.7216 (7216), pp. 1101–1104.

Reeb, V. and D. Bhattacharya (2010). "The Thermo-Acidophilic Cyanidiophyceae (Cyanidiales)". In: *Red Algae in the Genomic Age*. Ed. by J. Seckbach and D. J. Chapman. Cellular Origin, Life in Extreme Habitats and Astrobiology. Dordrecht: Springer Netherlands, pp. 409–426.

Reyes-Prieto, A., A. Moustafa, and D. Bhattacharya (2008). "Multiple Genes of Apparent Algal Origin Suggest Ciliates May Once Have Been Photosynthetic". In: *Current Biology* 18.13, pp. 956–962.

Reyes-Prieto, A., S. Russell, et al. (2018). "Chapter Four - Comparative Plastid Genomics of Glaucophytes". In: *Advances in Botanical Research*. Ed. by S.-M. Chaw and R. K. Jansen. Vol. 85. Plastid Genome Evolution. Academic Press, pp. 95–127.

Rice, D. W. and J. D. Palmer (2006). "An Exceptional Horizontal Gene Transfer in Plastids: Gene Replacement by a Distant Bacterial Paralog and Evidence That Haptophyte and Cryptophyte Plastids Are Sisters". In: *BMC Biology* 4.1 (1), pp. 1–15.

Roger, A. J. and L. A. Hug (2006). "The Origin and Diversification of Eukaryotes: Problems with Molecular Phylogenetics and Molecular Clock Estimation". In: *Philosophical Transactions of the Royal Society B: Biological Sciences* 361.1470, pp. 1039–1054.

Sagan, L. (1967). "On the Origin of Mitosing Cells". In: *Journal of Theoretical Biology* 14.3, 225–IN6.

Saldarriaga, J. F. et al. (2003). "Multiple Protein Phylogenies Show That *Oxyrrhis marina* and *Perkinsus marinus* Are Early Branches of the Dinoflagellate Lineage". In: *International Journal of Systematic and Evolutionary Microbiology* 53.1, pp. 355–365.

Sánchez-Baracaldo, P. et al. (2017). "Early Photosynthetic Eukaryotes Inhabited Low-Salinity Habitats". In: *Proceedings of the National Academy of Sciences of the United States of America* 114.37, E7737–E7745.

Sanchez-Puerta, M. V. and C. F. Delwiche (2008). "A Hypothesis For Plastid Evolution In Chromalveolates". In: *Journal of Phycology* 44.5, pp. 1097–1107.

Sarai, C. et al. (2020). "Dinoflagellates with Relic Endosymbiont Nuclei as Models for Elucidating Organellogenesis". In: *Proceedings of the National Academy of Sciences* 117.10, pp. 5364–5375.

Saunders, G. W. and M. H. Hommersand (2004). "Assessing Red Algal Supraordinal Diversity and Taxonomy in the Context of Contemporary Systematic Data". In: *American Journal of Botany* 91.10, pp. 1494–1507.

Sauquet, H. (2013). "A Practical Guide to Molecular Dating". In: *Comptes Rendus Palevol*. Systematics beyond Phylogenetics / La Systématique Au-Delà de La Phylogénétique 12.6, pp. 355–367.

Schenk, H. E. A. (1994). "*Cyanophora paradoxa*: Anagenetic Model Or Missing Link Of Plastid Evolution?" In: *Cell Res.* P. 20.

Schirrmeister, B. E., P. Sanchez-Baracaldo, and D. Wacey (2016). "Cyanobacterial Evolution during the Precambrian". In: *International Journal of Astrobiology* 15.3, pp. 187–204.

Schmidt, M. et al. (2012). "*Synchroma pusillum* Sp. Nov. and Other New Algal Isolates with Chloroplast Complexes Confirm the Synchromophyceae (Ochrophyta) as a Widely Distributed Group of Amoeboid Algae". In: *Protist* 163.4, pp. 544–559.

Schön, M. E. et al. (2021). "Single Cell Genomics Reveals Plastid-Lacking Picozoa Are Close Relatives of Red Algae". In: *Nature Communications* 12.1 (1), p. 6651.

Schvarcz, C. R. et al. (2022). "Overlooked and Widespread Pennate Diatom-Diazotroph Symbioses in the Sea". In: *Nature Communications* 13.1 (1), p. 799.

Scornavacca, C., F. Delsuc, and N. Galtier (2020). *Phylogenetics in the Genomic Era*. Ed. by C. Scornavacca, F. Delsuc, and N. Galtier. p.p. 1-568.

Scott, J. et al. (2011). "On the Genus *Rhodella*, the Emended Orders Dixoniellales and Rhodellales with a New Order Glaucosphaerales (Rhodellophyceae, Rhodophyta)". In: *Algae* 26.4, pp. 277–288.

Ševčíková, T. et al. (2015). "Updating Algal Evolutionary Relationships through Plastid Genome Sequencing: Did Alveolate Plastids Emerge through Endosymbiosis of an Ochrophyte?" In: *Scientific Reports* 5.1, pp. 1–12.

Shalchian-Tabrizi, K., W. Eikrem, et al. (2006). "Telonemia, a New Protist Phylum with Affinity to Chromist Lineages". In: *Proceedings of the Royal Society B: Biological Sciences* 273.1595, pp. 1833–1842.

Shalchian-Tabrizi, K., K. Reier-Røberg, et al. (2011). "Marine–Freshwater Colonizations of Haptophytes Inferred from Phylogeny of Environmental 18S rDNA Sequences". In: *Journal of Eukaryotic Microbiology* 58.4, pp. 315–318.

Shalchian-Tabrizi, K., M. Skånseng, et al. (2006). "Heterotachy Processes in Rhodophyte-Derived Secondhand Plastid Genes: Implications for Addressing the Origin and Evolution of Dinoflagellate Plastids". In: *Molecular Biology and Evolution* 23.8, pp. 1504–1515.

Shiflett, A. M. and P. J. Johnson (2010). "Mitochondrion-Related Organelles in Eukaryotic Protists". In: *Annual Review of Microbiology* 64, pp. 409–429.

Sibbald, S. J. and J. M. Archibald (2020). "Genomic Insights into Plastid Evolution". In: *Genome Biology and Evolution*.

Simon, M. et al. (2013). "New Haptophyte Lineages and Multiple Independent Colonizations of Freshwater Ecosystems". In: *Environmental Microbiology Reports* 5.2, pp. 322–332.

Sims, P. A., D. G. Mann, and L. K. Medlin (2006). "Evolution of the Diatoms: Insights from Fossil, Biological and Molecular Data". In: *Phycologia* 45.4, pp. 361–402.

Slater, G. S. C. and E. Birney (2005). "Automated Generation of Heuristics for Biological Sequence Comparison". In: *BMC Bioinformatics* 6.1, p. 31.

Smith, C. (2001). "The Challenge of Siphonous Green Algae". In: *American Scientist*.

Smith, D. R. (2018). "Chapter Two - Lost in the Light: Plastid Genome Evolution in Nonphotosynthetic Algae". In: *Advances in Botanical Research*. Ed. by S.-M. Chaw and R. K. Jansen. Vol. 85. Plastid Genome Evolution. Academic Press, pp. 29–53.

Smith, D. R. and R. W. Lee (2009). "The Mitochondrial and Plastid Genomes of Volvox Carteri: Bloated Molecules Rich in Repetitive DNA". In: *BMC Genomics* 10.1, p. 132.

Stiller, J. W. and B. D. Hall (1997). "The Origin of Red Algae: Implications for Plastid Evolution". In: *Proceedings of the National Academy of Sciences* 94.9, pp. 4520–4525.

Stiller, J. W., J. Huang, et al. (2009). "Are Algal Genes in Nonphotosynthetic Protists Evidence of Historical Plastid Endosymbioses?" In: *BMC Genomics* 10.1, p. 484.

Stiller, J. W., J. Schreiber, et al. (2014). "The Evolution of Photosynthesis in Chromist Algae through Serial Endosymbioses". In: *Nature Communications* 5, p. 5764.

Stoebe, B. and K. V. Kowallik (1999). "Gene-Cluster Analysis in Chloroplast Genomics". In: *Trends in Genetics* 15.9, pp. 344–347.

Strassert, J. F. H., M. Jamy, et al. (2019). "New Phylogenomic Analysis of the Enigmatic Phylum Telonemia Further Resolves the Eukaryote Tree of Life". In: *Molecular Biology and Evolution* 36.4, pp. 757–765.

Strassert, J. F. H., I. Irisarri, et al. (2021). "A Molecular Timescale for Eukaryote Evolution with Implications for the Origin of Red Algal-Derived Plastids". In: *Nature Communications* 12.1 (1), p. 1879.

Su, D. et al. (2021). "Large-Scale Phylogenomic Analyses Reveal the Monophyly of Bryophytes and Neoproterozoic Origin of Land Plants". In: *Molecular Biology and Evolution* 38.8, pp. 3332–3344.

Suzuki, S. et al. (2022). "What Happened before Losses of Photosynthesis in Cryptophyte Algae?" In: *Molecular Biology and Evolution* 39.2, msac001.

Takishita, K. et al. (2009). "A Hypothesis for the Evolution of Nuclear-Encoded, Plastid-Targeted Glyceraldehyde-3-Phosphate Dehydrogenase Genes in "Chromalveolate" Members". In: *PLOS ONE* 4.3, e4737.

Tang, Q. et al. (2020). "A One-Billion-Year-Old Multicellular Chlorophyte". In: *Nature Ecology & Evolution* 4.4 (4), pp. 543–549.

Tanifuji, G. and N. T. Onodera (2017). "Chapter Eight - Cryptomonads: A Model Organism Sheds Light on the Evolutionary History of Genome Reorganization in Secondary Endosymbioses". In: *Advances in Botanical Research*. Ed. by Y. Hirakawa. Vol. 84. Secondary Endosymbioses. Academic Press, pp. 263–320.

Tappan, H. (1976). "Possible Eucaryotic Algae (Bangiophycidae) among Early Proterozoic Microfossils". In: *GSA Bulletin* 87.4, pp. 633–639.

Taylor, D. L. (1970). "Chloroplasts as Symbiotic Organelles". In: *International Review of Cytology*. Ed. by G. H. Bourne, J. F. Danlelli, and K. W. Jeon. Vol. 27. Academic Press, pp. 29–64.

Teich, R. et al. (2007). "Origin and Distribution of Calvin Cycle Fructose and Sedoheptulose Bisphosphatases in Plantae and Complex Algae: A Single Secondary Origin of Complex Red Plastids and Subsequent Propagation via Tertiary Endosymbioses". In: *Protist* 158.3, pp. 263–276.

The UniProt Consortium (2015). "UniProt: A Hub for Protein Information". In: *Nucleic Acids Research* 43.D1, pp. D204–D212.

Theissen, U. and W. Martin (2006). "The Difference between Organelles and Endosymbionts". In: *Current Biology* 16.24, R1016–R1017.

Thornton, D. C. O. (2012). "Primary Production in the Ocean". In: *Advances in Photosynthesis–Fundamental Aspects. Intech, Rijeca, Croatia*, pp. 563–588.

Tillich, M. et al. (2017). "GeSeq – Versatile and Accurate Annotation of Organelle Genomes". In: *Nucleic Acids Research* 45.W1, W6–W11.

Tomas, R. N. and E. R. Cox (1973). "Observations On The Symbiosis Of *Peridinium balticum* And Its Intracellular Alga. I. Ultrastructure". In: *Journal of Phycology* 9.3, pp. 304–323.

Tsuji, Y. and M. Yoshida (2017). "Chapter Seven - Biology of Haptophytes: Complicated Cellular Processes Driving the Global Carbon Cycle". In: *Advances in Botanical Research*. Ed. by Y. Hirakawa. Vol. 84. Secondary Endosymbioses. Academic Press, pp. 219–261.

Turmel, M., C. Otis, and C. Lemieux (2016). "Mitochondrion-to-Chloroplast DNA Transfers and Intragenomic Proliferation of Chloroplast Group II Introns in *Gloeotilopsis* Green Algae (Ulotrichales, Ulvophyceae)". In: *Genome Biology and Evolution* 8.9, pp. 2789–2805.

Tyler, B. M. et al. (2006). "*Phytophthora* Genome Sequences Uncover Evolutionary Origins and Mechanisms of Pathogenesis". In: *Science (New York, N.Y.)* 313.5791, pp. 1261–1266.

Van Beveren, F. et al. (2022). "Independent Size Expansions and Intron Proliferation in Red Algal Plastid and Mitochondrial Genomes". In: *Genome Biology and Evolution* 14.4, evac037.

Van Dongen, S. M. (2000). "Graph Clustering by Flow Simulation". PhD thesis.

Van Dooren, G. G. and S. V. Hapuarachchi (2017). "Chapter Five - The Dark Side of the Chloroplast: Biogenesis, Metabolism and Membrane Biology of the Apicoplast". In: *Advances in Botanical Research*. Ed. by Y. Hirakawa. Vol. 84. Secondary Endosymbioses. Academic Press, pp. 145–185.

Verbruggen, H. et al. (2009). "A Multi-Locus Time-Calibrated Phylogeny of the Siphonous Green Algae". In: *Molecular Phylogenetics and Evolution* 50.3, pp. 642–653.

Volkman, J. K. et al. (1980). "Long-Chain Alkenes and Alkenones in the Marine Coccolithophorid *Emiliania huxleyi*". In: *Phytochemistry* 19.12, pp. 2619–2622.

Waller, R. F., S. G. Gornik, et al. (2016). "Metabolic Pathway Redundancy within the Apicomplexan-Dinoflagellate Radiation Argues against an Ancient Chromalveolate Plastid". In: *Communicative & Integrative Biology* 9.1, e1116653.

Waller, R. F. and L. Kořený (2017). "Chapter Four - Plastid Complexity in Dinoflagellates: A Picture of Gains, Losses, Replacements and Revisions". In: *Advances in Botanical Research*. Ed. by Y. Hirakawa. Vol. 84. Secondary Endosymbioses. Academic Press, pp. 105–143.

Watanabe, M. M., S. Suda, et al. (1990). "*Lepidodinium viride* Gen. Et Sp. Nov. (Gymnodinaiales, Dinophyta), a Green Dinoflagellate with a Chlorophyll *a*- and *b*-Containing Endosymbiont". In: *Journal of Phycology* 26.4, pp. 741–751.

Watanabe, M. M., Y. Takeda, et al. (1987). "A Green Dinoflagellate With Chlorophylls *a* And *b*: Morphology, Fine Structure Of The Chloroplast And Chlorophyll Composition". In: *Journal of Phycology* 23.s2, pp. 382–389.

Wernegreen, J. J. (2004). "Endosymbiosis: Lessons in Conflict Resolution". In: *PLOS Biology* 2.3, e68.

Wetherbee, R., T. T. Bringloe, et al. (2021). "New Pelagophytes Show a Novel Mode of Algal Colony Development and Reveal a Perforated Theca That May Define the Class". In: *Journal of Phycology* 57.2, pp. 396–411.

Wetherbee, R., C. J. Jackson, et al. (2019). "The Golden Paradox – a New Heterokont Lineage with Chloroplasts Surrounded by Two Membranes". In: *Journal of Phycology*.

Whatley, J. M. and F. R. Whatley (1981). "Chloroplast Evolution". In: *The New Phytologist* 87.2, pp. 233–247.

Whelan, S., I. Irisarri, and F. Burki (2018). "PREQUAL: Detecting Non-Homologous Characters in Sets of Unaligned Homologous Sequences". In: *Bioinformatics* 34.22, pp. 3929–3930.

Wick, R. R., L. M. Judd, et al. (2017). "Unicycler: Resolving Bacterial Genome Assemblies from Short and Long Sequencing Reads". In: *PLOS Computational Biology* 13.6, e1005595.

Wick, R. R., M. B. Schultz, et al. (2015). "Bandage: Interactive Visualization of de Novo Genome Assemblies". In: *Bioinformatics* 31.20, pp. 3350–3352.

Wickett, N. J. et al. (2014). "Phylotranscriptomic Analysis of the Origin and Early Diversification of Land Plants". In: *Proceedings of the National Academy of Sciences* 111.45, E4859–E4868.

Wikfors, G. H. and M. Ohno (2001). "Impact of Algal Research in Aquaculture". In: *Journal of Phycology* 37.6, pp. 968–974.

Wyman, S. K., R. K. Jansen, and J. L. Boore (2004). "Automatic Annotation of Organellar Genomes with DOGMA". In: *Bioinformatics* 20.17, pp. 3252–3255.

Xiao, S. et al. (2004). "Phosphatized Multicellular Algae in the Neoproterozoic Doushantuo Formation, China, and the Early Evolution of Florideophyte Red Algae". In: *American Journal of Botany* 91.2, pp. 214–227.

Xu, M. et al. (2016). "Synthesis of Chlorophyll-c Derivatives by Modifying Natural Chlorophyll-a". In: *Photosynthesis Research* 127.3, pp. 335–345.

Yabuki, A. et al. (2014). "*Palpitomonas bilix* Represents a Basal Cryptist Lineage: Insight into the Character Evolution in Cryptista". In: *Scientific Reports* 4, p. 4641.

Yamada, K., S. Sato, et al. (2020). "New Clade of Silicified Bolidophytes That Belong to *Triparma* (Bolidophyceae, Stramenopiles)". In: *Phycological Research* 68.2, pp. 178–182.

Yamada, N., J. J. Bolton, et al. (2019). "Discovery of a Kleptoplastic 'Dinotom' Dinoflagellate and the Unique Nuclear Dynamics of Converting Kleptoplastids to Permanent Plastids". In: *Scientific Reports* 9.1 (1), p. 10474.

Yamaguchi, H. et al. (2010). "Phylogeny and Taxonomy of the Raphidophyceae (Heterokontophyta) and *Chlorinimonas sublosa* Gen. et Sp. Nov., a New Marine Sand-Dwelling Raphidophyte". In: *Journal of Plant Research* 123.3, pp. 333–342.

Yang, E. C., G. H. Boo, et al. (2012). "Supermatrix Data Highlight the Phylogenetic Relationships of Photosynthetic Stramenopiles". In: *Protist* 163.2, pp. 217–231.

Yang, E. C., J. Scott, et al. (2010). "New Taxa of the Porphyridiophyceae (Rhodophyta): *Timspurckia Oligopyrenoides* Gen. et Sp. Nov. and *Erythrolobus madagascarensis* Sp. Nov". In: *Phycologia* 49.6, pp. 604–616.

Yazaki, E. et al. (2022). "The Closest Lineage of Archaeplastida Is Revealed by Phylogenomics Analyses That Include *Microheliella maris*". In: *Open Biology* 12.4, p. 210376.

Yoon, H. S., R. A. Andersen, et al. (2009). "Stramenopiles". In: *Encyclopedia of Microbiology (Third Edition)*. Ed. by M. Schaechter. Oxford: Academic Press, pp. 721–731.

Yoon, H. S., J. D. Hackett, G. Pinto, et al. (2002). "The Single, Ancient Origin of Chromist Plastids". In: *Journal of Phycology* 38.s1, pp. 40–40.

Yoon, H. S., J. D. Hackett, C. Ciniglia, et al. (2004). "A Molecular Timeline for the Origin of Photosynthetic Eukaryotes". In: *Molecular Biology and Evolution* 21.5, pp. 809–818.

Yoon, H. S., K. M. Muller, et al. (2006). "Defining the Major Lineages of Red Algae (Rhodophyta)". In: *Journal of Phycology* 42.2, pp. 482–492.

Yoon, H. S., G. C. Zuccarello, and D. Bhattacharya (2010). "Evolutionary History and Taxonomy of Red Algae". In: *Red Algae in the Genomic Age*. Ed. by J. Seckbach and D. J. Chapman. Vol. 13. Dordrecht: Springer Netherlands, pp. 25–42.

Young, J. R., H. Andruleit, and I. Probert (2009). "Coccolith Function and Morphogenesis: Insights from Appendage-Bearing Coccolithophores of the Family Syracosphaeraceae (Haptophyta)1". In: *Journal of Phycology* 45.1, pp. 213–226.

Yu, M. et al. (2018). "Chapter Five - Evolution of the Plastid Genomes in Diatoms". In: *Advances in Botanical Research*. Ed. by S.-M. Chaw and R. K. Jansen. Vol. 85. Plastid Genome Evolution. Academic Press, pp. 129–155.

Zapata, M. et al. (2004). "Photosynthetic Pigments in 37 Species (65 Strains) of Haptophyta: Implications for Oceanography and Chemotaxonomy". In: *Marine Ecology Progress Series* 270, pp. 83–102.

Zhang, X. et al. (2019). "The Mitochondrial and Chloroplast Genomes of the Green Alga *Haematococcus* Are Made up of Nearly Identical Repetitive Sequences". In: *Current Biology* 29.15, R736–R737.

Zuccarello, G. et al. (2000). "Molecular Phylogeny of *Rhodochaete Parvula* (Bangiophycidae, Rhodophyta)". In: *Phycologia* 39.1, pp. 75–81.

Supplementary material

Rights to licensed images

Figure	Copyright holder	License/Rights
1.1	Ponce-Toledo et al. 2019	Used with authorization from original authors
1.2B	Li et al. 2020	CC BY 4.0
1.2D	Oliver S.	CC BY-SA 3.0 DE
1.2E	Jon Houseman	CC BY-SA 4.0
1.3B	Gabriele Kothe-Heinrich	CC BY-SA 3.0
1.3C	Gabriele Kothe-Heinrich	CC BY-SA 3.0
1.3D	Hansen et al. 2019	CC BY-NC 3.0
1.3E	Meichtry Zaburlin et al. 2019	CC BY 4.0
1.3F	Ciniglia, Yang, et al. 2014	Permission received (from the journal)
1.3G	Muñoz-Gómez, Mejía-Franco, et al. 2017	Permission received (through Copyright Clearance Center)
1.3H	Muñoz-Gómez, Mejía-Franco, et al. 2017	Permission received (through Copyright Clearance Center)
1.4C	Wetherbee, Bringloe, et al. 2021	Permission received (through Copyright Clearance Center)
1.4D	Kristian Peters	CC BY-SA 3.0
1.4E	Kuwata et al. 2018	CC BY-SA 3.0
1.4F	Nocila Angeli / MUSE Science Museum	CC BY-SA 3.0
1.4G	Bradbury 2004	CC BY 4.0
1.5A-N	Monteiro et al. 2016	CC BY 4.0
1.5O	Kawachi, Nakayama, et al. 2021	CC BY-ND-ND 4.0
1.5P	Manning and La Claire 2010	CC BY 3.0
1.6B	Kim and Archibald 2013	Permission received (through Copyright Clearance Center)
1.6C	Martynenko et al. 2020	CC BY 4.0
1.7B	Sonja I. Rueckert	CC BY 3.0
1.7C	Oborník et al. 2012	Permission received (through Copyright Clearance Center)
1.7D-G	Waller and Kořený 2017	Permission received (through Copyright Clearance Center)

Size expansion in red algal organellar genomes

The supplementary material can be found in the following pages. Supplementary data are hosted on Figshare and can be accessed online at `https://doi.org/10.6084/m9.figshare.17111693.v2`.

Independent Size Expansions in Red Algal Plastid and Mitochondrial Genomes

Supplementary Material

Fabian van Beveren, Laura Eme, Purificación López-García, Maria Ciobanu, and David Moreira

Ecologie Systématique Evolution, Centre National de la Recherche Scientifique - CNRS, Université Paris-Saclay, AgroParisTech, Orsay, France

Table S1. Sequencing information of sequenced species and genome statistics of all plastid and mitochondrial genomes used in this study. Data generated in this study is shown in bold. Species name, Class, DNA extraction method and sequencing method are shown in purple. Plastid and mitochondrial genome statistics are shown in green and orange respectively. DNA extraction and sequencing methods are not shown for genomic data not generated in this study. Number of contigs, total genome size and GC content were determined using bbmap. Number of protein coding genes and exons are based on final annotations, and the number of introns and protein-coding genes is calculated as: total number of exons in protein coding genes - total number of protein-coding genes. Mitochondrial data not available is indicated with grey cells. For extremely fragmented (>100 contigs) mitochondrial genomes, the total assembly statistics are given, with in parentheses the statistics based only on contigs with protein-coding genes. Statistics shown for the plastid genome of Compsopogon caeruleus are based on a reannotation of the protein-coding genes, resulting in an equal number of proteins-encoding genes (193) but a higher number of predicted introns (from 20 to 85).

Fig. S1. Presence/absence of genes in red algal plastid genomes. The tree topology (cladogram) on the left side is based on fig. 1 but without meaningful branch lengths. Leaf names are colored by class of the relative species: Cyanidiophyceae (cyan), Rhodellophyceae (red), Stylonematophyceae (purple), Porphyridiophyceae (yellow), Compsopogonophyceae (blue), Bangiophyceae (brown) and Florideophyceae (lavender). Gene names are given underneath and are sorted from most- to least-well distributed. Presence and absence are indicated by blue and red respectively. Gene absence has to be considered cautiously as in some cases it may be due to genome sequence missannotation.

Fig. S2. Presence/absence of genes in red algal mitochondrial genomes. The tree topology (cladogram) on the left side is based on fig. 2 but without meaningful branch lengths. Leaf names are colored by class of the relative species: Cyanidiophyceae (cyan), Rhodellophyceae (red), Stylonematophyceae (purple), Porphyridiophyceae (yellow), Compsopogonophyceae (blue), Bangiophyceae (brown) and Florideophyceae (lavender). Gene names are given underneath and are sorted from most- to least-well distributed. Presence and absence are indicated by blue and red respectively. Gene absence has to be considered cautiously as in some cases it may be due to genome sequence missannotation.

Fig. S3. Maximum likelihood phylogenies of plastid and mitochondrial datasets. (A, B) Plastid phylogenies from maximum likelihood analysis using IQ-TREE. (A) Phylogeny using the cpREV+C60+F+R7 model and (B) a phylogeny using the cpREV+PMSF(C60)+F+R7 model with the tree from A as the guide tree. (C, D) Mitochondrial phylogenies from maximum likelihood analysis using IQ-TREE. (C) Phylogeny using the mtZOA+C60+F+R9 model and (D) a phylogeny using the mtZOA+PMSF(C60)+F+R9 model with the tree from C as the guide tree.

Fig. S4. Alternative tree topologies used for the AU-test and results. Leaf names are colored by class of the corresponding species: Cyanidiophyceae (cyan), Rhodellophyceae (red), Stylonematophyceae (purple), Porphyridiophyceae (yellow), Compsopogonophyceae (blue), Bangiophyceae (brown) and Florideophyceae (lavender). (A) Tree topology based on the plastid dataset, but constrained to the relationships among Proteorhodophytina as in the mitochondrial topology from fig. 2. We considered the relationships between Rhodellophyceae and Compsopogonophyceae as unresolved in fig. 2. (B) Tree topology based on the mitochondrial dataset, but constrained to the relationships among Proteorhodophytina as in the plastid topology from fig. 1. (C) AU-test results giving the two alternative topologies and the p-values calculated from the different datasets, showing the mitochondrial phylogeny is rejected by the plastid dataset.

Fig. S5. Examples of degenerate and complete intron encoded proteins in *Rhodella violacea*. For each genomic segment, the location of the start and end is shown on the side of the genomes. Coding sequences (exons) are shown in blue, introns are shown in light blue and intron-encoded proteins are shown in white, intergenic regions with no trace of an intron-encoded protein are shown as a black line. For intron-encoded proteins, in-frame stop codons are shown assuming no frame shifts occurred.

Fig. S6. Phylogenetic tree of intron-encoded proteins (IEPs) detected in red algal organellar genomes. The phylogeny was reconstructed using IQ-TREE with the model LG+C20+G+F. Ultrafast bootstrap support is shown above each branch. Leaf names are colored by class of the relative species: Cyanidiophyceae (cyan), Rhodellophyceae (red), Stylonematophyceae (purple), Porphyridiophyceae (yellow), Compsopogonophyceae (blue), Bangiophyceae (brown) and Florideophyceae (lavender). Organellar genome type is indicated with green squares and orange circles for plastid and mitochondrial, respectively. For readability, large clades with the same phylogenetic affinity and organellar location were collapsed, with text besides giving the corresponding information (phylogenetic affinity, number of IEPs and number of organellar genomes).

Fig. S7. Plastid genome rearrangements between representatives of the classes Cyanidiophyceae, Stylonematophyceae, Compspogonophyceae, Porphyridiophyceae and Rhodellophyceae. Colors are arbitrary and are used to identify the differently identified syntenic regions (Locally Collinear Blocks) by Mauve.

Fig. S8. Mitochondrial genome rearrangements between representatives of the classes Cyanidiophyceae, Compspogonophyceae, Porphyridiophyceae and Rhodellophyceae. Colors are arbitrary and are used to identify the differently identified syntenic regions (Locally Collinear Blocks) by Mauve. Mitogenomes of Cyanidioschyzonaceae sp. ACUF 421 and representatives of Stylonematophyceae are not shown due to their fragmented assembly.

Fig. S9. Plot of trace parameters from Bayesian inference of the plastid dataset. An alignment consisting of 189 plastid-encoded protein sequences (37,256 characters) of 69 species was used for phylogenetic construction with Bayesian inference. Phylobayes was run using the CAT+GTR model with four independent chains, each run for at least 10,000 iterations. The labels visualize the log likelihood, the mean site entropy (statent), the total tree length (length), the alpha parameter of the gamma distribution (alpha), the number of components of the mixture model (Nmode) and the entropy of exchange rates (rrent).

Fig. S10. Plot of trace parameters from Bayesian inference of the mitochondrial dataset. An alignment consisting of 189 plastid-encoded protein sequences (4,537 characters) of 59 species was used for phylogenetic construction with Bayesian inference. Phylobayes was run using the CAT+GTR model with four independent chains, each run for at least 10,000 iterations. The labels visualize the log likelihood, the mean site entropy (statent), the total tree length (length), the alpha parameter of the gamma distribution (alpha), the number of components of the mixture model (Nmode) and the entropy of exchange rates (rrent).

Fig. S11. Phylogenetic reconstruction of RbcL protein sequences of red algae. A maximum likelihood phylogeny reconstructed with IQ-TREE using the cpREV+C40+F+R4 model and rooted between the Cyanidiophyta and all other red algae. The RbcL sequences coming from species included in the phylogeny of fig. 2 are shown in bold, all other sequences are from the Uniprot database and include their accession number. SH-aLRT/ultrafast bootstrap support is shown above each branch. Colors are based on taxonomic affiliation given in the name of the sequence, sequence names with no affiliation to any red algal class are shown in grey.

Table S1

General				
Species	**Class**	**Collection strain ID**	**DNA extraction method**	**Sequencing**
Bangia atropurpurea	**Bangiophyceae**	**SAG 1351-1**	**Heatshock + Dneasy PowerBiofilm (short reads)**	**Illumina HiSeq 2500 v2, 2x150**
Madagascaria erythrocladioides	**Compsopogonophyceae**	**CCAP 1342/2**	**Heatshock + Dneasy PowerBiofilm (short reads)**	**Illumina HiSeq 2500 v4, 2x100**
Pseudoerythrocladia kornmannii	**Compsopogonophyceae**	**CCMP 3220**	**Heatshock + Dneasy PowerBiofilm (short reads)**	**Illumina HiSeq 2500 v4, 2x100**
Sahlingia subintegra	**Compsopogonophyceae**	**CCMP 3223**	**Heatshock + Dneasy PowerBiofilm (short reads); CTAB extraction (long**	**Illumina HiSeq 2500 v4, 2x100; Nanopore Minion (SQK-LSK109), flow cell R9.4.1**
Pulvinaster venetus	**Compsopogonophyceae**	**CCAP 1386/1**	**Heatshock + Dneasy PowerBiofilm (short reads)**	**Illumina NovaSeq 6000 S2, 2x150**
Cyanidioschyzonaceae sp. ACUF421	**Cyanidiophyceae**	**ACUF 421**	**Heatshock + Dneasy PowerBiofilm + WGA-X (short reads)**	**Illumina NovaSeq 6000 S2, 2x150**
Cyanidioschyzonaceae sp. ACUF627	**Cyanidiophyceae**	**ACUF 627**	**Heatshock + Dneasy PowerBiofilm (short reads)**	**Illumina NovaSeq 6000 S2, 2x150**
Galdieria sp. ACUF613	**Cyanidiophyceae**	**ACUF 613**	**Heatshock + Dneasy PowerBiofilm + WGA-X (short reads)**	**Illumina NovaSeq 6000 S2, 2x150**
Erythrolobus coxiae	**Porphyridiophyceae**	**CCAP 1393/6**	**Heatshock + Dneasy PowerBiofilm (short reads)**	**Illumina NovaSeq 6000 S2, 2x150**
Porphyridium aerugineum	**Porphyridiophyceae**	**SAG 1380-2**	**Heatshock + Dneasy PowerBiofilm (short reads)**	**Illumina HiSeq 2500 v2, 2x150**
Timspurckia oligopyrenoides	**Porphyridiophyceae**	**CCAP 1393/1**	**Heatshock + Dneasy PowerBiofilm (short reads)**	**Illumina NovaSeq 6000 S2, 2x150**
Dixoniella grisea	**Rhodellophyceae**	**SAG 72.90**	**Heatshock + Dneasy PowerBiofilm (short reads)**	**Illumina HiSeq 2500 v2, 2x150**
Glaucosphaera vacuolata	**Rhodellophyceae**	**SAG 13.82**	**Heatshock + Dneasy PowerBiofilm (short reads)**	**Illumina HiSeq 2500 v4, 2x100**
Neorhodella cyanea	**Rhodellophyceae**	**CCAP 1346/1**	**Heatshock + Dneasy PowerBiofilm (short reads)**	**Illumina HiSeq 2500 v4, 2x100**
Rhodella violacea	**Rhodellophyceae**	**SAG 115.79**	**Heatshock + Dneasy PowerBiofilm (short and long reads)**	**Illumina HiSeq 2500 v2, 2x150; Nanopore Minion (SQK-LSK109), flow cell R9.4.1**
Chroodactylon ornatum	**Stylonematophyceae**	**SAG 103.79**	**Heatshock + Dneasy PowerBiofilm (short reads)**	**Illumina HiSeq 2500 v2, 2x150**
Chroothece richteriana	**Stylonematophyceae**	**SAG 104.79**	**Heatshock + Dneasy PowerBiofilm (short reads)**	**Illumina HiSeq 2500 v2, 2x150**
Goniotrichopsis reniformis	**Stylonematophyceae**	**CCMP 3252**	**Heatshock + Dneasy PowerBiofilm (short reads)**	**Illumina HiSeq 2500 v4, 2x100**
Boldiaceae sp. CCMP3255	**Compsopogonophyceae**	**CCMP 3255**	**Heatshock + Dneasy PowerBiofilm (short reads)**	**Illumina HiSeq 2500 v4, 2x100**
Rhodaphanes brevistipitata	**Stylonematophyceae**	**CCAP 1387/1**	**Heatshock + Dneasy PowerBiofilm (short reads)**	**Illumina HiSeq 2500 v4, 2x100**
Rhodospora sordida	**Stylonematophyceae**	**SAG 50.94**	**Heatshock + Dneasy PowerBiofilm (short reads)**	**Illumina HiSeq 2500 v2, 2x150**
Stylonema alsidii	**Stylonematophyceae**	**SAG 2.94**	**Heatshock + Dneasy PowerBiofilm (short reads)**	**Illumina HiSeq 2500 v2, 2x150**
Tsunamia transpacifica	**Stylonematophyceae**	**CCMP 3460**	**Heatshock + Dneasy PowerBiofilm (short reads)**	**Illumina HiSeq 2500 v4, 2x100**
Cyanidioschyzonaceae sp. CHI1	**Cyanidiophyceae**	**NA**	**Heatshock + Dneasy PowerBiofilm (short reads)**	**Illumina NovaSeq 6000 S2, 2x150**
Cyanidiaceae sp. CHI2	**Cyanidiophyceae**	**NA**	**Heatshock + Dneasy PowerBiofilm (short reads)**	**Illumina NovaSeq 6000 S2, 2x150**
Galdieria sulphuraria	Cyanidiophyceae			
Cyanidium caldarium	Cyanidiophyceae			
Cyanidioschyzon merolae	Cyanidiophyceae			
Cyanidiaceae sp. MX-AZ01	Cyanidiophyceae			
Corynoplastis japonica	Rhodellophyceae			
Bulboplastis apyrenoidosa	Rhodellophyceae			
Flintiella sanguinaria	Porphyridiophyceae			
Porphyridium purpureum	Porphyridiophyceae			
Porphyridium sordidum	Porphyridiophyceae			
Bangiopsis subsimplex	Stylonematophyceae			
Boldia erythrosiphon	Compsopogonophyceae			
Compsopogon caeruleus	Compsopogonophyceae			
Erythrotrichia carnea	Compsopogonophyceae			
Rhodochaete parvula	Compsopogonophyceae			
Bangia fuscopurpurea	Bangiophyceae			
Porphyra umbilicalis	Bangiophyceae			
Porphyra purpurea	Bangiophyceae			
Pyropia endiviifolia	Bangiophyceae			
Neopyropia fucicola	Bangiophyceae			
Neoporphyra haitanensis	Bangiophyceae			
Pyropia kanakaensis	Bangiophyceae			
Neoporphyra perforata	Bangiophyceae			
Pyropia pulchra	Bangiophyceae			
Neopyropia yezoensis	Bangiophyceae			
Wildemania schizophylla	Bangiophyceae			
Gracilaria chilensis	Florideophyceae			
Ahnfeltia plicata	Florideophyceae			
Asparagopsis taxiformis	Florideophyceae			
Calliarthron tuberculosum	Florideophyceae			
Chondrus crispus	Florideophyceae			
Palmaria palmata	Florideophyceae			
Gelidium vagum	Florideophyceae			
Grateloupia filicina	Florideophyceae			
Hildenbrandia rubra	Florideophyceae			
Lithothamnion sp.	Florideophyceae			
Plocamium cartilagineum	Florideophyceae			
Rhodogorgon sp.	Florideophyceae			
Rhodymenia pseudopalmata	Florideophyceae			
Schimmelmannia schousboei	Florideophyceae			
Schizymenia dubyi	Florideophyceae			
Sebdenia flabellata	Florideophyceae			
Sheathia arcuata	Florideophyceae			
Thorea hispida	Florideophyceae			
Vertebrata lanosa	Florideophyceae			

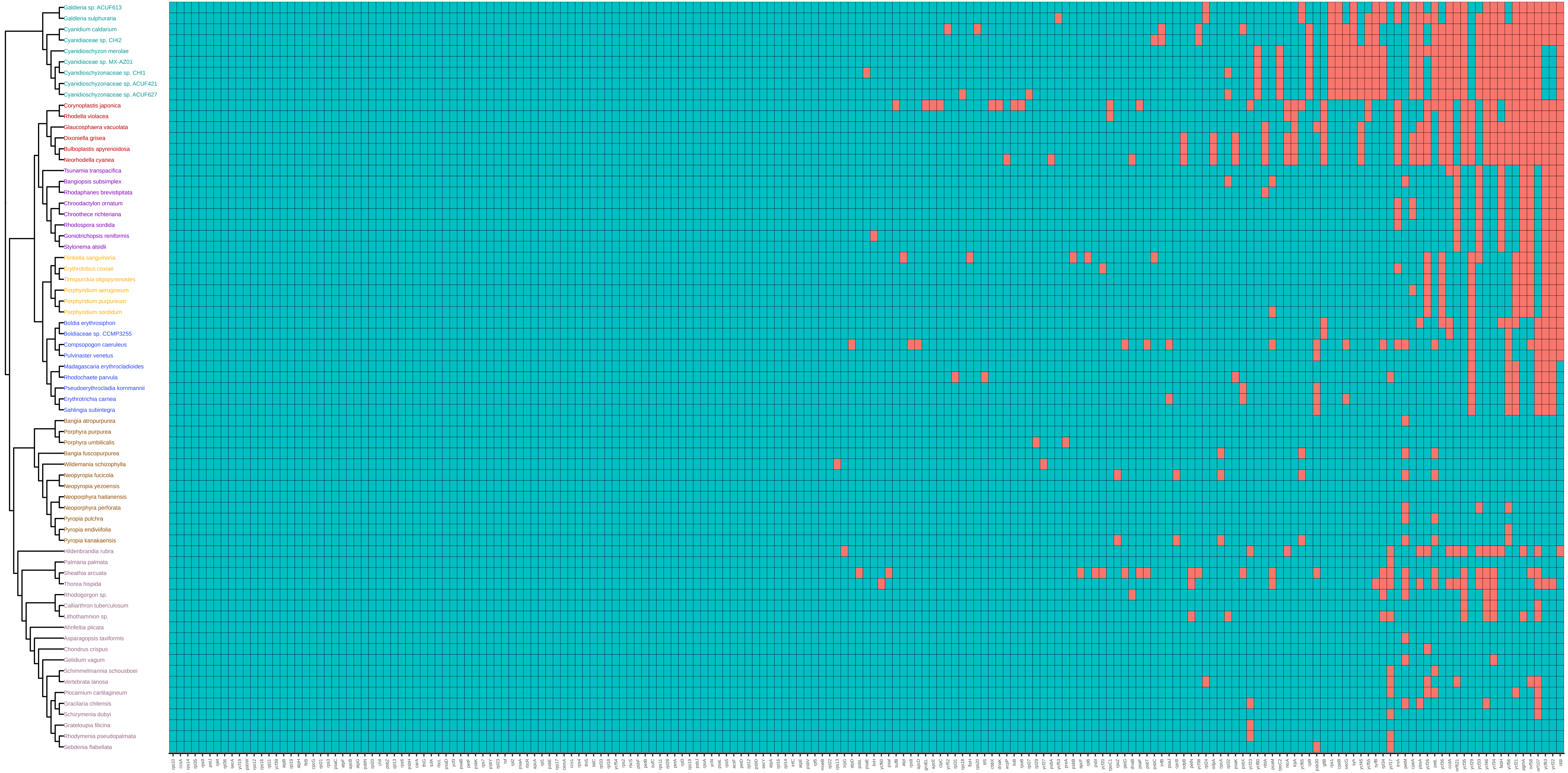
Table S1 (continued)

Plastid								
Source	Circlular assembly?	Assembly software	Number of contigs	Total size (bp)	GC content proportion	Number of protein coding genes	Number of exons	Number of introns in protein-coding genes
MZ569052	**Yes**	**NOVOPlasty**	**1**	**194293**	**0.32713**	**208**	**208**	**0**
MZ681963	**Yes**	**NOVOPlasty**	**1**	**216426**	**0.30089**	**196**	**256**	**60**
MZ681964	**Yes**	**NOVOPlasty**	**1**	**216464**	**0.30625**	**193**	**256**	**63**
MZ681967	**Yes**	**Unicycler (--conservative)**	**1**	**211447**	**0.31218**	**195**	**256**	**61**
MZ681965	**Yes**	**NOVOPlasty**	**1**	**224653**	**0.26747**	**199**	**289**	**90**
MZ681959	**Yes**	**NOVOPlasty**	**1**	**147875**	**0.39642**	**201**	**201**	**0**
MZ681960	**Yes**	**NOVOPlasty**	**1**	**147983**	**0.39633**	**197**	**197**	**0**
MZ681961	**Yes**	**NOVOPlasty**	**1**	**167118**	**0.28688**	**190**	**190**	**0**
MZ853148	**Yes**	**NOVOPlasty**	**1**	**297871**	**0.28223**	**191**	**303**	**112**
MZ853149	**Yes**	**NOVOPlasty**	**1**	**201607**	**0.28329**	**190**	**224**	**34**
MZ748306	**Yes**	**NOVOPlasty**	**1**	**291973**	**0.35638**	**189**	**332**	**143**
MZ681966	**Yes**	**NOVOPlasty**	**1**	**516819**	**0.23165**	**172**	**342**	**170**
MZ895126	**Yes**	**NOVOPlasty**	**1**	**475182**	**0.26384**	**179**	**354**	**175**
Supplementary Data	**No**	**GetOrganelle**	**32**	**682745**	**0.24557**	**165**	**378**	**213**
MZ853150	**Yes**	**Unicycler (--bold)**	**1**	**903688**	**0.28141**	**175**	**378**	**203**
MZ603803	**Yes**	**NOVOPlasty**	**1**	**207095**	**0.31651**	**197**	**224**	**27**
MZ680365	**Yes**	**NOVOPlasty**	**1**	**207333**	**0.31483**	**196**	**224**	**28**
MZ603804	**Yes**	**NOVOPlasty**	**1**	**240595**	**0.28897**	**199**	**261**	**62**
MZ681962	**Yes**	**NOVOPlasty**	**1**	**231003**	**0.26826**	**193**	**288**	**95**
MZ603618	**No**	**GetOrganelle**	**1**	**227519**	**0.2959**	**201**	**256**	**55**
MZ603805	**Yes**	**NOVOPlasty**	**1**	**208685**	**0.31688**	**198**	**231**	**33**
MZ603806	**Yes**	**NOVOPlasty**	**1**	**204775**	**0.29646**	**200**	**232**	**32**
MZ680366	**Yes**	**NOVOPlasty**	**1**	**207429**	**0.3174**	**195**	**228**	**33**
MZ681957	**Yes**	**NOVOPlasty**	**1**	**145899**	**0.37445**	**207**	**207**	**0**
MZ681958	**Yes**	**NOVOPlasty**	**1**	**165057**	**0.32943**	**191**	**191**	**0**
KJ700459.1	Yes		1	167741	0.28486	182	182	0
AF022186.2	Yes		1	164921	0.32731	197	197	0
AB002583.1	Yes		1	149987	0.37629	193	193	0
KJ569775.1	Yes		1	145533	0.37486	210	210	0
KY709210.1	No		1	1127474	0.23697	168	380	212
KY709209.1	Yes		1	610063	0.23255	169	389	220
KY709211.1	No		1	370672	0.25463	186	341	155
MF401423.1	Yes		1	220483	0.30403	198	240	42
KX284720.1	Yes		1	259429	0.30719	191	256	65
KX284718.1	Yes		1	204784	0.29248	194	231	37
KY709208.1	Yes		1	226658	0.27047	187	281	94
KY083067.1	Yes		1	221013	0.26001	193	278	85
KX284721.2	Yes		1	210691	0.31868	191	249	58
KY709212.1	No		1	221656	0.28279	192	262	70
KP714733.1	Yes		1	196913	0.33541	205	205	0
MF385003.1	Yes		1	190173	0.32878	207	207	0
U38804.1	Yes		1	191028	0.32986	204	204	0
KT716756.1	Yes		1	195784	0.3329	210	210	0
KJ776837.1	No		1	187282	0.32732	203	203	0
KC464603.1	Yes		1	195597	0.32984	211	211	0
KJ776836.1	No		1	189931	0.32854	206	206	0
KC904971.1	Yes		1	189789	0.32909	209	209	0
KT266789.1	Yes		1	194175	0.3334	207	207	0
KC517072.1	Yes		1	191975	0.33089	213	213	0
KR020505.1	Yes		1	193008	0.34416	211	211	0
MF401963.1	Yes		1	185640	0.29341	202	202	0
KX284715.1	Yes		1	190451	0.32541	205	205	0
KX284717.1	Yes		1	177091	0.294	203	203	0
KC153978.1	Yes		1	178981	0.29157	201	202	1
HF562234.1	Yes		1	180086	0.28728	204	204	0
KX284726.1	Yes		1	192960	0.33906	204	205	1
KT266787.1	Yes		1	179853	0.29874	200	200	0
MG598531.1	Yes		1	195990	0.30671	233	233	0
KX284724.1	Yes		1	180141	0.31371	190	191	1
MH281627.1	Yes		1	183822	0.31083	200	201	1
KX284727.1	Yes		1	171392	0.27212	197	197	0
MH281630.1	Yes		1	190860	0.32931	202	203	1
KX284709.1	Yes		1	194153	0.3203	201	201	0
KX284711.1	Yes		1	181030	0.28613	202	202	0
KX284712.1	Yes		1	183959	0.29968	204	204	0
KX284713.1	Yes		1	192140	0.29225	205	205	0
NC_035231.1	Yes		1	187354	0.29884	185	186	1
KY083065.1	No		1	175278	0.28327	193	194	1
KP308097.1	Yes		1	167158	0.29984	192	192	0

Table S1 (continued)

Mitochondrial								
Source	Circlular assembly?	Assembly software	Number of contigs	Total size (bp)	GC content proportion	Number of protein coding genes	Number of exons	Number of introns in protein-coding genes
MZ748311	Yes	GetOrganelle	1	35668	0.30904	25	28	3
MZ748308	Yes	NOVOPlasty	1	75292	0.23989	20	51	31
MZ748318	No	SPAdes	1	49042	0.29774	20	28	8
MZ748320	Yes	Unicycler (--conservative)	1	52388	0.30039	21	26	5
MZ748315	Yes	GetOrganelle	1	100252	0.18896	21	79	58
Supplementary Data	No	SPAdes	26	29117	0.28056	15	17	2
MZ748314	Yes	GetOrganelle	1	32386	0.26397	33	34	1
MZ748319	Yes	SPAdes	1	19912	0.44466	17	17	0
MZ748317	No	NOVOPlasty	1	102701	0.3003	23	57	34
MZ748310	Yes	NOVOPlasty	1	26688	0.23902	17	18	1
Supplementary Data	No	GetOrganelle	2	34163	0.29391	16	26	10
MZ748316	No	GetOrganelle	1	58206	0.18771	18	29	11
MZ748307	Yes	NOVOPlasty	1	28895	0.28676	17	29	12
MZ748309	Yes	NOVOPlasty	1	47861	0.24141	17	27	10
MZ748321	No	Unicycler (--conservative)	1	45191	0.25186	19	24	5
Supplementary Data	No	GetOrganelle	119 (10)	42435 (18440)	0.44242 (0.4532)	9	10	1
Supplementary Data	No	metaSPAdes	118 (10)	49787 (29017)	0.4467 (0.46032)	12	13	1
Supplementary Data	No	GetOrganelle	310 (13)	67170 (21489)	0.45392 (0.47443)	11	12	1
Supplementary Data	No	SPAdes	2	25352	0.25103	19	20	1
Supplementary Data	No	GetOrganelle	314 (13)	75719 (27725)	0.41251 (0.45421)	11	12	1
Supplementary Data	No	metaSPAdes	282 (14)	54060 (23238)	0.58141 (0.55448)	12	15	3
Supplementary Data	No	GetOrganelle	350 (13)	96682 (34528)	0.44957 (0.49925)	12	13	1
Supplementary Data	No	SPAdes	957 (14)	100358 (22596)	0.36475 (0.44388)	12	13	1
MZ748312	Yes	GetOrganelle	1	32663	0.26862	35	36	1
MZ748313	Yes	GetOrganelle	1	34774	0.26365	36	36	0
KJ700460.1	Yes		1	21428	0.43994	19	19	0
NC_000887.3	Yes		1	32211	0.27065	34	34	0
KJ569774.1	Yes		1	32620	0.2668	33	33	0
KY083068.1	Yes		1	26807	0.27377	21	21	0
KP710961.1	Yes		1	43517	0.32989	30	33	3
JQ388471.1	Yes		1	29123	0.31875	26	26	0
AF114794.1	Yes		1	36753	0.33486	29	29	0
KU356193.1	Yes		1	34603	0.3046	26	26	0
KJ708762.1	Yes		1	35035	0.32459	31	31	0
JQ736808.1	Yes		1	37023	0.30678	28	30	2
KJ708765.1	Yes		1	39300	0.3	30	30	0
KF515971.1	Yes		1	33919	0.3144	28	28	0
MT588076.1	Yes		1	33190	0.30669	26	27	1
NC_017837.1	Yes		1	41688	0.32719	32	35	3
KJ689442.1	Yes		1	29156	0.33221	26	26	0
MF401962.1	Yes		1	26897	0.27557	25	25	0
KF649303.1	Yes		1	32878	0.33386	26	28	2
KJ398158.1	Yes		1	26097	0.26746	25	25	0
KR005619.1	Yes		1	26469	0.27311	25	25	0
NC_001677.2	Yes		1	25836	0.27887	28	28	0
KF649305.1	Yes		1	29735	0.32163	26	26	0
KC875854.1	Yes		1	24901	0.30466	23	23	0
MG598532.1	Yes		1	29274	0.31946	25	26	1
KF649304.1	Yes		1	33066	0.32198	24	25	1
MH281621.1	Yes		1	25605	0.27167	25	25	0
KJ398160.1	No		1	26431	0.23569	24	25	1
MH281625.1	Yes		1	30547	0.26025	26	28	2
KC875852.1	Yes		1	26166	0.29546	24	24	0
KJ398162.1	No		1	25906	0.26668	23	23	0
KJ398163.1	No		1	26438	0.25856	24	24	0
KJ398164.1	No		1	26767	0.28466	23	23	0
KY083064.1	Yes		1	25086	0.29674	23	23	0
KY083066.1	Yes		1	25380	0.28255	25	25	0
KX687880.1	Yes		1	25337	0.23606	23	23	0

Fig. S1



Galdieria sp. ACUF613
Galdieria sulphuraria
Cyanidium caldarium
Cyanidiaceae sp. CHI2
Cyanidioschyzon merolae
Cyanidiaceae sp. MX-AZ01
Cyanidioschyzonaceae sp. CHI1
Cyanidioschyzonaceae sp. ACUF421
Cyanidioschyzonaceae sp. ACUF627
Corynoplastis japonica
Rhodella violacea
Glaucosphaera vacuolata
Dixoniella grisea
Bulboplastis apyrenoidosa
Neorhodella cyanea
Tsunamia transpacifica
Bangiopsis subsimplex
Rhodaphanes brevistipitata
Chroodactylon ornatum
Chroothece richteriana
Rhodospora sordida
Goniotrichopsis reniformis
Stylonema alsidii
Flintiella sanguinaria
Erythrolobus coxiae
Timspurckia oligopyrenoides
Porphyridium aerugineum
Porphyridium purpureum
Porphyridium sordidum
Boldia erythrosiphon
Boldiaceae sp. CCMP3255
Compsopogon caeruleus
Pulvinaster venetus
Madagascaria erythrocladioides
Rhodochaete parvula
Pseudoerythrocladia kornmannii
Erythrotrichia carnea
Sahlingia subintegra
Bangia atropurpurea
Porphyra purpurea
Porphyra umbilicalis
Bangia fuscopurpurea
Wildemania schizophylla
Neopyropia fucicola
Neopyropia yezoensis
Neoporphyra haitanensis
Neoporphyra perforata
Pyropia pulchra
Pyropia endiviifolia
Pyropia kanakaensis
Hildenbrandia rubra
Palmaria palmata
Sheathia arcuata
Thorea hispida
Rhodogorgon sp.
Calliarthron tuberculosum
Lithothamnion sp.
Ahnfeltia plicata
Asparagopsis taxiformis
Chondrus crispus
Gelidium vagum
Schimmelmannia schousboei
Vertebrata lanosa
Plocamium cartilagineum
Gracilaria chilensis
Schizymenia dubyi
Grateloupia filicina
Rhodymenia pseudopalmata
Sebdenia flabellata

Fig. S2

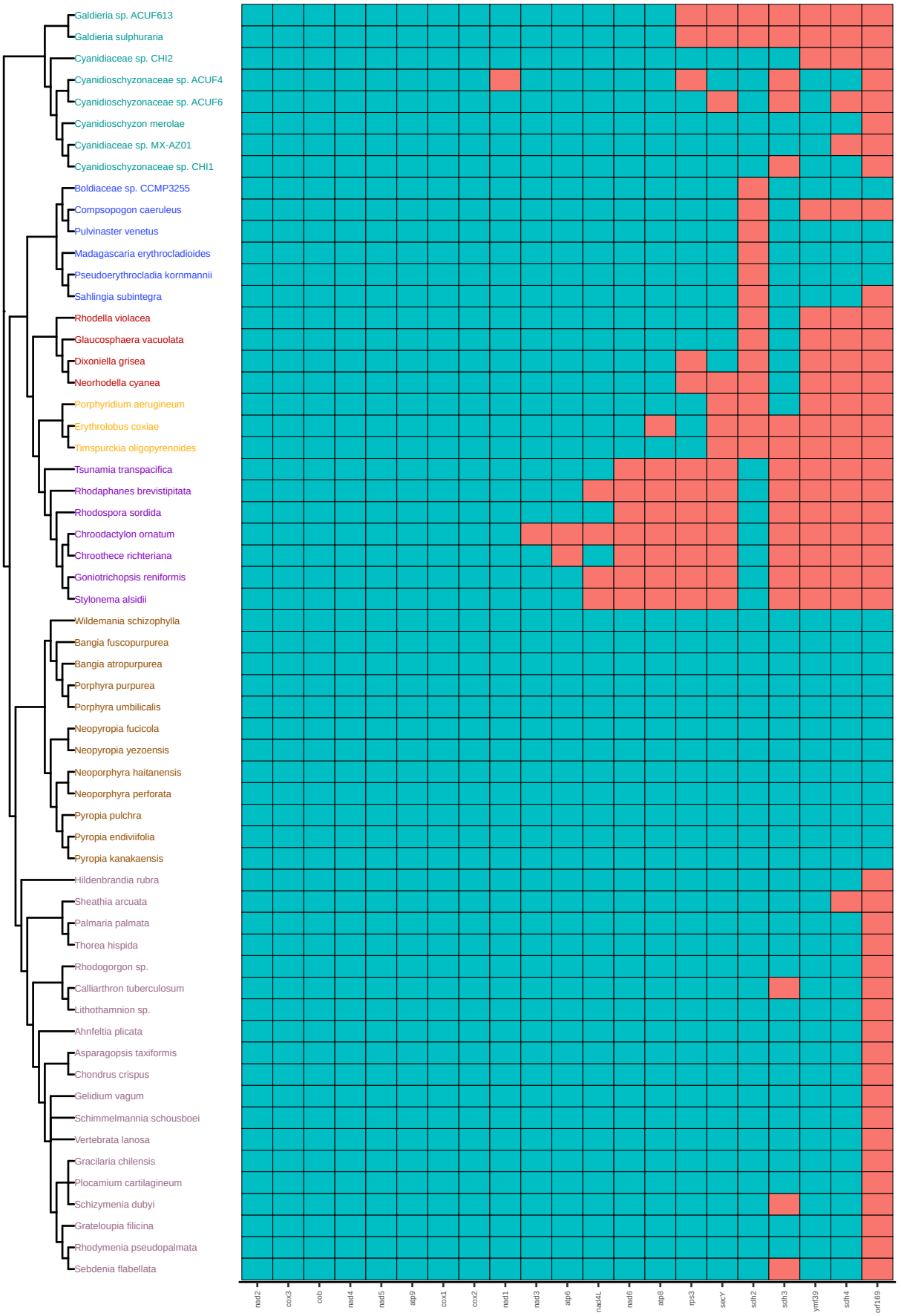
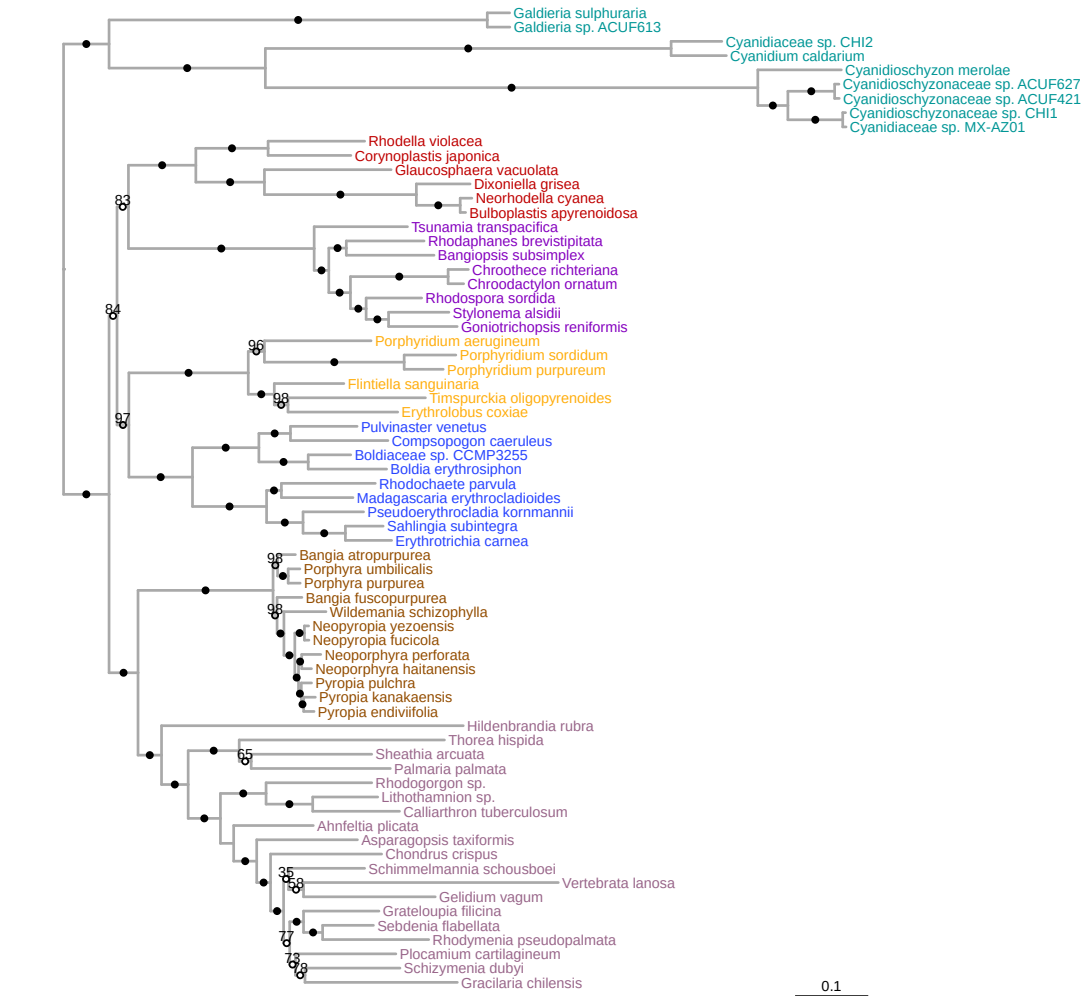
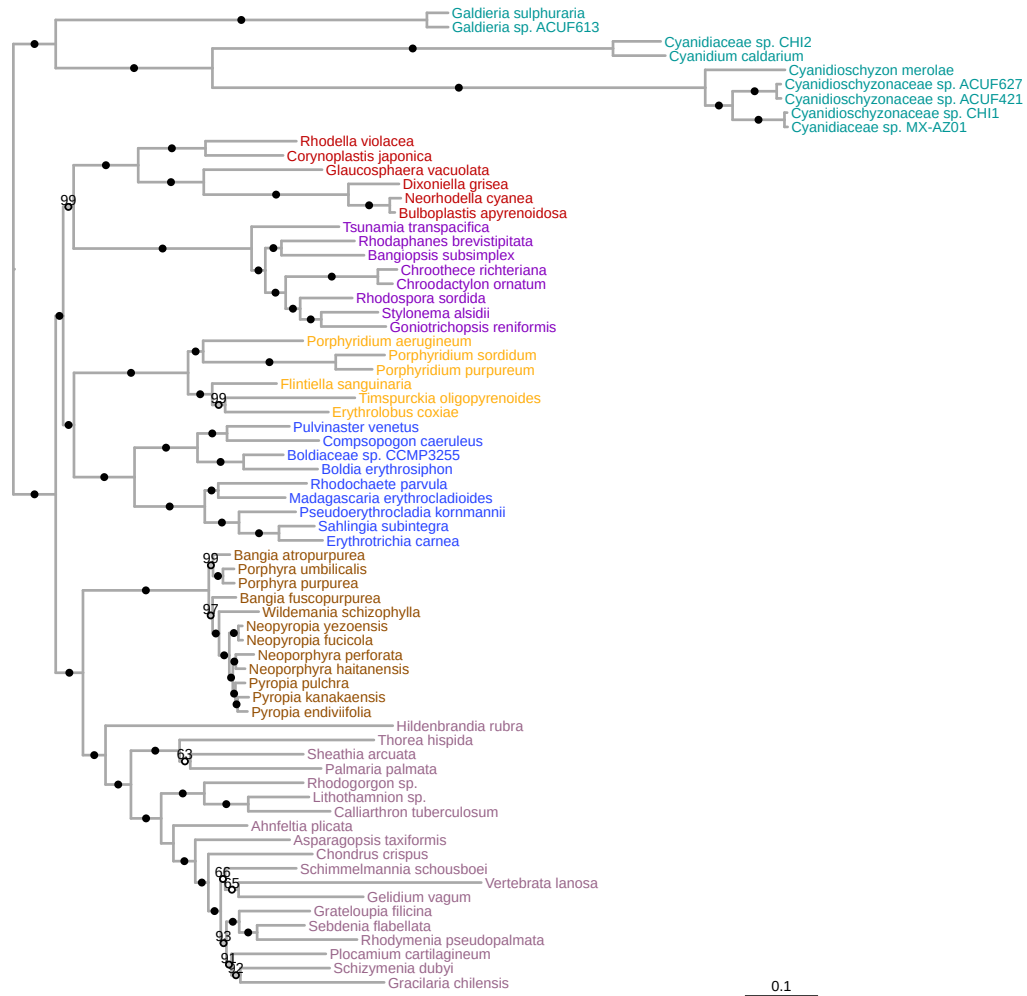

Galdieria sp. ACUF613
Galdieria sulphuraria
Cyanidiaceae sp. CHI2
Cyanidioschyzonaceae sp. ACUF4
Cyanidioschyzonaceae sp. ACUF6
Cyanidioschyzon merolae
Cyanidiaceae sp. MX-AZ01
Cyanidioschyzonaceae sp. CHI1
Boldiaceae sp. CCMP3255
Compsopogon caeruleus
Pulvinaster venetus
Madagascaria erythrocladioides
Pseudoerythrocladia kornmannii
Sahlingia subintegra
Rhodella violacea
Glaucosphaera vacuolata
Dixoniella grisea
Neorhodella cyanea
Porphyridium aerugineum
Erythrolobus coxiae
Timspurckia oligopyrenoides
Tsunamia transpacifica
Rhodaphanes brevistipitata
Rhodospora sordida
Chroodactylon ornatum
Chroothece richteriana
Goniotrichopsis reniformis
Stylonema alsidii
Wildemania schizophylla
Bangia fuscopurpurea
Bangia atropurpurea
Porphyra purpurea
Porphyra umbilicalis
Neopyropia fucicola
Neopyropia yezoensis
Neoporphyra haitanensis
Neoporphyra perforata
Pyropia pulchra
Pyropia endiviifolia
Pyropia kanakaensis
Hildenbrandia rubra
Sheathia arcuata
Palmaria palmata
Thorea hispida
Rhodogorgon sp.
Calliarthron tuberculosum
Lithothamnion sp.
Ahnfeltia plicata
Asparagopsis taxiformis
Chondrus crispus
Gelidium vagum
Schimmelmannia schousboei
Vertebrata lanosa
Gracilaria chilensis
Plocamium cartilagineum
Schizymenia dubyi
Grateloupia filicina
Rhodymenia pseudopalmata
Sebdenia flabellata
nad2
cox3
cob
nad4
nad5
atp9
cox1
cox2
nad1
nad3
atp6
nad4L
nad6
atp8
rps3
secY
sdh2
sdh3
ymf39
sdh4
orf169

Fig. S3

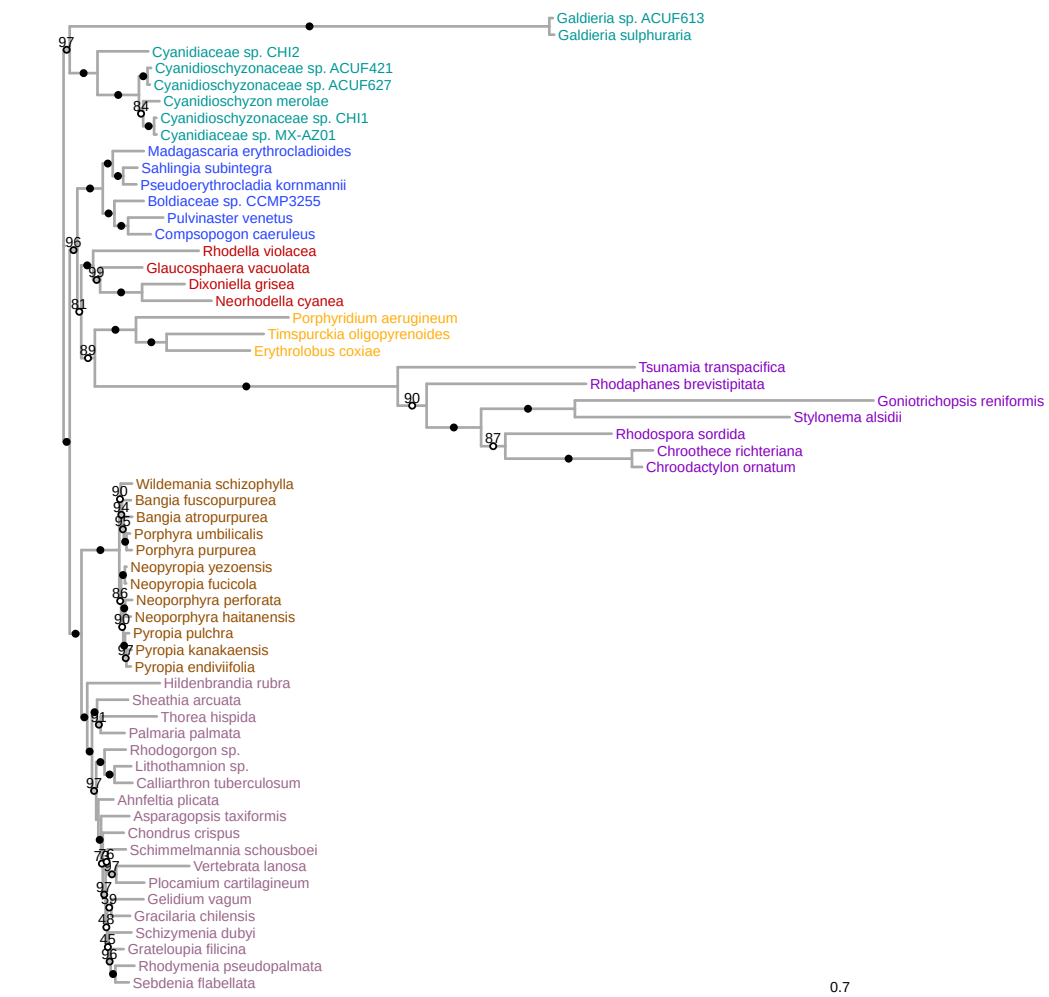
A


Galdieria sulphuraria
Galdieria sp. ACUF613
Cyanidiaceae sp. CHI2
Cyanidium caldarium
Cyanidioschyzon merolae
Cyanidioschyzonaceae sp. ACUF627
Cyanidioschyzonaceae sp. ACUF421
Cyanidioschyzonaceae sp. CHI1
Cyanidiaceae sp. MX-AZ01
Rhodella violacea
Corynoplastis japonica
Glaucosphaera vacuolata
Dixoniella grisea
Neorhodella cyanea
Bulboplastis apyrenoidosa
Tsunamia transpacifica
Rhodaphanes brevistipitata
Bangiopsis subsimplex
Chroothece richteriana
Chroodactylon ornatum
Rhodospora sordida
Stylonema alsidii
Goniotrichopsis reniformis
Porphyridium aerugineum
Porphyridium sordidum
Porphyridium purpureum
Flintiella sanguinaria
Timspurckia oligopyrenoides
Erythrolobus coxiae
Pulvinaster venetus
Compsopogon caeruleus
Boldiaceae sp. CCMP3255
Boldia erythrosiphon
Rhodochaete parvula
Madagascaria erythrocladioides
Pseudoerythrocladia kornmannii
Sahlingia subintegra
Erythrotrichia carnea
Bangia atropurpurea
Porphyra umbilicalis
Porphyra purpurea
Bangia fuscopurpurea
Wildemania schizophylla
Neopyropia yezoensis
Neopyropia fucicola
Neoporphyra perforata
Neoporphyra haitanensis
Pyropia pulchra
Pyropia kanakaensis
Pyropia endiviifolia
Hildenbrandia rubra
Thorea hispida
Sheathia arcuata
Palmaria palmata
Rhodogorgon sp.
Lithothamnion sp.
Calliarthron tuberculosum
Ahnfeltia plicata
Asparagopsis taxiformis
Chondrus crispus
Schimmelmannia schousboei
Vertebrata lanosa
Gelidium vagum
Grateloupia filicina
Sebdenia flabellata
Rhodymenia pseudopalmata
Plocamium cartilagineum
Schizymenia dubyi
Gracilaria chilensis
0.1

B


Galdieria sulphuraria
Galdieria sp. ACUF613
Cyanidiaceae sp. CHI2
Cyanidium caldarium
Cyanidioschyzon merolae
Cyanidioschyzonaceae sp. ACUF627
Cyanidioschyzonaceae sp. ACUF421
Cyanidioschyzonaceae sp. CHI1
Cyanidiaceae sp. MX-AZ01
Rhodella violacea
Corynoplastis japonica
Glaucosphaera vacuolata
Dixoniella grisea
Neorhodella cyanea
Bulboplastis apyrenoidosa
Tsunamia transpacifica
Rhodaphanes brevistipitata
Bangiopsis subsimplex
Chroothece richteriana
Chroodactylon ornatum
Rhodospora sordida
Stylonema alsidii
Goniotrichopsis reniformis
Porphyridium aerugineum
Porphyridium sordidum
Porphyridium purpureum
Flintiella sanguinaria
Timspurckia oligopyrenoides
Erythrolobus coxiae
Pulvinaster venetus
Compsopogon caeruleus
Boldiaceae sp. CCMP3255
Boldia erythrosiphon
Rhodochaete parvula
Madagascaria erythrocladioides
Pseudoerythrocladia kornmannii
Sahlingia subintegra
Erythrotrichia carnea
Bangia atropurpurea
Porphyra umbilicalis
Porphyra purpurea
Bangia fuscopurpurea
Wildemania schizophylla
Neopyropia yezoensis
Neopyropia fucicola
Neoporphyra perforata
Neoporphyra haitanensis
Pyropia pulchra
Pyropia kanakaensis
Pyropia endiviifolia
Hildenbrandia rubra
Thorea hispida
Sheathia arcuata
Palmaria palmata
Rhodogorgon sp.
Lithothamnion sp.
Calliarthron tuberculosum
Ahnfeltia plicata
Asparagopsis taxiformis
Chondrus crispus
Schimmelmannia schousboei
Vertebrata lanosa
Gelidium vagum
Grateloupia filicina
Sebdenia flabellata
Rhodymenia pseudopalmata
Plocamium cartilagineum
Schizymenia dubyi
Gracilaria chilensis
0.1

C


Galdieria sp. ACUF613
Galdieria sulphuraria
Cyanidiaceae sp. CHI2
Cyanidioschyzonaceae sp. ACUF421
Cyanidioschyzonaceae sp. ACUF627
Cyanidioschyzon merolae
Cyanidioschyzonaceae sp. CHI1
Cyanidiaceae sp. MX-AZ01
Madagascaria erythrocladioides
Sahlingia subintegra
Pseudoerythrocladia kornmannii
Boldiaceae sp. CCMP3255
Pulvinaster venetus
Compsopogon caeruleus
Rhodella violacea
Glaucosphaera vacuolata
Dixoniella grisea
Neorhodella cyanea
Porphyridium aerugineum
Timspurckia oligopyrenoides
Erythrolobus coxiae
Tsunamia transpacifica
Rhodaphanes brevistipitata
Goniotrichopsis reniformis
Stylonema alsidii
Rhodospora sordida
Chroothece richteriana
Chroodactylon ornatum
Wildemania schizophylla
Bangia fuscopurpurea
Bangia atropurpurea
Porphyra umbilicalis
Porphyra purpurea
Neopyropia yezoensis
Neopyropia fucicola
Neoporphyra perforata
Neoporphyra haitanensis
Pyropia pulchra
Pyropia kanakaensis
Pyropia endiviifolia
Hildenbrandia rubra
Sheathia arcuata
Thorea hispida
Palmaria palmata
Rhodogorgon sp.
Lithothamnion sp.
Calliarthron tuberculosum
Ahnfeltia plicata
Asparagopsis taxiformis
Chondrus crispus
Schimmelmannia schousboei
Vertebrata lanosa
Plocamium cartilagineum
Gelidium vagum
Gracilaria chilensis
Schizymenia dubyi
Grateloupia filicina
Rhodymenia pseudopalmata
Sebdenia flabellata
0.7

D

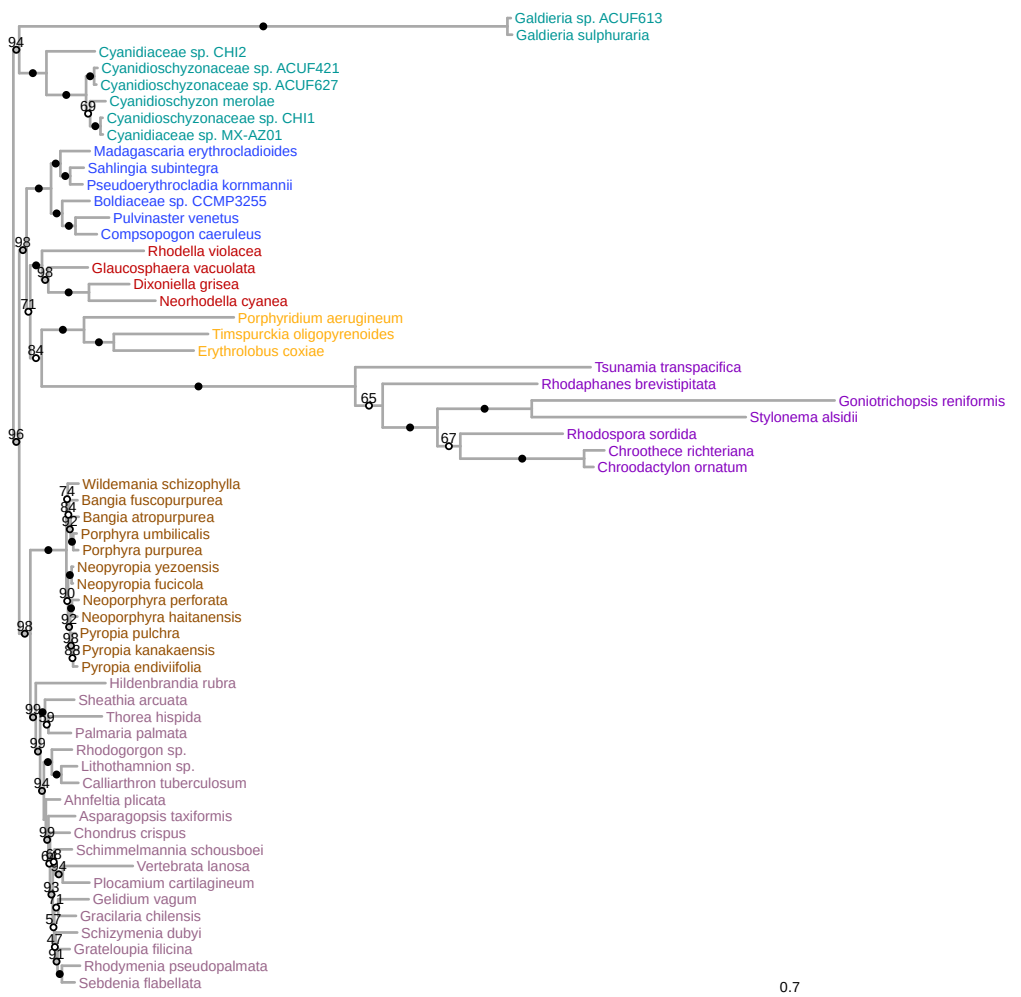
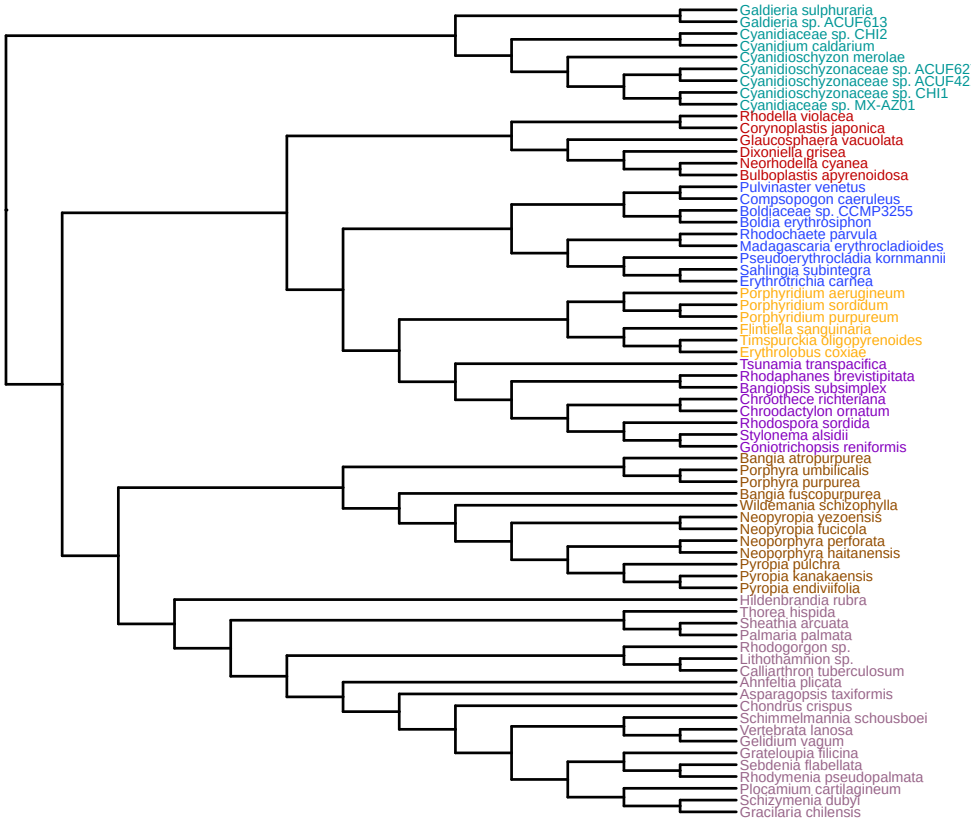
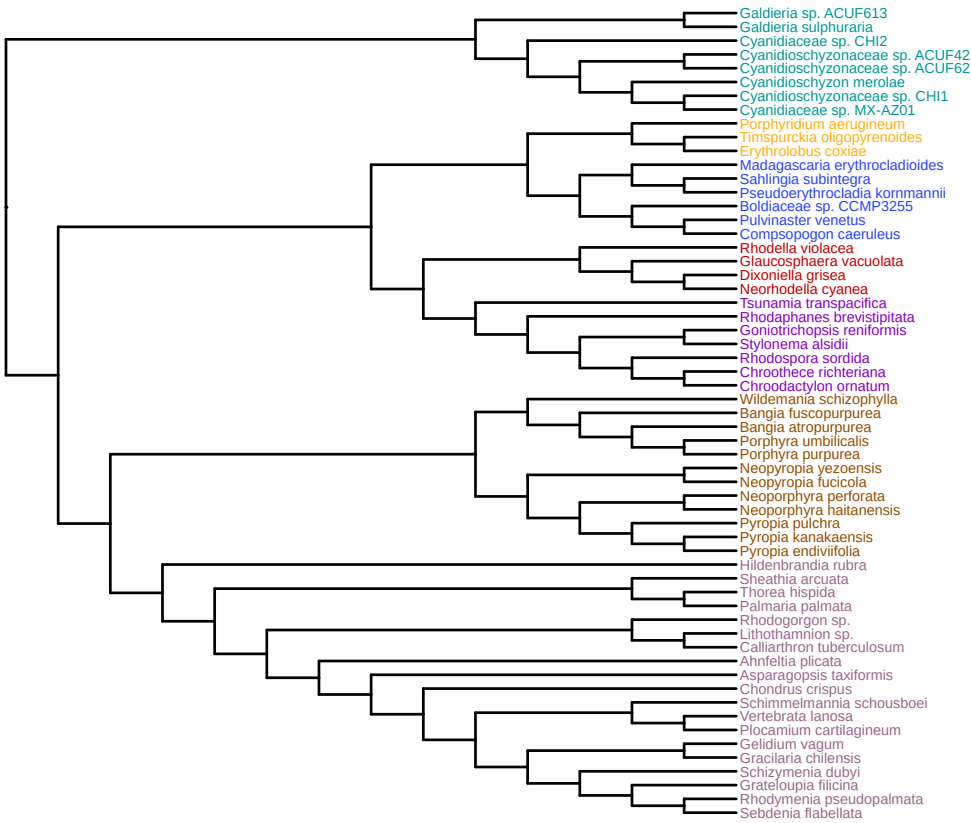

Galdieria sp. ACUF613
Galdieria sulphuraria
Cyanidiaceae sp. CHI2
Cyanidioschyzonaceae sp. ACUF421
Cyanidioschyzonaceae sp. ACUF627
Cyanidioschyzon merolae
Cyanidioschyzonaceae sp. CHI1
Cyanidiaceae sp. MX-AZ01
Madagascaria erythrocladioides
Sahlingia subintegra
Pseudoerythrocladia kornmannii
Boldiaceae sp. CCMP3255
Pulvinaster venetus
Compsopogon caeruleus
Rhodella violacea
Glaucosphaera vacuolata
Dixoniella grisea
Neorhodella cyanea
Porphyridium aerugineum
Timspurckia oligopyrenoides
Erythrolobus coxiae
Tsunamia transpacifica
Rhodaphanes brevistipitata
Goniotrichopsis reniformis
Stylonema alsidii
Rhodospora sordida
Chroothece richteriana
Chroodactylon ornatum
Wildemania schizophylla
Bangia fuscopurpurea
Bangia atropurpurea
Porphyra umbilicalis
Porphyra purpurea
Neopyropia yezoensis
Neopyropia fucicola
Neoporphyra perforata
Neoporphyra haitanensis
Pyropia pulchra
Pyropia kanakaensis
Pyropia endiviifolia
Hildenbrandia rubra
Sheathia arcuata
Thorea hispida
Palmaria palmata
Rhodogorgon sp.
Lithothamnion sp.
Calliarthron tuberculosum
Ahnfeltia plicata
Asparagopsis taxiformis
Chondrus crispus
Schimmelmannia schousboei
Vertebrata lanosa
Plocamium cartilagineum
Gelidium vagum
Gracilaria chilensis
Schizymenia dubyi
Grateloupia filicina
Rhodymenia pseudopalmata
Sebdenia flabellata
0.7

Fig. S4

A



B



C

Hypothesis	Plastid dataset	Mitochondrial dataset
Porphyridiophyceae Compsopogonophyceae Rhodellophyceae Stylonematophyceae	0.780	0.400
Compsopogonophyceae Rhodellophyceae Porphyridiophyceae Stylonematophyceae	**0.00238***	0.967

* = P-value < 0.05, hypothesis rejected

Fig. S5

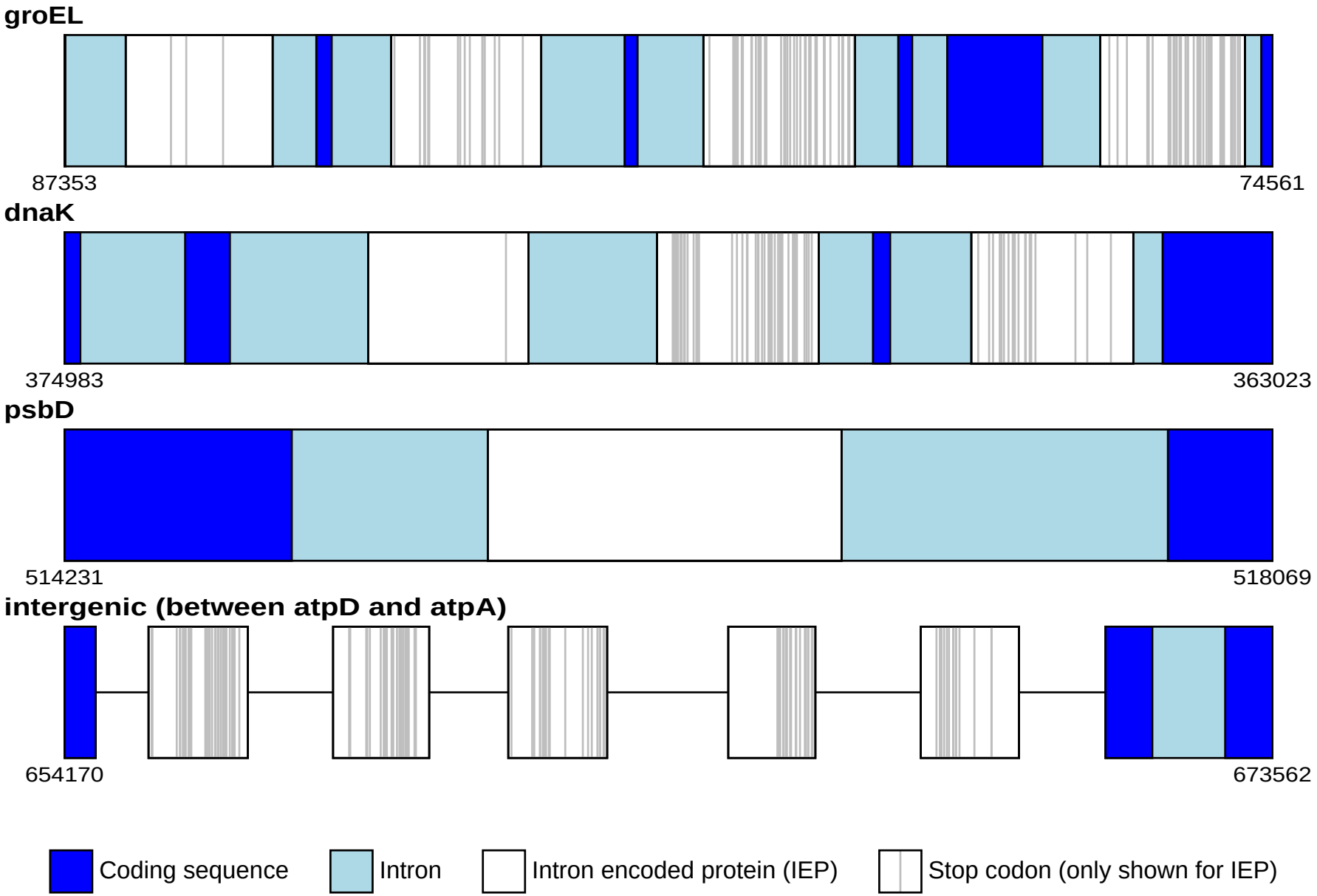

groEL
87353
74561
dnaK
374983
363023
psbD
514231
518069
intergenic (between atpD and atpA)
654170
673562
Coding sequence
Intron
Intron encoded protein (IEP)
Stop codon (only shown for IEP)

Fig. S6

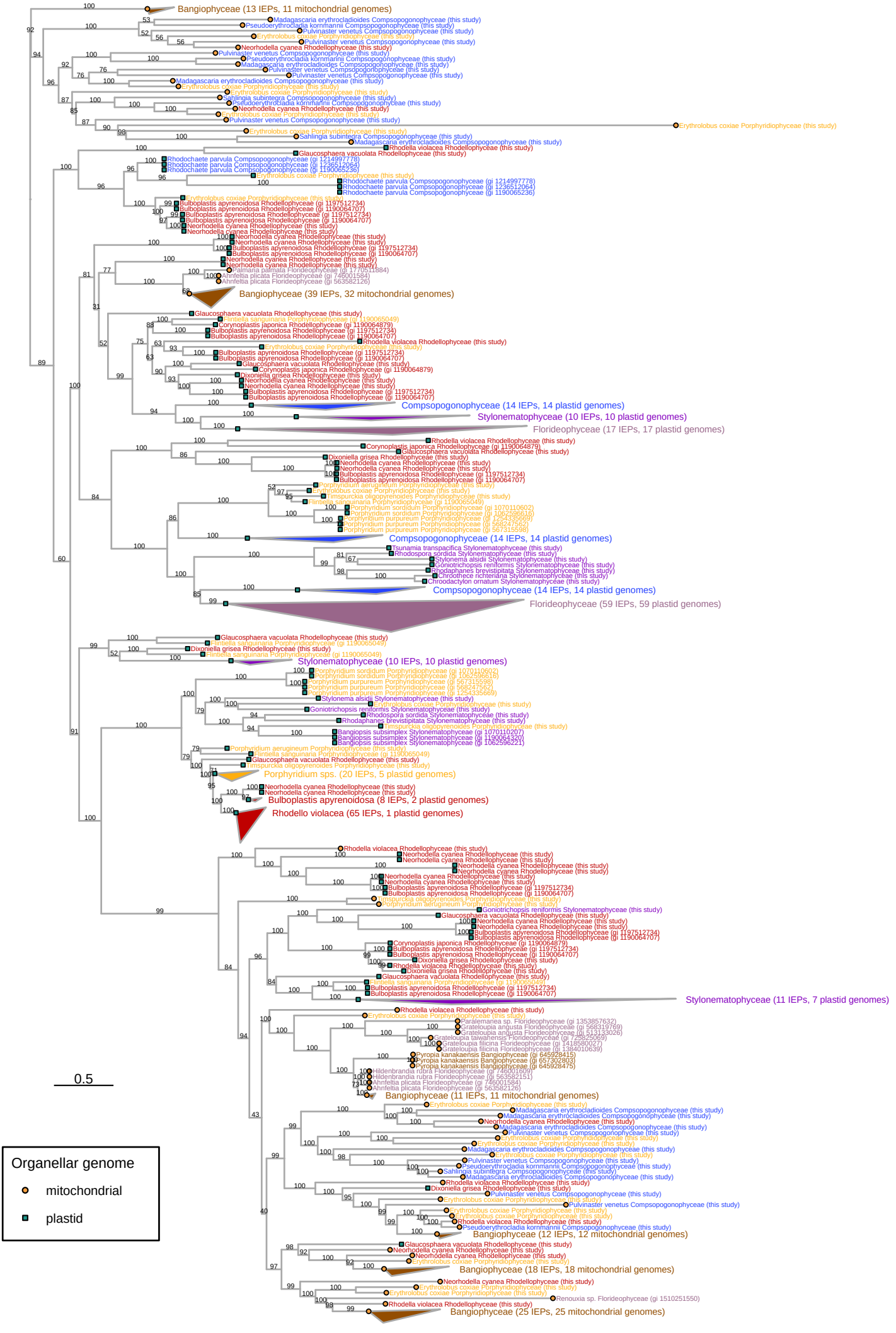
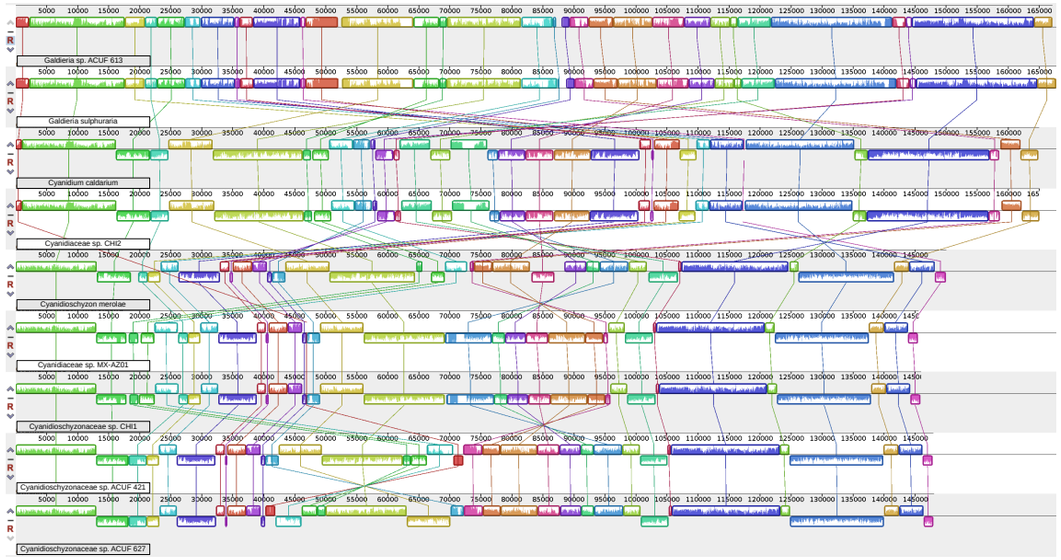

Bangiophyceae (13 IEPs, 11 mitochondrial genomes)
Bangiophyceae (39 IEPs, 32 mitochondrial genomes)
Compsopogonophyceae (14 IEPs, 14 plastid genomes)
Stylonematophyceae (10 IEPs, 10 plastid genomes)
Florideophyceae (17 IEPs, 17 plastid genomes)
Compsopogonophyceae (14 IEPs, 14 plastid genomes)
Compsopogonophyceae (14 IEPs, 14 plastid genomes)
Florideophyceae (59 IEPs, 59 plastid genomes)
Stylonematophyceae (10 IEPs, 10 plastid genomes)
Porphyridium sps. (20 IEPs, 5 plastid genomes)
Bulboplastis apyrenoidosa (8 IEPs, 2 plastid genomes)
Rhodello violacea (65 IEPs, 1 plastid genomes)
Stylonematophyceae (11 IEPs, 7 plastid genomes)
Bangiophyceae (11 IEPs, 11 mitochondrial genomes)
Bangiophyceae (12 IEPs, 12 mitochondrial genomes)
Bangiophyceae (18 IEPs, 18 mitochondrial genomes)
Bangiophyceae (25 IEPs, 25 mitochondrial genomes)
0.5
Organellar genome
mitochondrial
plastid

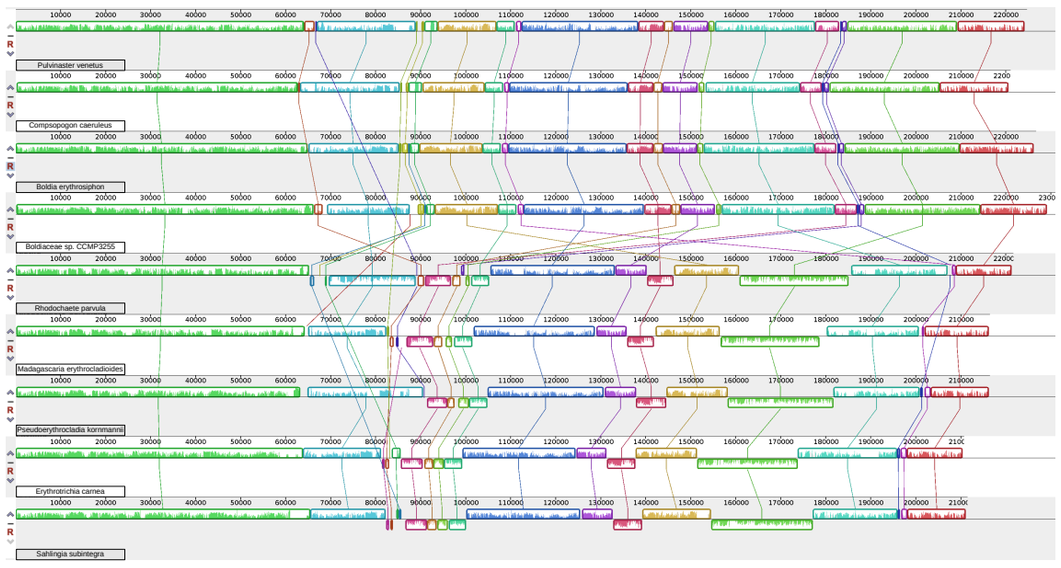
Fig. S7

Cyanidiophyceae

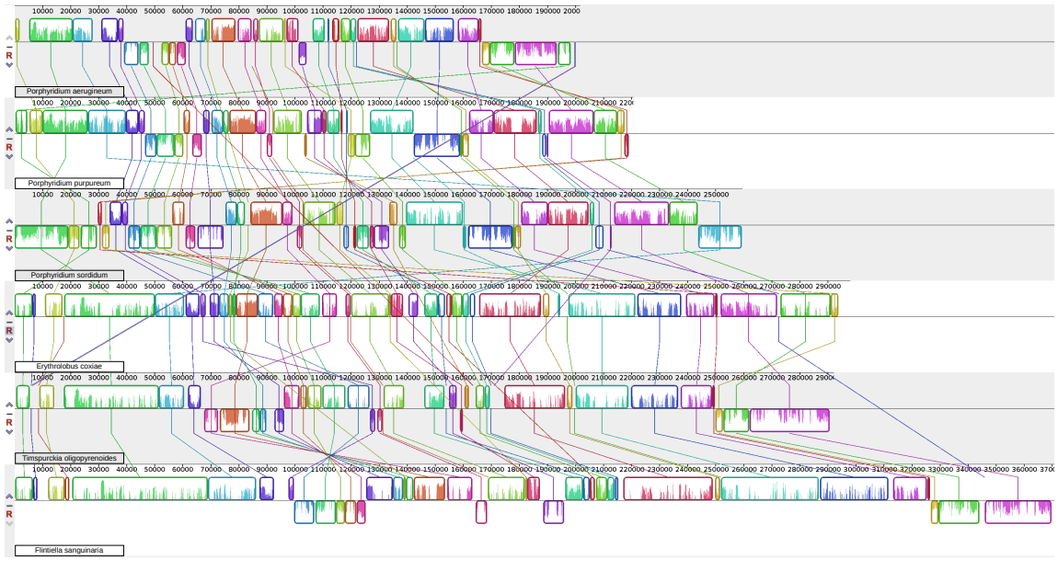


Stylonematophyceae

Compsopogonophyceae



Porphyridiophyceae



Rhodellophyceae

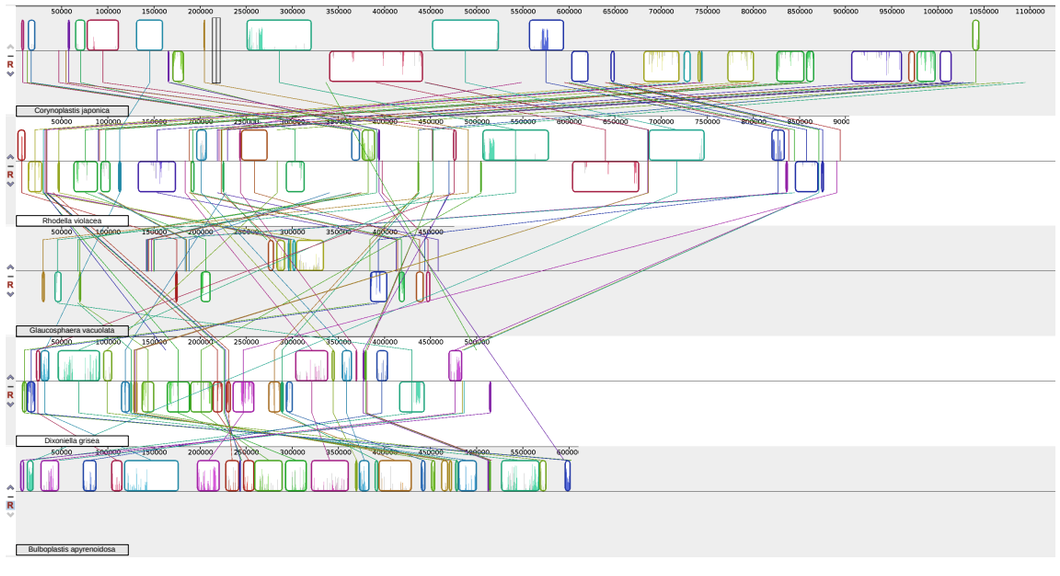
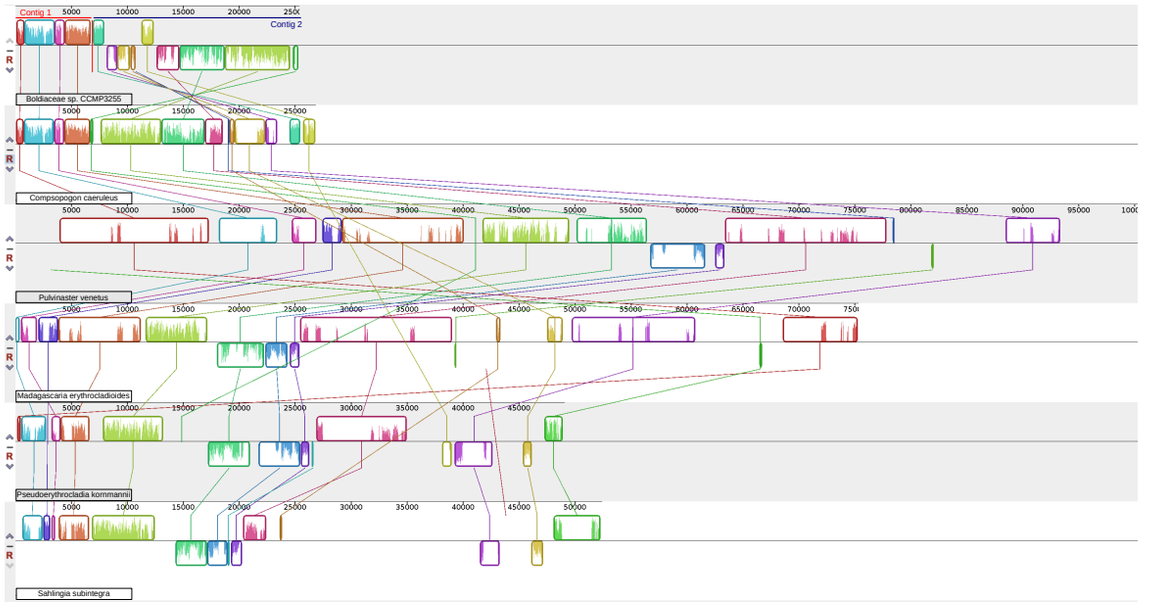


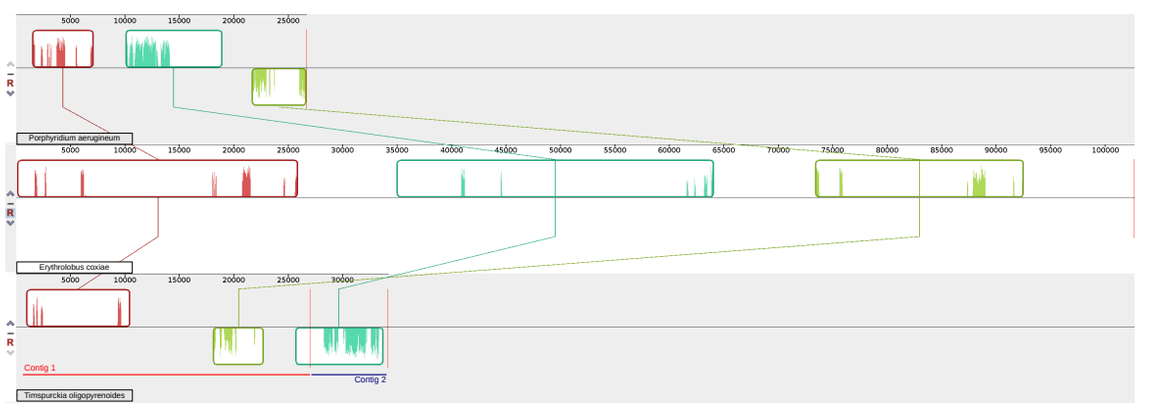
Fig. S8

Cyanidiophyceae

Compsopogonophyceae



Porphyridiophyceae



Rhodellophyceae

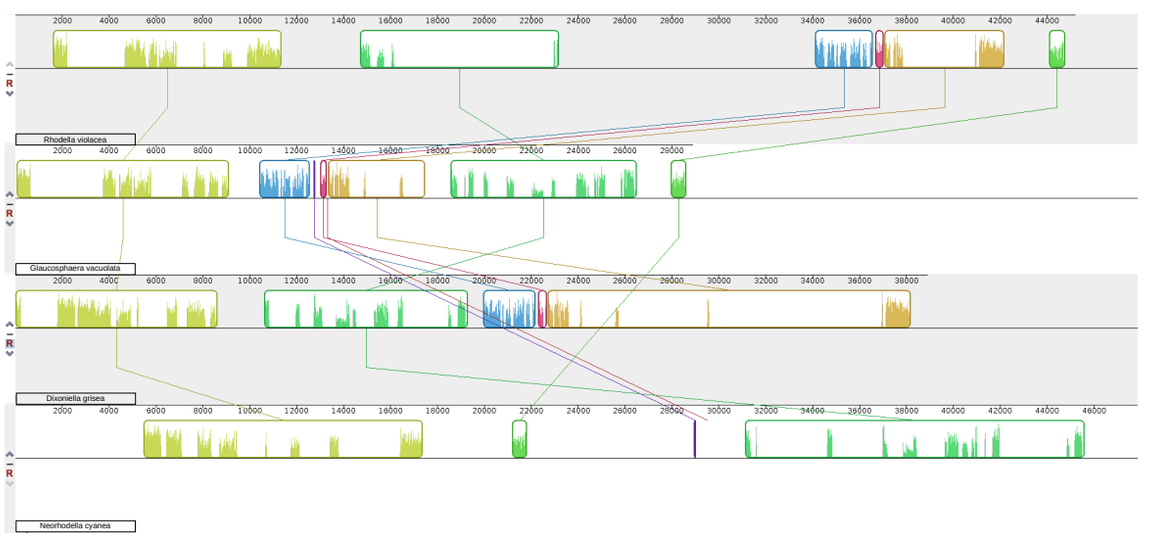


Fig. S9

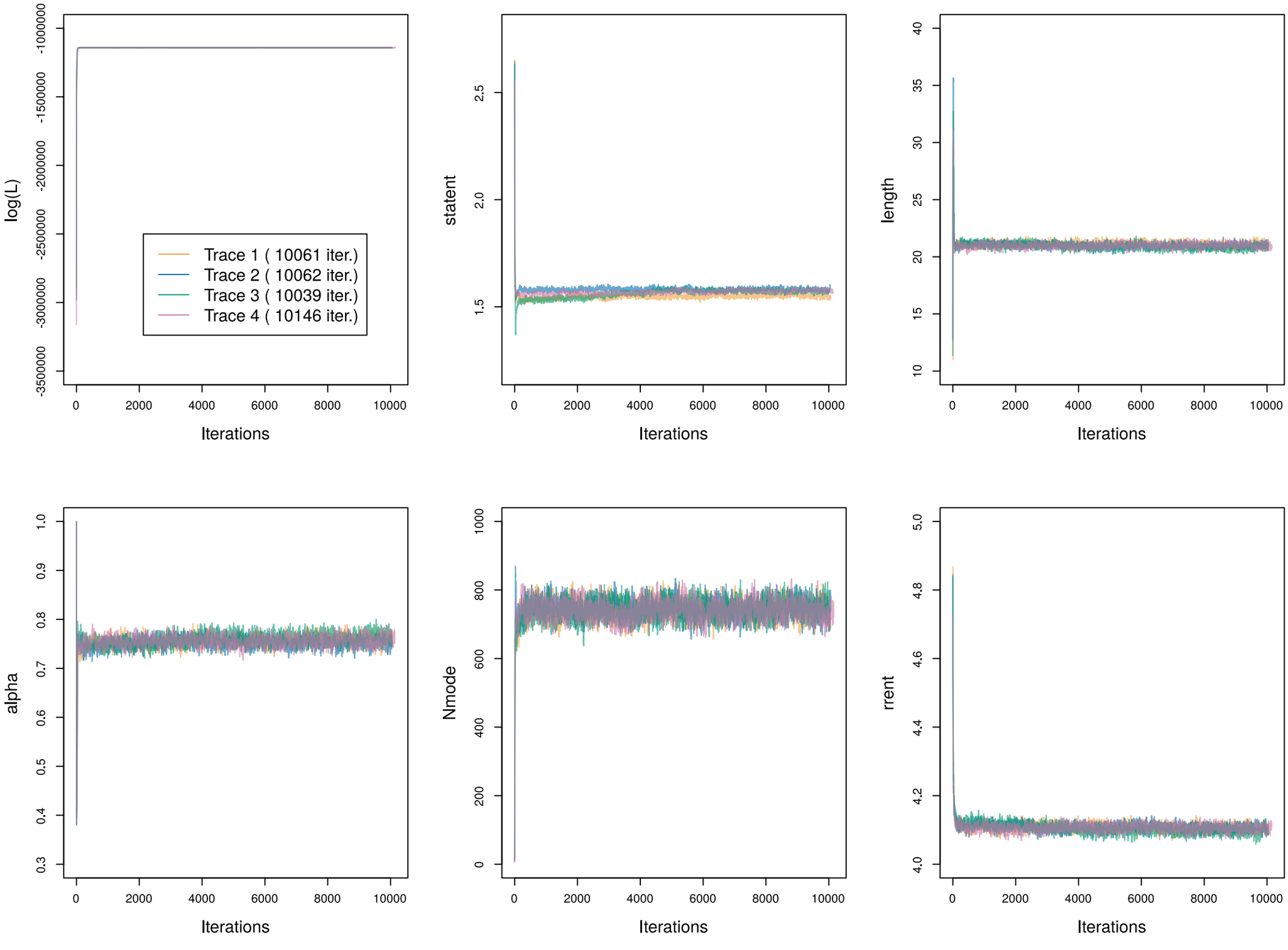

log(L)
-1000000
-1500000
-2000000
-2500000
-3000000
-3500000
Trace 1 (10061 iter.)
Trace 2 (10062 iter.)
Trace 3 (10039 iter.)
Trace 4 (10146 iter.)
0
2000
4000
6000
8000
10000
Iterations
statent
2.5
2.0
1.5
Iterations
length
40
35
30
25
20
15
10
Iterations
alpha
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
Iterations
Nmode
1000
800
600
400
200
0
Iterations
rrent
5.0
4.8
4.6
4.4
4.2
4.0
Iterations

Fig. S10

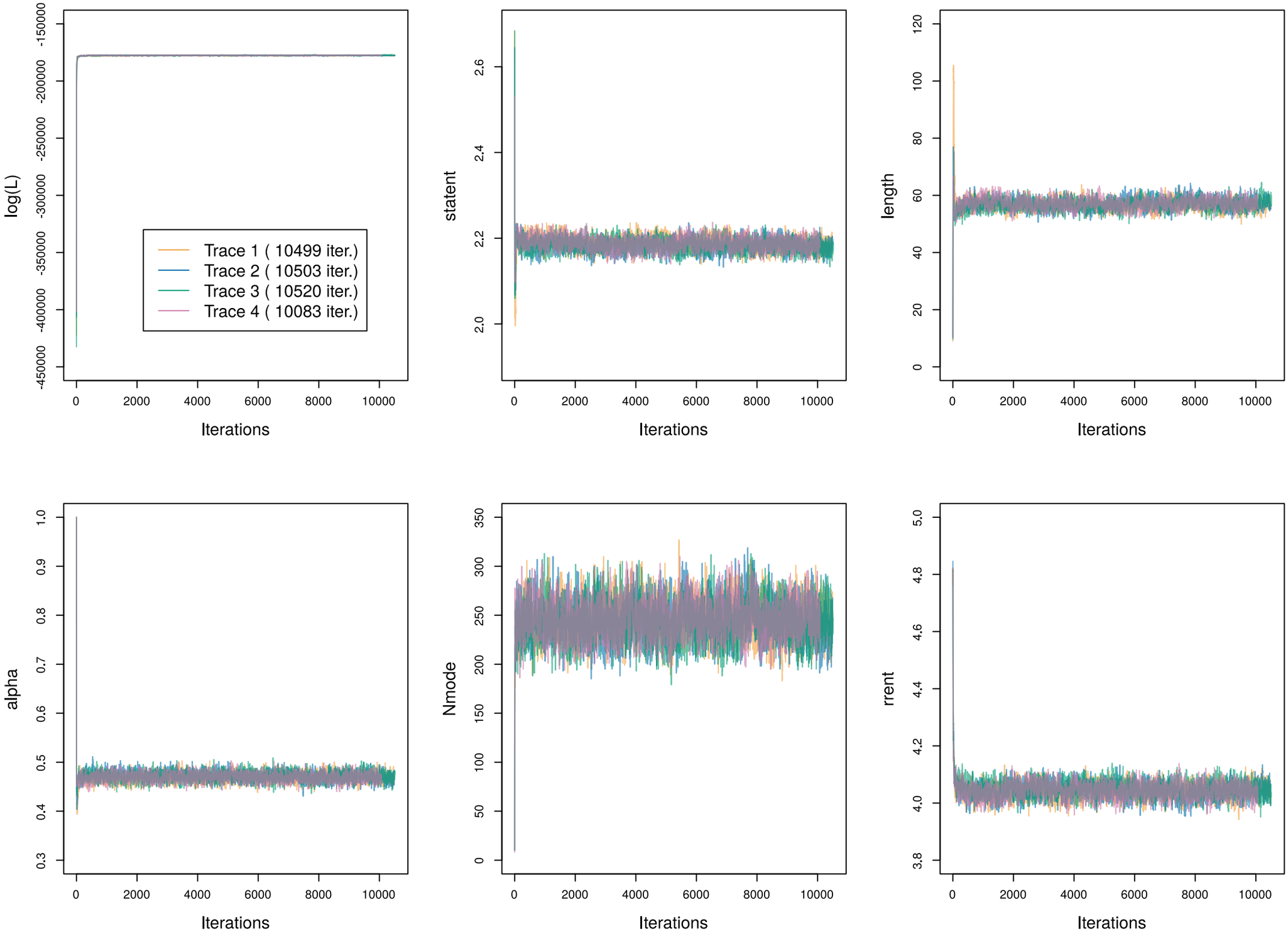

log(L)
-150000
-200000
-250000
-300000
-350000
-400000
-450000
Trace 1 (10499 iter.)
Trace 2 (10503 iter.)
Trace 3 (10520 iter.)
Trace 4 (10083 iter.)
0
2000
4000
6000
8000
10000
Iterations
statent
2.6
2.4
2.2
2.0
Iterations
length
120
100
80
60
40
20
0
Iterations
alpha
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
Iterations
Nmode
350
300
250
200
150
100
50
0
Iterations
rrent
5.0
4.8
4.6
4.4
4.2
4.0
3.8
Iterations

Fig. S11

Evaluating serial endosymbiosis hypotheses

The supplementary material can be found in the following pages.

Supplementary Material for "Evaluating serial endosymbiosis hypotheses in the light of red plastid phylogenomics".
Van Beveren et al, in prep.

Supplementary information S1. Calibrations

Default minimum calibrations

Node: Corallinophycidae – Ahnfeltiophycidae, Rhodymeniophycidae

Fossils: *Thallophyca corrugata*, *T. ramosa* and *Paramecia incognata* (Xiao et al. 2004)
Minimum age: 550.4 Ma
Phylogenetic justification: The fossils *Thallophyca* and *Paramecia* have characteristic pseudoparenchymatous tissue, but they lack (bio)calcification - a synapomorphy of modern coralline algae - and thus do not belong to any crown-group of Corallinophycidea, and are here instead considered stem group Corallinophycidea (Xiao et al. 2004). We are unaware of any disagreement on the affinity of these fossils as stem-group Corallinophycidae.
Age justification: The age of fossiliferous Doushantuo phosphorites have been estimated using Pb-Pb geochronometry and Lu-Hf dating to be ~599 and ~584 Ma (Barfod et al. 2002). The lower Doushantuo formation is estimated at 635.2 ± 0.6 Ma (Condon et al. 2005). The age of the top of the uppermost Doushantuo Formation, the Miaohe Member is estimated at 551.1 ± 0.7 Ma, bordering the Dengying Formation (Condon et al. 2005), thus we make use of a minimum age of 550.4 Ma.

Node: Sporolithales – Corallinales, Hapaldiales

Fossil: *Sporolithon* sp. (Chatalov et al. 2015)
Minimum age: 133.3 Ma
Phylogenetic justification: The first bona fide *Sporolithon* fossil is from the Alkiki Limestones (interval B/II) in northern Greece (Chatalov et al. 2015).This fossil is older than *Sporolithon rude* and *S. phylloideum* from the Hauterivian (Ghosh and Maithy 1996; Tomás et al. 2007).
Age justification: The fossil is from the Aliki Limestones, which is considered of lower Valanginian age. The Valanginian ranges from 139.4 ± 0.7 Ma to 133.9 ± 0.6 Ma (Gradstein 2012), thus we use a minimum age of 133.3 Ma. Older fossils have been claimed but are controversial (Peña et al. 2020).

Node: Corallinales

Fossil: *Karpathia sphaerocellulosa* (Aguirre et al. 2007)
Minimum age: 63.8 Ma
Phylogenetic justification: Based on the oldest fossil of the Lithophyllaceae, *Karpathia sphaerocellulosa*, which shows characters specific to the subfamily Hydrolithoideae (Aguirre et al. 2007; Rösler et al. 2017). Our analysis does not include genomic data of Hydrolithoideae, but does of the family Lithophyllaceae to which it belongs (Rösler et al. 2017; Peña et al. 2020).

Age justification: *Karpathia sphaerocellulosa* is from the Iberian Peninsula, from the Maastrichtian-Danian depositional sequence, shortly after the Cretaceous-Tertiary mass extinction that occurred at ~66 Ma (Aguirre et al. 2007; Rösler et al. 2017). Dated at 66 ± 2.23 (Rösler et al. 2017), we make use of a minimum age of 63.8 Ma.

Node: Phaeocystales – Isochrysidales, Coccolithales

Fossils: *Crucirhabdus primilus*, *C. minutus* and *Archaeozygodiscus koessenensis* (Bown 1998)
Minimum age: 209.5 Ma
Phylogenetic justification: Heterococcoliths are defining features of coccolithophorids. Some members of Isochrysidales lack heterococcoliths, but these are assumed to be secondary losses (Henderiks et al. 2022). The oldest heterococcolith fossils are from the fossils *Crucirhabdus primulus*, *C. minutus* and *Archaeozygodiscus koessenensis* (Bown 1998). Coccolithophorids are represented in our study by Isochrysidales and Coccolithales, with Phaeocystales as the nearest sister lineage.
Age justification: We follow Medlin et al (2008), based on Bown (1998). First occurence in the Norian stage of the Late Triassic, based on the fossils *Crucirhabdus primulus*, *C. minutus* and *Archaeozygodiscus koessenensis*. As no specific age is given, we take the range of the Norian (228.4 - 209.5), and use a minimum age of 209.5 Ma (Gradstein et al, 2012). Older claimed coccolith fossils are likely artifacts or contamination (Bown 1998; Medlin et al. 2008).

Node: Coccolithales – Isochrysidales

Fossil: *Anfractus* spp. (First holococcoliths) (Farinacci and Howe 2022)
Minimum age: 180.4 Ma
Phylogenetic justification: Holococcoliths are found in a subgroup of coccolithophorids, specifically excluding the Isochrysidales among other lineages, whilst including the Coccolithales (the two orders represented in this study) (Hagino and Young 2015). Isochrysidales is often found as a sister group to the holococcolith-bearing lineages, making it possible that holococcoliths evolved after the divergence between Isochrysidales and other coccolithophorid groups (including Coccolithales), although it is not impossible that holococcoliths evolved first but did not fossilize well (Henderiks et al. 2022). The earliest holococcolith fossils are from the Early Toarcian (Bown, 1998), specifically early ranges are known for *Anfractus parvus* and *A. youngii* (Bown 1998; Farinacci and Howe 2022). The original description of both are by Bown (1993), which we could not access so we make use of the Catalog of Calcareous Nannofossils (Farinacci and Howe 2022)
Age justification: First occurence of *A. younggi* is in the NJ6 zone, base in the Toarcian stage (Farinacci and Howe 2022). The top of NJ6 zone is dated at least 180.49 Ma (Gradstein et al, 2012). We make use of a minimum age of 180.4 Ma.

Node: Rhizosoleniales – Coscinodiscophycidae

Fossil: Highly branched isoprenoids (Damsté et al. 2004)
Minimum age: 90.0 Ma
Phylogenetic justification: Highly branched isoprenoids (HBI) are only known to be produced by rhizosolenid diatoms, which fall in two different clades: one consisting of species of the genus *Rhizosolenia*, the other group includes multiple species of Naviculales (raphid diatoms) (Damsté et al. 2004). The maximum could be used for both groups, as both

produce the C25 HBI that are found ~91.5 million years ago, but the minimum age for Naviculales can instead be better constrained with the first rapid diatom fossils around ~73 Ma (Witkowski, Harwood, et al. 2011). In our tree we only include one *Rhizosolenia* species (*R. setigera*) that is known to produce HBI. However, *R. fallax* is also known to produce it (Damsté et al. 2004), and in our preliminary trees it is more closely related to the species *R. imbricata* which we include in our final tree. Thus, HBI production occurred before the common ancestor of *R. setigera* and *R. imbricata* (and *Guinardia striata*, another rhizosolenid diatom that does not produce HBI). The node before their divergence is defined by the split between Rhizosoleniales and Coscinodiscophycidae.
Age justification: The earliest date is at 91.5 ± 1.5, thus we make use of a minimum age of 90.0 Ma (Damsté et al, 2004).

Node: *Eunotogramma – Thalassiosira*

Fossil: *Eunotogramma fueloepi* (Hajós 1975)
Minimum age: 65.9 Ma
Phylogenetic justification: *Eunotogramma fueloepi* was described by Hajós from the deep sea drilling project (1975). We include one representative of the genus *Eunotogramma* (*Eunotogramma* sp. MG755797) in our study, which is sister to *Thalassiosira weissflogii*.
Age justification: The fossil is from the deep sea drilling project (DSDP Leg 29, Sample 275-2-1, 130-132 cm) and is considered part of the Upper Cretaceous (Hajós 1975; Sims et al. 2006). The border between the Cretaceous and Paleogene is dated at 66.0 ± 0.05 (Gradstein 2012). We make use of a minimum age of 65.9 Ma.

Node: Hemiaulales – Chaetocerotales

Fossil: *Hemiaulus* sp. from thai amber sample KM01 (Girard et al, 2020)
Minimum age: 145 Ma
Phylogenetic justification: The fossil from sample KM01 is considered of the genus *Hemiaulus*, with no further specification possible for the exact species (Girard et al, 2020). Classification is based on the presence of characters such as "a bipolar frustule, an elliptical valve with long polar elevations" (Girard et al, 2020). Our analysis includes two representatives of the order Hemiaulales (*Cerataulina daemon* and *Eucampia zodiacus*), which is recovered as sister to Chaetocerotales.
Age justification: Estimates based on thermogravimetric analyses suggest an age range between 104 to 165 Ma. Based on literature of the Khlong Min Formation, the layer is likely of Jurassic age (Girard et al. 2020), which overlaps with the estimates but suggests an older minimum age (145 Ma). We follow Girard et al (2020) and use a minimum age of 145 Ma. The oldest often accepted fossil is from the Early Jurassic (Rothpletz, 1896). However, the finding has been difficult to reproduce, the age of the layer is uncertain and the original material is lost (Falkowski et al, 2004; Girard et al, 2020). *Hemiaulus* sp. is a bipolar diatom, whereas the first diatoms are assumed to be radial centrics based on phylogenetics (Kooistra et al. 2007; Medlin 2016a). However, *Hemiaulus* sp. is the oldest bona fide diatom fossil.

Node: Urneidophycidae

Fossil: *Incisoria lanceolata*, *Sceptroneis dimorpha* and *S. gracilis* (Witkowski, Harwood, et al. 2011)

Minimum age: 73 Ma
Phylogenetic justification: Here, we refer to the pennate diatom fossils from the Kanguk Formation described by Witkowski, Hardwood et al (2011). Three araphid pennate species were described: *Incisoria lanceolata*, *Sceptroneis dimorpha* and *S. gracilis*. *Sceptronis* is assigned to the Rhaphoneidales, which is part of the subclass Urneidophycidae (Medlin 2016b). We include one representative of Rhaphoneidales (*Asterionellopsis glacialis*), which is the first to branch off at the Urneidophycidae node, which we calibrate with these fossils.
Age justification: The oldest pennate diatoms are found around ~75 Ma (Sims et al. 2006). The oldest certain fossils we refer to here are dated based on the presence of palynomorphs (Witkowski, Harwood, et al. 2011; Witkowski, Sims, et al. 2011). The palynomorph *Translucentipolis plicatilis* is present at the upper layer of the Eidsbotn graben, and suggests a time interval of 76-73 Ma, which we use here (Witkowski, Harwood, et al. 2011). The decision to consider this as the oldest pennate fossil follows Witkowski, Sims et al (2011).

Node: Bacillariophycidae - Fragilariophycidae

Fossil: *Lyrella* spp.
Originally described by Witt (1886) and Pantocsek (1889), but we base this calibration on Sims et al. (2006).
Minimum age: 56 Ma
Phylogenetic justification: *Lyrella* spp. fossils are pennate fossils with a raphe, which is a synapomorphy of the raphid pennate diatoms (Bacillariophycidae). Three representatives are considered here, *L. praetexta*, *L. hennedyi* and *L. simbirskiana*, which were originally described by Witt (1886) and Pantocsek (1889). They were initially described in the genus *Navicula* (Sims et al. 2006). Although *Lyrella* is not considered an early-branching lineage (Jones et al. 2005), we have no representatives of Lyrellales in this study, thus we cannot calibrate a more recent node. We find representatives of Fragilariophycidae as sister to Bacillariophycidae.
Age justification: The three *Lyrella* fossils are from sediments of the Paleocene area (Ross and Sims 1985; Jones et al. 2005; Sims et al. 2006). Here, we use the ages of the Paleocene, which ranges from 66.0 Ma to 56.0 Ma (Gradstein 2012).

Node: *Synura – Mallomonas*

Fossil: *Synura uvella* and *Mallomonas insignis* (Siver and Wolfe 2005)
Minimum age: 46.4 Ma
Phylogenetic justification: Fossils of four chrysophyte genera, including *Mallomonas* and *Synura*, are described from the Giraffe Pipe kimberlite location near Lac de Gras, Northwest Territories, Canada (Siver and Wolfe 2005). As both species are present (moreover, monophyletic with this taxon sampling) we can date the split between these two sister genera.
Age justification: The age of the Giraffe Pipe was estimated using Rb-Sr at 47.8 ± 1.4 Ma (Creaser et al. 2004). We use a minimum age of 46.4 Ma.

Node: Dasycladales – Oltmannsiellopsidales

Fossil: *Palaeocymopolia silurica* (Mastik and Tinn 2015).
Minimum age: 438.3 Ma

Phylogenetic justification: The fossil *Palaeocymopolia silurica* is considered to be part of the order Dasycladales, as it is comparable to *Cympolia* species (Mastik and Tinn 2015). See the supplementary of Morris et al (2018a) for details. Morris et al (2018a) consider the fossil to belong to the Chlorophyceae (as the sister lineage to Prasinophyaceae). The terminology refers to the Chlorophytina instead in this study. However, we include representatives of the Dasycladales, thus we can use this for the split between Dasycladales and Oltmannsiellopsidales. This placement of Dasycladales is similar to a previous study (Gulbrandsen et al. 2021)
Age justification: The fossil is from the Kalana Lagerstätte (Mastik and Tinn 2015), within the *Pribylogratus leptotheca* graptolite Biozone (Männik et al. 2016). The upper border of this biozone is dated at 439.21 ± 0.9031 (Morris et al. 2018a). We made use of a minimum age of 438.3 Ma.

Node: *Caulerpa – Halimeda*, *Pseudocodium*, *Rhipilia*

Fossil: *Caulerpa* sp. (Gustavson and Delevoryas 1992)
Minimum age: 271.8 Ma
Phylogenetic justification: The fossil from the Permian Basin (West Texas, USA) is characterized by "numerous short-stalked bulbous branchlets that are attached to the main axis in a helical pattern", comparable to *Caulerpa racemosa* var. *clavifera* (Gustavson and Delevoryas 1992). In our study, *Caulerpa okamurae* is the only representative of *Caulerpa* (and of Caulerpaceae), and is sister to a clade consisting of *Halimeda*, *Pseudocodium* and *Rhipilia*. We thus use this fossil as a minimum calibration for the node defining the divergence between *Caulerpa* and the three sister species.
Age justification: The sediment the fossil was found in is considered of the Wolfcampian (Gustavson and Delevoryas, 1992), which corresponds to the early Cisuralian (Gradstein 2012). However, it is unclear what the border is between the Wolfcampian and the Leonardian (Gradstein 2012). Verburggen et al used an age of 280 Ma (Verbruggen et al. 2009). To be conservative, we here use a similar but younger age corresponding to the border between the Cisuralian and the Guadalupian at 272.3 ± 0.5 Ma (Gradstein 2012), giving a minimum age of 271.8 Ma.

Node: *Halimeda – Pseudocodium*, *Rhipilia*

Fossils: *Halimeda* fossils of Norian age (Schlagintweit and Pavlik 2008)
Minimum age: 209.5 Ma
Phylogenetic justification: Multiple fossils are considered here, specifically those described and mentioned by Schlagintweit & Pavlik (2008). Fossils of *Bouein*, later merged with *Halimeda* are included, such as *B. marondei* and *B. hochstetteri* (Flügel 1988). The oldest fossil attributed to *Halimeda* is *Halimeda soltanensis*, which in the literature is considered “a 'phylloid alga', not a true *Halimeda*” (Vachard et al. 2001) and "a possible case of homeomorphism" (Bucur 1994). The node calibrated is defined by the divergence of *Halimeda minima* and sister genera *Pseudocodium* and *Rhipilia*.
Age justification: Several *Halimeda* fossils are of Norian age, including *Boueina marondei* (now *Halimeda*; Flügel 1988) and *Halimeda* species from Dachstein Limestone (Northern Calcerous Alps; Schlagintweit & Pavlik, 2008). The border between Rhaetian and Norian is dated at 209.5 Ma (Gradstein et al, 2012), which is the age we use. This is similar to the usage by Verbruggen et al (2009), who used an age of 203.6 Ma.

Node: Embryophyta – Zygnematophyceae

Fossil: *Tetrahedraletes* cf. *medinensis* (Rubinstein et al. 2010)
Minimum age: 466.2 Ma
Phylogenetic justification: See the supplementary of Morris et al (2018a) for details. As the fossil is of Embryophyte affinity but not to any crown group, we use it to calibrate the node defined by the divergence of Embryophyta and the sister streptophyte algal group Zygnematophyceae.
Age justification: The fossil is from the Zanjón Formation, and is considered of the Dapingian age, which ranges from 470.0 ± 1.4 to 467.3 ± 1.1 Ma (Gradstein et al, 2012). We thus use a minimum age of 466.2 Ma.

Node: Embryophyta

Fossil: *Cooksonia* cf. *pertoni* (Edwards and Feehan 1980; Edwards et al. 1983)
Minimum age: 426.9 Ma
Phylogenetic justification: See the supplementary of Morris et al (2018a) for details. *Cooksonia* is considered a member of the Tracheophyta. The node where Tracheophyta diverge from their sisters (Bryophyta) is the last common ancestor of extant Embryophyta members.
Age justification: The fossil is of the Holmerian Stage, specifically in the *Colongraptus ludensis* zone (Morris et al, 2018). The range of this zone is from 427.92 ± 0.4949 Ma to 427.36 ± 0.5306 Ma (Gradstein et al, 2012), thus we use a minimum age of 426.9 Ma.

Node: Jungermanniopsida - Marchantiopsida

Fossil: *Riccardiothallus devonicus* (Guo et al. 2012)
Minimum age: 405.0 Ma
Phylogenetic justification: Considered affiliated to Jungermanniopsida, see the supplementary of Morris et al (2018a) for details. The sister group of Jungermanniopsida is Marchantiopsida, the divergence between these two groups can thus be minimally calibrated with this fossil.
Age justification: The fossil was collected from the Posongchong formation, which is considered to be of Pragian in age (Morris et al. 2018a). The top of the Pragian is dated at 407.6 ± 2.6, thus we use a minimum age of 405.0 Ma (Gradstein 2012; Morris et al. 2018a).

Node: Marchantiales - Lunulariales

Fossil: *Marchantites cyathodoides* (Anderson 1976)
Minimum age: 226.4 Ma
Phylogenetic justification: *Marchantites* is considered a representative of Marchantiales, see the supplementary of Morris et al (2018a) for details. We find Lunulariales as the sister group, so we use this fossil to calibrate the node defined by the divergence between Lunulariales and Marchantiales.
Age justification: *Marchantites cyathodoides* is from the Molteno Formation, which is considered to be from the Carnian Period. Morris et al (2018a) uses the age of 227 Ma for the top of the Carnian. As we follow Gradstein et al (2012) for most dates in this study, we use the age 228.4 ± 2 Ma, which gives a minimum age of 226.4 Ma.

Node: Bryophyta

Fossil: Fossilized moss remains assigned to Sphagnales (Hübers and Kerp 2012)
Minimum age: 330.7 Ma
Phylogenetic justification: See the supplementary of Morris et al (2018a) for details. In our analysis, we include a member of *Eosphagnum*, which are also considered members of Sphagnopsida. We recover Sphagnopsida as sister of all other Bryophyta, thus the calibrated node is defined by the divergence between Sphagnopsida and all other members of Bryophyta.
Age justification: Sphagnales fossils were collected from the Ortelsdorf formation, which is considered of the Visean age (Hübers and Kerp 2012; Morris et al. 2018a). The minimum age would be the age of the top of the Visean, which is dated at 330.9 ± 0.2 Ma (Gradstein et al, 2012; Morris et al, 2018a). Thus we use a minimum age of 330.7 Ma.

Node: Lycopodiophytina - Polypodiophytina, Spermatophytina

Fossil: *Zosterophyllum* sp. (Kotyk et al. 2002)
Minimum age: 420.7 Ma
Phylogenetic justification: *Zosterophyllum* is considered part of Lycopodiophytina. See the supplementary of Morris et al (2018a) and Clarke et al (2011) for details. The sister group of Lycopodiophytina are euphyllophytes, i.e. Polypodiophytina and Spermatophytina.
Age justification: Following Medlin et al (2018), the calibration is based on the co-occurence of *Ozarkodina douroensis*, which is restricted to the mid-late Luwlow Epoch. As the Ludlow Epoch is short and mid-late does not refer to a clear distinction, we make use of the whole age range of the Luwlow Epoch, which ranges from 427.4 ± 0.5 to 423.0 ± 2.3 Ma (Gradstein et al, 2012). We thus use a minimum age of 420.7 Ma.

Node: Isoetales, Selaginellales – Lycopodiales

Fossil: *Leclercqia complexa* (Gensel and Albright 2006)
Minimum age: 392.1 Ma
Phylogenetic justification: See the supplementary of Morris et al (2018a) for details. Morris et al (2018a) refer to Iseotopsida, which is equivalent to Isoetales plus Selaginellales. *Leclerqia complexa* “had a ligule on adaxial side of lycophylls. This structure is common in extant Isoetales, Selaginellales or in fossil Lepidodendrales” (Hrabovskỳ 2020). We calibrate the node defining the divergence between these two groups and their sister Lycopodiales.
Age justification: The fossil is considered of a mid Emsian age (Morris et al, 2018), which is considered to range from 407.6 ± 2.6 to 393.3 ± 1.2 Ma (Gradstein et al, 2012). The minimum age used is 392.1 Ma.

Node: Spermatophytina - Polypodiophytina

Fossil: *Rellimia thomsonii* (Bonamo 1977)
Minimum age: 385.5 Ma
Phylogenetic justification: *Rellimia thomsonii* is considered the oldest fossil of stem Spermatophytina (progymnosperm) (Clarke et al. 2011), see the supplementary of Morris et al (2018a) for details. Morris et al (2018a) uses the name Spermatophyta, which is equivalent to subdivision Spermatophytina as used by the ITIS database that we follow here.

We use this fossil to minimally constrain the node defined by the divergence between Spermatophytina and their sister lineage Polypodiophytina.
Age justification: See Morris et al (2018a). Collected from the Panther Mountain Formation, the top of which is considered to be of age between the rhenanus-ansatus border at 386.25 ± 0.679 (Morris et al, 2018; Gradstein et al, 2012).

Node: Ophioglossidae, Marattiidae - Polypodiidae

Fossil: *Psaronius simplicicaulis* (DiMichele and Phillips 1977)
Minimum age: 318.7 Ma
Phylogenetic justification: This fossil is considered a marattialean tree fern (DiMichele and Phillips 1977; Morris et al. 2018a). In Morris et al (2018a), the fossil is used for the node Psilotopsida + Marattiopsida splitting from Polypodiopsida. Marattiopsida is equivalent to Marattiidae in the ITIS database. The groups Ophioglossidae and Polypodiidae are synonyms of Psilotopsida and Polypodiopsida respectively, and are sister to Marattiidae.
Age justification: *Psaronius simplicicaulis* is from the Caseyville Formation (DiMichele and Phillips 1977), which is considered of Marrowan age (Morris et al, 2018; Gradstein et al, 2012). We use the same minimal age of 319.09 ± 0.38 Ma, based on *Declinograptus margindosus* Biozone and the *Idiognathodus sinuosus* Biozone. The used minimum age is 318.7 Ma.

Node: Pinopsida, Gnetopsida – Cycadopsida, Ginkgoopsida

Fossil: *Cordaixylon iowensis* (Trivett 2015)
Minimum age: 308.1 Ma
Phylogenetic justification: *Cordaixylon* is the earliest fossil placed in Pinopsida (Clarke et al. 2011). See the supplementary of Morris et al (2018a) for details. We use this fossil to calibrate the node between Pinopsida (Pines), and their sister group Gnetopsida (represented by *Welwitschia* and *Ephedra*).
Age justification: See Morris et al (2018a). From the Laddsdale Coal, equivalent to the Bluejacket Coal. We use their ages for the pc10 bracket, which ranges from 312.01 ± 0.37 Ma to 308.5 ± 0.36 Ma. We use a minimum age of 308.1 Ma.

Node: Eudicots – Monocots

Fossil: "Tricolpate pollen" (Hughes and McDougall 1990)
Minimum age: 125.9 Ma
Phylogenetic justification: See the supplementary of Morris et al (2018a) for details. In our taxonomic system, there is no name for Eudicots, but it is represented by Rosanae, Asteranae and Ranunculaceae (i.e. *Corydalis*, *Apium*, *Vicia*). Two monocots are resolved as their sister in this study: *Lilium* and *Acorus*. We find eudicots as sister to monocots, to the exclusion of magnoliids. This allows us to use this calibration between eudicots and monocots, similar to figure 2D from Clarke et al (2011).
Age justification: See Morris et al (2018a) for details. Barremian ranges from 130.8 ± 0.5 Ma to 126.3 ± 0.4 Ma. We use a minimum age of 125.9 Ma.

Node: *Calycanthus – Chimonanthus*

Fossil: *Jerseyanthus calycanthoides* (Crepet et al. 2005)

Minimum age: 85.8 Ma
Phylogenetic justification: This fossil represents the Calcycanthoideae (Massoni et al. 2015). This lineage is represented by *Calycanthus chinensis*, and the fossil calibrates the node defined by the divergence between *C. chinensis* and *Chimonanthus grammatus*.
Age justification: The minimum age of the fossil is 85.8 Ma (Massoni et al. 2015).

Node: Annonaceae, Magnoliaceae – *Myristica*

Fossils: *Endressinia brasiliana* and *Schenkeriphyllum glanduliferum* (Doyle and Endress 2010; Mohr et al. 2013)
Minimum age: 112.6 Ma
Phylogenetic justification: Fossils considered as crown group Magnoliineae, following Massoni et al (2015). We include the families Annonaceae (*Annona reticulata*) and Magnoliaceae (*Magnolia virginiana*) representing Magnoliineae, and the fossil calibrates their divergence from *Myristica yunnanensis*.
Age justification: The minimum age of both fossils is 112.6 Ma (Massoni et al. 2015).

Node: Canellales – Piperales

Fossil: *Walkeripolis gabonensis* (Doyle and Endress 2010)
Minimum age: 125.9 Ma
Phylogenetic justification: This fossil represents the crown group Canellales, specifically belonging to the Winteraceae (Massoni et al. 2015). We include no Winteraceae representatives in this study, so instead we use this to calibrate the node between Canellales and its sister group Piperales.
Age justification: The minimum age of the fossil is 125.9 Ma (Massoni et al. 2015).

Node: Ranunculales – Fabales, Apiales

Fossil: *Hyrcantha decussata* (Dilcher et al. 2007)
Minimum age: 119.6 Ma
Phylogenetic justification: Following Morris et al (2018a), this fossil is a stem-group representative of the Ranunculales. Ranunculales (represented by *Corydalis chrysosphaera*) is sister to the remaining eudicots included in this study (*Vicia faba*, *Apium graveolens*).
Age justification: The minimum age of both fossils is 119.6 Ma, based on an estimated age of 120.3 ± 0.7 of the Jiufotang Formation which overlies the Yixian Formation (Morris et al, 2018).

Default maximum calibrations

Root

Calibration source: Putative maximum age for the origin of cyanobacteria, allowing for the possibility that plastids originated before the great oxidation event (GOE).
Maximum age: 3000 Ma
Justification: The great oxidation event (~2400 Ma) is an undisputed mark for the existence of cyanobacteria. However, the actual origin of cyanobacteria may have been earlier. A potential first evidence of oxygenic photosynthesis could be isotopic evidence in the Fig Tree

Group at roughly 3200 Ma (Satkoski et al. 2015), although it is possible that oxygenic photosynthesis originated in cyanobacteria (Sánchez-Baracaldo and Cardona 2020). For these reasons, we cannot exclude that plastids evolved before the great oxidation event. Thus, we use a conservative maximum age of 3000 Ma.

Corallinophycidae

Calibration source: *Thallophyca* and *Paramecia* (Xiao et al. 2004) and absence of calcified coralline algal fossils (Zhang et al. 1998)
Maximum age: 635.8 Ma
Phylogenetic justification: The fossils *Thallophyca* and *Paramecia* are considered stem Coralliniophycidae (Xiao et al. 2004). Importantly, they lack the calcification present in all extent members of this group (Xiao et al. 2004; Peña et al. 2020). Calcified species should fossilize more easily, yet the oldest bona fide fossils of Sporolithales are found around 130 Ma (Tomás et al. 2007). A very conservative maximum could thus be the oldest possible age of the Caoralliniophycidae.
Age justification: The lower Doushantuo formation is estimated at 635.2 ± 0.6 Ma (Condon et al. 2005), so we make use of a maximum age of 635.8 Ma.

Dasycladales

Calibration source: Macroalgae fossils of uppermost Doushantuo Formation lacking Dasycladales-like (and Bryopsidales-like) algae (Zhang et al. 1998)
Maximum age: 635.8 Ma
Phylogenetic justification: The Ediacaran Konservat-Lagerstätten contain no fossils of Dasycladales or Bryopsidales (unicellular siphonous algae), but macroalgae are abundant (Zhang et al. 1998; Xiao et al. 2002; Verbruggen et al. 2009).
Age justification: The lower Doushantuo formation is estimated at 635.2 ± 0.6 Ma (Condon et al. 2005), so we make use of a maximum age of 635.8 Ma.

Bryopsidales

Calibration source: Macroalgae fossils of uppermost Doushantuo Formation lacking Bryopsidales-like (and Dasycladales-like) algae.
Maximum age: 635.8 Ma
Phylogenetic justification: See above (Dasycladales)
Age justification: See above (Dasycladales)

Embryophyta

Calibration source: “Precambrian sediments of the Torridon Group, Scotland, which represent an environment in which embryophytes would be expected to have flourished were their lineage established.” (Clarke et al. 2011).
Maximum age: 1042 Ma
Phylogenetic justification: The earliest traces of embryophytes come from cryptospores. The oldest cryptospores not from marine environments are much younger than this age, but the use of this calibration has been controversial (Hedges et al. 2018; Morris et al. 2018b). As the cryptospore-based maximum calibration also leaves little space for the molecular

data and models to estimate the age (Su et al. 2021), we prefer this age based on Precambrian sediments as our default calibration.
Age justification: See Clarke et al. (2011).

Angiosperms

Calibration source: Sediments below the oldest occurrence of angiosperm-like pollen. (Hughes and McDougall 1990)
Maximum age: 247.2 Ma
Phylogenetic justification: Although ages based on absence of fossilized matter are not optimal, this age is not much contested in contrast to other maximum calibrations of deeper land plant nodes (Hedges et al. 2018).
Age justification: See Morris et al (2018a).

Extra maximum calibrations

Coccolithales – Isochrysidales

Calibration source: Absence of coccolith fossils in the Permian period (De vargas et al. 2007)
Maximum age: 252.2 Ma
Justification: The oldest coccolith fossils are from the Triassic, with no Permian fossils known. It is possible that the first coccoliths did not fossilize well (Henderiks et al. 2022), which is why we do not use this calibration by default. Using the age of the border between Triassic and Permian, we allow for the possibility that no coccolith forming algae were present until the Triassic.

Bacillariophyta

Calibration source: Absence of diatom fossils in the Permian period (Medlin 2015)
Maximum age: 252.2 Ma
Justification: The oldest diatom fossils are from the Triassic, although they remain controversial. Using the age of the border between Triassic and Permian, we allow for the possibility that no diatoms were present until the Triassic.

Bacillariophyceae

Calibration source: The maximum age of *Hemiaulus* sp. (Girard et al. 2020), the first bona fide fossil of diatoms, ~100 million years before the first pennate fossils (Bacillariophyceae).
Maximum age: 165.0 Ma
Justification: The earliest bona fide diatoms are found at 165-145 Ma (Girard et al. 2020), but the first pennate diatom is found at only 76-73 Ma (Witkowski, Harwood, et al. 2011). As centric diatoms appear earlier in the fossil record, and pennate diatoms are clearly sister to a subset of centric diatoms (Medlin 2016a), we can potentially use this maximum to constrain the earliest possible origin of pennate diatoms.

Bacillariophycidae

Calibration source: The maximum age of the earliest bona fide pennate diatom fossils (Witkowski, Harwood, et al. 2011), which appear before the raphid pennate diatoms.
Maximum age: 76.00
Justification: The earliest fossils of pennate diatoms have no raphe, and raphid diatoms are sister to only a subset of pennate diatoms (Medlin 2016a). The oldest possible age of the earliest certain pennate fossils can provide a good soft maximum calibration of the raphid pennate diatoms (Bacillariophycidae).

Embryophyta

Calibration source: Oldest (non-marine) cryptospores (Clarke et al. 2011; Morris et al. 2018a)
Maximum age: 515.5 Ma
Justification: See Morris et al (2018a) for details. The layer below the earliest potential cryptospores may constitute a good maximum of Embryophyta, although note that early spores may not have fossilized well or have simply not yet been found. By default we make use of a different maximum constraint (of 1042 Ma, see Clarke et al. 2011), but this is an interesting alternative maximum to test.

Spermatophytina

Calibration source: First seeds (Clarke et al. 2011; Morris et al. 2018a)
Maximum age: 365.6 Ma
Justification: See Morris et al (2018a) for details. Based on the first certain records of seeds.

Pinopsida – Gnetopsida

Calibration source: Sediments below *Cordaixylon iowensis* bearing strata.
Maximum age: 321.4 Ma
Justification: See Morris et al (2018a) for details. Age based on geological layers below *Cordaixylon* without any Pinopsida-like features.

Bangiomorpha

Cyanidiophytina; Eurhodophytina + Proteorhodophytina

Fossil: *Bangiomorpha pubescens* (Butterfield et al. 1990; Butterfield 2000)
Minimum age: 1030 Ma
Phylogenetic justification: Although considered more specifically as a Bangiophyceae species, the fossil is typically used to represent a stem group of Eurhodophytina or perhaps similar to certain Proteorhodophytina. Nevertheless, its old age and simple characteristics may suggest a stem-group eukaryotic affiliation or even a prokaryotic one (Cavalier-Smith 2006). The radial symmetrical arrangement of the cells within a filament is similar to modern Bangiophytes, but (1) this may be relatively easy to evolve for cyanobacteria and (2) other characteristic red algal features, such as pit plugs, are missing. Older fossils (*Rafatazmia* and *Ramathallus*) may be putatively red algae, but the presence of chloroplasts and pit plugs

is controversial (Bengtson et al. 2017; Carlisle et al. 2021). As it is possible that *Bangiomorpha* does represent an ancestor of Eurhodophytina and Proteorhodophytina, it can be used to calibrate the node defined by the divergence of Cyanidiophytina and the other red algae.
Age justification: Previous dating measurements suggested an age of ~1.2 Ga, based on Pb-Pb geochronology, but a more recent study using Re-Os isotopic ages suggests an age of 1047 +0.013/-0.017 Ma (Gibson et al. 2018).

Proterocladus

Ulvophyceae (incl. Chlorarachniophyceae) – Trebouxiophyceae

Fossil: *Proterocladus antiquus* (Tang et al. 2020)
Minimum age: 940.4 Ma
Phylogenetic justification: The fossil species is suggested to be multicellular, siphonous alga and while siphonous growth is known from other algal groups, Tang et al (2020) suggest the best attribution is to the green algae. Multicellular siphonous green algae are common in the Ulvophyceae, and this fossil may thus be attributed to a (stem) lineage of Ulvophyceae (Tang et al. 2020). This fossil could then be used to minimally calibrate the node defined by the split between Ulvophyceae and Trebouxiophyceae. Note that we include in our analyses green complex plastids from chlorarachniophytes, the plastids of which are of ulvophyte origin (Jackson et al. 2018).
Age justification: The age of *Proteocladus* is estimated at 947.8 ± 7.4 Ma, thus we use an age of 940.4 Ma (Tang et al. 2020).

Supplementary figures

Fig. S1. Phylogenetic reconstruction of Archaeplastida and multiple lineages with red and green complex plastids based on plastid genome data (‘all plastid’ data set). The tree was reconstructed with a trimmed alignment of 120 plastid-encoded protein sequences (22,209 amino acid positions), using both Bayesian inference (CAT+GTR+G4) and maximum likelihood (cpREV+C60+F+R9 with ultrafast bootstrap support and cpREV+C60+R9+PMSF with nonparametric bootstrap support). Support values are shown on the branches as posterior probability (0.0 - 1.0), ultrafast bootstrap support (when using model cpREV+C60+R9; 0-100) and standard bootstrap support (model cpREV+C60+R9+PMSF; 0-100).

Fig. S2. Phylogenetic reconstruction of red algae and red complex plastids based on plastid genome data (‘red-only’ data set) using a supertree approach with ASTRAL-III. Tree is based on single-gene trees with weakly supported nodes (<10 ultrafast bootstrap support) collapsed. Support values shown are the posterior probabilities as calculated by ASTRAL.

Fig. S3. Phylogenetic reconstruction of Archaeplastida and multiple lineages with red and green complex plastids based on plastid genome data (‘all plastid’ data set) using a supertree approach with ASTRAL-III. The tree is based on single-gene trees with weakly supported nodes (<10 ultrafast bootstrap support) collapsed. Support values shown are the posterior probabilities as calculated by ASTRAL.

Fig. S4. Phylogenetic reconstruction of red algae and red complex plastids based on plastid genome data (‘red-only’ data set). The tree was reconstructed with maximum likelihood using the model cpREV+C60+F+R8. Support is given as SH-aLRT support (%) / ultrafast bootstrap support (%).

Fig. S5. Phylogenetic reconstruction of red algae and red complex plastids based on plastid genome data (‘red-only’ data set). The tree was reconstructed with maximum likelihood using the model cpREV+C60+F+R8+PMSF. Support given is standard, non-parametric bootstrap support (%).

Fig. S6. Phylogenetic reconstruction of red algae and red complex plastids based on plastid genome data (‘red-only’ data set). The tree was reconstructed with maximum likelihood using the model LG+C60+F+R8. Support is given as SH-aLRT support (%) / ultrafast bootstrap support (%).

Fig. S7. Phylogenetic reconstruction of Archaeplastida and multiple lineages with red and green complex plastids based on plastid genome data (‘all plastid’ data set). The tree was reconstructed with maximum likelihood using the model cpREV+C60+F+R9. Support is given as SH-aLRT support (%) / ultrafast bootstrap support (%).

Fig. S8. Phylogenetic reconstruction of Archaeplastida and multiple lineages with red and green complex plastids based on plastid genome data (‘all plastid’). The tree was

reconstructed with maximum likelihood using the model cpREV+C60+F+R9+PMSF. Support given is standard, non-parametric bootstrap support (%).

Fig. S9. Phylogenetic reconstruction of Archaeplastida and multiple lineages with red and green complex plastids based on plastid genome data ('all plastid' data set). The tree was reconstructed with maximum likelihood using the model LG+C60+F+R9. Support is given as SH-aLRT support (%) / ultrafast bootstrap support (%).

Fig. S10. Phylogenetic reconstruction of red algae and red complex plastids based on plastid genome data ('red-only' data set) after removal of the fast-evolving ESC group (Eustigmatophyceae, Synurophyceae, Chrysophyceae). The tree was reconstructed with maximum likelihood using the model cpREV+C60+F+R8. Support is given as SH-aLRT support (%) / ultrafast bootstrap support (%).

Fig. S11. Principal component analysis showing compositional biases of the 'red-only' data set after filtering out all positions in the alignment where more than three taxa have a gap. Results are similar to those without filtering any positions (see Fig. 3A in main text).

Fig. S12. Phylogenetic tree based on the 'red-only' data set after Dayhoff recoding with four categories. The tree was reconstructed with maximum likelihood using the model GTR+G4. Support is given as SH-aLRT support (%) / ultrafast bootstrap support (%).

Fig. S13. Phylogenetic tree based on the 'red-only' data set after recoding with four categories using minmax-Chisq. The tree was reconstructed with maximum likelihood using the model GTR+G4. Support is given as SH-aLRT support (%) / ultrafast bootstrap support (%).

Fig. S14. Cladogram showing the relationships among Archaeplastida and multiple lineages with red and green complex plastids (same topology as supplementary Fig. S1). Our calibrations for each calibrated node are shown next to it. For each potential calibration, five lines are shown. Grayed out lines mean that the calibration was not used in one of the five data sets, which are listed in the following order: Default, Default+B, Default+BP, Extended, Extended+BP (see main text for details). Ages are given in Ma (million years ago).

Fig. S15. Cladogram showing the relationships among Archaeplastida and multiple lineages with red and green complex plastids (same topology as supplementary Fig. S1). The calibrations from Strassert et al. (2021) that were applicable to our data set are indicated at the relevant nodes. Ages are given in Ma (million years ago).

Fig. S16. Phylogeny of haptophytes based on a protein alignment of RbcL, showing the position of *Chrysotila dentata* (MZ819921.1) within Coccolithales. The phylogeny was reconstructed using IQ-TREE with the model cpREV+F+G4.

Fig. S17. Phylogeny of haptophytes based on a protein alignment of TufA, showing the position of *Chrysotila dentata* (MZ819921.1) within Coccolithales. The phylogeny was reconstructed using IQ-TREE with the model cpREV+F+G4.

Fig. S18. Preliminary phylogeny of plastid-bearing lineages based on a protein alignment of RbcL+PsaA. The phylogeny was reconstructed using IQ-TREE with the model cpREV+F+G4, support given is ultrafast bootstrap (%).

Fig. S19-S21. Phylogenetic reconstruction of red algae and red complex plastids based on plastid genome data (‘red-only’ data set) using Bayesian inference with three independent chains. Posterior probability of each branch is shown (0.0-1.0).

Fig. S22-S24. Phylogenetic reconstruction of Archaeplastida and multiple lineages with red and green complex plastids based on plastid genome data (‘all plastid’ data set) using Bayesian inference with three independent chains. Posterior probability of each branch is shown (0.0-1.0).


Fig. S1
Cyanophora_paradoxa_KM198929.1
Gloeochaete_wittrockiana_MF167426.1
Glaucocystis_incrassata_MF167425.1
Prasinoderma_coloniale_KJ746598.1
Micromonas_sp._RCC299_FJ858267.1
Pyramimonas_parkeae_KX013546.1
Strombomonas_acuminata_JN674637.1
Eutreptiella_gymnastica_NC_017754.2
Nephroselmis_pyriformis_MW077730.1
Chloropicon_sieburthii_MK085986.1
Chlamydomonas_reinhardtii_BK000554.2
Scherffelia_dubia_KU167098.1
Pedinomonas_minor_FJ968740.1
Planctonema_lauterborni_KM462880.1
Gymnochlora_stellata_AP014947.1
Lotharella_sp._CCMP622_KF438023.1
Bigelowiella_natans_DQ851108.1
Ostreobium_quekettii_LT593849.1
Bryopsis_plumosa_LN810504.1
Avrainvillea_mazei_KY509313.1
Caulerpa_okamurae_KX809677.1
Halimeda_minima_MH591101.1
Rhipilia_penicilloides_KY819065.1
Pseudocodium_devriesii_MH591109.1
Ignatius_tetrasporus_KY407659.1
Ulva_compressa_MW548841.1
Trichosarcina_mucosa_KY407656.1
Neodangemannia_microcystis_KY407660.1
Parvocaulis_polyphysoides_XXXK
Neomeris_sp._HV02668_XXXB
Acetabularia_acetabulum_XXXA
Halicoryne_sp._HV04044_XXXD
Bornetella_oligospora_XXXI
Dasycladus_sp_HV04045_XXXC
Batophora_oerstedii_XXXF
Chlorokybus_atmophyticus_DQ422812.2
Entransia_fimbriata_KU646490.1
Chara_vulgaris_DQ229107.1
Chaetosphaeridium_globosum_AF494278.1
Mesotaenium_endlicherianum_KJ461682.1
Cosmarium_botrytis_KU646492.1
Leiosporoceros_dussii_MH577299.1
Anthoceros_agrestis_MK087646.1
Haplomitrium_blumei_MH064516.1
Schistochila_macrodonta_MH064506.1
Ptilidium_pulcherrimum_HM222519.1
Sphaerocarpos_texanus_MT023020.1
Lunularia_cruciata_MW429511.1
Wiesnerella_denudata_MT712073.1
Marchantia_polymorpha_MH635409.1
Sphagnum_wulfianum_KU725459.1
Eosphagnum_sp._BIS-2016_KU725456.1
Andreaea_rothii_MW561627.1
Tetraphis_pellucida_KJ817846.1
Polytrichum_commune_MW528408.1
Buxbaumia_aphylla_MN496310.1
Syntrichia_filaris_MK852705.1
Bartramia_pomiformis_MT024676.1
Lycopodium_clavatum_MH549642.1
Selaginella_lepidophylla_MK089531.1
Isoetes_malinverniana_MH549640.1
Equisetum_arvense_GU191334.1
Tmesipteris_elongata_KJ569699.1
Mankyua_chejuensis_KP205433.1
Hymenophyllum_coreanum_MN583335.1
Danaea_sellowiana_MN412588.1
Christensenia_aesculifolia_MN056350.1
Ginkgo_biloba_JN867585.1
Cycas_debaoensis_KU743927.1
Taxus_baccata_KR476375.1
Sciadopitys_verticillata_KT601210.1
Welwitschia_mirabilis_EU342371.1
Ephedra_distachya_MG594460.1
Amborella_trichopoda_AJ506156.2
Ceratophyllum_demersum_EF614270.1
Lilium_candidum_MK753244.1
Acorus_calamus_AJB79453.1
Corydalis_chrysosphaera_MZ983401.1
Vicia_faba_MT120813.1
Apium_graveolens_MK036045.1
Chloranthus_henryi_MK922064.1
Tasmannia_lanceolata_MT555078.1
Saruma_henryi_MG520100.1
Cinnamomum_aromaticum_MN698963.1
Chimonanthus_grammatus_MW166217.1
Calycanthus_chinensis_MG561304.1
Myristica_yunnanensis_MN495964.1
Magnolia_virginiana_MN990608.1
Annona_reticulata_MT742547.1
Galdieria_sulphuraria_KJ700459.1
Cyanidium_caldarium_AF022186.2
Cyanidioschyzon_merolae_AB002583.1
Porphyridium_purpureum_MF401423.1
Flintiella_sanguinaria_KY709211.1
Erythrolobus_coxiae_MZ853148.1
Madagascaria_erythrocladioides_MZ681963.1
Erythrotrichia_carnea_KX284721.2
Pulvinaster_venetus_MZ681965.1
Boldiaceae_sp._CCMP3255_MZ681962.1
Tsunamia_transpacifica_MZ680366.1
Rhodospora_sordida_MZ603805.1
Bangiopsis_subsimplex_KX284718.1
Rhodella_violacea_MZ853150.1
Corynoplastis_japonica_KY709210.1
Glaucosphaera_vacuolata_MZ895126.1
Bulboplastis_apyrenoidosa_KY709209.1
Porphyra_umbilicalis_MF385003.1
Neopyropia_fucicola_KJ776837.1
Neoporphyra_perforata_KC904971.1
Hildenbrandia_rivularis_KX284723.1
Apophlaea_sinclairii_KX284716.1
Sheathia_arcuata_NC_035231.1
Palmaria_palmata_KX284726.1
Ahnfeltia_plicata_KX284715.1
Asparagopsis_taxiformis_KX284717.1
Gelidium_vagum_KT266787.1
Chondrus_crispus_HF562234.1
Rhodogorgon_sp._MH281630_MH281630.1
Renouxia_sp._MH281629_MH281629.1
Sporolithon_durum_KT266785.1
Mesophyllum_expansum_MEX45
Lithothamnion_sp._MH281627_MH281627.1
Porolithon_onkodes_KY212106.1
Lithophyllum_sp._LIT44_LIT44
Neogoniolithon_spectabile_MH281628.1
Corallina_officinalis_MT211887.1
Storeatula_sp._CCMP1868_KY856940.1
Guillardia_theta_AF041468_AF041468.1
Cryptomonas_curvata_CCAP979_52_LC484192.1
Chroomonas_mesostigmatica_CCMP1168_KY860574.1
Pavlomulina_rannuculiformis_LC564893.1
Phaeocystis_globosa_CNS00067_MT471328.1
Chrysochromulina_parva_MG520331.1
Chrysotila_dentata_MZ819921.1
Isochrysis_galbana_MT304829.1
Emiliania_huxleyi_CCMP_373_AY741371.1
Heterosigma_akashiwo_EU168190.1
Vaucheria_litorea_EU912438.1
Ishige_okamurae_MW762687.1
Egregia_menziesii_MZ156039.1
Akkesiphycus_lubricus_MZ156027.1
Trachydiscus_minutus_KJ624065.1
Monodopsis_sp._MarTras21_KX839260.1
Ochromonas_sp._CCMP1393_KJ877675.1
Neotessella_volvocina_MH795132.1
Mallomonas_splendens_MH795131.1
Synura_uvella_MH795130.1
Synura_petersenii_MH795128.1
Rhizochromulina_marina_MK561360.1
Aureoumbra_lagunensis_GQ231542.1
Triparma_laevis_AP014625.1
Leptocylindrus_danicus_KC509524.1
Proboscia_sp._MG755791_MG755791.1
Coscinodiscus_granii_MW561225.1
Rhizosolenia_setigera_MG755793.1
Rhizosolenia_imbricata_KJ958482.1
Guinardia_striata_MG755796.1
Lithodesmium_undulatum_KC509525.1
Thalassiosira_weissflogii_KJ958485.1
Eunotogramma_sp._MG755797_MG755797.1
Toxarium_undulatum_KX619437.1
Plagiogrammopsis_vanheurckii_MG755794.1
Chaetoceros_tenuissimus_MW845778.1
Chaetoceros_laevisporus_MW845779.1
Eucampia_zodiacus_MW412838.1
Cerataulina_daemon_KJ958484.1
Attheya_longicornis_MG755798.1
Psammoneis_obaidii_MG755803.1
Asterionellopsis_glacialis_KC509520.1
Nanofrustulum_shiloi_MN276191.1
Asterionella_formosa_KC509519.1
Seminavis_robusta_MH356727.1
Phaeodactylum_tricornutum_EF067920.1
Nitzschia_supralitorea_MT383638.1
Trybilonella_apiculata_MT383642.1
Cylindrotheca_closterium_KC509522.1
0.125

Fig. S2

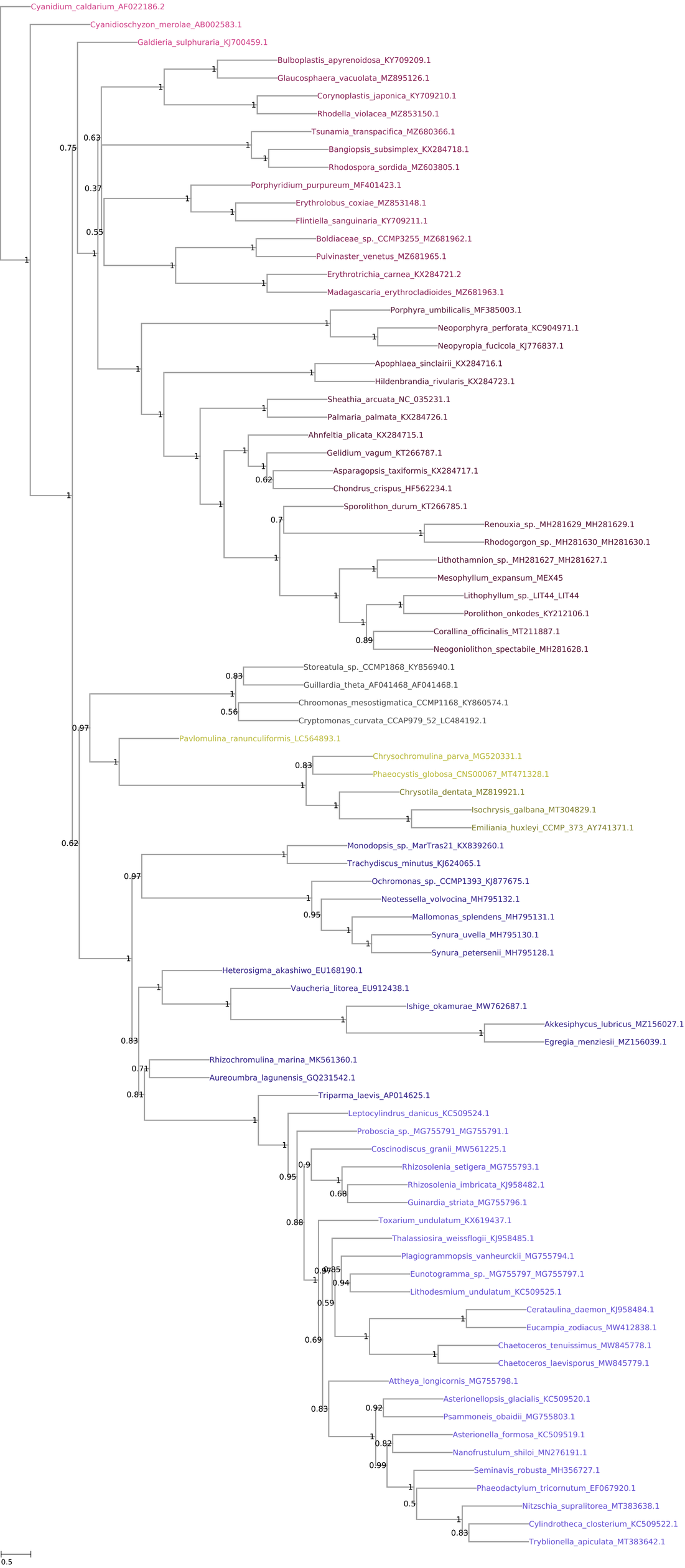

Cyanidium_caldarium_AF022186.2
Cyanidioschyzon_merolae_AB002583.1
Galdieria_sulphuraria_KJ700459.1
Bulboplastis_apyrenoidosa_KY709209.1
Glaucosphaera_vacuolata_MZ895126.1
Corynoplastis_japonica_KY709210.1
Rhodella_violacea_MZ853150.1
Tsunamia_transpacifica_MZ680366.1
Bangiopsis_subsimplex_KX284718.1
Rhodospora_sordida_MZ603805.1
Porphyridium_purpureum_MF401423.1
Erythrolobus_coxiae_MZ853148.1
Flintiella_sanguinaria_KY709211.1
Boldiaceae_sp._CCMP3255_MZ681962.1
Pulvinaster_venetus_MZ681965.1
Erythrotrichia_carnea_KX284721.2
Madagascaria_erythrocladioides_MZ681963.1
Porphyra_umbilicalis_MF385003.1
Neoporphyra_perforata_KC904971.1
Neopyropia_fucicola_KJ776837.1
Apophlaea_sinclairii_KX284716.1
Hildenbrandia_rivularis_KX284723.1
Sheathia_arcuata_NC_035231.1
Palmaria_palmata_KX284726.1
Ahnfeltia_plicata_KX284715.1
Gelidium_vagum_KT266787.1
Asparagopsis_taxiformis_KX284717.1
Chondrus_crispus_HF562234.1
Sporolithon_durum_KT266785.1
Renouxia_sp._MH281629_MH281629.1
Rhodogorgon_sp._MH281630_MH281630.1
Lithothamnion_sp._MH281627_MH281627.1
Mesophyllum_expansum_MEX45
Lithophyllum_sp._LIT44_LIT44
Porolithon_onkodes_KY212106.1
Corallina_officinalis_MT211887.1
Neogoniolithon_spectabile_MH281628.1
Storeatula_sp._CCMP1868_KY856940.1
Guillardia_theta_AF041468_AF041468.1
Chroomonas_mesostigmatica_CCMP1168_KY860574.1
Cryptomonas_curvata_CCAP979_52_LC484192.1
Pavlomulina_ranunculiformis_LC564893.1
Chrysochromulina_parva_MG520331.1
Phaeocystis_globosa_CNS00067_MT471328.1
Chrysotila_dentata_MZ819921.1
Isochrysis_galbana_MT304829.1
Emiliania_huxleyi_CCMP_373_AY741371.1
Monodopsis_sp._MarTras21_KX839260.1
Trachydiscus_minutus_KJ624065.1
Ochromonas_sp._CCMP1393_KJ877675.1
Neotessella_volvocina_MH795132.1
Mallomonas_splendens_MH795131.1
Synura_uvella_MH795130.1
Synura_petersenii_MH795128.1
Heterosigma_akashiwo_EU168190.1
Vaucheria_litorea_EU912438.1
Ishige_okamurae_MW762687.1
Akkesiphycus_lubricus_MZ156027.1
Egregia_menziesii_MZ156039.1
Rhizochromulina_marina_MK561360.1
Aureoumbra_lagunensis_GQ231542.1
Triparma_laevis_AP014625.1
Leptocylindrus_danicus_KC509524.1
Proboscia_sp._MG755791_MG755791.1
Coscinodiscus_granii_MW561225.1
Rhizosolenia_setigera_MG755793.1
Rhizosolenia_imbricata_KJ958482.1
Guinardia_striata_MG755796.1
Toxarium_undulatum_KX619437.1
Thalassiosira_weissflogii_KJ958485.1
Plagiogrammopsis_vanheurckii_MG755794.1
Eunotogramma_sp._MG755797_MG755797.1
Lithodesmium_undulatum_KC509525.1
Cerataulina_daemon_KJ958484.1
Eucampia_zodiacus_MW412838.1
Chaetoceros_tenuissimus_MW845778.1
Chaetoceros_laevisporus_MW845779.1
Attheya_longicornis_MG755798.1
Asterionellopsis_glacialis_KC509520.1
Psammoneis_obaidii_MG755803.1
Asterionella_formosa_KC509519.1
Nanofrustulum_shiloi_MN276191.1
Seminavis_robusta_MH356727.1
Phaeodactylum_tricornutum_EF067920.1
Nitzschia_supralitorea_MT383638.1
Cylindrotheca_closterium_KC509522.1
Tryblionella_apiculata_MT383642.1
0.5

Fig. S3

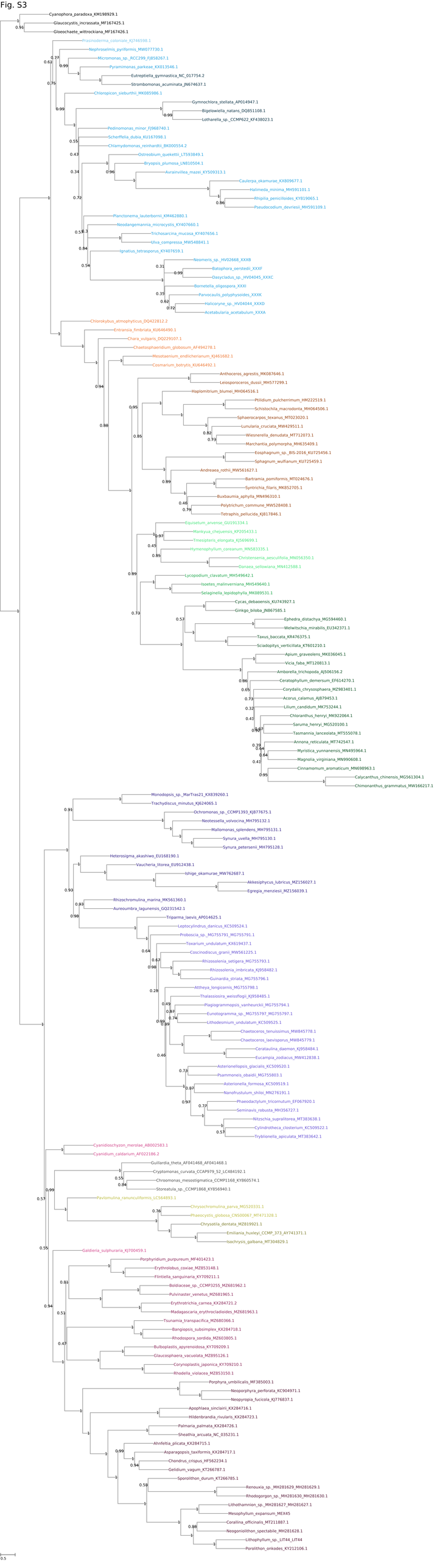

Cyanophora_paradoxa_KM198929.1
Glaucocystis_incrassata_MF167425.1
Gloeochaete_wittrockiana_MF167426.1
Prasinoderma_coloniale_KJ746598.1
Nephroselmis_pyriformis_MW077730.1
Micromonas_sp._RCC299_FJ858267.1
Pyramimonas_parkeae_KX013546.1
Eutreptiella_gymnastica_NC_017754.2
Strombomonas_acuminata_JN674637.1
Chloropicon_sieburthii_MK085986.1
Gymnochlora_stellata_AP014947.1
Bigelowiella_natans_DQ851108.1
Lotharella_sp._CCMP622_KF438023.1
Pedinomonas_minor_FJ968740.1
Scherffelia_dubia_KU167098.1
Chlamydomonas_reinhardtii_BK000554.2
Ostreobium_queekettii_LT593849.1
Bryopsis_plumosa_LN810504.1
Avrainvillea_mazei_KY509313.1
Caulerpa_okamurae_KX809677.1
Halimeda_minima_MH591101.1
Rhipilia_penicillioides_KY819065.1
Pseudocodium_devriesii_MH591109.1
Planctonema_lauterbornii_KM462880.1
Neodangemannia_microcystis_KY407660.1
Trichosarcina_mucosa_KY407656.1
Ulva_compressa_MW548841.1
Ignatius_tetrasporus_KY407659.1
Neomeris_sp._HV02668_XXXB
Batophora_oerstedii_XXXF
Dasycladus_sp._HV04045_XXXC
Bornetella_oligospora_XXXI
Parvocaulis_polyphysoides_XXXK
Halicoryne_sp._HV04044_XXXD
Acetabularia_acetabulum_XXXA
Chlorokybus_atmophyticus_DQ422812.2
Entransia_fimbriata_KU646490.1
Chara_vulgaris_DQ229107.1
Chaetosphaeridium_globosum_AF494278.1
Mesotaenium_endlicherianum_KJ461682.1
Cosmarium_botrytis_KU646492.1
Anthoceros_agrestis_MK087646.1
Leiosporoceros_dussii_MH577299.1
Haplomitrium_blumei_MH064516.1
Ptilidium_pulcherrimum_HM222519.1
Schistochila_macrodonta_MH064506.1
Sphaerocarpos_texanus_MT023020.1
Lunularia_cruciata_MW429511.1
Wiesnerella_denudata_MT712073.1
Marchantia_polymorpha_MH635409.1
Eosphagnum_sp._BIS-2016_KU725456.1
Sphagnum_wulfianum_KU725459.1
Andreaea_rothii_MW561627.1
Bartramia_pomiformis_MT024676.1
Syntrichia_filaris_MK852705.1
Buxbaumia_aphylla_MN496310.1
Polytrichum_commune_MW528408.1
Tetraphis_pellucida_KJ817846.1
Equisetum_arvense_GU191334.1
Mankyua_chejuensis_KP205433.1
Tmesipteris_elongata_KJ569699.1
Hymenophyllum_coreanum_MN583335.1
Christensenia_aesculifolia_MN056350.1
Danaea_sellowiana_MN412588.1
Lycopodium_clavatum_MH549642.1
Isoetes_malinverniana_MH549640.1
Selaginella_lepidophylla_MK089531.1
Cycas_debaoensis_KU743927.1
Ginkgo_biloba_JN867585.1
Ephedra_distachya_MG594460.1
Welwitschia_mirabilis_EU342371.1
Taxus_baccata_KR476375.1
Sciadopitys_verticillata_KT601210.1
Apium_graveolens_MK036045.1
Vicia_faba_MT120813.1
Amborella_trichopoda_AJ506156.2
Ceratophyllum_demersum_EF614270.1
Corydalis_chrysosphaera_MZ983401.1
Acorus_calamus_AJ879453.1
Lilium_candidum_MK753244.1
Chloranthus_henryi_MK922064.1
Saruma_henryi_MG520100.1
Tasmannia_lanceolata_MT555078.1
Annona_reticulata_MT742547.1
Myristica_yunnanensis_MN495964.1
Magnolia_virginiana_MN990608.1
Cinnamomum_aromaticum_MN698963.1
Calycanthus_chinensis_MG561304.1
Chimonanthus_grammatus_MW166217.1
Monodopsis_sp._MarTras21_KX839260.1
Trachydiscus_minutus_KJ624065.1
Ochromonas_sp._CCMP1393_KJB77675.1
Neotessella_volvocina_MH795132.1
Mallomonas_splendens_MH795131.1
Synura_uvella_MH795130.1
Synura_petersenii_MH795128.1
Heterosigma_akashiwo_EU168190.1
Vaucheria_litorea_EU912438.1
Ishige_okamurae_MW762687.1
Akkesiphycus_lubricus_MZ156027.1
Egregia_menziesii_MZ156039.1
Rhizochromulina_marina_MK561360.1
Aureoumbra_lagunensis_GQ231542.1
Triparma_laevis_AP014625.1
Leptocylindrus_danicus_KC509524.1
Proboscia_sp._MG755791_MG755791.1
Toxarium_undulatum_KX619437.1
Coscinodiscus_granii_MW561225.1
Rhizosolenia_setigera_MG755793.1
Rhizosolenia_imbricata_KJ958482.1
Guinardia_striata_MG755796.1
Attheya_longicornis_MG755798.1
Thalassiosira_weissflogii_KJ958485.1
Plagiogrammopsis_vanheurckii_MG755794.1
Eunotogramma_sp._MG755797_MG755797.1
Lithodesmium_undulatum_KC509525.1
Chaetoceros_tenuissimus_MW845778.1
Chaetoceros_laevisporus_MW845779.1
Cerataulina_daemon_KJ958484.1
Eucampia_zodiacus_MW412838.1
Asterionellopsis_glacialis_KC509520.1
Psammoneis_obaidii_MG755803.1
Asterionella_formosa_KC509519.1
Nanofrustulum_shiloi_MN276191.1
Phaeodactylum_tricornutum_EF067920.1
Seminavis_robusta_MH356727.1
Nitzschia_supralitorea_MT383638.1
Cylindrotheca_closterium_KC509522.1
Tryblionella_apiculata_MT383642.1
Cyanidioschyzon_merolae_AB002583.1
Cyanidium_caldarium_AF022186.2
Guillardia_theta_AF041468_AF041468.1
Cryptomonas_curvata_CCAP979_52_LC484192.1
Chroomonas_mesostigmatica_CCMP1168_KY860574.1
Storeatula_sp._CCMP1868_KY856940.1
Pavlomulina_ranunculiformis_LC564893.1
Chrysochromulina_parva_MG520331.1
Phaeocystis_globosa_CNS00067_MT471328.1
Chrysotila_dentata_MZ819921.1
Emiliania_huxleyi_CCMP_373_AY741371.1
Isochrysis_galbana_MT304829.1
Galdieria_sulphuraria_KJ700459.1
Porphyridium_purpureum_MF401423.1
Erythrolobus_coxiae_MZ853148.1
Flintiella_sanguinaria_KY709211.1
Boldiaceae_sp._CCMP3255_MZ681962.1
Pulvinaster_venetus_MZ681965.1
Erythrotrichia_carnea_KX284721.2
Madagascaria_erythrocladioides_MZ681963.1
Tsunamia_transpacifica_MZ680366.1
Bangiopsis_subsimplex_KX284718.1
Rhodospora_sordida_MZ603805.1
Bulboplastis_apyrenoidosa_KY709209.1
Glaucosphaera_vacuolata_MZ895126.1
Corynoplastis_japonica_KY709210.1
Rhodella_violacea_MZ853150.1
Porphyra_umbilicalis_MF385003.1
Neoporphyra_perforata_KC904971.1
Neopyropia_fucicola_KJ776837.1
Apophlaea_sinclairii_KX284716.1
Hildenbrandia_rivularis_KX284723.1
Palmaria_palmata_KX284726.1
Sheathia_arcuata_NC_035231.1
Ahnfeltia_plicata_KX284715.1
Asparagopsis_taxiformis_KX284717.1
Chondrus_crispus_HF562234.1
Gelidium_vagum_KT266787.1
Sporolithon_durum_KT266785.1
Renouxia_sp._MH281629_MH281629.1
Rhodogorgon_sp._MH281630_MH281630.1
Lithothamnion_sp._MH281627_MH281627.1
Mesophyllum_expansum_MEX45
Corallina_officinalis_MT211887.1
Neogoniolithon_spectabile_MH281628.1
Lithophyllum_sp._LIT44_LIT44
Porolithon_onkodes_KY212106.1
0.5

Fig. S4

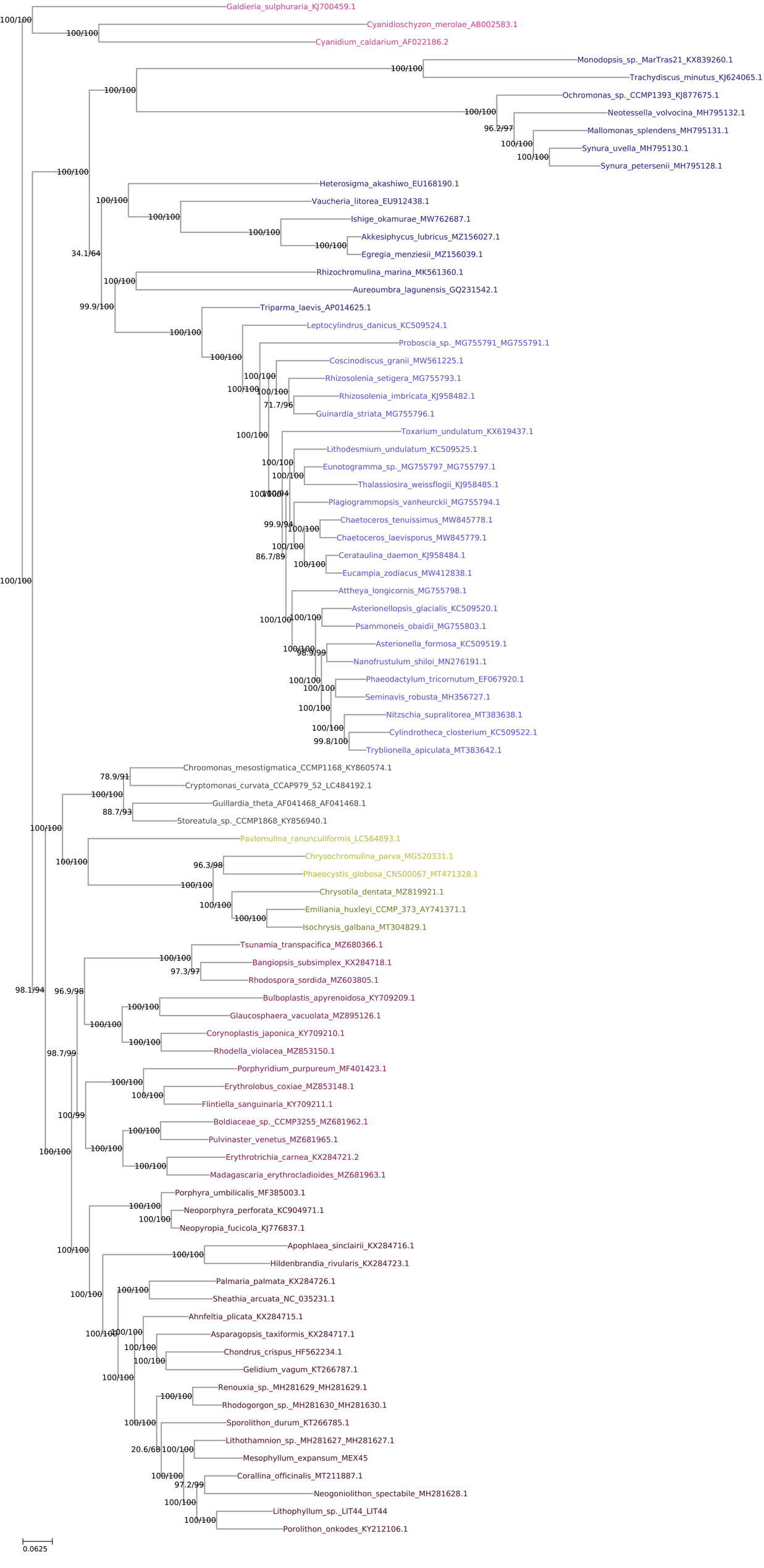

Galdieria_sulphuraria_KJ700459.1
Cyanidioschyzon_merolae_AB002583.1
Cyanidium_caldarium_AF022186.2
Monodopsis_sp._MarTras21_KX839260.1
Trachydiscus_minutus_KJ624065.1
Ochromonas_sp._CCMP1393_KJ877675.1
Neotessella_volvocina_MH795132.1
Mallomonas_splendens_MH795131.1
Synura_uvella_MH795130.1
Synura_petersenii_MH795128.1
Heterosigma_akashiwo_EU168190.1
Vaucheria_litorea_EU912438.1
Ishige_okamurae_MW762687.1
Akkesiphycus_lubricus_MZ156027.1
Egregia_menziesii_MZ156039.1
Rhizochromulina_marina_MK561360.1
Aureoumbra_lagunensis_GQ231542.1
Triparma_laevis_AP014625.1
Leptocylindrus_danicus_KC509524.1
Proboscia_sp._MG755791_MG755791.1
Coscinodiscus_granii_MW561225.1
Rhizosolenia_setigera_MG755793.1
Rhizosolenia_imbricata_KJ958482.1
Guinardia_striata_MG755796.1
Toxarium_undulatum_KX619437.1
Lithodesmium_undulatum_KC509525.1
Eunotogramma_sp._MG755797_MG755797.1
Thalassiosira_weissflogii_KJ958485.1
Plagiogrammopsis_vanheurckii_MG755794.1
Chaetoceros_tenuissimus_MW845778.1
Chaetoceros_laevisporus_MW845779.1
Cerataulina_daemon_KJ958484.1
Eucampia_zodiacus_MW412838.1
Attheya_longicornis_MG755798.1
Asterionellopsis_glacialis_KC509520.1
Psammoneis_obaidii_MG755803.1
Asterionella_formosa_KC509519.1
Nanofrustulum_shiloi_MN276191.1
Phaeodactylum_tricornutum_EF067920.1
Seminavis_robusta_MH356727.1
Nitzschia_supralitorea_MT383638.1
Cylindrotheca_closterium_KC509522.1
Tryblionella_apiculata_MT383642.1
Chroomonas_mesostigmatica_CCMP1168_KY860574.1
Cryptomonas_curvata_CCAP979_52_LC484192.1
Guillardia_theta_AF041468_AF041468.1
Storeatula_sp._CCMP1868_KY856940.1
Pavlomulina_ranunculiformis_LC564893.1
Chrysochromulina_parva_MG520331.1
Phaeocystis_globosa_CNS00067_MT471328.1
Chrysotila_dentata_MZ819921.1
Emiliania_huxleyi_CCMP_373_AY741371.1
Isochrysis_galbana_MT304829.1
Tsunamia_transpacifica_MZ680366.1
Bangiopsis_subsimplex_KX284718.1
Rhodospora_sordida_MZ603805.1
Bulboplastis_apyrenoidosa_KY709209.1
Glaucosphaera_vacuolata_MZ895126.1
Corynoplastis_japonica_KY709210.1
Rhodella_violacea_MZ853150.1
Porphyridium_purpureum_MF401423.1
Erythrolobus_coxiae_MZ853148.1
Flintiella_sanguinaria_KY709211.1
Boldiaceae_sp._CCMP3255_MZ681962.1
Pulvinaster_venetus_MZ681965.1
Erythrotrichia_carnea_KX284721.2
Madagascaria_erythrocladioides_MZ681963.1
Porphyra_umbilicalis_MF385003.1
Neoporphyra_perforata_KC904971.1
Neopyropia_fucicola_KJ776837.1
Apophlaea_sinclairii_KX284716.1
Hildenbrandia_rivularis_KX284723.1
Palmaria_palmata_KX284726.1
Sheathia_arcuata_NC_035231.1
Ahnfeltia_plicata_KX284715.1
Asparagopsis_taxiformis_KX284717.1
Chondrus_crispus_HF562234.1
Gelidium_vagum_KT266787.1
Renouxia_sp._MH281629_MH281629.1
Rhodogorgon_sp._MH281630_MH281630.1
Sporolithon_durum_KT266785.1
Lithothamnion_sp._MH281627_MH281627.1
Mesophyllum_expansum_MEX45
Corallina_officinalis_MT211887.1
Neogoniolithon_spectabile_MH281628.1
Lithophyllum_sp._LIT44_LIT44
Porolithon_onkodes_KY212106.1
0.0625

Fig. S5

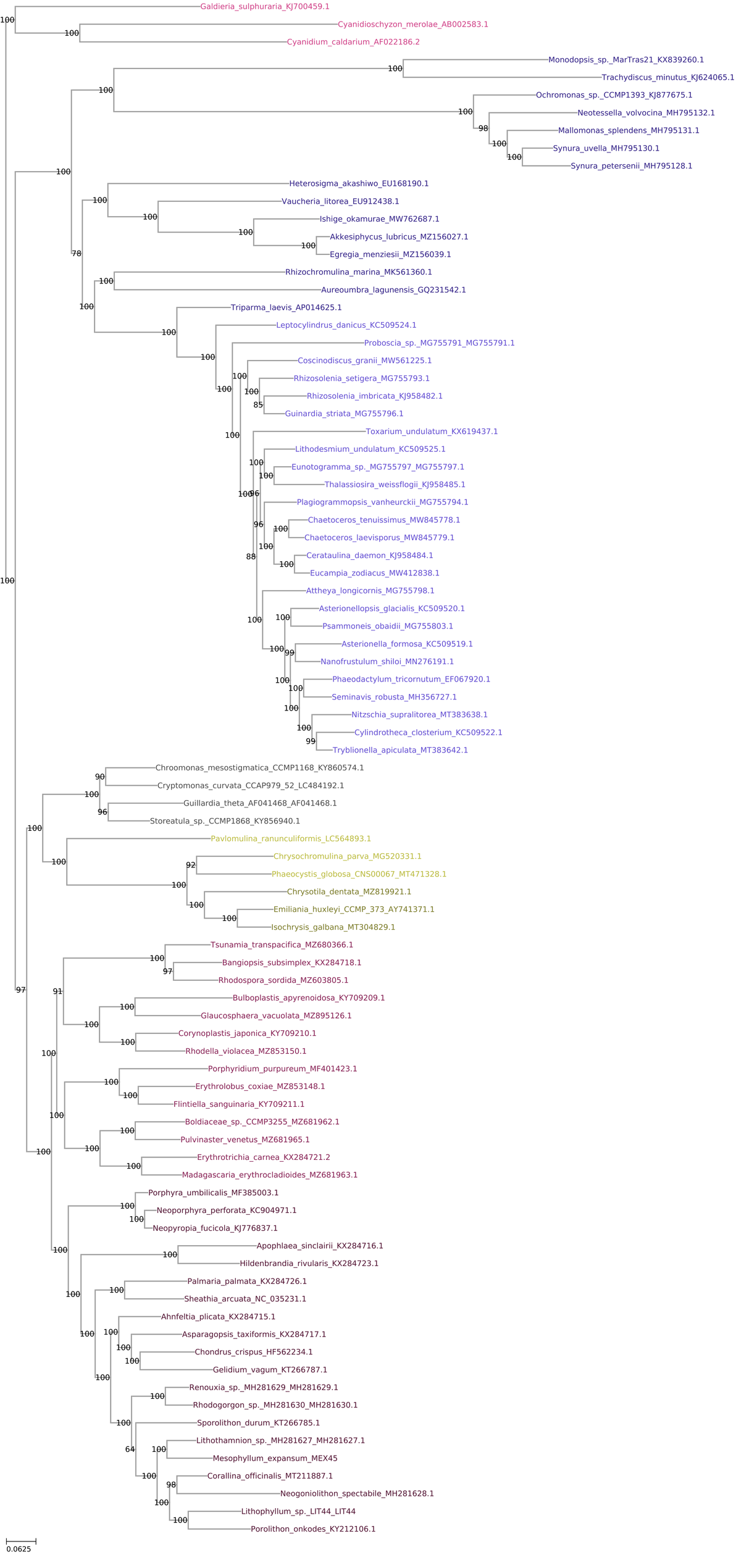

Galdieria_sulphuraria_KJ700459.1
Cyanidioschyzon_merolae_AB002583.1
Cyanidium_caldarium_AF022186.2
Monodopsis_sp._MarTras21_KX839260.1
Trachydiscus_minutus_KJ624065.1
Ochromonas_sp._CCMP1393_KJ877675.1
Neotessella_volvocina_MH795132.1
Mallomonas_splendens_MH795131.1
Synura_uvella_MH795130.1
Synura_petersenii_MH795128.1
Heterosigma_akashiwo_EU168190.1
Vaucheria_litorea_EU912438.1
Ishige_okamurae_MW762687.1
Akkesiphycus_lubricus_MZ156027.1
Egregia_menziesii_MZ156039.1
Rhizochromulina_marina_MK561360.1
Aureoumbra_lagunensis_GQ231542.1
Triparma_laevis_AP014625.1
Leptocylindrus_danicus_KC509524.1
Proboscia_sp._MG755791_MG755791.1
Coscinodiscus_granii_MW561225.1
Rhizosolenia_setigera_MG755793.1
Rhizosolenia_imbricata_KJ958482.1
Guinardia_striata_MG755796.1
Toxarium_undulatum_KX619437.1
Lithodesmium_undulatum_KC509525.1
Eunotogramma_sp._MG755797_MG755797.1
Thalassiosira_weissflogii_KJ958485.1
Plagiogrammopsis_vanheurckii_MG755794.1
Chaetoceros_tenuissimus_MW845778.1
Chaetoceros_laevisporus_MW845779.1
Cerataulina_daemon_KJ958484.1
Eucampia_zodiacus_MW412838.1
Attheya_longicornis_MG755798.1
Asterionellopsis_glacialis_KC509520.1
Psammoneis_obaidii_MG755803.1
Asterionella_formosa_KC509519.1
Nanofrustulum_shiloi_MN276191.1
Phaeodactylum_tricornutum_EF067920.1
Seminavis_robusta_MH356727.1
Nitzschia_supralitorea_MT383638.1
Cylindrotheca_closterium_KC509522.1
Tryblionella_apiculata_MT383642.1
Chroomonas_mesostigmatica_CCMP1168_KY860574.1
Cryptomonas_curvata_CCAP979_52_LC484192.1
Guillardia_theta_AF041468_AF041468.1
Storeatula_sp._CCMP1868_KY856940.1
Pavlomulina_ranunculiformis_LC564893.1
Chrysochromulina_parva_MG520331.1
Phaeocystis_globosa_CNS00067_MT471328.1
Chrysotila_dentata_MZ819921.1
Emiliania_huxleyi_CCMP_373_AY741371.1
Isochrysis_galbana_MT304829.1
Tsunamia_transpacifica_MZ680366.1
Bangiopsis_subsimplex_KX284718.1
Rhodospora_sordida_MZ603805.1
Bulboplastis_apyrenoidosa_KY709209.1
Glaucosphaera_vacuolata_MZ895126.1
Corynoplastis_japonica_KY709210.1
Rhodella_violacea_MZ853150.1
Porphyridium_purpureum_MF401423.1
Erythrolobus_coxiae_MZ853148.1
Flintiella_sanguinaria_KY709211.1
Boldiaceae_sp._CCMP3255_MZ681962.1
Pulvinaster_venetus_MZ681965.1
Erythrotrichia_carnea_KX284721.2
Madagascaria_erythrocladioides_MZ681963.1
Porphyra_umbilicalis_MF385003.1
Neoporphyra_perforata_KC904971.1
Neopyropia_fucicola_KJ776837.1
Apophlaea_sinclairii_KX284716.1
Hildenbrandia_rivularis_KX284723.1
Palmaria_palmata_KX284726.1
Sheathia_arcuata_NC_035231.1
Ahnfeltia_plicata_KX284715.1
Asparagopsis_taxiformis_KX284717.1
Chondrus_crispus_HF562234.1
Gelidium_vagum_KT266787.1
Renouxia_sp._MH281629_MH281629.1
Rhodogorgon_sp._MH281630_MH281630.1
Sporolithon_durum_KT266785.1
Lithothamnion_sp._MH281627_MH281627.1
Mesophyllum_expansum_MEX45
Corallina_officinalis_MT211887.1
Neogoniolithon_spectabile_MH281628.1
Lithophyllum_sp._LIT44_LIT44
Porolithon_onkodes_KY212106.1
0.0625

Fig. S6

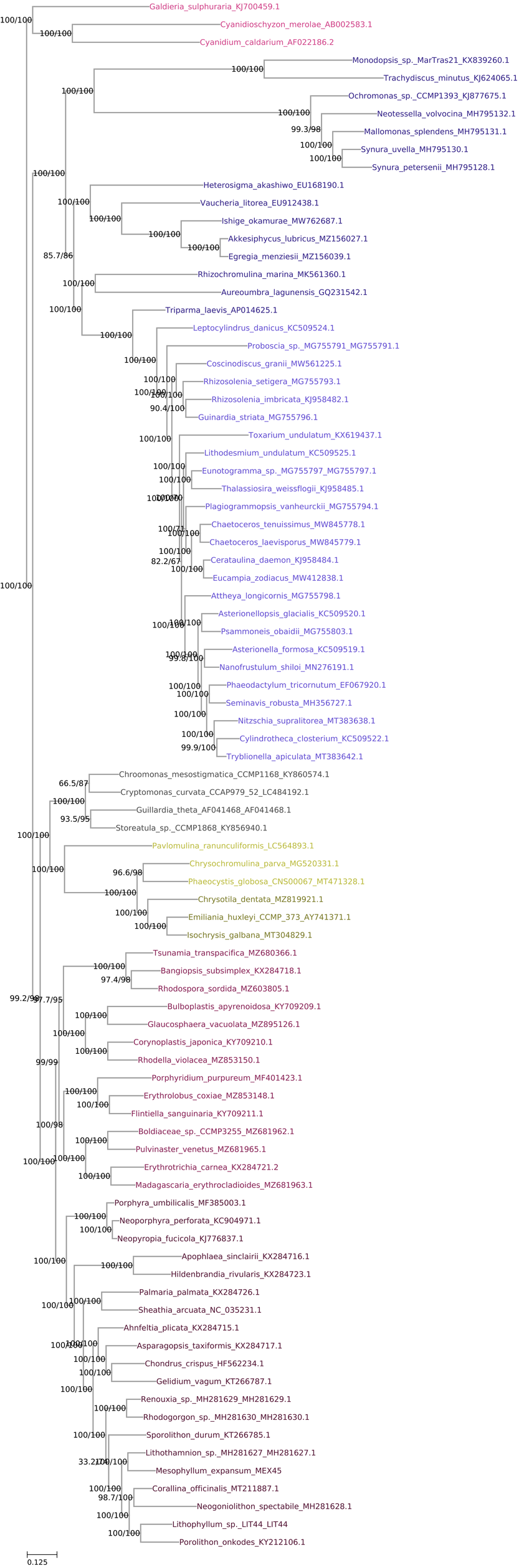

Galdieria_sulphuraria_KJ700459.1
Cyanidioschyzon_merolae_AB002583.1
Cyanidium_caldarium_AF022186.2
Monodopsis_sp._MarTras21_KX839260.1
Trachydiscus_minutus_KJ624065.1
Ochromonas_sp._CCMP1393_KJ877675.1
Neotessella_volvocina_MH795132.1
Mallomonas_splendens_MH795131.1
Synura_uvella_MH795130.1
Synura_petersenii_MH795128.1
Heterosigma_akashiwo_EU168190.1
Vaucheria_litorea_EU912438.1
Ishige_okamurae_MW762687.1
Akkesiphycus_lubricus_MZ156027.1
Egregia_menziesii_MZ156039.1
Rhizochromulina_marina_MK561360.1
Aureoumbra_lagunensis_GQ231542.1
Triparma_laevis_AP014625.1
Leptocylindrus_danicus_KC509524.1
Proboscia_sp._MG755791_MG755791.1
Coscinodiscus_granii_MW561225.1
Rhizosolenia_setigera_MG755793.1
Rhizosolenia_imbricata_KJ958482.1
Guinardia_striata_MG755796.1
Toxarium_undulatum_KX619437.1
Lithodesmium_undulatum_KC509525.1
Eunotogramma_sp._MG755797_MG755797.1
Thalassiosira_weissflogii_KJ958485.1
Plagiogrammopsis_vanheurckii_MG755794.1
Chaetoceros_tenuissimus_MW845778.1
Chaetoceros_laevisporus_MW845779.1
Cerataulina_daemon_KJ958484.1
Eucampia_zodiacus_MW412838.1
Attheya_longicornis_MG755798.1
Asterionellopsis_glacialis_KC509520.1
Psammoneis_obaidii_MG755803.1
Asterionella_formosa_KC509519.1
Nanofrustulum_shiloi_MN276191.1
Phaeodactylum_tricornutum_EF067920.1
Seminavis_robusta_MH356727.1
Nitzschia_supralitorea_MT383638.1
Cylindrotheca_closterium_KC509522.1
Tryblionella_apiculata_MT383642.1
Chroomonas_mesostigmatica_CCMP1168_KY860574.1
Cryptomonas_curvata_CCAP979_52_LC484192.1
Guillardia_theta_AF041468_AF041468.1
Storeatula_sp._CCMP1868_KY856940.1
Pavlomulina_ranunculiformis_LC564893.1
Chrysochromulina_parva_MG520331.1
Phaeocystis_globosa_CNS00067_MT471328.1
Chrysotila_dentata_MZ819921.1
Emiliania_huxleyi_CCMP_373_AY741371.1
Isochrysis_galbana_MT304829.1
Tsunamia_transpacifica_MZ680366.1
Bangiopsis_subsimplex_KX284718.1
Rhodospora_sordida_MZ603805.1
Bulboplastis_apyrenoidosa_KY709209.1
Glaucosphaera_vacuolata_MZ895126.1
Corynoplastis_japonica_KY709210.1
Rhodella_violacea_MZ853150.1
Porphyridium_purpureum_MF401423.1
Erythrolobus_coxiae_MZ853148.1
Flintiella_sanguinaria_KY709211.1
Boldiaceae_sp._CCMP3255_MZ681962.1
Pulvinaster_venetus_MZ681965.1
Erythrotrichia_carnea_KX284721.2
Madagascaria_erythrocladioides_MZ681963.1
Porphyra_umbilicalis_MF385003.1
Neoporphyra_perforata_KC904971.1
Neopyropia_fucicola_KJ776837.1
Apophlaea_sinclairii_KX284716.1
Hildenbrandia_rivularis_KX284723.1
Palmaria_palmata_KX284726.1
Sheathia_arcuata_NC_035231.1
Ahnfeltia_plicata_KX284715.1
Asparagopsis_taxiformis_KX284717.1
Chondrus_crispus_HF562234.1
Gelidium_vagum_KT266787.1
Renouxia_sp._MH281629_MH281629.1
Rhodogorgon_sp._MH281630_MH281630.1
Sporolithon_durum_KT266785.1
Lithothamnion_sp._MH281627_MH281627.1
Mesophyllum_expansum_MEX45
Corallina_officinalis_MT211887.1
Neogoniolithon_spectabile_MH281628.1
Lithophyllum_sp._LIT44_LIT44
Porolithon_onkodes_KY212106.1
0.125

Fig. S7

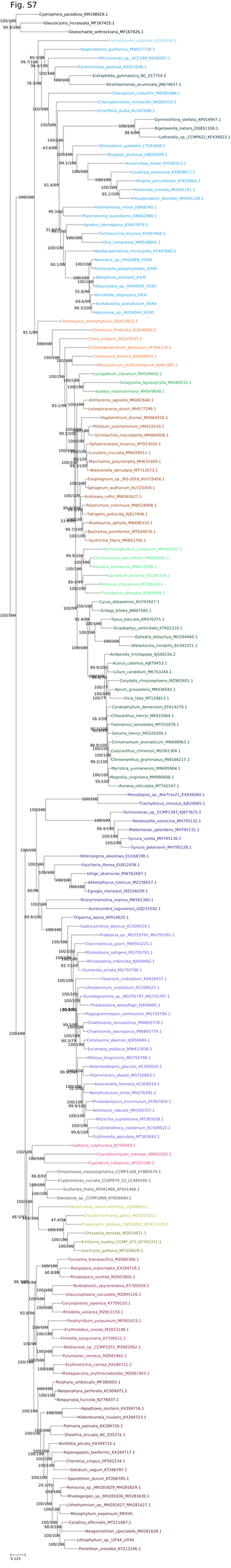
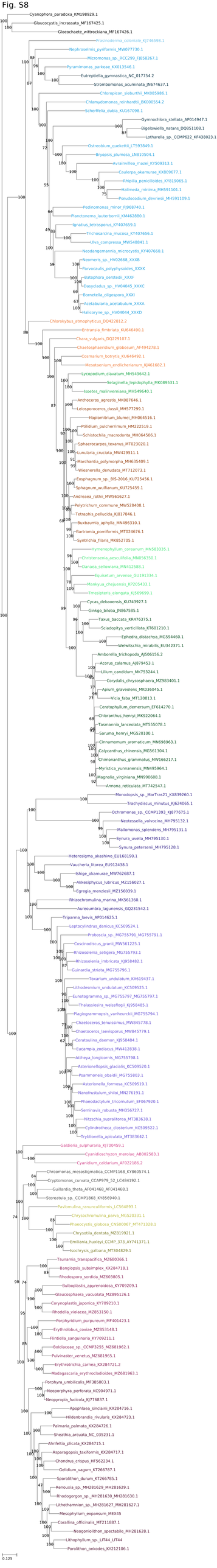

Cyanophora_paradoxa_KM198929.1
Glaucocystis_incrassata_MF167425.1
Gloeochaete_wittrockiana_MF167426.1
Prasinoderma_coloniale_KJ746598.1
Nephroselmis_pyriformis_MW077730.1
Micromonas_sp._RCC299_FJ858267.1
Pyramimonas_parkeae_KX013546.1
Eutreptiella_gymnastica_NC_017754.2
Strombomonas_acuminata_JN674637.1
Chloropicon_sieburthii_MK085986.1
Chlamydomonas_reinhardtii_BK000554.2
Scherffelia_dubia_KU167098.1
Gymnochlora_stellata_AP014947.1
Bigelowiella_natans_DQ851108.1
Lotharella_sp._CCMP622_KF438023.1
Ostreobium_quekettii_LT593849.1
Bryopsis_plumosa_LN810504.1
Avrainvillea_mazei_KY509313.1
Caulerpa_okamurae_KX809677.1
Rhipilia_penicilliodes_KY819065.1
Halimeda_minima_MH591101.1
Pseudocodium_devriesii_MH591109.1
Pedinomonas_minor_FJ968740.1
Planctonema_lauterbornii_KM462880.1
Ignatius_tetrasporus_KY407659.1
Trichosarcina_mucosa_KY407656.1
Ulva_compressa_MW548841.1
Neodangemannia_microcystis_KY407660.1
Neomeris_sp._HV02668_XXXB
Parvocalulis_polyphysoides_XXXK
Batophora_oerstedii_XXXF
Dasycladus_sp._HV04045_XXXC
Bornetella_oligospora_XXXI
Acetabularia_acetabulum_XXXA
Halicoryne_sp._HV04044_XXXD
Chlorokybus_atmophyticus_DQ422812.2
Entransia_fimbriata_KU646490.1
Chara_vulgaris_DQ229107.1
Chaetosphaeridium_globosum_AF494278.1
Cosmarium_botrytis_KU646492.1
Mesotaenium_endlicherianum_KJ461682.1
Lycopodium_clavatum_MH549642.1
Selaginella_lepidophylla_MK089531.1
Isoetes_malinverniana_MH549640.1
Anthoceros_agrestis_MK087646.1
Leiosporoceros_dussii_MH577299.1
Haplomitrium_blumei_MH064516.1
Ptilidium_pulcherrimum_HM222519.1
Schistochila_macrodonta_MH064506.1
Sphaerocarpos_texanus_MT023020.1
Lunularia_cruciata_MW429511.1
Marchantia_polymorpha_MH635409.1
Wiesneriella_denudata_MT712073.1
Eosphagnum_sp._BIS-2016_KU725456.1
Sphagnum_wulfianum_KU725459.1
Andreaea_rothii_MW561627.1
Polytrichum_commune_MW528408.1
Tetraphis_pellucida_KJ817846.1
Buxbaumia_aphylla_MN496310.1
Bartramia_pomiformis_MT024676.1
Syntrichia_filaris_MK852705.1
Hymenophyllum_coreanum_MN583335.1
Christensenia_aesculifolia_MN056350.1
Danaea_sellowiana_MN412588.1
Equisetum_arvense_GU191334.1
Mankyua_chejuensis_KP205433.1
Tmesipteris_elongata_KJ569699.1
Cycas_debaoensis_KU743927.1
Ginkgo_biloba_JN867585.1
Taxus_baccata_KR476375.1
Sciadopitys_verticillata_KT601210.1
Ephedra_distachya_MG594460.1
Welwitschia_mirabilis_EU342371.1
Amborella_trichopoda_AJ506156.2
Acorus_calamus_AJ879453.1
Lilium_candidum_MK753244.1
Corydalis_chrysosphaera_MZ983401.1
Apium_graveolens_MK036045.1
Vicia_faba_MT120813.1
Ceratophyllum_demersum_EF614270.1
Chloranthus_henryi_MK922064.1
Tasmannia_lanceolata_MT555078.1
Saruma_henryi_MG520100.1
Cinnamomum_aromaticum_MN689863.1
Calycanthus_chinensis_MG561304.1
Chimonanthus_grammatus_MW166217.1
Myristica_yunnanensis_MN495964.1
Magnolia_virginiana_MN990608.1
Annona_reticulata_MT742547.1
Monodopsis_sp._MarTras21_KX839260.1
Trachydiscus_minutus_KJ624065.1
Ochromonas_sp._CCMP1393_KJ877675.1
Neotessela_volvocina_MH795132.1
Mallomonas_splendens_MH795131.1
Synura_uvella_MH795130.1
Synura_petersenii_MH795128.1
Heterosigma_akashiwo_EU168190.1
Vaucheria_litorea_EU912438.1
Ishige_okamurae_MW762687.1
Akkesiphycus_lubricus_MZ156027.1
Egregia_menziesii_MZ156039.1
Rhizochromulina_marina_MK561360.1
Aureoumbra_lagunensis_GQ231542.1
Triparma_laevis_AP014625.1
Leptocylindrus_danicus_KC509524.1
Proboscia_sp._MG755791_MG755791.1
Coscinodiscus_granii_MW561225.1
Rhizosolenia_setigera_MG755793.1
Rhizosolenia_imbricata_KJ958482.1
Guinardia_striata_MG755796.1
Toxarium_undulatum_KX619437.1
Lithodesmium_undulatum_KC509525.1
Eunotogramma_sp._MG755797_MG755797.1
Thalassiosira_weissflogii_KJ958485.1
Plagiogrammopsis_vanheurckii_MG755794.1
Chaetoceros_tenuissimus_MW845778.1
Chaetoceros_laevisporus_MW845779.1
Cerataulina_daemon_KJ958484.1
Eucampia_zodiacus_MW412838.1
Attheya_longicornis_MG755798.1
Asterionellopsis_glacialis_KC509520.1
Psammoneis_obaidii_MG755803.1
Asterionella_formosa_KC509519.1
Nanofrustulum_shiloi_MN276191.1
Phaeodactylum_tricornutum_EF067920.1
Seminavis_robusta_MH356727.1
Nitzschia_supralitorea_MT383638.1
Cylindrotheca_closterium_KC509522.1
Tryblionella_apiculata_MT383642.1
Galdieria_sulphuraria_KJ700459.1
Cyanidioschyzon_merolae_AB002583.1
Cyanidium_caldarium_AF022186.2
Chroomonas_mesostigmatica_CCMP1168_KY860574.1
Cryptomonas_curvata_CCAP979_52_LC484192.1
Guillardia_theta_AF041468_AF041468.1
Storeatula_sp._CCMP1868_KY856940.1
Pavlomulina_rannunculiformis_LC564893.1
Chrysochromulina_parva_MG520331.1
Phaeocystis_globosa_CNS00067_MT471328.1
Chrysotila_dentata_MZ819921.1
Emiliania_huxleyi_CCMP_373_AY741371.1
Isochrysis_galbana_MT304829.1
Tsunamia_transpacifica_MZ680366.1
Bangiopsis_subsimplex_KX284718.1
Rhodospora_sordida_MZ603805.1
Bulboplastis_apyrenoidosa_KY709209.1
Glaucosphaera_vacuolata_MZ895126.1
Corynoplastis_japonica_KY709210.1
Rhodella_violacea_MZ853150.1
Porphyridium_purpureum_MF401423.1
Erythrolobus_coxiae_MZ853148.1
Flintiella_sanguinaria_KY709211.1
Boldiaceae_sp._CCMP3255_MZ681962.1
Pulvinaster_venetus_MZ681965.1
Erythrotrichia_carnea_KX284721.2
Madagascaria_erythrocladioides_MZ681963.1
Porphyra_umbilicalis_MF385003.1
Neoporphyra_perforata_KC904971.1
Neoporphyra_fucicola_KJ776837.1
Apophlaea_sinclairii_KX284716.1
Hildenbrandia_rivularis_KX284723.1
Palmaria_palmata_KX284726.1
Sheathia_arcuata_NC_035231.1
Ahnfeltia_plicata_KX284715.1
Asparagopsis_taxiformis_KX284717.1
Chondrus_crispus_HF562234.1
Gelidium_vagum_KT266787.1
Sporolithon_durum_KT266785.1
Renouxia_sp._MH281629_MH281629.1
Rhodogorgon_sp._MH281630_MH281630.1
Lithothamnion_sp._MH281627_MH281627.1
Mesophyllum_expansum_MEX45
Corallina_officinalis_MT211887.1
Neogoniolithon_spectabile_MH281628.1
Lithophyllum_sp._LIT44_LIT44
Porolithon_onkodes_KY212106.1
0.125

Fig. S8


Cyanophora_paradoxa_KM198929.1
Glaucocystis_incrassata_MF167425.1
Gloeochaete_wittrockiana_MF167426.1
Prasinoderma_coloniale_KJ746598.1
Nephroselmis_pyriformis_MW077730.1
Micromonas_sp._RCC299_FJB58267.1
Pyramimonas_parkeae_KX013546.1
Eutreptiella_gymnastica_NC_017754.2
Strombomonas_acuminata_JN674637.1
Chloropicon_sieburthii_MK085986.1
Chlamydomonas_reinhardtii_BK000554.2
Scherffelia_dubia_KU167098.1
Gymnochlora_stellata_AP014947.1
Bigelowiella_natans_DQ851108.1
Lotharella_sp._CCMP622_KF438023.1
Ostreobium_quekettii_LT593849.1
Bryopsis_plumosa_LN810504.1
Avrainvillea_mazei_KY509313.1
Caulerpa_okamurae_KX809677.1
Rhipilia_penicilioides_KY819065.1
Halimeda_minima_MH591101.1
Pseudocodium_devriesii_MH591109.1
Pedinomonas_minor_FJ968740.1
Planctonema_lauterbornii_KM462880.1
Ignatius_tetrasporus_KY407659.1
Trichosarcina_mucosa_KY407656.1
Ulva_compressa_MW548841.1
Neodangemannia_microcystis_KY407660.1
Neomeris_sp._HV02668_XXXB
Parvocalis_polyphysoides_XXXX
Batophora_oerstedii_XXXF
Dasyclades_sp._HV04045_XXXC
Bornetella_oligospora_XXXI
Acetabularia_acetabulum_XXXA
Halicoryne_sp._HV04044_XXXD
Chlorokybus_atmophyticus_DQ422812.2
Entransia_fimbriata_KU646490.1
Chara_vulgaris_DQ229107.1
Chaetosphaeridium_globosum_AF494278.1
Cosmarium_botrytis_KU646492.1
Mesotaenium_endlicherianum_KJ461682.1
Lycopodium_clavatum_MH549642.1
Selaginella_lepidophylla_MK089531.1
Isoetes_malinverniana_MH549640.1
Anthoceros_agrestis_MK087646.1
Leiosporoceros_dussii_MH577299.1
Haplomitrium_blumei_MH064516.1
Ptilidium_pulcherrimum_HM222519.1
Schistochila_macrodonta_MH064506.1
Sphaerocarpos_texanus_MT023020.1
Lunularia_cruciata_MW429511.1
Marchantia_polymorpha_MH635409.1
Wiesnerella_denudata_MT712073.1
Eosphagnum_sp._BIS-2016_KU725456.1
Sphagnum_wulfianum_KU725459.1
Andreaea_rothii_MW561627.1
Polytrichum_commune_MW528408.1
Tetraphis_pellucida_KJ817846.1
Buxbaumia_aphylla_MN496310.1
Bartramia_pomiformis_MT024676.1
Syntrichia_filaris_MK852705.1
Hymenophyllum_coreanum_MN583335.1
Christensenia_aesculifolia_MN056350.1
Danaea_sellowiana_MN412588.1
Equisetum_arvense_GU191334.1
Mankyua_chejuensis_KP205433.1
Tmesipteris_elongata_KJ569699.1
Cycas_debaoensis_KU743927.1
Ginkgo_biloba_JN867585.1
Taxus_baccata_KR476375.1
Sciadopitys_verticillata_KT601210.1
Ephedra_distachya_MG594460.1
Welwitschia_mirabilis_EU342371.1
Amborella_trichopoda_AJ506156.2
Acorus_calamus_AJ879453.1
Lilium_candidum_MK753244.1
Corydalis_chrysosphaera_MZ983401.1
Apium_graveolens_MK036045.1
Vicia_faba_MT120813.1
Ceratophyllum_demersum_EF614270.1
Chloranthus_henryi_MK922064.1
Tasmannia_lanceolata_MT555078.1
Saruma_henryi_MG520100.1
Cinnamomum_aromaticum_MN698963.1
Calycanthus_chinensis_MG561304.1
Chimonanthus_grammatus_MW166217.1
Myristica_yunnanensis_MN495964.1
Magnolia_virginiana_MN990608.1
Annona_reticulata_MT742547.1
Monodopsis_sp._MarTras21_KX839260.1
Trachydiscus_minutus_KJ624065.1
Ochromonas_sp._CCMP1393_KJ877675.1
Neotessella_volvocina_MH795132.1
Mallomonas_splendens_MH795131.1
Synura_uvella_MH795130.1
Synura_petersenii_MH795128.1
Heterosigma_akashiwo_EU168190.1
Vaucheria_litorea_EU912438.1
Ishige_okamurae_MW762687.1
Akkesiphycus_lubricus_MZ156027.1
Egregia_menziesii_MZ156039.1
Rhizochromulina_marina_MK561360.1
Aureoumbra_lagunensis_GQ231542.1
Triparma_laevis_AP014625.1
Leptocylindrus_danicus_KC509524.1
Proboscia_sp._MG755791_MG755791.1
Coscinodiscus_granii_MW561225.1
Rhizosolenia_setigera_MG755793.1
Rhizosolenia_imbricata_KJ958482.1
Guinardia_striata_MG755796.1
Toxarium_undulatum_KX619437.1
Lithodesmium_undulatum_KC509525.1
Eunotogramma_sp._MG755797_MG755797.1
Thalassiosira_weissflogii_KJ958485.1
Plagiogrammopsis_vanheurckii_MG755794.1
Chaetoceros_tenuissimus_MW845778.1
Chaetoceros_laevisporus_MW845779.1
Cerataulina_daemon_KJ958484.1
Eucampia_zodiacus_MW412838.1
Attheya_longicornis_MG755798.1
Asterionellopsis_glacialis_KC509520.1
Psammoneis_obaidii_MG755803.1
Asterionella_formosa_KC509519.1
Nanofrustulum_shiloi_MN276191.1
Phaeodactylum_tricornutum_EF067920.1
Seminavis_robusta_MH356727.1
Nitzschia_supralitorea_MT383638.1
Cylindrotheca_closterium_KC509522.1
Trybilonella_apiculata_MT383642.1
Galdieria_sulphuraria_KJ700459.1
Cyanidioschyzon_merolae_AB002583.1
Cyanidium_caldarium_AF022186.2
Chroomonas_mesostigmatica_CCMP1168_KY860574.1
Cryptomonas_curvata_CCAP979_52_LC484192.1
Guillardia_theta_AF041468_AF041468.1
Storeatula_sp._CCMP1868_KY856940.1
Pavlomulina_rannuculiformis_LC564893.1
Chrysochromulina_parva_MG520331.1
Phaeocystis_globosa_CNS00067_MT471328.1
Chrysotila_dentata_MZ819921.1
Emiliania_huxleyi_CCMP_373_AY741371.1
Isochrysis_galbana_MT304829.1
Tsunamia_transpacifica_MZ680366.1
Bangiopsis_subsimplex_KX284718.1
Rhodosorsa_sordida_MZ603805.1
Bulboplastis_apyrenoidosa_KY709209.1
Glaucosphaera_vacuolata_MZ895126.1
Corynoplastis_japonica_KY709210.1
Rhodella_violacea_MZ853150.1
Porphyridium_purpureum_MF401423.1
Erythrolobus_coxiae_MZ853148.1
Flintiella_sanguinaria_KY709211.1
Boldiaceae_sp._CCMP3255_MZ681962.1
Pulvinaster_venetus_MZ681965.1
Erythrotrichia_carnea_KX284721.2
Madagascaria_erythrocladioides_MZ681963.1
Porphyra_umbilicalis_MF385003.1
Neoporphyra_perforata_KC904971.1
Neopyropia_fucicola_KJ776837.1
Apophlaea_sinclairii_KX284716.1
Hildenbrandia_rivularis_KX284723.1
Palmaria_palmata_KX284726.1
Sheathia_arcuata_NC_035231.1
Ahnfeltia_plicata_KX284715.1
Asparagopsis_taxiformis_KX284717.1
Chondrus_crispus_MF562234.1
Gelidium_vagum_KT266787.1
Sporolithon_durum_KT266785.1
Renouxia_sp._MH281629_MH281629.1
Rhodogorgon_sp._MH281630_MH281630.1
Lithothamnion_sp._MH281627_MH281627.1
Mesophyllum_expansum_MEX45
Corallina_officinalis_MT211887.1
Neogoniolithon_spectabile_MH281628.1
Lithophyllum_sp._LIT44_LIT44
Porolithon_onkodes_KY212106.1
0.125

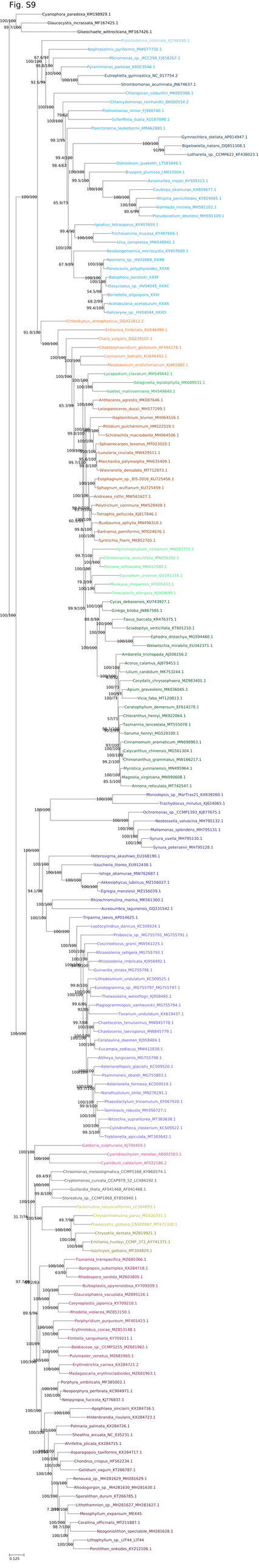


Fig. S9
Cyanophora_paradoxa_KM198929.1
Glaucocystis_incrassata_MF167425.1
Gloeochaete_wittrockiana_MF167426.1
Prasinoderma_coloniale_KJ746598.1
Nephroselmis_pyriformis_MW077730.1
Micromonas_sp._RCC299_FJ858267.1
Pyramimonas_parkeae_KX013546.1
Eutreptiella_gymnastica_NC_017754.2
Strombomonas_acuminata_JN674637.1
Chloropicon_sieburthii_MK085986.1
Chlamydomonas_reinhardtii_BK000554.2
Pedinomonas_minor_FJ968740.1
Scherffelia_dubia_KU167098.1
Planctonema_lauterbornii_KM462880.1
Gymnochlora_stellata_AP014947.1
Bigelowiella_natans_DQ851108.1
Lotharella_sp._CCMP622_KF438023.1
Ostreobium_quekettii_LT593849.1
Bryopsis_plumosa_LN810504.1
Avrainvillea_mazei_KY509313.1
Caulerpa_okamurae_KX809677.1
Rhipilia_penicilloides_KY819065.1
Halimeda_minima_MH591101.1
Pseudocodium_devriesii_MH591109.1
Ignatius_tetrasporus_KY407659.1
Trichosarcina_mucosa_KY407656.1
Ulva_compressa_MW548841.1
Neodangemannia_microcystis_KY407660.1
Neomeris_sp._HV02668_XXXB
Parvocalis_polyphysoides_XXXK
Batophora_oerstedii_XXXF
Dasycladus_sp._HV04045_XXXC
Bornetella_oligospora_XXXI
Acetabularia_acetabulum_XXXA
Halicoryne_sp._HV04044_XXXD
Chlorokybus_atmophyticus_DQ422812.2
Entransia_fimbriata_KU646490.1
Chara_vulgaris_DQ229107.1
Chaetosphaeridium_globosum_AF494278.1
Cosmarium_botrytis_KU646492.1
Mesotaenium_endlicherianum_KJ461682.1
Lycopodium_clavatum_MH549642.1
Selaginella_lepidophylla_MK089531.1
Isoetes_malinverniana_MH549640.1
Anthoceros_agrestis_MK087646.1
Leiosporoceros_dussii_MH577299.1
Haplomitrium_blumei_MH064516.1
Pilidium_pulcherrimum_HM222519.1
Schistochila_macrodonta_MH064506.1
Sphaerocarpos_texanus_MT023020.1
Lunularia_cruciata_MW429511.1
Marchantia_polymorpha_MH635409.1
Wiesnerella_denudata_MT712073.1
Eosphagnum_sp._BIS-2016_KU725456.1
Sphagnum_wulfianum_KU725459.1
Andreaea_rothii_MW561627.1
Polytrichum_commune_MW528408.1
Tetraphis_pellucida_KJ817846.1
Buxbaumia_aphylla_MN496310.1
Bartramia_pomiformis_MT024676.1
Syntrichia_filaris_MK852705.1
Hymenophyllum_coreanum_MN583335.1
Christensenia_aesculifolia_MN056350.1
Danaea_sellowiana_MN412588.1
Equisetum_arvense_GU191334.1
Mankyua_chejuensis_KP205433.1
Tmesipteris_elongata_KJ569699.1
Cycas_debaoensis_KU743927.1
Ginkgo_biloba_JN867585.1
Taxus_baccata_KR476375.1
Sciadopitys_verticillata_KT601210.1
Ephedra_distachya_MG594460.1
Welwitschia_mirabilis_EU342371.1
Amborella_trichopoda_AJ506156.2
Acorus_calamus_AJ879453.1
Lilium_candidum_MK753244.1
Corydalis_chrysosphaera_MZ983401.1
Apium_graveolens_MK036045.1
Vicia_faba_MT120813.1
Ceratophyllum_demersum_EF614270.1
Chloranthus_henryi_MK922064.1
Tasmannia_lanceolata_MT555078.1
Saruma_henryi_MG520100.1
Cinnamomum_aromaticum_MN698963.1
Calycanthus_chinensis_MG561304.1
Chimonanthus_grammatus_MW166217.1
Myristica_yunnanensis_MN495964.1
Magnolia_virginiana_MN990608.1
Annona_reticulata_MT742547.1
Monodopsis_sp._MarTras21_KX839260.1
Trachydiscus_minutus_KJ624065.1
Ochromonas_sp._CCMP1393_KJ877675.1
Neotessella_volvocina_MH795132.1
Mallomonas_splendens_MH795131.1
Synura_uvella_MH795130.1
Synura_petersenii_MH795128.1
Heterosigma_akashiwo_EU168190.1
Vaucheria_litorea_EU912438.1
Ishige_okamurae_MW762687.1
Akkesiphycus_lubricus_MZ156027.1
Egregia_menziesii_MZ156039.1
Rhizochromulina_marina_MK561360.1
Aureoumbra_lagunensis_GQ231542.1
Triparma_laevis_AP014625.1
Leptocylindrus_danicus_KC509524.1
Proboscia_sp._MG755791_MG755791.1
Coscinodiscus_granii_MW561225.1
Rhizosolenia_setigera_MG755793.1
Rhizosolenia_imbricata_KJ958482.1
Guinardia_striata_MG755796.1
Lithodesmium_undulatum_KC509525.1
Eunotogramma_sp._MG755797_MG755797.1
Thalassiosira_weissflogii_KJ958485.1
Plagiogrammopsis_vanheurckii_MG755794.1
Toxarium_undulatum_KX619437.1
Chaetoceros_tenuissimus_MW845778.1
Chaetoceros_laevisporus_MW845779.1
Cerataulina_daemon_KJ958484.1
Eucampia_zodiacus_MW412838.1
Attheya_longicornis_MG755798.1
Asterionellopsis_glacialis_KC509520.1
Psammoneis_obaidii_MG755803.1
Asterionella_formosa_KC509519.1
Nanofrustulum_shiloi_MN276191.1
Phaeodactylum_tricornutum_EF067920.1
Seminavis_robusta_MH356727.1
Nitzschia_supralitorea_MT383638.1
Cylindrotheca_closterium_KC509522.1
Tryblionella_apiculata_MT383642.1
Galdieria_sulphuraria_KJ700459.1
Cyanidioschyzon_merolae_AB002583.1
Cyanidium_caldarium_AF022186.2
Chroomonas_mesostigmatica_CCMP1168_KY860574.1
Cryptomonas_curvata_CCAP979_52_LC484192.1
Guillardia_theta_AF041468_AF041468.1
Storeatula_sp._CCMP1868_KY856940.1
Pavlomulina_ranunculiformis_LC564893.1
Chrysochromulina_parva_MG520331.1
Phaeocystis_globosa_CNS00067_MT471328.1
Chrysotila_dentata_MZ819921.1
Emiliania_huxleyi_CCMP_373_AY741371.1
Isochrysis_galbana_MT304829.1
Tsunamia_transpacifica_MZ680366.1
Bangiopsis_subsimplex_KX284718.1
Rhodospora_sordida_MZ603805.1
Bulboplastis_apyrenoidosa_KY709209.1
Glaucosphaera_vacuolata_MZ895126.1
Corynoplastis_japonica_KY709210.1
Rhodella_violacea_MZ853150.1
Porphyridium_purpureum_MF401423.1
Erythrolobus_coxiae_MZ853148.1
Flintiella_sanguinaria_KY709211.1
Boldiaceae_sp._CCMP3255_MZ681962.1
Pulvinaster_venetus_MZ681965.1
Erythrotrichia_carnea_KX284721.2
Madagascaria_erythrocladioides_MZ681963.1
Porphyra_umbilicalis_MF385003.1
Neoporphyra_perforata_KC904971.1
Neopyropia_fucicola_KJ776837.1
Apophlaea_sinclairii_KX284716.1
Hildenbrandia_rivularis_KX284723.1
Palmaria_palmata_KX284726.1
Sheathia_arcuata_NC_035231.1
Ahnfeltia_plicata_KX284715.1
Asparagopsis_taxiformis_KX284717.1
Chondrus_crispus_HF562234.1
Gelidium_vagum_KT266787.1
Renouxia_sp._MH281629_MH281629.1
Rhodogorgon_sp._MH281630_MH281630.1
Sporolithon_durum_KT266785.1
Lithothamnion_sp._MH281627_MH281627.1
Mesophyllum_expansum_MEX45
Corallina_officinalis_MT211887.1
Neogoniolithon_spectabile_MH281628.1
Lithophyllum_sp._LIT44_LIT44
Porolithon_onkodes_KY212106.1
0.125

Fig. S10

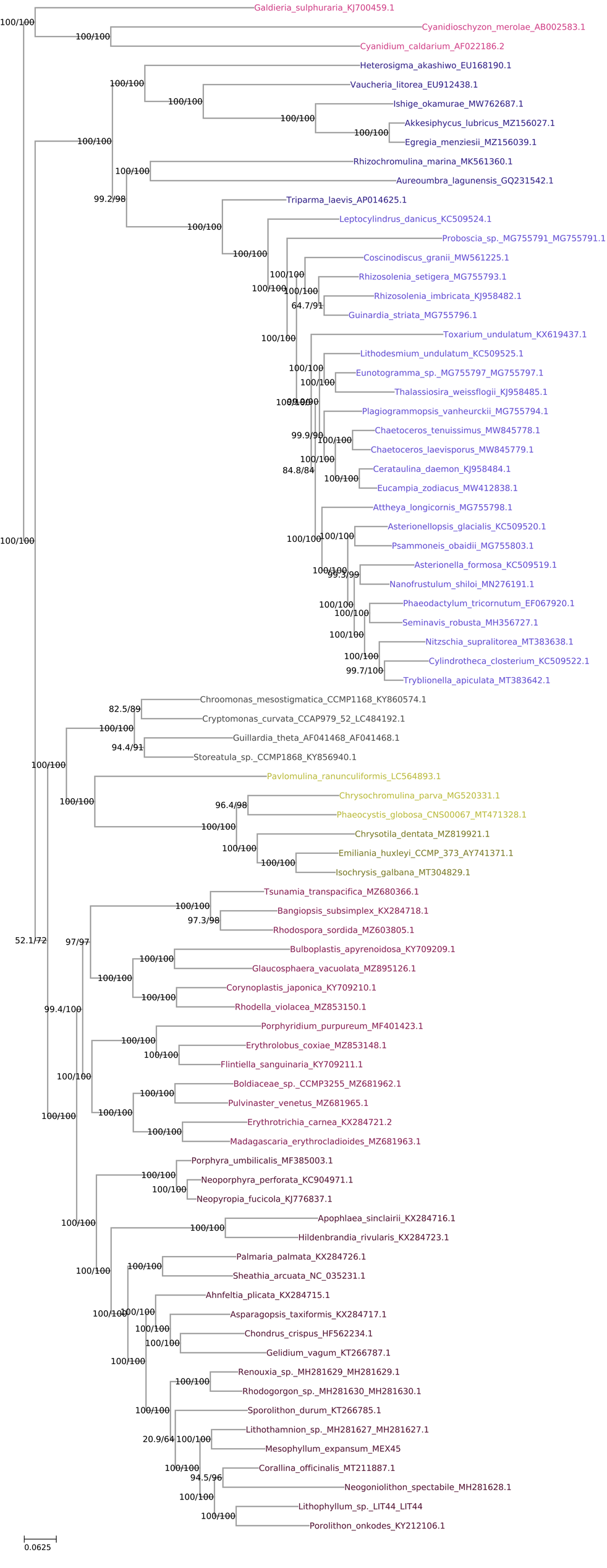

Galdieria_sulphuraria_KJ700459.1
Cyanidioschyzon_merolae_AB002583.1
Cyanidium_caldarium_AF022186.2
Heterosigma_akashiwo_EU168190.1
Vaucheria_litorea_EU912438.1
Ishige_okamurae_MW762687.1
Akkesiphycus_lubricus_MZ156027.1
Egregia_menziesii_MZ156039.1
Rhizochromulina_marina_MK561360.1
Aureoumbra_lagunensis_GQ231542.1
Triparma_laevis_AP014625.1
Leptocylindrus_danicus_KC509524.1
Proboscia_sp._MG755791_MG755791.1
Coscinodiscus_granii_MW561225.1
Rhizosolenia_setigera_MG755793.1
Rhizosolenia_imbricata_KJ958482.1
Guinardia_striata_MG755796.1
Toxarium_undulatum_KX619437.1
Lithodesmium_undulatum_KC509525.1
Eunotogramma_sp._MG755797_MG755797.1
Thalassiosira_weissflogii_KJ958485.1
Plagiogrammopsis_vanheurckii_MG755794.1
Chaetoceros_tenuissimus_MW845778.1
Chaetoceros_laevisporus_MW845779.1
Cerataulina_daemon_KJ958484.1
Eucampia_zodiacus_MW412838.1
Attheya_longicornis_MG755798.1
Asterionellopsis_glacialis_KC509520.1
Psammoneis_obaidii_MG755803.1
Asterionella_formosa_KC509519.1
Nanofrustulum_shiloi_MN276191.1
Phaeodactylum_tricornutum_EF067920.1
Seminavis_robusta_MH356727.1
Nitzschia_supralitorea_MT383638.1
Cylindrotheca_closterium_KC509522.1
Tryblionella_apiculata_MT383642.1
Chroomonas_mesostigmatica_CCMP1168_KY860574.1
Cryptomonas_curvata_CCAP979_52_LC484192.1
Guillardia_theta_AF041468_AF041468.1
Storeatula_sp._CCMP1868_KY856940.1
Pavlomulina_ranunculiformis_LC564893.1
Chrysochromulina_parva_MG520331.1
Phaeocystis_globosa_CNS00067_MT471328.1
Chrysotila_dentata_MZ819921.1
Emiliania_huxleyi_CCMP_373_AY741371.1
Isochrysis_galbana_MT304829.1
Tsunamia_transpacifica_MZ680366.1
Bangiopsis_subsimplex_KX284718.1
Rhodospora_sordida_MZ603805.1
Bulboplastis_apyrenoidosa_KY709209.1
Glaucosphaera_vacuolata_MZ895126.1
Corynoplastis_japonica_KY709210.1
Rhodella_violacea_MZ853150.1
Porphyridium_purpureum_MF401423.1
Erythrolobus_coxiae_MZ853148.1
Flintiella_sanguinaria_KY709211.1
Boldiaceae_sp._CCMP3255_MZ681962.1
Pulvinaster_venetus_MZ681965.1
Erythrotrichia_carnea_KX284721.2
Madagascaria_erythrocladioides_MZ681963.1
Porphyra_umbilicalis_MF385003.1
Neoporphyra_perforata_KC904971.1
Neopyropia_fucicola_KJ776837.1
Apophlaea_sinclairii_KX284716.1
Hildenbrandia_rivularis_KX284723.1
Palmaria_palmata_KX284726.1
Sheathia_arcuata_NC_035231.1
Ahnfeltia_plicata_KX284715.1
Asparagopsis_taxiformis_KX284717.1
Chondrus_crispus_HF562234.1
Gelidium_vagum_KT266787.1
Renouxia_sp._MH281629_MH281629.1
Rhodogorgon_sp._MH281630_MH281630.1
Sporolithon_durum_KT266785.1
Lithothamnion_sp._MH281627_MH281627.1
Mesophyllum_expansum_MEX45
Corallina_officinalis_MT211887.1
Neogoniolithon_spectabile_MH281628.1
Lithophyllum_sp._LIT44_LIT44
Porolithon_onkodes_KY212106.1
0.0625

Fig. S11

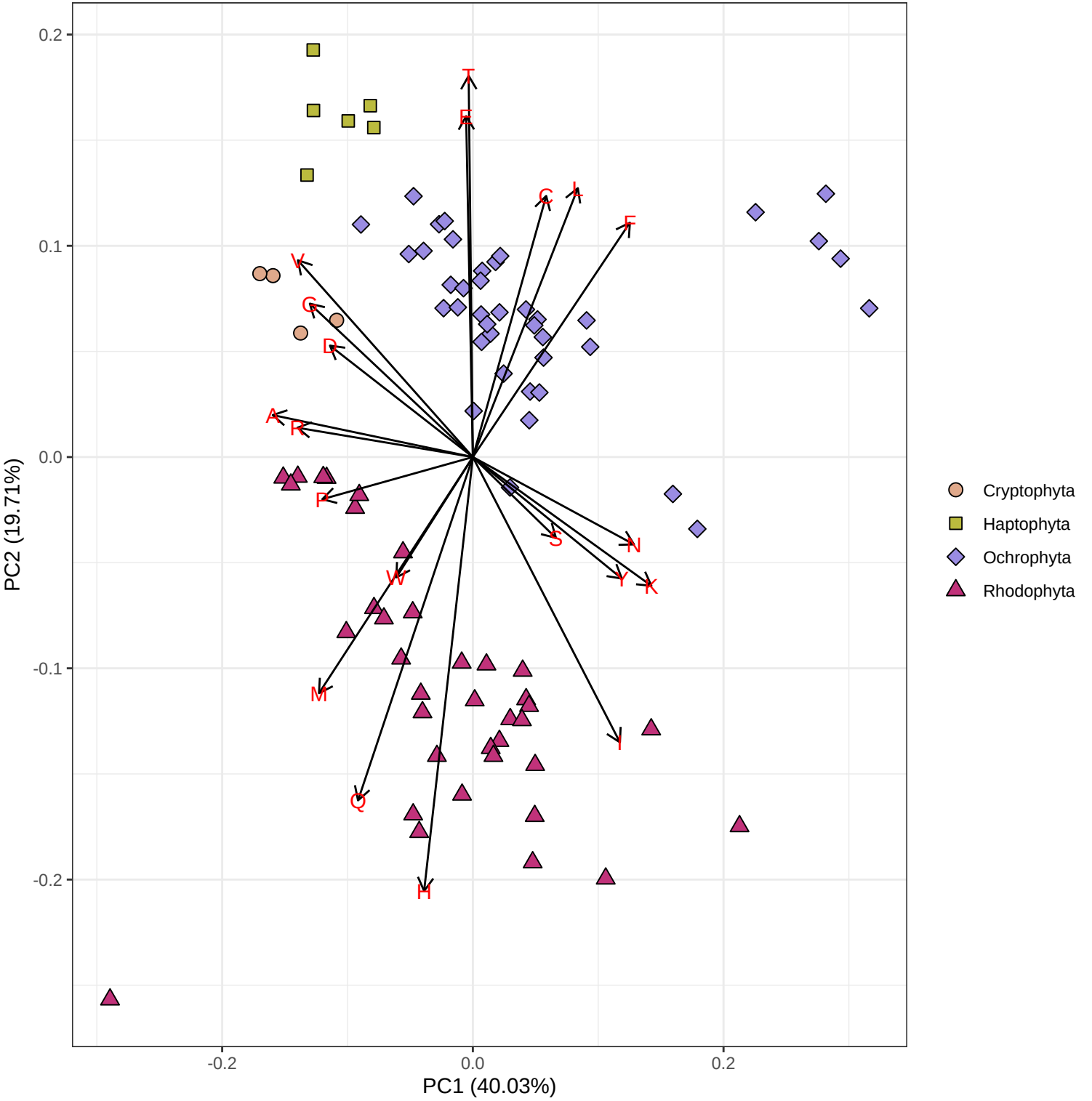

0.2
0.1
0.0
-0.1
-0.2
PC2 (19.71%)
-0.2
0.0
0.2
PC1 (40.03%)
T
E
C
I
F
V
G
D
A
R
P
S
N
W
Y
K
M
L
Q
H
Cryptophyta
Haptophyta
Ochrophyta
Rhodophyta

Fig. S12

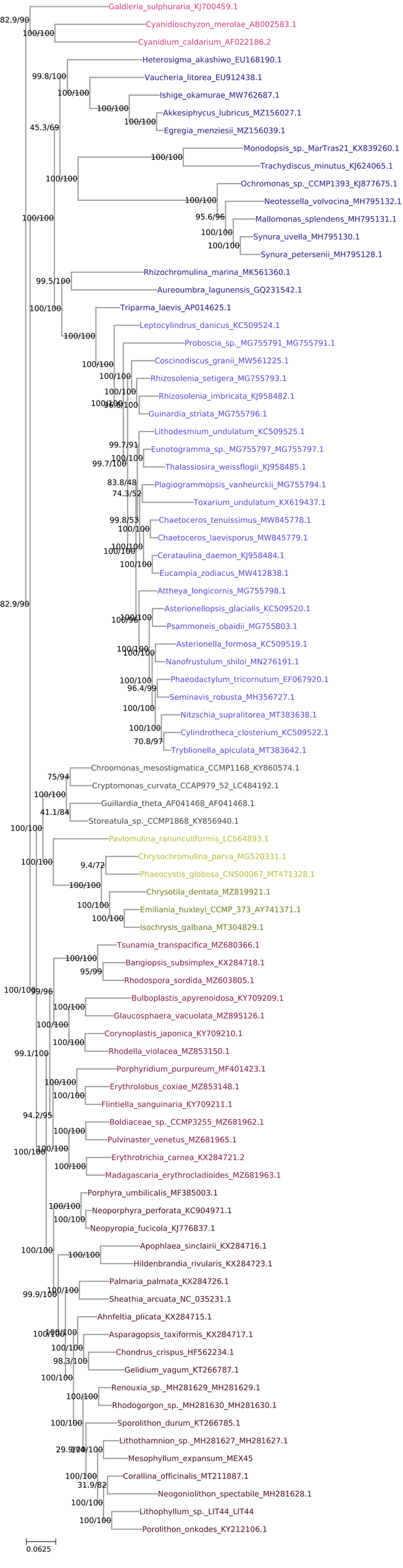

Galdieria_sulphuraria_KJ700459.1
Cyanidioschyzon_merolae_AB002583.1
Cyanidium_caldarium_AF022186.2
Heterosigma_akashiwo_EU168190.1
Vaucheria_litorea_EU912438.1
Ishige_okamurae_MW762687.1
Akkesiphycus_lubricus_MZ156027.1
Egregia_menziesii_MZ156039.1
Monodopsis_sp._MarTras21_KX839260.1
Trachydiscus_minutus_KJ624065.1
Ochromonas_sp._CCMP1393_KJ877675.1
Neotessella_volvocina_MH795132.1
Mallomonas_splendens_MH795131.1
Synura_uvella_MH795130.1
Synura_petersenii_MH795128.1
Rhizochromulina_marina_MK561360.1
Aureoumbra_lagunensis_GQ231542.1
Triparma_laevis_AP014625.1
Leptocylindrus_danicus_KC509524.1
Proboscia_sp._MG755791_MG755791.1
Coscinodiscus_granii_MW561225.1
Rhizosolenia_setigera_MG755793.1
Rhizosolenia_imbricata_KJ958482.1
Guinardia_striata_MG755796.1
Lithodesmium_undulatum_KC509525.1
Eunotogramma_sp._MG755797_MG755797.1
Thalassiosira_weissflogii_KJ958485.1
Plagiogrammopsis_vanheurckii_MG755794.1
Toxarium_undulatum_KX619437.1
Chaetoceros_tenuissimus_MW845778.1
Chaetoceros_laevisporus_MW845779.1
Cerataulina_daemon_KJ958484.1
Eucampia_zodiacus_MW412838.1
Attheya_longicornis_MG755798.1
Asterionellopsis_glacialis_KC509520.1
Psammoneis_obaidii_MG755803.1
Asterionella_formosa_KC509519.1
Nanofrustulum_shiloi_MN276191.1
Phaeodactylum_tricornutum_EF067920.1
Seminavis_robusta_MH356727.1
Nitzschia_supralitorea_MT383638.1
Cylindrotheca_closterium_KC509522.1
Tryblionella_apiculata_MT383642.1
Chroomonas_mesostigmatica_CCMP1168_KY860574.1
Cryptomonas_curvata_CCAP979_52_LC484192.1
Guillardia_theta_AF041468_AF041468.1
Storeatula_sp._CCMP1868_KY856940.1
Pavlomulina_ranunculiformis_LC564893.1
Chrysochromulina_parva_MG520331.1
Phaeocystis_globosa_CNS00067_MT471328.1
Chrysotila_dentata_MZ819921.1
Emiliania_huxleyi_CCMP_373_AY741371.1
Isochrysis_galbana_MT304829.1
Tsunamia_transpacifica_MZ680366.1
Bangiopsis_subsimplex_KX284718.1
Rhodospora_sordida_MZ603805.1
Bulboplastis_apyrenoidosa_KY709209.1
Glaucosphaera_vacuolata_MZ895126.1
Corynoplastis_japonica_KY709210.1
Rhodella_violacea_MZ853150.1
Porphyridium_purpureum_MF401423.1
Erythrolobus_coxiae_MZ853148.1
Flintiella_sanguinaria_KY709211.1
Boldiaceae_sp._CCMP3255_MZ681962.1
Pulvinaster_venetus_MZ681965.1
Erythrotrichia_carnea_KX284721.2
Madagascaria_erythrocladioides_MZ681963.1
Porphyra_umbilicalis_MF385003.1
Neoporphyra_perforata_KC904971.1
Neopyropia_fucicola_KJ776837.1
Apophlaea_sinclairii_KX284716.1
Hildenbrandia_rivularis_KX284723.1
Palmaria_palmata_KX284726.1
Sheathia_arcuata_NC_035231.1
Ahnfeltia_plicata_KX284715.1
Asparagopsis_taxiformis_KX284717.1
Chondrus_crispus_HF562234.1
Gelidium_vagum_KT266787.1
Renouxia_sp._MH281629_MH281629.1
Rhodogorgon_sp._MH281630_MH281630.1
Sporolithon_durum_KT266785.1
Lithothamnion_sp._MH281627_MH281627.1
Mesophyllum_expansum_MEX45
Corallina_officinalis_MT211887.1
Neogoniolithon_spectabile_MH281628.1
Lithophyllum_sp._LIT44_LIT44
Porolithon_onkodes_KY212106.1
0.0625

Fig. S13

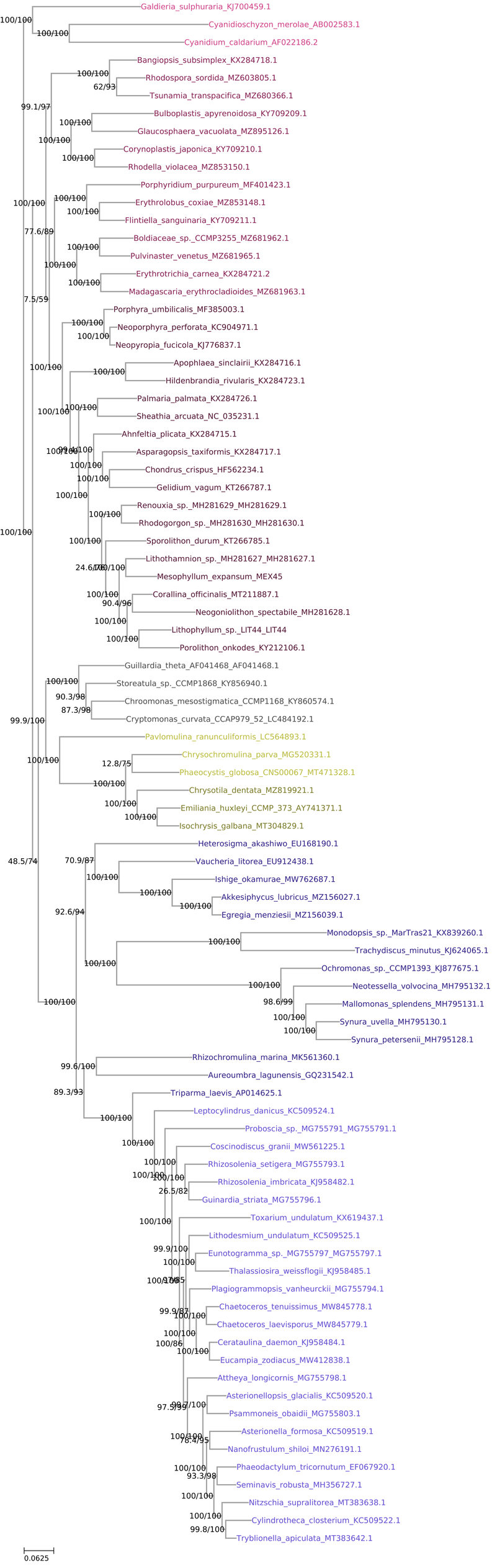

Galdieria_sulphuraria_KJ700459.1
Cyanidioschyzon_merolae_AB002583.1
Cyanidium_caldarium_AF022186.2
Bangiopsis_subsimplex_KX284718.1
Rhodospora_sordida_MZ603805.1
Tsunamia_transpacifica_MZ680366.1
Bulboplastis_apyrenoidosa_KY709209.1
Glaucosphaera_vacuolata_MZ895126.1
Corynoplastis_japonica_KY709210.1
Rhodella_violacea_MZ853150.1
Porphyridium_purpureum_MF401423.1
Erythrolobus_coxiae_MZ853148.1
Flintiella_sanguinaria_KY709211.1
Boldiaceae_sp._CCMP3255_MZ681962.1
Pulvinaster_venetus_MZ681965.1
Erythrotrichia_carnea_KX284721.2
Madagascaria_erythrocladioides_MZ681963.1
Porphyra_umbilicalis_MF385003.1
Neoporphyra_perforata_KC904971.1
Neopyropia_fucicola_KJ776837.1
Apophlaea_sinclairii_KX284716.1
Hildenbrandia_rivularis_KX284723.1
Palmaria_palmata_KX284726.1
Sheathia_arcuata_NC_035231.1
Ahnfeltia_plicata_KX284715.1
Asparagopsis_taxiformis_KX284717.1
Chondrus_crispus_HF562234.1
Gelidium_vagum_KT266787.1
Renouxia_sp._MH281629_MH281629.1
Rhodogorgon_sp._MH281630_MH281630.1
Sporolithon_durum_KT266785.1
Lithothamnion_sp._MH281627_MH281627.1
Mesophyllum_expansum_MEX45
Corallina_officinalis_MT211887.1
Neogoniolithon_spectabile_MH281628.1
Lithophyllum_sp._LIT44_LIT44
Porolithon_onkodes_KY212106.1
Guillardia_theta_AF041468_AF041468.1
Storeatula_sp._CCMP1868_KY856940.1
Chroomonas_mesostigmatica_CCMP1168_KY860574.1
Cryptomonas_curvata_CCAP979_52_LC484192.1
Pavlomulina_ranunculiformis_LC564893.1
Chrysochromulina_parva_MG520331.1
Phaeocystis_globosa_CNS00067_MT471328.1
Chrysotila_dentata_MZ819921.1
Emiliania_huxleyi_CCMP_373_AY741371.1
Isochrysis_galbana_MT304829.1
Heterosigma_akashiwo_EU168190.1
Vaucheria_litorea_EU912438.1
Ishige_okamurae_MW762687.1
Akkesiphycus_lubricus_MZ156027.1
Egregia_menziesii_MZ156039.1
Monodopsis_sp._MarTras21_KX839260.1
Trachydiscus_minutus_KJ624065.1
Ochromonas_sp._CCMP1393_KJ877675.1
Neotessella_volvocina_MH795132.1
Mallomonas_splendens_MH795131.1
Synura_uvella_MH795130.1
Synura_petersenii_MH795128.1
Rhizochromulina_marina_MK561360.1
Aureoumbra_lagunensis_GQ231542.1
Triparma_laevis_AP014625.1
Leptocylindrus_danicus_KC509524.1
Proboscia_sp._MG755791_MG755791.1
Coscinodiscus_granii_MW561225.1
Rhizosolenia_setigera_MG755793.1
Rhizosolenia_imbricata_KJ958482.1
Guinardia_striata_MG755796.1
Toxarium_undulatum_KX619437.1
Lithodesmium_undulatum_KC509525.1
Eunotogramma_sp._MG755797_MG755797.1
Thalassiosira_weissflogii_KJ958485.1
Plagiogrammopsis_vanheurckii_MG755794.1
Chaetoceros_tenuissimus_MW845778.1
Chaetoceros_laevisporus_MW845779.1
Cerataulina_daemon_KJ958484.1
Eucampia_zodiacus_MW412838.1
Attheya_longicornis_MG755798.1
Asterionellopsis_glacialis_KC509520.1
Psammoneis_obaidii_MG755803.1
Asterionella_formosa_KC509519.1
Nanofrustulum_shiloi_MN276191.1
Phaeodactylum_tricornutum_EF067920.1
Seminavis_robusta_MH356727.1
Nitzschia_supralitorea_MT383638.1
Cylindrotheca_closterium_KC509522.1
Tryblionella_apiculata_MT383642.1
0.0625

Fig. S14

Gloeochaete_wittrockiana_MF167426.1
Glaucocystis_incrassata_MF167425.1
Cyanophora_paradoxa_KM198929.1
Galdieria_sulphuraria_KJ700459.1
Cyanidium_caldarium_AF022186.2
Cyanidioschyzon_merolae_AB002583.1
Pavlomulina_ranunculiformis_LC564893.1
Minimum haptophyte2 180.4
Minimum haptophyte2 180.4
Minimum haptophyte2 180.4
Minimum haptophyte2 180.4
Minimum haptophyte2 180.4
Isochrysis_galbana_MT304829.1
Emiliania_huxleyi_CCMP_373_AY741371.1
Chrysotila_dentata_MZ819921.1
Maximum haptophyte_max1 252.2
Maximum haptophyte_max1 252.2
Maximum haptophyte_max1 252.2
Maximum haptophyte_max1 252.2
Maximum haptophyte_max1 252.2
Minimum haptophyte1 209.5
Minimum haptophyte1 209.5
Minimum haptophyte1 209.5
Minimum haptophyte1 209.5
Minimum haptophyte1 209.5
Phaeocystis_globosa_CNS00067_MT471328.1
Chrysochromulina_parva_MG520331.1
Storeatula_sp._CCMP1868_KY856940.1
Guillardia_theta_AF041468_AF041468.1
Cryptomonas_curvata_CCAP979_52_LC484192.1
Chroomonas_mesostigmatica_CCMP1168_KY860574.1
Rhizochromulina_marina_MK561360.1
Aureoumbra_lagunensis_GQ231542.1
Triparma_laevis_AP014625.1
Leptocylindrus_danicus_KC509524.1
Proboscia_sp._MG755791_MG755791.1
Maximum ochrophyte_max4 252.2
Maximum ochrophyte_max4 252.2
Maximum ochrophyte_max4 252.2
Maximum ochrophyte_max4 252.2
Maximum ochrophyte_max4 252.2
Minimum ochrophyte6 90.0
Minimum ochrophyte6 90.0
Minimum ochrophyte6 90.0
Minimum ochrophyte6 90.0
Minimum ochrophyte6 90.0
Rhizosolenia_setigera_MG755793.1
Rhizosolenia_imbricata_KJ958482.1
Guinardia_striata_MG755796.1
Coscinodiscus_granii_MW561225.1
Lithodesmium_undulatum_KC509525.1
Minimum ochrophyte4 65.9
Minimum ochrophyte4 65.9
Minimum ochrophyte4 65.9
Minimum ochrophyte4 65.9
Minimum ochrophyte4 65.9
Thalassiosira_weissflogii_KJ958485.1
Eunotogramma_sp._MG755797_MG755797.1
Toxarium_undulatum_KX619437.1
Plagiogrammopsis_vanheurckii_MG755794.1
Minimum ochrophyte1 145.0
Minimum ochrophyte1 145.0
Minimum ochrophyte1 145.0
Minimum ochrophyte1 145.0
Minimum ochrophyte1 145.0
Chaetoceros_tenuissimus_MW845778.1
Chaetoceros_laevisporus_MW845779.1
Eucampia_zodiacus_MW412838.1
Cerataulina_daemon_KJ958484.1
Attheya_longicornis_MG755798.1
Maximum root_max 3000.0
Maximum root_max 3000.0
Maximum root_max 3000.0
Maximum root_max 3000.0
Maximum root_max 3000.0
Minimum ochrophyte2 73.0
Minimum ochrophyte2 73.0
Minimum ochrophyte2 73.0
Minimum ochrophyte2 73.0
Minimum ochrophyte2 73.0
Psammoneis_obaidii_MG755803.1
Asterionellopsis_glacialis_KC509520.1
Maximum ochrophyte_max1 165.0
Maximum ochrophyte_max1 165.0
Maximum ochrophyte_max1 165.0
Maximum ochrophyte_max1 165.0
Maximum ochrophyte_max1 165.0
Maximum ochrophyte_max2 76.0
Maximum ochrophyte_max2 76.0
Maximum ochrophyte_max2 76.0
Maximum ochrophyte_max2 76.0
Maximum ochrophyte_max2 76.0
Seminavis_robusta_MH356727.1
Phaeodactylum_tricornutum_EF067920.1
Nitzschia_supralitorea_MT383638.1
Tryblionella_apiculata_MT383642.1
Cylindrotheca_closterium_KC509522.1
Minimum ochrophyte3 56.0
Minimum ochrophyte3 56.0
Minimum ochrophyte3 56.0
Minimum ochrophyte3 56.0
Minimum ochrophyte3 56.0
Nanofrustulum_shiloi_MN276191.1
Asterionella_formosa_KC509519.1
Minimum rhodophyte1 1030.0
Minimum rhodophyte1 1030.0
Minimum rhodophyte1 1030.0
Minimum rhodophyte1 1030.0
Minimum rhodophyte1 1030.0
Trachydiscus_minutus_KJ624065.1
Monodopsis_sp._MarTras21_KX839260.1
Ochromonas_sp._CCMP1393_KJ877675.1
Neotessella_volvocina_MH795132.1
Minimum ochrophyte5 46.4
Minimum ochrophyte5 46.4
Minimum ochrophyte5 46.4
Minimum ochrophyte5 46.4
Minimum ochrophyte5 46.4
Synura_uvella_MH795130.1
Synura_petersenii_MH795128.1
Mallomonas_splendens_MH795131.1
Heterosigma_akashiwo_EU168190.1
Vaucheria_litorea_EU912438.1
Ishige_okamurae_MW762687.1
Egregia_menziesii_MZ156039.1
Akkesiphycus_lubricus_MZ156027.1
Porphyridium_purpureum_MF401423.1
Flintiella_sanguinaria_KY709211.1
Erythrolobus_coxiae_MZ853148.1
Madagascaria_erythrocladioides_MZ681963.1
Erythrotrichia_carnea_KX284721.2
Pulvinaster_venetus_MZ681965.1
Boldiaceae_sp._CCMP3255_MZ681962.1
Rhodella_violacea_MZ853150.1
Corynoplastis_japonica_KY709210.1
Glaucosphaera_vacuolata_MZ895126.1
Bulboplastis_apyrenoidosa_KY709209.1
Tsunamia_transpacifica_MZ680366.1
Rhodospora_sordida_MZ603805.1
Bangiopsis_subsimplex_KX284718.1
Porphyra_umbilicalis_MF385003.1
Neopyropia_fucicola_KJ776837.1
Neoporphyra_perforata_KC904971.1
Hildenbrandia_rivularis_KX284723.1
Apophlaea_sinclairii_KX284716.1
Sheathia_arcuata_NC_035231.1
Palmaria_palmata_KX284726.1
Maximum rhodophyte_max1 635.8
Maximum rhodophyte_max1 635.8
Maximum rhodophyte_max1 635.8
Maximum rhodophyte_max1 635.8
Maximum rhodophyte_max1 635.8
Rhodogorgon_sp._MH281630_MH281630.1
Renouxia_sp._MH281629_MH281629.1
Minimum rhodophyte3 130.3
Minimum rhodophyte3 130.3
Minimum rhodophyte3 130.3
Minimum rhodophyte3 130.3
Minimum rhodophyte3 130.3
Sporolithon_durum_KT266785.1
Mesophyllum_expansum_MEX45
Lithothamnion_sp._MH281627_MH281627.1
Minimum rhodophyte4 63.8
Minimum rhodophyte4 63.8
Minimum rhodophyte4 63.8
Minimum rhodophyte4 63.8
Minimum rhodophyte4 63.8
Porolithon_onkodes_KY212106.1
Lithophyllum_sp._LIT44_LIT44
Neogoniolithon_spectabile_MH281628.1
Corallina_officinalis_MT211887.1
Minimum rhodophyte2 550.4
Minimum rhodophyte2 550.4
Minimum rhodophyte2 550.4
Minimum rhodophyte2 550.4
Minimum rhodophyte2 550.4
Gelidium_vagum_KT266787.1
Chondrus_crispus_HF562234.1
Asparagopsis_taxiformis_KX284717.1
Ahnfeltia_plicata_KX284715.1
Chlorokybus_atmophyticus_DQ422812.2
Entransia_fimbriata_KU646490.1
Chara_vulgaris_DQ229107.1
Chaetosphaeridium_globosum_AF494278.1
Mesotaenium_endlicherianum_KJ461682.1
Cosmarium_botrytis_KU646492.1
Leiosporoceros_dussii_MH577299.1
Anthoceros_agrestis_MK087646.1
Minimum viridiplantae10 405.0
Minimum viridiplantae10 405.0
Minimum viridiplantae10 405.0
Minimum viridiplantae10 405.0
Minimum viridiplantae10 405.0
Schistochila_macrodonta_MH064506.1
Ptilidium_pulcherrimum_HM222519.1
Sphaerocarpos_texanus_MT023020.1
Minimum viridiplantae12 226.4
Minimum viridiplantae12 226.4
Minimum viridiplantae12 226.4
Minimum viridiplantae12 226.4
Minimum viridiplantae12 226.4
Wiesnerella_denudata_MT712073.1
Marchantia_polymorpha_MN635409.1
Lunularia_cruciata_MW429511.1
Minimum viridiplantae2 466.2
Minimum viridiplantae2 466.2
Minimum viridiplantae2 466.2
Minimum viridiplantae2 466.2
Minimum viridiplantae2 466.2
Haplomitrium_blumei_MH064516.1
Minimum viridiplantae13 426.9
Minimum viridiplantae13 426.9
Minimum viridiplantae13 426.9
Minimum viridiplantae13 426.9
Minimum viridiplantae13 426.9
Sphagnum_wulfianum_KU725459.1
Eosphagnum_sp._BIS-2016_KU725456.1
Tetraphis_pellucida_KJ817846.1
Polytrichum_commune_MW528408.1
Buxbaumia_aphylla_MN496310.1
Syntrichia_filaris_MK852705.1
Bartramia_pomiformis_MT024676.1
Andreaea_rothii_MW561627.1
Minimum viridiplantae11 330.7
Minimum viridiplantae11 330.7
Minimum viridiplantae11 330.7
Minimum viridiplantae11 330.7
Minimum viridiplantae11 330.7
Maximum viridiplantae_max1 1042.0
Maximum viridiplantae_max1 1042.0
Maximum viridiplantae_max1 1042.0
Maximum viridiplantae_max1 1042.0
Maximum viridiplantae_max1 1042.0
Maximum viridiplantae_max2 515.5
Maximum viridiplantae_max2 515.5
Maximum viridiplantae_max2 515.5
Maximum viridiplantae_max2 515.5
Maximum viridiplantae_max2 515.5
Minimum viridiplantae4 392.1
Minimum viridiplantae4 392.1
Minimum viridiplantae4 392.1
Minimum viridiplantae4 392.1
Minimum viridiplantae4 392.1
Lycopodium_clavatum_MH549642.1
Selaginella_lepidophylla_MK089531.1
Isoetes_malinverniana_MH549640.1
Minimum viridiplantae6 384.7
Minimum viridiplantae6 384.7
Minimum viridiplantae6 384.7
Minimum viridiplantae6 384.7
Minimum viridiplantae6 384.7
Tmesipteris_elongata_KJ569699.1
Mankyua_chejuensis_KP205433.1
Equisetum_arvense_GU191334.1
Minimum viridiplantae3 420.7
Minimum viridiplantae3 420.7
Minimum viridiplantae3 420.7
Minimum viridiplantae3 420.7
Minimum viridiplantae3 420.7
Minimum viridiplantae9 318.7
Minimum viridiplantae9 318.7
Minimum viridiplantae9 318.7
Minimum viridiplantae9 318.7
Minimum viridiplantae9 318.7
Hymenophyllum_coreanum_MN583335.1
Danaea_sellowiana_MN412588.1
Christensenia_aesculifolia_MN056350.1
Minimum viridiplantae5 385.5
Minimum viridiplantae5 385.5
Minimum viridiplantae5 385.5
Minimum viridiplantae5 385.5
Minimum viridiplantae5 385.5
Maximum viridiplantae_max4 321.4
Maximum viridiplantae_max4 321.4
Maximum viridiplantae_max4 321.4
Maximum viridiplantae_max4 321.4
Maximum viridiplantae_max4 321.4
Minimum viridiplantae7 308.1
Minimum viridiplantae7 308.1
Minimum viridiplantae7 308.1
Minimum viridiplantae7 308.1
Minimum viridiplantae7 308.1
Taxus_baccata_KR476375.1
Sciadopitys_verticillata_KT601210.1
Welwitschia_mirabilis_EU342371.1
Ephedra_distachya_MG594460.1
Maximum viridiplantae_max3 365.6
Maximum viridiplantae_max3 365.6
Maximum viridiplantae_max3 365.6
Maximum viridiplantae_max3 365.6
Maximum viridiplantae_max3 365.6
Ginkgo_biloba_JN867585.1
Cycas_debaoensis_KU743927.1
Amborella_trichopoda_AJ506156.2
Ceratophyllum_demersum_EF614270.1
Chloranthus_henryi_MK922064.1
Maximum viridiplantae_max5 247.2
Maximum viridiplantae_max5 247.2
Maximum viridiplantae_max5 247.2
Maximum viridiplantae_max5 247.2
Maximum viridiplantae_max5 247.2
Minimum viridiplantae17 125.9
Minimum viridiplantae17 125.9
Minimum viridiplantae17 125.9
Minimum viridiplantae17 125.9
Minimum viridiplantae17 125.9
Tasmannia_lanceolata_MT555078.1
Saruma_henryi_MG520100.1
Cinnamomum_aromaticum_MN698963.1
Minimum viridiplantae18 85.8
Minimum viridiplantae18 85.8
Minimum viridiplantae18 85.8
Minimum viridiplantae18 85.8
Minimum viridiplantae18 85.8
Chimonanthus_grammatus_MW166217.1
Calycanthus_chinensis_MG561304.1
Minimum viridiplantae19 112.6
Minimum viridiplantae19 112.6
Minimum viridiplantae19 112.6
Minimum viridiplantae19 112.6
Minimum viridiplantae19 112.6
Myristica_yunnanensis_MN495964.1
Magnolia_virginiana_MN990608.1
Annona_reticulata_MT742547.1
Minimum viridiplantae16 119.6
Minimum viridiplantae16 119.6
Minimum viridiplantae16 119.6
Minimum viridiplantae16 119.6
Minimum viridiplantae16 119.6
Corydalis_chrysosphaera_MZ983401.1
Vicia_faba_MT120813.1
Apium_graveolens_MK036045.1
Minimum viridiplantae8 125.9
Minimum viridiplantae8 125.9
Minimum viridiplantae8 125.9
Minimum viridiplantae8 125.9
Minimum viridiplantae8 125.9
Lilium_candidum_MK753244.1
Acorus_calamus_AJ879453.1
Prasinoderma_coloniale_KJ746598.1
Micromonas_sp._RCC299_FJ858267.1
Pyramimonas_parkeae_KX013546.1
Strombomonas_acuminata_JN674637.1
Eutreptiella_gymnastica_NC_017754.2
Nephroselmis_pyriformis_MW077730.1
Chloropicon_sieburthii_MK085986.1
Chlamydomonas_reinhardtii_BK000554.2
Scherffelia_dubia_KU167098.1
Pedinomonas_minor_FJ968740.1
Planctonema_lauterbornii_KM462880.1
Gymnochlora_stellata_AP014947.1
Lotharella_sp._CCMP622_KF438023.1
Bigelowiella_natans_DQ851108.1
Minimum viridiplantae20 940.4
Minimum viridiplantae20 940.4
Minimum viridiplantae20 940.4
Minimum viridiplantae20 940.4
Minimum viridiplantae20 940.4
Maximum viridiplantae_max7 635.8
Maximum viridiplantae_max7 635.8
Maximum viridiplantae_max7 635.8
Maximum viridiplantae_max7 635.8
Maximum viridiplantae_max7 635.8
Ostreobium_quekettii_LT593849.1
Bryopsis_plumosa_LN810504.1
Minimum viridiplantae15 209.5
Minimum viridiplantae15 209.5
Minimum viridiplantae15 209.5
Minimum viridiplantae15 209.5
Minimum viridiplantae15 209.5
Rhipilia_penicilloides_KY819065.1
Pseudocodium_devriesii_MH591109.1
Halimeda_minima_MH591101.1
Minimum viridiplantae14 271.8
Minimum viridiplantae14 271.8
Minimum viridiplantae14 271.8
Minimum viridiplantae14 271.8
Minimum viridiplantae14 271.8
Caulerpa_okamurae_KX809677.1
Avrainvillea_mazei_KY509313.1
Ulva_compressa_MW548841.1
Trichosarcina_mucosa_KY407656.1
Ignatius_tetrasporus_KY407659.1
Minimum viridiplantae1 439.2
Minimum viridiplantae1 439.2
Minimum viridiplantae1 439.2
Minimum viridiplantae1 439.2
Minimum viridiplantae1 439.2
Neodangemannia_microcystis_KY407660.1
Maximum viridiplantae_max6 635.8
Maximum viridiplantae_max6 635.8
Maximum viridiplantae_max6 635.8
Maximum viridiplantae_max6 635.8
Maximum viridiplantae_max6 635.8
Parvocaulis_polyphysoides_XXXK
Neomeris_sp._HV02668_XXXB
Bornetella_oligospora_XXXI
Dasycladus_sp._HV04045_XXXC
Batophora_oerstedii_XXXF
Acetabularia_acetabulum_XXXA
Halicoryne_sp._HV04044_XXXD
0.722951

Fig. S15

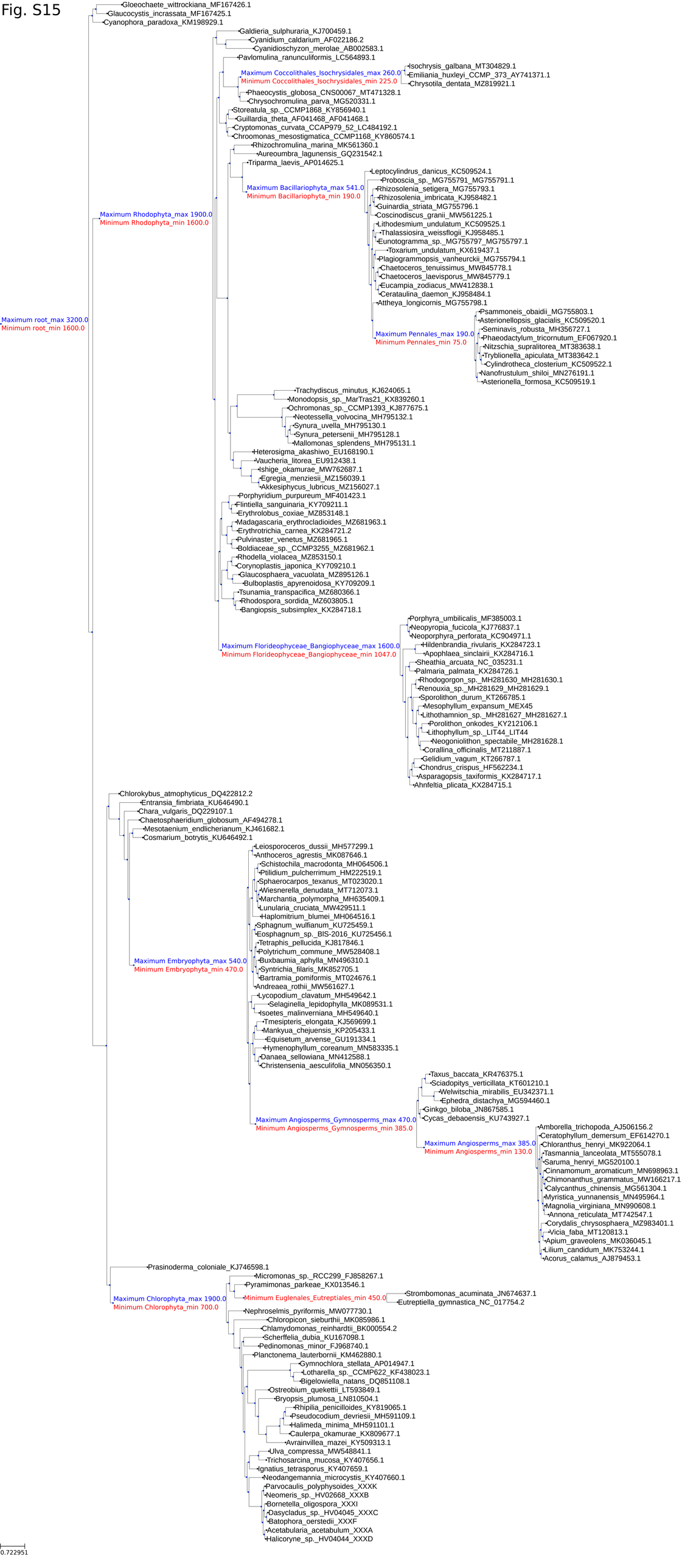

Gloeochaete_wittrockiana_MF167426.1
Glaucocystis_incrassata_MF167425.1
Cyanophora_paradoxa_KM198929.1
Maximum root_max 3200.0
Minimum root_min 1600.0
Galdieria_sulphuraria_KJ700459.1
Cyanidium_caldarium_AF022186.2
Cyanidioschyzon_merolae_AB002583.1
Pavlomulina_ranunculiformis_LC564893.1
Maximum Coccolithales_Isochrysidales_max 260.0
Minimum Coccolithales_Isochrysidales_min 225.0
Isochrysis_galbana_MT304829.1
Emiliania_huxleyi_CCMP_373_AY741371.1
Chrysotila_dentata_MZ819921.1
Phaeocystis_globosa_CNS00067_MT471328.1
Chrysochromulina_parva_MG520331.1
Storeatula_sp._CCMP1868_KY856940.1
Guillardia_theta_AF041468_AF041468.1
Cryptomonas_curvata_CCAP979_52_LC484192.1
Chroomonas_mesostigmatica_CCMP1168_KY860574.1
Rhizochromulina_marina_MK561360.1
Aureoumbra_lagunensis_GQ231542.1
Triparma_laevis_AP014625.1
Maximum Rhodophyta_max 1900.0
Minimum Rhodophyta_min 1600.0
Maximum Bacillariophyta_max 541.0
Minimum Bacillariophyta_min 190.0
Leptocylindrus_danicus_KC509524.1
Proboscia_sp._MG755791_MG755791.1
Rhizosolenia_setigera_MG755793.1
Rhizosolenia_imbricata_KJ958482.1
Guinardia_striata_MG755796.1
Coscinodiscus_granii_MW561225.1
Lithodesmium_undulatum_KC509525.1
Thalassiosira_weissflogii_KJ958485.1
Eunotogramma_sp._MG755797_MG755797.1
Toxarium_undulatum_KX619437.1
Plagiogrammopsis_vanheurckii_MG755794.1
Chaetoceros_tenuissimus_MW845778.1
Chaetoceros_laevisporus_MW845779.1
Eucampia_zodiacus_MW412838.1
Cerataulina_daemon_KJ958484.1
Attheya_longicornis_MG755798.1
Maximum Pennales_max 190.0
Minimum Pennales_min 75.0
Psammoneis_obaidii_MG755803.1
Asterionellopsis_glacialis_KC509520.1
Seminavis_robusta_MH356727.1
Phaeodactylum_tricornutum_EF067920.1
Nitzschia_supralitorea_MT383638.1
Tryblionella_apiculata_MT383642.1
Cylindrotheca_closterium_KC509522.1
Nanofrustulum_shiloi_MN276191.1
Asterionella_formosa_KC509519.1
Trachydiscus_minutus_KJ624065.1
Monodopsis_sp._MarTras21_KX839260.1
Ochromonas_sp._CCMP1393_KJ877675.1
Neotessella_volvocina_MH795132.1
Synura_uvella_MH795130.1
Synura_petersenii_MH795128.1
Mallomonas_splendens_MH795131.1
Heterosigma_akashiwo_EU168190.1
Vaucheria_litorea_EU912438.1
Ishige_okamurae_MW762687.1
Egregia_menziesii_MZ156039.1
Akkesiphycus_lubricus_MZ156027.1
Porphyridium_purpureum_MF401423.1
Flintiella_sanguinaria_KY709211.1
Erythrolobus_coxiae_MZ853148.1
Madagascaria_erythrocladioides_MZ681963.1
Erythrotrichia_carnea_KX284721.2
Pulvinaster_venetus_MZ681965.1
Boldiaceae_sp._CCMP3255_MZ681962.1
Rhodella_violacea_MZ853150.1
Corynoplastis_japonica_KY709210.1
Glaucosphaera_vacuolata_MZ895126.1
Bulboplastis_apyrenoidosa_KY709209.1
Tsunamia_transpacifica_MZ680366.1
Rhodospora_sordida_MZ603805.1
Bangiopsis_subsimplex_KX284718.1
Maximum Florideophyceae_Bangiophyceae_max 1600.0
Minimum Florideophyceae_Bangiophyceae_min 1047.0
Porphyra_umbilicalis_MF385003.1
Neopyropia_fucicola_KJ776837.1
Neoporphyra_perforata_KC904971.1
Hildenbrandia_rivularis_KX284723.1
Apophlaea_sinclairii_KX284716.1
Sheathia_arcuata_NC_035231.1
Palmaria_palmata_KX284726.1
Rhodogorgon_sp._MH281630_MH281630.1
Renouxia_sp._MH281629_MH281629.1
Sporolithon_durum_KT266785.1
Mesophyllum_expansum_MEX45
Lithothamnion_sp._MH281627_MH281627.1
Porolithon_onkodes_KY212106.1
Lithophyllum_sp._LIT44_LIT44
Neogoniolithon_spectabile_MH281628.1
Corallina_officinalis_MT211887.1
Gelidium_vagum_KT266787.1
Chondrus_crispus_HF562234.1
Asparagopsis_taxiformis_KX284717.1
Ahnfeltia_plicata_KX284715.1
Chlorokybus_atmophyticus_DQ422812.2
Entransia_fimbriata_KU646490.1
Chara_vulgaris_DQ229107.1
Chaetosphaeridium_globosum_AF494278.1
Mesotaenium_endlicherianum_KJ461682.1
Cosmarium_botrytis_KU646492.1
Maximum Embryophyta_max 540.0
Minimum Embryophyta_min 470.0
Leiosporoceros_dussii_MH577299.1
Anthoceros_agrestis_MK087646.1
Schistochila_macrodonta_MH064506.1
Ptilidium_pulcherrimum_HM222519.1
Sphaerocarpos_texanus_MT023020.1
Wiesnerella_denudata_MT712073.1
Marchantia_polymorpha_MH635409.1
Lunularia_cruciata_MW429511.1
Haplomitrium_blumei_MH064516.1
Sphagnum_wulfianum_KU725459.1
Eosphagnum_sp._BIS-2016_KU725456.1
Tetraphis_pellucida_KJ817846.1
Polytrichum_commune_MW528408.1
Buxbaumia_aphylla_MN496310.1
Syntrichia_filaris_MK852705.1
Bartramia_pomiformis_MT024676.1
Andreaea_rothii_MW561627.1
Lycopodium_clavatum_MH549642.1
Selaginella_lepidophylla_MK089531.1
Isoetes_malinverniana_MH549640.1
Tmesipteris_elongata_KJ569699.1
Mankyua_chejuensis_KP205433.1
Equisetum_arvense_GU191334.1
Hymenophyllum_coreanum_MN583335.1
Danaea_sellowiana_MN412588.1
Christensenia_aesculifolia_MN056350.1
Maximum Angiosperms_Gymnosperms_max 470.0
Minimum Angiosperms_Gymnosperms_min 385.0
Taxus_baccata_KR476375.1
Sciadopitys_verticillata_KT601210.1
Welwitschia_mirabilis_EU342371.1
Ephedra_distachya_MG594460.1
Ginkgo_biloba_JN867585.1
Cycas_debaoensis_KU743927.1
Maximum Angiosperms_max 385.0
Minimum Angiosperms_min 130.0
Amborella_trichopoda_AJ506156.2
Ceratophyllum_demersum_EF614270.1
Chloranthus_henryi_MK922064.1
Tasmannia_lanceolata_MT555078.1
Saruma_henryi_MG520100.1
Cinnamomum_aromaticum_MN698963.1
Chimonanthus_grammatus_MW166217.1
Calycanthus_chinensis_MG561304.1
Myristica_yunnanensis_MN495964.1
Magnolia_virginiana_MN990608.1
Annona_reticulata_MT742547.1
Corydalis_chrysosphaera_MZ983401.1
Vicia_faba_MT120813.1
Apium_graveolens_MK036045.1
Lilium_candidum_MK753244.1
Acorus_calamus_AJ879453.1
Prasinoderma_coloniale_KJ746598.1
Micromonas_sp._RCC299_FJ858267.1
Pyramimonas_parkeae_KX013546.1
Maximum Chlorophyta_max 1900.0
Minimum Chlorophyta_min 700.0
Minimum Euglenales_Eutreptiales_min 450.0
Strombomonas_acuminata_JN674637.1
Eutreptiella_gymnastica_NC_017754.2
Nephroselmis_pyriformis_MW077730.1
Chloropicon_sieburthii_MK085986.1
Chlamydomonas_reinhardtii_BK000554.2
Scherffelia_dubia_KU167098.1
Pedinomonas_minor_FJ968740.1
Planctonema_lauterbornii_KM462880.1
Gymnochlora_stellata_AP014947.1
Lotharella_sp._CCMP622_KF438023.1
Bigelowiella_natans_DQ851108.1
Ostreobium_quekettii_LT593849.1
Bryopsis_plumosa_LN810504.1
Rhipilia_penicilloides_KY819065.1
Pseudocodium_devriesii_MH591109.1
Halimeda_minima_MH591101.1
Caulerpa_okamurae_KX809677.1
Avrainvillea_mazei_KY509313.1
Ulva_compressa_MW548841.1
Trichosarcina_mucosa_KY407656.1
Ignatius_tetrasporus_KY407659.1
Neodangemannia_microcystis_KY407660.1
Parvocaulis_polyphysoides_XXXK
Neomeris_sp._HV02668_XXXB
Bornetella_oligospora_XXXI
Dasycladus_sp._HV04045_XXXC
Batophora_oerstedii_XXXF
Acetabularia_acetabulum_XXXA
Halicoryne_sp._HV04044_XXXD
0.722951

Fig. S16

Haptophyta_environmental_rbcL_Q946S4
Haptophyta_environmental_rbcL_Q946L7
Haptophyta_environmental_rbcL_Q946S3
Haptophyta_environmental_rbcL_Q946S2
Haptophyta_environmental_rbcL_Q93WD4
Haptophyceae_sp._rbcL_A0A7R7AKA1
Exanthemachrysis_gayraliae_rbcL_Q9GFW9
Rebecca_salina_rbcL_Q37266
Pavlova_pinguis_rbcL_K4ERQ5
Pavlova_gyrans_rbcL_Q8MA96
Haptophyta_environmental_rbcL_Q946S1
(branch length shorted for readability)
Diacronema_lutheri_rbcL_Q8MA95
Diacronema_lutheri_rbcL_K4ESR4
Pavlova_sp._rbcL_A0A7D3U365
Diacronema_lutheri_rbcL_M1K554
Other Haptophyta
Chrysocampanula_spinifera_rbcL_Q9GFX0
Haptophyceae_sp._rbcL_G9M5I9
Cruciplacolithus_neohelis_rbcL_Q9GFY0
Haptophyta_environmental_rbcL_Q946K1
Haptophyta_environmental_rbcL_Q946K2
Haptophyta_environmental_rbcL_Q946K4
Calcidiscus_leptoporus_rbcL_Q9GFX9
Coccolithus_pelagicus_rbcL_K4ERQ6
Coccolithus_pelagicus_rbcL_Q9TK49
Haptophyceae_sp._rbcL_G9M5H6
Haptophyceae_sp._rbcL_G9M5L3
Haptophyceae_sp._rbcL_G9M5B7
Haptophyceae_sp._rbcL_G9M5D4
Haptophyceae_sp._rbcL_G9M5G6
Coccolithales
Haptophyceae_sp._rbcL_G9M5F3
Haptophyceae_sp._rbcL_G9M5J0
Haptophyceae_sp._rbcL_G9M5G5
Chrysotila_scherffelii_rbcL_D3JAI8
Pleurochrysis_haptonemofera_rbcL_Q9GFY1
Chrysotila_dentata_MZ819921.1|rbcL__101422_102889
C. dentata plastid genome annotated in this study (MZ819921.1)
Chrysotila_carterae_rbcL_Q08051
Coccolithus_pelagicus_rbcL_A7UMF6
Pleurochrysis_elongata_rbcL_K4ERE8
Chrysotila_placolithoides_rbcL_K4ES22
Calyptrosphaera_sphaeroidea_rbcL_P48687
Calyptrosphaera_sphaeroidea_rbcL_Q54AE2
Umbilicosphaera_sibogae_rbcL_Q7JCS8
Umbilicosphaera_sibogae_rbcL_Q37290
Helicosphaera_carteri_rbcL_Q9GFX8 (Zygodiscales)
Chrysotila_lamellosa_rbcL_K4ERE7
Isochrysis_galbana_rbcL_K4ESA9
Isochrysis_sp._rbcL_Q8MA97
Haptophyta_sp._rbcL_O20380
Isochrysis_galbana_rbcL_Q9GFX7
Isochrysis_galbana_rbcL_A0A7D4X9S6
Tisochrysis_lutea_rbcL_X5CVA0
Isochrysidales
Haptophyceae_sp._rbcL_G9M5L6
Haptophyceae_sp._rbcL_G9M5G8
Haptophyceae_sp._rbcL_G9M5E1
Haptophyceae_sp._rbcL_G9M5K2
Haptophyceae_sp._rbcL_G9M5F4
Haptophyceae_sp._rbcL_G9M5L8
Haptophyceae_sp._rbcL_G9M5C5
Chrysochromulina_tobinii_rbcL_A0A2H4C5W6
Chrysochromulina_parva_rbcL_A0A2H4C5X3
Chrysochromulina_parva_rbcL_Q9GFX6
Chrysochromulina_parva_rbcL_A0A2I7YW98
Chrysochromulina_tobinii_rbcL_A0A075DW54
Haptophyceae_sp._rbcL_G9M5D5
Haptophyceae_sp._rbcL_G9M5E4
Haptophyta_environmental_rbcL_Q946L3
Haptophyceae_sp._rbcL_G9M5D6
Haptophyceae_sp._rbcL_G9M5I0
Haptophyceae_sp._rbcL_G9M5F7
Haptophyceae_sp._rbcL_G9M5H7
Haptophyta_environmental_rbcL_Q946L6
Haptophyceae_sp._rbcL_G9M5J9
Haptophyta_environmental_rbcL_Q946M1
Haptophyceae_sp._rbcL_G9M5M6
Haptophyceae_sp._rbcL_G9M5L4
Haptophyceae_sp._rbcL_G9M5E9
Haptophyceae_sp._rbcL_G9M5B6
Chrysochromulina_alifera_rbcL_Q9GFX5
Haptophyceae_sp._rbcL_G9M5L1
Haptophyceae_sp._rbcL_G9M5D8
Haptophyceae_sp._rbcL_G9M5J2
Haptophyceae_sp._rbcL_G9M5E2
Haptophyceae_sp._rbcL_G9M5H2
Haptophyceae_sp._rbcL_G9M5L2
Haptophyceae_sp._rbcL_G9M5E8
Haptophyceae_sp._rbcL_G9M5F6
Haptophyceae_sp._rbcL_G9M5C6
Chrysochromulina_sp._rbcL_D7UPR7
Haptophyceae_sp._rbcL_G9M5J4
Haptophyceae_sp._rbcL_G9M5M5
Haptophyceae_sp._rbcL_G9M5B0
Haptophyceae_sp._rbcL_G9M5G4
Chrysochromulina_simplex_rbcL_D7UPR9
Haptophyta_environmental_rbcL_Q946L4
Haptophyta_environmental_rbcL_Q946K6
Haptophyceae_sp._rbcL_G9M5E0
Haptophyceae_sp._rbcL_G9M5K3
Haptophyceae_sp._rbcL_G9M5L0
Haptophyta_environmental_rbcL_Q946K9
Haptophyceae_sp._rbcL_G9M5C4
Haptophyceae_sp._rbcL_G9M5H3
Haptophyceae_sp._rbcL_G9M5I4
Haptophyceae_sp._rbcL_G9M5B5
Haptophyceae_sp._rbcL_G9M5G3
Haptophyceae_sp._rbcL_G9M5D1
Haptophyceae_gen._rbcL_Q9GGR2
Haptophyceae_sp._rbcL_G9M5I5
Haptophyceae_sp._rbcL_G9M5H9
Haptophyceae_sp._rbcL_G9M5F8
Haptophyceae_sp._rbcL_G9M5M3
Haptophyceae_sp._rbcL_G9M5I8
Haptophyceae_sp._rbcL_G9M5G0
Haptophyceae_sp._rbcL_G9M5B2
Haptophyceae_sp._rbcL_G9M5K7
Haptophyceae_sp._rbcL_G9M5C7
Haptophyta_environmental_rbcL_Q946K5
Haptophyceae_sp._rbcL_G9M5K6
Haptophyceae_sp._rbcL_G9M5H1
Haptophyceae_sp._rbcL_G9M5C1
Haptophyceae_sp._rbcL_G9M5B3
Haptophyceae_sp._rbcL_G9M5C9
Haptophyceae_sp._rbcL_G9M5J6
Haptophyceae_sp._rbcL_G9M5I1
Haptophyceae_sp._rbcL_G9M5J5
Haptophyceae_sp._rbcL_G9M5F2
Haptophyceae_gen._rbcL_Q9GGR1
Haptophyceae_gen._rbcL_Q9GGR7
Haptophyceae_sp._rbcL_G9M5E3
Haptophyceae_sp._rbcL_G9M5K4
Haptophyceae_sp._rbcL_G9M5E7
Haptophyceae_sp._rbcL_G9M5D2
Haptophyceae_sp._rbcL_G9M5I7
Haptophyceae_sp._rbcL_G9M5J8
Prymnesium_parvum_rbcL_K4ESB0
Haptophyceae_sp._rbcL_G9M5J7
Haptophyceae_sp._rbcL_G9M5J1
Platychrysis_sp._rbcL_Q9GFX1
Prymnesium_parvum_rbcL_Q9GFX2
Chrysochromulina_sp._rbcL_Q9GFX3
Haptophyceae_sp._rbcL_G9M5F9
Haptophyceae_sp._rbcL_G9M5I3
Haptophyceae_sp._rbcL_G9M5I2
Haptophyceae_sp._rbcL_G9M5H0
Haptophyceae_sp._rbcL_G9M5B4
Haptophyta_environmental_rbcL_Q946L0
Prymnesium_neolepis_rbcL_Q3LHM7
Haptophyceae_sp._rbcL_G9M5F5
Haptophyta_environmental_rbcL_Q946L8
Haptophyceae_sp._rbcL_G9M5D7
Haptolina_hirta_rbcL_P48692
Haptolina_hirta_rbcL_Q54AE1
Haptophyceae_sp._rbcL_G9M5C0
Dicrateria_rotunda_rbcL_Q9GFX4
Haptophyceae_sp._rbcL_G9M5K5
Haptophyceae_sp._rbcL_G9M5B9
Chrysoculter_rhomboideus_rbcL_Q6I6Y4
Haptophyta_environmental_rbcL_Q946L2
Haptophyceae_sp._rbcL_G9M5D9
Haptophyceae_sp._rbcL_G9M5B1
Haptophyceae_sp._rbcL_G9M5H4
Haptophyta_environmental_rbcL_Q946L9
Gephyrocapsa_oceanica_rbcL_V9NGF6
Emiliania_huxleyi_rbcL_L7QI74
Emiliania_huxleyi_rbcL_Q4G3F4
Emiliania_huxleyi_rbcL_H9LS33
Gephyrocapsa_oceanica_rbcL_Q37210
Emiliania_huxleyi_rbcL_Q37197
Haptophyta_environmental_rbcL_Q946L1
Emiliania_huxleyi_rbcL_V9NF84
Isochrysidales
Haptophyta_environmental_rbcL_Q946K7
Phaeocystis_jahnii_rbcL_A0A0D5CE41
Phaeocystis_sp._rbcL_K7SF89
Phaeocystis_sp._rbcL_K7SF72
Phaeocystis_sp._rbcL_K7S6W5
Phaeocystis_sp._rbcL_K7SF39
Phaeocystis_sp._rbcL_K7SUS9
Phaeocystis_sp._rbcL_K7SUQ8
Phaeocystis_sp._rbcL_K7RY98
Haptophyceae_gen._rbcL_Q9GGR6
Phaeocystis_sp._rbcL_K7RY70
Phaeocystis_sp._rbcL_K7SUP5
Phaeocystis_sp._rbcL_K7S301
Phaeocystis_sp._rbcL_K7SF44
Phaeocystis_sp._rbcL_K7RY56
Phaeocystis_sp._rbcL_K7S6Q4
Phaeocystis_sp._rbcL_K7RY88
Phaeocystis_sp._rbcL_K7SF03
Phaeocystis_sp._rbcL_K7SF06
Haptophyceae_sp._rbcL_G9M5K8
Haptophyceae_sp._rbcL_G9M5D0
Haptophyceae_sp._rbcL_G9M5K0
Haptophyta_environmental_rbcL_Q93WJ0
Haptophyta_environmental_rbcL_Q946K8
Phaeocystis_rex_rbcL_A0A0D5CD69
Phaeocystis_sp._rbcL_K7RY81
Phaeocystis_sp._rbcL_K7S321
Phaeocystis_sp._rbcL_K7SUP9
Phaeocystis_sp._rbcL_K7SF22
Phaeocystis_sp._rbcL_K7SF12
Phaeocystis_sp._rbcL_K7SUR7
Phaeocystis_sp._rbcL_K7S6S2
Phaeocystis_sp._rbcL_K7SUR3
Phaeocystis_sp._rbcL_K7S308
Phaeocystis_sp._rbcL_K7SF17
Phaeocystis_sp._rbcL_K7RY85
Phaeocystis_sp._rbcL_K7RY65
Phaeocystis_sp._rbcL_K7RY76
Phaeocystis_sp._rbcL_K7S6S7
Haptophyceae_sp._rbcL_G9M5M4
Phaeocystis_sp._rbcL_K7S6R4
Phaeocystis_cordata_rbcL_K7SUV9
Haptophyta_environmental_rbcL_Q946M0
Haptophyceae_sp._rbcL_G9M5B8
Haptophyceae_sp._rbcL_G9M5J3
Phaeocystis_sp._rbcL_K7RY94
Haptophyceae_sp._rbcL_G9M5F1
Haptophyta_environmental_rbcL_Q946L5
Haptophyceae_sp._rbcL_G9M5G1
Phaeocystis_rex_rbcL_A0A0D5CE01
Haptophyceae_sp._rbcL_G9M5K9
Phaeocystis_sp._rbcL_K7RYD0
Phaeocystis_globosa_rbcL_K7S6X5
Phaeocystis_globosa_rbcL_K7S362
Phaeocystis_sp._rbcL_K7S6X2
Phaeocystis_sp._rbcL_K7S6W9
Phaeocystis_globosa_rbcL_K7RYD5
Phaeocystis_globosa_rbcL_A0A8A1RJL4
Phaeocystis_globosa_rbcL_K4ESR5
Phaeocystis_pouchetii_rbcL_A8CCX0
Phaeocystis_globosa_rbcL_R9ZSA7
Phaeocystis_sp._rbcL_K7RYC1
Phaeocystis_globosa_rbcL_K7SUW9
Phaeocystis_sp._rbcL_K7S6Q0
Phaeocystis_sp._rbcL_K7S2Z6
Phaeocystis_globosa_rbcL_K7SF94
Phaeocystis_globosa_rbcL_K7S365
Phaeocystis_sp._rbcL_K7SUW4
Phaeocystis_sp._rbcL_K7SF79
Phaeocystis_globosa_rbcL_K7RYC6
Phaeocystis_globosa_rbcL_K7SUX3
Phaeocystis_sp._rbcL_K7SF84
Phaeocystis_sp._rbcL_K7S357
Haptophyta_environmental_rbcL_Q946K3
Phaeocystis_sp._rbcL_K7S324
Phaeocystis_antarctica_rbcL_A0A0D5CD64
Phaeocystis_sp._rbcL_K7SF33
Phaeocystis_sp._rbcL_K7SEZ8
Phaeocystis_sp._rbcL_K7S6Q9
Phaeocystis_antarctica_rbcL_G9FID8
Phaeocystis_sp._rbcL_K7RY52
Phaeocystis_sp._rbcL_K7SUP0
Phaeocystis_sp._rbcL_K7SEZ3
Phaeocystis_sp._rbcL_K7SUN5
Phaeocystis_sp._rbcL_K7S2Z0
Phaeocystis_rex_rbcL_A0A0D5CE66
Phaeocystis_globosa_rbcL_A0A0D5CE46
0.7

Fig. S17

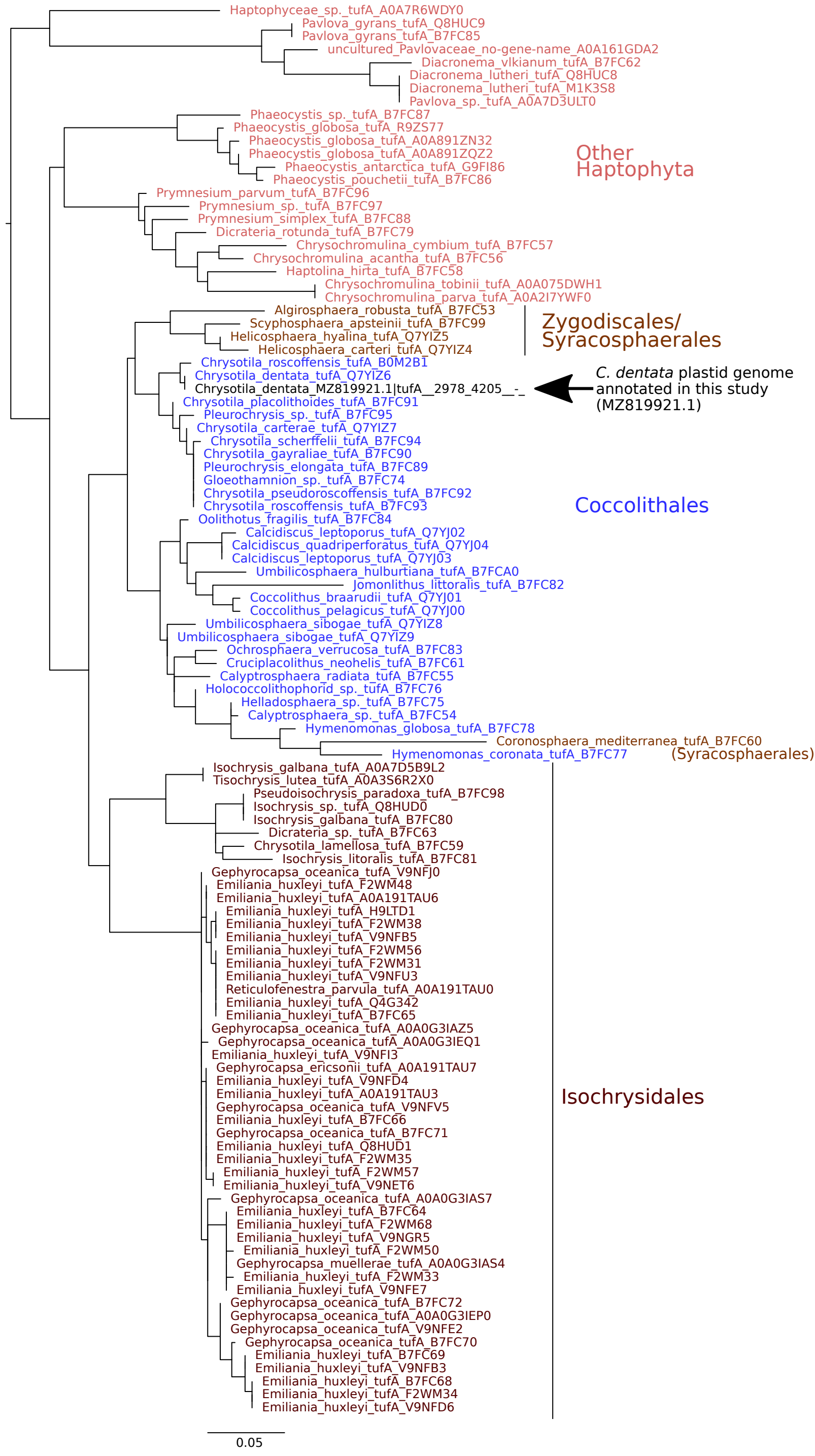

Haptophyceae_sp._tufA_A0A7R6WDY0
Pavlova_gyrans_tufA_Q8HUC9
Pavlova_gyrans_tufA_B7FC85
uncultured_Pavlovaceae_no-gene-name_A0A161GDA2
Diacronema_vlkianum_tufA_B7FC62
Diacronema_lutheri_tufA_Q8HUC8
Diacronema_lutheri_tufA_M1K3S8
Pavlova_sp._tufA_A0A7D3ULT0
Phaeocystis_sp._tufA_B7FC87
Phaeocystis_globosa_tufA_R9ZS77
Phaeocystis_globosa_tufA_A0A891ZN32
Phaeocystis_globosa_tufA_A0A891ZQZ2
Phaeocystis_antarctica_tufA_G9FI86
Phaeocystis_pouchetii_tufA_B7FC86
Other Haptophyta
Prymnesium_parvum_tufA_B7FC96
Prymnesium_sp._tufA_B7FC97
Prymnesium_simplex_tufA_B7FC88
Dicrateria_rotunda_tufA_B7FC79
Chrysochromulina_cymbium_tufA_B7FC57
Chrysochromulina_acantha_tufA_B7FC56
Haptolina_hirta_tufA_B7FC58
Chrysochromulina_tobinii_tufA_A0A075DWH1
Chrysochromulina_parva_tufA_A0A2I7YWF0
Algirosphaera_robusta_tufA_B7FC53
Scyphosphaera_apsteinii_tufA_B7FC99
Helicosphaera_hyalina_tufA_Q7YIZ5
Helicosphaera_carteri_tufA_Q7YIZ4
Zygodiscales/ Syracosphaerales
Chrysotila_roscoffensis_tufA_B0M2B1
Chrysotila_dentata_tufA_Q7YIZ6
Chrysotila_dentata_MZ819921.1|tufA__2978_4205__-_
C. dentata plastid genome annotated in this study (MZ819921.1)
Chrysotila_placolithoides_tufA_B7FC91
Pleurochrysis_sp._tufA_B7FC95
Chrysotila_carterae_tufA_Q7YIZ7
Chrysotila_scherffelii_tufA_B7FC94
Chrysotila_gayraliae_tufA_B7FC90
Pleurochrysis_elongata_tufA_B7FC89
Gloeothamnion_sp._tufA_B7FC74
Chrysotila_pseudoroscoffensis_tufA_B7FC92
Chrysotila_roscoffensis_tufA_B7FC93
Coccolithales
Oolithotus_fragilis_tufA_B7FC84
Calcidiscus_leptoporus_tufA_Q7YJ02
Calcidiscus_quadriperforatus_tufA_Q7YJ04
Calcidiscus_leptoporus_tufA_Q7YJ03
Umbilicosphaera_hulburtiana_tufA_B7FCA0
Jomonlithus_littoralis_tufA_B7FC82
Coccolithus_braarudii_tufA_Q7YJ01
Coccolithus_pelagicus_tufA_Q7YJ00
Umbilicosphaera_sibogae_tufA_Q7YIZ8
Umbilicosphaera_sibogae_tufA_Q7YIZ9
Ochrosphaera_verrucosa_tufA_B7FC83
Cruciplacolithus_neohelis_tufA_B7FC61
Calyptrosphaera_radiata_tufA_B7FC55
Hololithophorid_sp._tufA_B7FC76
Helladosphaera_sp._tufA_B7FC75
Calyptrosphaera_sp._tufA_B7FC54
Hymenomonas_globosa_tufA_B7FC78
Coronosphaera_mediterranea_tufA_B7FC60
Hymenomonas_coronata_tufA_B7FC77
(Syracosphaerales)
Isochrysis_galbana_tufA_A0A7D5B9L2
Tisochrysis_lutea_tufA_A0A3S6R2X0
Pseudoisochrysis_paradoxa_tufA_B7FC98
Isochrysis_sp._tufA_Q8HUD0
Isochrysis_galbana_tufA_B7FC80
Dicrateria_sp._tufA_B7FC63
Chrysotila_lamellosa_tufA_B7FC59
Isochrysis_litoralis_tufA_B7FC81
Gephyrocapsa_oceanica_tufA_V9NFJ0
Emiliania_huxleyi_tufA_F2WM48
Emiliania_huxleyi_tufA_A0A191TAU6
Emiliania_huxleyi_tufA_H9LTD1
Emiliania_huxleyi_tufA_F2WM38
Emiliania_huxleyi_tufA_V9NFB5
Emiliania_huxleyi_tufA_F2WM56
Emiliania_huxleyi_tufA_F2WM31
Emiliania_huxleyi_tufA_V9NFU3
Reticulofenestra_parvula_tufA_A0A191TAU0
Emiliania_huxleyi_tufA_Q4G342
Emiliania_huxleyi_tufA_B7FC65
Gephyrocapsa_oceanica_tufA_A0A0G3IAZ5
Gephyrocapsa_oceanica_tufA_A0A0G3IEQ1
Emiliania_huxleyi_tufA_V9NFI3
Gephyrocapsa_ericsonii_tufA_A0A191TAU7
Emiliania_huxleyi_tufA_V9NFD4
Emiliania_huxleyi_tufA_A0A191TAU3
Gephyrocapsa_oceanica_tufA_V9NFV5
Emiliania_huxleyi_tufA_B7FC66
Gephyrocapsa_oceanica_tufA_B7FC71
Emiliania_huxleyi_tufA_Q8HUD1
Emiliania_huxleyi_tufA_F2WM35
Emiliania_huxleyi_tufA_F2WM57
Emiliania_huxleyi_tufA_V9NET6
Gephyrocapsa_oceanica_tufA_A0A0G3IAS7
Emiliania_huxleyi_tufA_B7FC64
Emiliania_huxleyi_tufA_F2WM68
Emiliania_huxleyi_tufA_V9NGR5
Emiliania_huxleyi_tufA_F2WM50
Gephyrocapsa_muellerae_tufA_A0A0G3IAS4
Emiliania_huxleyi_tufA_F2WM33
Emiliania_huxleyi_tufA_V9NFE7
Gephyrocapsa_oceanica_tufA_B7FC72
Gephyrocapsa_oceanica_tufA_A0A0G3IEP0
Gephyrocapsa_oceanica_tufA_V9NFE2
Gephyrocapsa_oceanica_tufA_B7FC70
Emiliania_huxleyi_tufA_B7FC69
Emiliania_huxleyi_tufA_V9NFB3
Emiliania_huxleyi_tufA_B7FC68
Emiliania_huxleyi_tufA_F2WM34
Emiliania_huxleyi_tufA_V9NFD6
Isochrysidales
0.05

Fig. S19

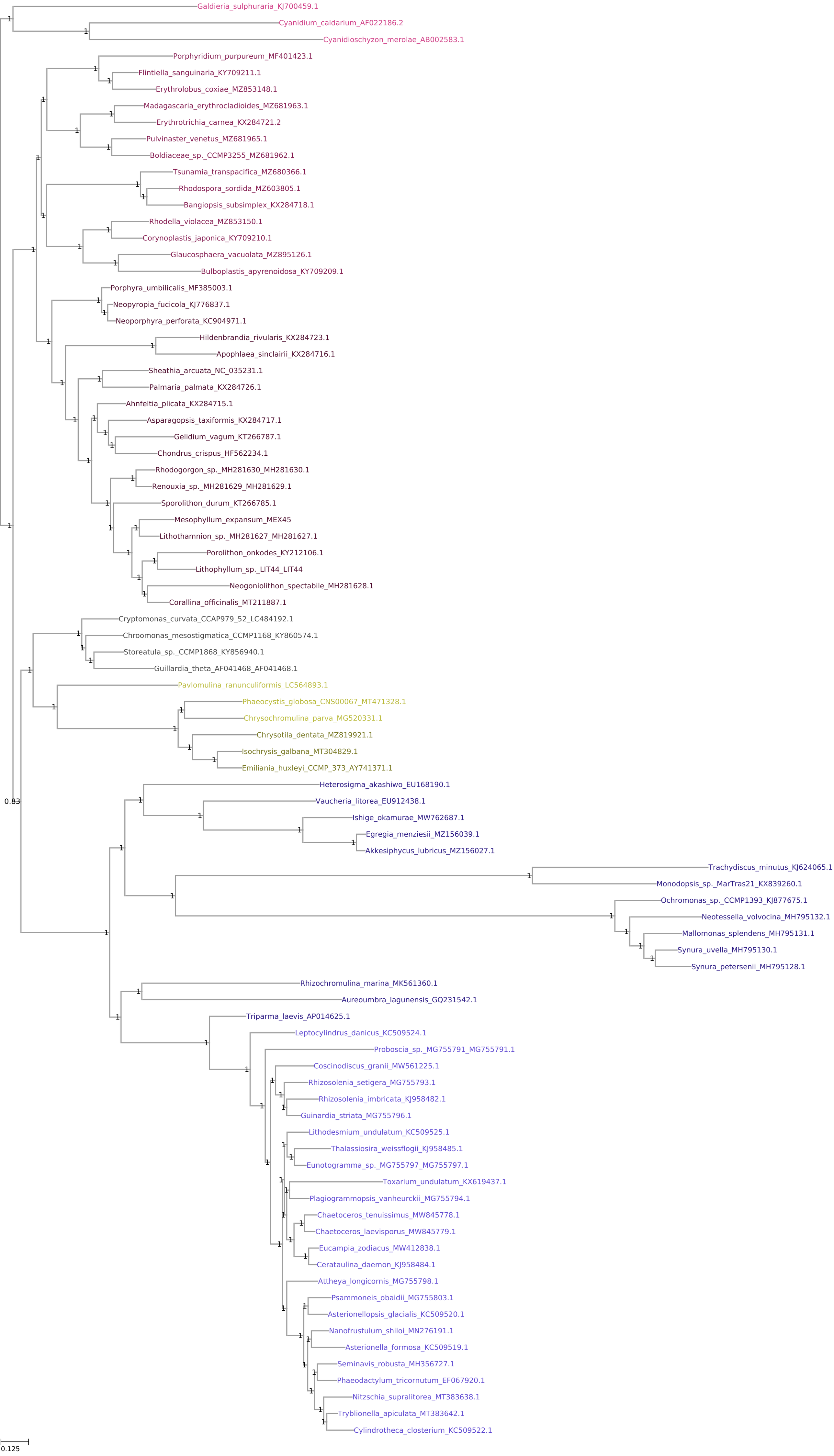

Galdieria_sulphuraria_KJ700459.1
Cyanidium_caldarium_AF022186.2
Cyanidioschyzon_merolae_AB002583.1
Porphyridium_purpureum_MF401423.1
Flintiella_sanguinaria_KY709211.1
Erythrolobus_coxiae_MZ853148.1
Madagascaria_erythrocladioides_MZ681963.1
Erythrotrichia_carnea_KX284721.2
Pulvinaster_venetus_MZ681965.1
Boldiaceae_sp._CCMP3255_MZ681962.1
Tsunamia_transpacifica_MZ680366.1
Rhodospora_sordida_MZ603805.1
Bangiopsis_subsimplex_KX284718.1
Rhodella_violacea_MZ853150.1
Corynoplastis_japonica_KY709210.1
Glaucosphaera_vacuolata_MZ895126.1
Bulboplastis_apyrenoidosa_KY709209.1
Porphyra_umbilicalis_MF385003.1
Neopyropia_fucicola_KJ776837.1
Neoporphyra_perforata_KC904971.1
Hildenbrandia_rivularis_KX284723.1
Apophlaea_sinclairii_KX284716.1
Sheathia_arcuata_NC_035231.1
Palmaria_palmata_KX284726.1
Ahnfeltia_plicata_KX284715.1
Asparagopsis_taxiformis_KX284717.1
Gelidium_vagum_KT266787.1
Chondrus_crispus_HF562234.1
Rhodogorgon_sp._MH281630_MH281630.1
Renouxia_sp._MH281629_MH281629.1
Sporolithon_durum_KT266785.1
Mesophyllum_expansum_MEX45
Lithothamnion_sp._MH281627_MH281627.1
Porolithon_onkodes_KY212106.1
Lithophyllum_sp._LIT44_LIT44
Neogoniolithon_spectabile_MH281628.1
Corallina_officinalis_MT211887.1
Cryptomonas_curvata_CCAP979_52_LC484192.1
Chroomonas_mesostigmatica_CCMP1168_KY860574.1
Storeatula_sp._CCMP1868_KY856940.1
Guillardia_theta_AF041468_AF041468.1
Pavlomulina_ranunculiformis_LC564893.1
Phaeocystis_globosa_CNS00067_MT471328.1
Chrysochromulina_parva_MG520331.1
Chrysotila_dentata_MZ819921.1
Isochrysis_galbana_MT304829.1
Emiliania_huxleyi_CCMP_373_AY741371.1
Heterosigma_akashiwo_EU168190.1
Vaucheria_litorea_EU912438.1
Ishige_okamurae_MW762687.1
Egregia_menziesii_MZ156039.1
Akkesiphycus_lubricus_MZ156027.1
Trachydiscus_minutus_KJ624065.1
Monodopsis_sp._MarTras21_KX839260.1
Ochromonas_sp._CCMP1393_KJ877675.1
Neotessella_volvocina_MH795132.1
Mallomonas_splendens_MH795131.1
Synura_uvella_MH795130.1
Synura_petersenii_MH795128.1
Rhizochromulina_marina_MK561360.1
Aureoumbra_lagunensis_GQ231542.1
Triparma_laevis_AP014625.1
Leptocylindrus_danicus_KC509524.1
Proboscia_sp._MG755791_MG755791.1
Coscinodiscus_granii_MW561225.1
Rhizosolenia_setigera_MG755793.1
Rhizosolenia_imbricata_KJ958482.1
Guinardia_striata_MG755796.1
Lithodesmium_undulatum_KC509525.1
Thalassiosira_weissflogii_KJ958485.1
Eunotogramma_sp._MG755797_MG755797.1
Toxarium_undulatum_KX619437.1
Plagiogrammopsis_vanheurckii_MG755794.1
Chaetoceros_tenuissimus_MW845778.1
Chaetoceros_laevisporus_MW845779.1
Eucampia_zodiacus_MW412838.1
Cerataulina_daemon_KJ958484.1
Attheya_longicornis_MG755798.1
Psammoneis_obaidii_MG755803.1
Asterionellopsis_glacialis_KC509520.1
Nanofrustulum_shiloi_MN276191.1
Asterionella_formosa_KC509519.1
Seminavis_robusta_MH356727.1
Phaeodactylum_tricornutum_EF067920.1
Nitzschia_supralitorea_MT383638.1
Tryblionella_apiculata_MT383642.1
Cylindrotheca_closterium_KC509522.1
0.83
0.125

Fig. S20

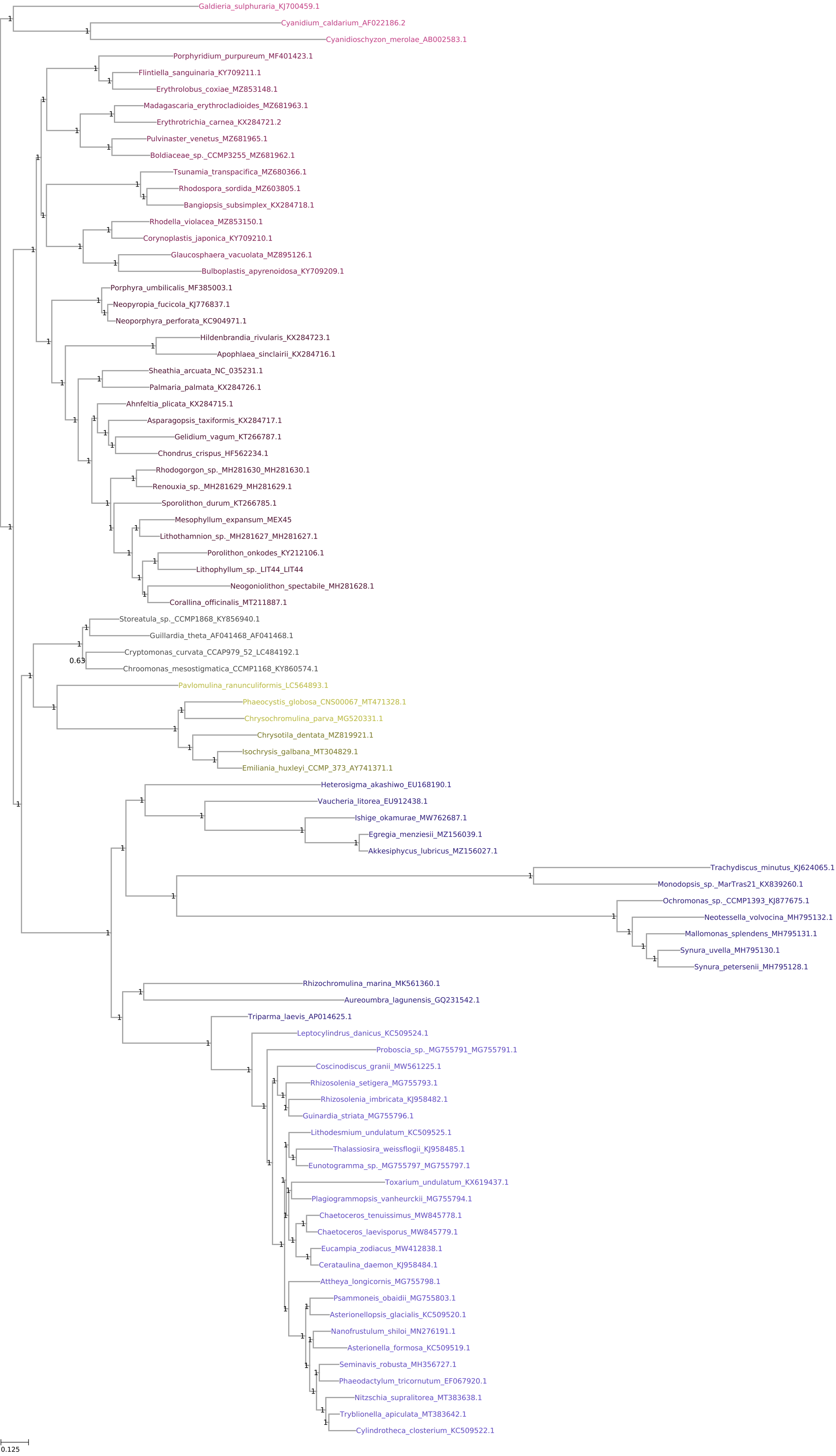

Galdieria_sulphuraria_KJ700459.1
Cyanidium_caldarium_AF022186.2
Cyanidioschyzon_merolae_AB002583.1
Porphyridium_purpureum_MF401423.1
Flintiella_sanguinaria_KY709211.1
Erythrolobus_coxiae_MZ853148.1
Madagascaria_erythrocladioides_MZ681963.1
Erythrotrichia_carnea_KX284721.2
Pulvinaster_venetus_MZ681965.1
Boldiaceae_sp._CCMP3255_MZ681962.1
Tsunamia_transpacifica_MZ680366.1
Rhodospora_sordida_MZ603805.1
Bangiopsis_subsimplex_KX284718.1
Rhodella_violacea_MZ853150.1
Corynoplastis_japonica_KY709210.1
Glaucosphaera_vacuolata_MZ895126.1
Bulboplastis_apyrenoidosa_KY709209.1
Porphyra_umbilicalis_MF385003.1
Neopyropia_fucicola_KJ776837.1
Neoporphyra_perforata_KC904971.1
Hildenbrandia_rivularis_KX284723.1
Apophlaea_sinclairii_KX284716.1
Sheathia_arcuata_NC_035231.1
Palmaria_palmata_KX284726.1
Ahnfeltia_plicata_KX284715.1
Asparagopsis_taxiformis_KX284717.1
Gelidium_vagum_KT266787.1
Chondrus_crispus_HF562234.1
Rhodogorgon_sp._MH281630_MH281630.1
Renouxia_sp._MH281629_MH281629.1
Sporolithon_durum_KT266785.1
Mesophyllum_expansum_MEX45
Lithothamnion_sp._MH281627_MH281627.1
Porolithon_onkodes_KY212106.1
Lithophyllum_sp._LIT44_LIT44
Neogoniolithon_spectabile_MH281628.1
Corallina_officinalis_MT211887.1
Storeatula_sp._CCMP1868_KY856940.1
Guillardia_theta_AF041468_AF041468.1
Cryptomonas_curvata_CCAP979_52_LC484192.1
Chroomonas_mesostigmatica_CCMP1168_KY860574.1
Pavlomulina_ranunculiformis_LC564893.1
Phaeocystis_globosa_CNS00067_MT471328.1
Chrysochromulina_parva_MG520331.1
Chrysotila_dentata_MZ819921.1
Isochrysis_galbana_MT304829.1
Emiliania_huxleyi_CCMP_373_AY741371.1
Heterosigma_akashiwo_EU168190.1
Vaucheria_litorea_EU912438.1
Ishige_okamurae_MW762687.1
Egregia_menziesii_MZ156039.1
Akkesiphycus_lubricus_MZ156027.1
Trachydiscus_minutus_KJ624065.1
Monodopsis_sp._MarTras21_KX839260.1
Ochromonas_sp._CCMP1393_KJ877675.1
Neotessella_volvocina_MH795132.1
Mallomonas_splendens_MH795131.1
Synura_uvella_MH795130.1
Synura_petersenii_MH795128.1
Rhizochromulina_marina_MK561360.1
Aureoumbra_lagunensis_GQ231542.1
Triparma_laevis_AP014625.1
Leptocylindrus_danicus_KC509524.1
Proboscia_sp._MG755791_MG755791.1
Coscinodiscus_granii_MW561225.1
Rhizosolenia_setigera_MG755793.1
Rhizosolenia_imbricata_KJ958482.1
Guinardia_striata_MG755796.1
Lithodesmium_undulatum_KC509525.1
Thalassiosira_weissflogii_KJ958485.1
Eunotogramma_sp._MG755797_MG755797.1
Toxarium_undulatum_KX619437.1
Plagiogrammopsis_vanheurckii_MG755794.1
Chaetoceros_tenuissimus_MW845778.1
Chaetoceros_laevisporus_MW845779.1
Eucampia_zodiacus_MW412838.1
Cerataulina_daemon_KJ958484.1
Attheya_longicornis_MG755798.1
Psammoneis_obaidii_MG755803.1
Asterionellopsis_glacialis_KC509520.1
Nanofrustulum_shiloi_MN276191.1
Asterionella_formosa_KC509519.1
Seminavis_robusta_MH356727.1
Phaeodactylum_tricornutum_EF067920.1
Nitzschia_supralitorea_MT383638.1
Tryblionella_apiculata_MT383642.1
Cylindrotheca_closterium_KC509522.1
0.63
0.125

Fig. S21

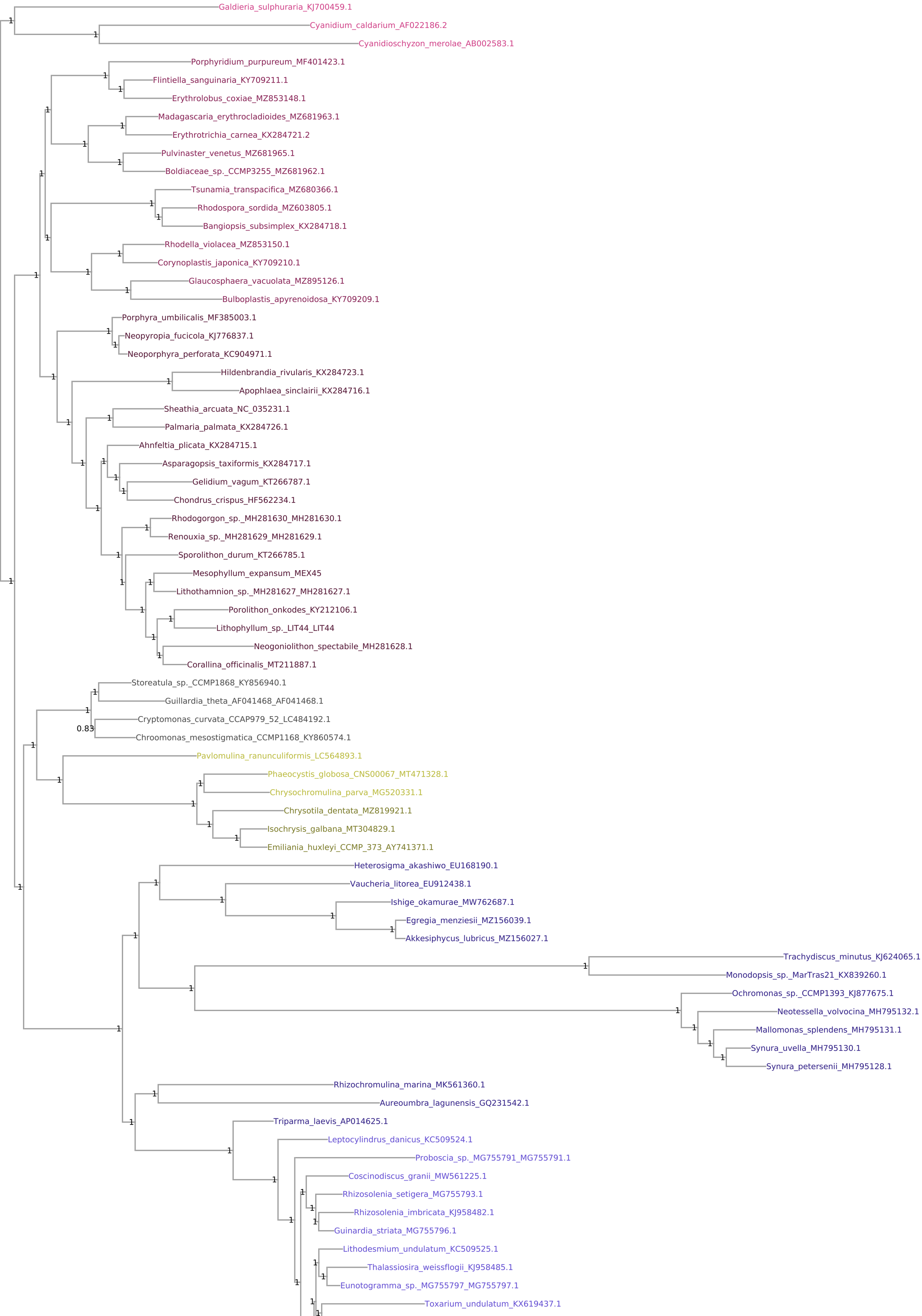
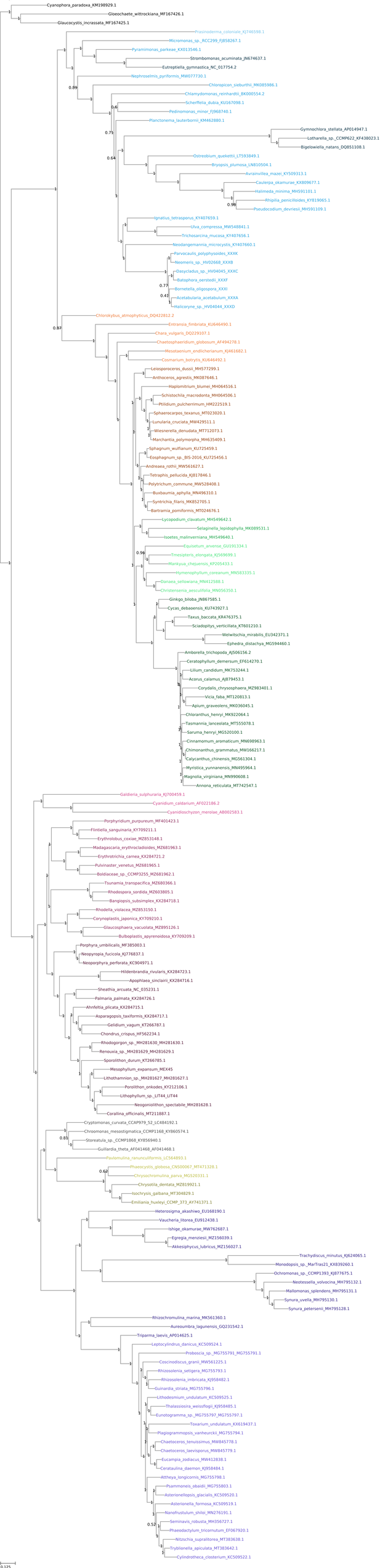

Galdieria_sulphuraria_KJ700459.1
Cyanidium_caldarium_AF022186.2
Cyanidioschyzon_merolae_AB002583.1
Porphyridium_purpureum_MF401423.1
Flintiella_sanguinaria_KY709211.1
Erythrolobus_coxiae_MZ853148.1
Madagascaria_erythrocladioides_MZ681963.1
Erythrotrichia_carnea_KX284721.2
Pulvinaster_venetus_MZ681965.1
Boldiaceae_sp._CCMP3255_MZ681962.1
Tsunamia_transpacifica_MZ680366.1
Rhodospora_sordida_MZ603805.1
Bangiopsis_subsimplex_KX284718.1
Rhodella_violacea_MZ853150.1
Corynoplastis_japonica_KY709210.1
Glaucosphaera_vacuolata_MZ895126.1
Bulboplastis_apyrenoidosa_KY709209.1
Porphyra_umbilicalis_MF385003.1
Neopyropia_fucicola_KJ776837.1
Neoporphyra_perforata_KC904971.1
Hildenbrandia_rivularis_KX284723.1
Apophlaea_sinclairii_KX284716.1
Sheathia_arcuata_NC_035231.1
Palmaria_palmata_KX284726.1
Ahnfeltia_plicata_KX284715.1
Asparagopsis_taxiformis_KX284717.1
Gelidium_vagum_KT266787.1
Chondrus_crispus_HF562234.1
Rhodogorgon_sp._MH281630_MH281630.1
Renouxia_sp._MH281629_MH281629.1
Sporolithon_durum_KT266785.1
Mesophyllum_expansum_MEX45
Lithothamnion_sp._MH281627_MH281627.1
Porolithon_onkodes_KY212106.1
Lithophyllum_sp._LIT44_LIT44
Neogoniolithon_spectabile_MH281628.1
Corallina_officinalis_MT211887.1
Storeatula_sp._CCMP1868_KY856940.1
Guillardia_theta_AF041468_AF041468.1
Cryptomonas_curvata_CCAP979_52_LC484192.1
Chroomonas_mesostigmatica_CCMP1168_KY860574.1
Pavlomulina_ranunculiformis_LC564893.1
Phaeocystis_globosa_CNS00067_MT471328.1
Chrysochromulina_parva_MG520331.1
Chrysotila_dentata_MZ819921.1
Isochrysis_galbana_MT304829.1
Emiliania_huxleyi_CCMP_373_AY741371.1
Heterosigma_akashiwo_EU168190.1
Vaucheria_litorea_EU912438.1
Ishige_okamurae_MW762687.1
Egregia_menziesii_MZ156039.1
Akkesiphycus_lubricus_MZ156027.1
Trachydiscus_minutus_KJ624065.1
Monodopsis_sp._MarTras21_KX839260.1
Ochromonas_sp._CCMP1393_KJ877675.1
Neotessella_volvocina_MH795132.1
Mallomonas_splendens_MH795131.1
Synura_uvella_MH795130.1
Synura_petersenii_MH795128.1
Rhizochromulina_marina_MK561360.1
Aureoumbra_lagunensis_GQ231542.1
Triparma_laevis_AP014625.1
Leptocylindrus_danicus_KC509524.1
Proboscia_sp._MG755791_MG755791.1
Coscinodiscus_granii_MW561225.1
Rhizosolenia_setigera_MG755793.1
Rhizosolenia_imbricata_KJ958482.1
Guinardia_striata_MG755796.1
Lithodesmium_undulatum_KC509525.1
Thalassiosira_weissflogii_KJ958485.1
Eunotogramma_sp._MG755797_MG755797.1
Toxarium_undulatum_KX619437.1
Plagiogrammopsis_vanheurckii_MG755794.1
Chaetoceros_tenuissimus_MW845778.1
Chaetoceros_laevisporus_MW845779.1
Eucampia_zodiacus_MW412838.1
Cerataulina_daemon_KJ958484.1
Attheya_longicornis_MG755798.1
Psammoneis_obaidii_MG755803.1
Asterionellopsis_glacialis_KC509520.1
Nanofrustulum_shiloi_MN276191.1
Asterionella_formosa_KC509519.1
Seminavis_robusta_MH356727.1
Phaeodactylum_tricornutum_EF067920.1
Nitzschia_supralitorea_MT383638.1
Tryblionella_apiculata_MT383642.1
Cylindrotheca_closterium_KC509522.1
0.83
0.125

Fig. S22


Cyanophora_paradoxa_KM198929.1
Gloeochaete_wittrockiana_MF167426.1
Glaucocystis_incrassata_MF167425.1
Prasinoderma_coloniale_KJ746598.1
Micromonas_sp._RCC299_FJ858267.1
Pyramimonas_parkeae_KX013546.1
Strombomonas_acuminata_JN674637.1
Eutreptiella_gymnastica_NC_017754.2
Nephroselmis_pyriformis_MW077730.1
Chloropicon_sieburthii_MK085986.1
Chlamydomonas_reinhardtii_BK000554.2
Scherffelia_dubia_KU167098.1
Pedinomonas_minor_FJ968740.1
Planctonema_lauterbornii_KM462880.1
Gymnochlora_stellata_AP014947.1
Lotharella_sp._CCMP622_KF438023.1
Bigelowiella_natans_DQ851108.1
Ostreobium_quekettii_LT593849.1
Bryopsis_plumosa_LN810504.1
Avrainvillea_mazei_KY509313.1
Caulerpa_okamurae_KX809677.1
Halimeda_minima_MH591101.1
Rhipilia_penicilloides_KY819065.1
Pseudocodium_devriesii_MH591109.1
Ignatius_tetrasporus_KY407659.1
Ulva_compressa_MW548841.1
Trichosarcina_mucosa_KY407656.1
Neodangemannia_microcystis_KY407660.1
Parvocaulis_polyphysoides_XXXK
Neomeris_sp._HV02668_XXXB
Dasycladus_sp._HV04045_XXXC
Batophora_oerstedii_XXXF
Bornetella_oligospora_XXXI
Acetabularia_acetabulum_XXXA
Halicoryne_sp._HV04044_XXXD
Chlorokybus_atmophyticus_DQ422812.2
Entransia_fimbriata_KU646490.1
Chara_vulgaris_DQ229107.1
Chaetosphaeridium_globosum_AF494278.1
Mesotaenium_endlicherianum_KJ461682.1
Cosmarium_botrytis_KU646492.1
Leiosporoceros_dussii_MH577299.1
Anthoceros_agrestis_MK087646.1
Haplomitrium_blumei_MH064516.1
Schistochila_macrodonta_MH064506.1
Ptilidium_pulcherrimum_HM222519.1
Sphaerocarpos_texanus_MT023020.1
Lunularia_cruciata_MW429511.1
Wiesnerella_denudata_MT712073.1
Marchantia_polymorpha_MH635409.1
Sphagnum_wulfianum_KU725459.1
Eosphagnum_sp._BIS-2016_KU725456.1
Andreaea_rothii_MW561627.1
Tetraphis_pellucida_KJ817846.1
Polytrichum_commune_MW528408.1
Buxbaumia_aphylla_MN496310.1
Syntrichia_filaris_MK852705.1
Bartramia_pomiformis_MT024676.1
Lycopodium_clavatum_MH549642.1
Selaginella_lepidophylla_MK089531.1
Isoetes_malinverniana_MH549640.1
Equisetum_arvense_GU191334.1
Tmesipteris_elongata_KJ569699.1
Mankyua_chejuensis_KP205433.1
Hymenophyllum_coreanum_MN583335.1
Danaea_sellowiana_MN412588.1
Christensenia_aesculifolia_MN056350.1
Ginkgo_biloba_JN867585.1
Cycas_debaoensis_KU743927.1
Taxus_baccata_KR476375.1
Sciadopitys_verticillata_KT601210.1
Welwitschia_mirabilis_EU342371.1
Ephedra_distachya_MG594460.1
Amborella_trichopoda_AJ506156.2
Ceratophyllum_demersum_EF614270.1
Lilium_candidum_MK753244.1
Acorus_calamus_AJ879453.1
Corydalis_chrysosophaera_MZ983401.1
Vicia_faba_MT120813.1
Apium_graveolens_MK036045.1
Chloranthus_henryi_MK922064.1
Tasmannia_lanceolata_MT555078.1
Saruma_henryi_MG520100.1
Cinnamomum_aromaticum_MN698963.1
Chimonanthus_grammatus_MW166217.1
Calycanthus_chinensis_MG561304.1
Myristica_yunnanensis_MN495964.1
Magnolia_virginiana_MN990608.1
Annona_reticulata_MT742547.1
Galdieria_sulphuraria_KJ700459.1
Cyanidium_caldarium_AF022186.2
Cyanidioschyzon_merolae_AB002583.1
Porphyridium_purpureum_MF401423.1
Flintiella_sanguinaria_KY709211.1
Erythrolobus_coxiae_MZ853148.1
Madagascaria_erythrocladioides_MZ681963.1
Erythrotrichia_carnea_KX284721.2
Pulvinaster_venetus_MZ681965.1
Boldiaceae_sp._CCMP3255_MZ681962.1
Tsunamia_transpacifica_MZ680366.1
Rhodospora_sordida_MZ603805.1
Bangiopsis_subsimplex_KX284718.1
Rhodella_violacea_MZ853150.1
Corynoplastis_japonica_KY709210.1
Glaucosphaera_vacuolata_MZ895126.1
Bulboplastis_apyrenoidosa_KY709209.1
Porphyra_umbilicalis_MF385003.1
Neopyropia_fucicola_KJ776837.1
Neoporphyra_perforata_KC904971.1
Hildenbrandia_rivularis_KX284723.1
Apophlaea_sinclairii_KX284716.1
Sheathia_arcuata_NC_035231.1
Palmaria_palmata_KX284726.1
Ahnfeltia_plicata_KX284715.1
Asparagopsis_taxiformis_KX284717.1
Gelidium_vagum_KT266787.1
Chondrus_crispus_HF562234.1
Rhodogorgon_sp._MH281630_MH281630.1
Renouxia_sp._MH281629_MH281629.1
Sporolithon_durum_KT266785.1
Mesophyllum_expansum_MEX45
Lithothamnion_sp._MH281627_MH281627.1
Porolithon_onkodes_KY212106.1
Lithophyllum_sp._LIT44_LIT44
Neogoniolithon_spectabile_MH281628.1
Corallina_officinalis_MT211887.1
Cryptomonas_curvata_CCAP979_52_LC484192.1
Chroomonas_mesostigmatica_CCMP1168_KY860574.1
Storeatula_sp._CCMP1868_KY856940.1
Guillardia_theta_AF041468_AF041468.1
Pavlomulina_ranunculiformis_LC564893.1
Phaeocystis_globosa_CNS00067_MT471328.1
Chrysochromulina_parva_MG520331.1
Chrysotila_dentata_MZ819921.1
Isochrysis_galbana_MT304829.1
Emiliania_huxleyi_CCMP_373_AY741371.1
Heterosigma_akashiwo_EU168190.1
Vaucheria_litorea_EU912438.1
Ishige_okamurae_MW762687.1
Egregia_menziesii_MZ156039.1
Akkesiphycus_lubricus_MZ156027.1
Trachydiscus_minutus_KJ624065.1
Monodopsis_sp._MarTras21_KX839260.1
Ochromonas_sp._CCMP1393_KJ877675.1
Neotessella_volvocina_MH795132.1
Mallomonas_splendens_MH795131.1
Synura_uvella_MH795130.1
Synura_petersenii_MH795128.1
Rhizochromulina_marina_MK561360.1
Aureoumbra_lagunensis_GQ231542.1
Triparma_laevis_AP014625.1
Leptocylindrus_danicus_KC509524.1
Proboscia_sp._MG755791_MG755791.1
Coscinodiscus_granii_MW561225.1
Rhizosolenia_setigera_MG755793.1
Rhizosolenia_imbricata_KJ958482.1
Guinardia_striata_MG755796.1
Lithodesmium_undulatum_KC509525.1
Thalassiosira_weissflogii_KJ958485.1
Eunotogramma_sp._MG755797_MG755797.1
Toxarium_undulatum_KX619437.1
Plagiogrammopsis_vanheurckii_MG755794.1
Chaetoceros_tenuissimus_MW845778.1
Chaetoceros_laevisporus_MW845779.1
Eucampia_zodiacus_MW412838.1
Cerataulina_daemon_KJ958484.1
Attheya_longicornis_MG755798.1
Psammoneis_obaidii_MG755803.1
Asterionellopsis_glacialis_KC509520.1
Asterionella_formosa_KC509519.1
Nanofrustulum_shiloi_MN276191.1
Seminavis_robusta_MH356727.1
Phaeodactylum_tricornutum_EF067920.1
Nitzschia_supralitorea_MT383638.1
Tryblionella_apiculata_MT383642.1
Cylindrotheca_closterium_KC509522.1
0.125

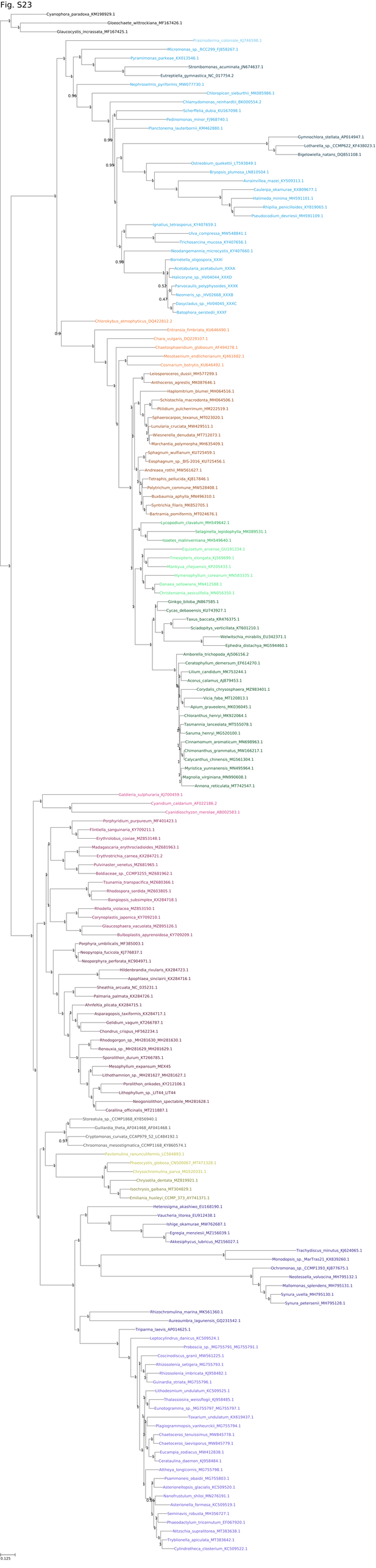
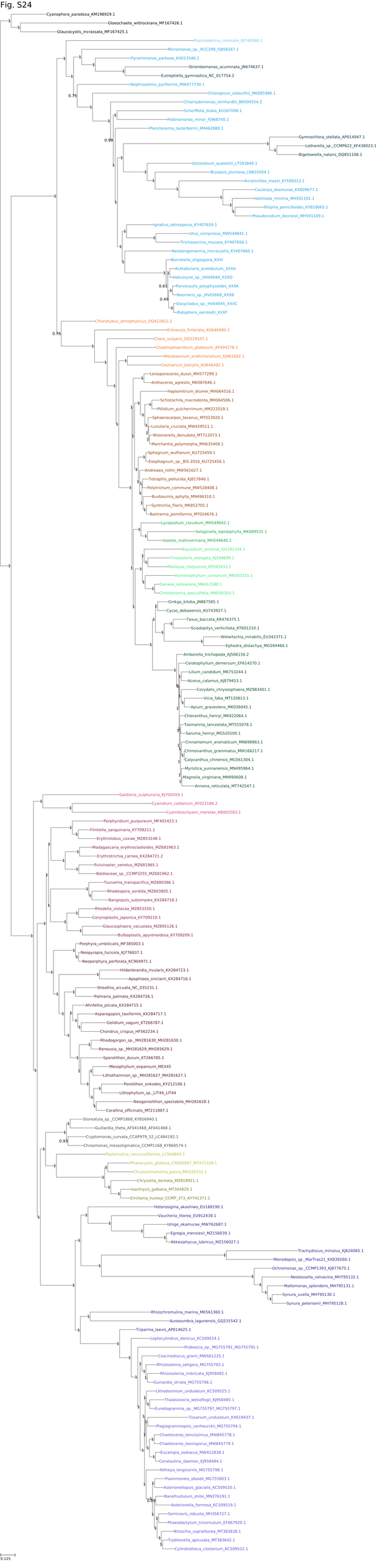


Fig. S23
Cyanophora_paradoxa_KM198929.1
Gloeochaete_wittrockiana_MF167426.1
Glaucocystis_incrassata_MF167425.1
Prasinoderma_coloniale_KJ746598.1
Micromonas_sp._RCC299_FJB58267.1
Pyramimonas_parkeae_KX013546.1
Strombomonas_acuminata_JN674637.1
Eutreptiella_gymnastica_NC_017754.2
Nephroselmis_pyriformis_MW077730.1
Chloropicon_sieburthii_MK085986.1
Chlamydomonas_reinhardtii_BK000554.2
Scherffelia_dubia_KU167098.1
Pedinomonas_minor_FJ968740.1
Planctonema_lauterbornii_KM462880.1
Gymnochlora_stellata_AP014947.1
Lotharella_sp._CCMP622_KF438023.1
Bigelowiella_natans_DQ851108.1
Ostreobium_quekettii_LT593849.1
Bryopsis_plumosa_LN810504.1
Avrainvillea_mazei_KY509313.1
Caulerpa_okamurae_KX809677.1
Halimeda_minima_MH591101.1
Rhipilia_penicilloides_KY819065.1
Pseudocodium_devriesii_MH591109.1
Ignatius_tetrasporus_KY407659.1
Ulva_compressa_MW548841.1
Trichosarcina_mucosa_KY407656.1
Neodangemannia_microcystis_KY407660.1
Bornetella_oligospora_XXXI
Acetabularia_acetabulum_XXXA
Halicoryne_sp._HV04044_XXXD
Parvocaulis_polyphysoides_XXXK
Neomeris_sp._HV02668_XXXB
Dasycladus_sp._HV04045_XXXC
Batophora_oerstedii_XXXF
Chlorokybus_atmophyticus_DQ422812.2
Entransia_fimbriata_KU646490.1
Chara_vulgaris_DQ229107.1
Chaetosphaeridium_globosum_AF494278.1
Mesotaenium_endlicherianum_KJ461682.1
Cosmarium_botrytis_KU646492.1
Leiosporoceros_dussii_MH577299.1
Anthoceros_agrestis_MK087646.1
Haplomitrium_blumei_MH064516.1
Schistochila_macrodonta_MH064506.1
Ptilidium_pulcherrimum_HM222519.1
Sphaerocarpos_texanus_MT023020.1
Lunularia_cruciata_MW429511.1
Wiesnerella_denudata_MT712073.1
Marchantia_polymorpha_MH635409.1
Sphagnum_wulfianum_KU725459.1
Eosphagnum_sp._BIS-2016_KU725456.1
Andreaea_rothii_MW561627.1
Tetraphis_pellucida_KJ817846.1
Polytrichum_commune_MW528408.1
Buxbaumia_aphylla_MN496310.1
Syntrichia_filaris_MK852705.1
Bartramia_pomiformis_MT024676.1
Lycopodium_clavatum_MH549642.1
Selaginella_lepidophylla_MK089531.1
Isoetes_malinverniana_MH549640.1
Equisetum_arvense_GU191334.1
Tmesipteris_elongata_KJ569699.1
Mankyua_chejuensis_KP205433.1
Hymenophyllum_coreanum_MN583335.1
Danaea_seilowiana_MN412588.1
Christensenia_aesculifolia_MN056350.1
Ginkgo_biloba_JN867585.1
Cycas_debaoensis_KU743927.1
Taxus_baccata_KR476375.1
Sciadopitys_verticillata_KT601210.1
Welwitschia_mirabilis_EU342371.1
Ephedra_distachya_MG594460.1
Amborella_trichopoda_AJ506156.2
Ceratophyllum_demersum_EF614270.1
Lilium_candidum_MK753244.1
Acorus_calamus_AJB79453.1
Corydalis_chrysosphaera_MZ983401.1
Vicia_faba_MT120813.1
Apium_graveolens_MK036045.1
Chloranthus_henryi_MK922064.1
Tasmannia_lanceolata_MT555078.1
Saruma_henryi_MG520100.1
Cinnamomum_aromaticum_MN698963.1
Chimonanthus_grammatus_MW166217.1
Calycanthus_chinensis_MG561304.1
Myristica_yunnanensis_MN495964.1
Magnolia_virginiana_MN990608.1
Annona_reticulata_MT742547.1
Galdieria_sulphuraria_KJ700459.1
Cyanidium_caldarium_AF022186.2
Cyanidioschyzon_merolae_AB002583.1
Porphyridium_purpureum_MF401423.1
Flintiella_sanguinaria_KY709211.1
Erythrolobus_coxiae_MZ853148.1
Madagascaria_erythrocladioides_MZ681963.1
Erythrotrichia_carnea_KX284721.2
Pulvinaster_venetus_MZ681965.1
Boldiaceae_sp._CCMP3255_MZ681962.1
Tsunamia_transpacifica_MZ680366.1
Rhodospora_sordida_MZ603805.1
Bangiopsis_subsimplex_KX284718.1
Rhodella_violacea_MZ853150.1
Corynoplastis_japonica_KY709210.1
Glaucosphaera_vacuolata_MZ895126.1
Bulboplastis_apyrenoidosa_KY709209.1
Porphyra_umbilicalis_MF385003.1
Neopyropia_fucicola_KJ776837.1
Neoporphyra_perforata_KC904971.1
Hildenbrandia_rivularis_KX284723.1
Apophlaea_sinclairii_KX284716.1
Sheathia_arcuata_NC_035231.1
Palmaria_palmata_KX284726.1
Ahnfeltia_plicata_KX284715.1
Asparagopsis_taxiformis_KX284717.1
Gelidium_vagum_KT266787.1
Chondrus_crispus_HF562234.1
Rhodogorgon_sp._MH281630_MH281630.1
Renouxia_sp._MH281629_MH281629.1
Sporolithon_durum_KT266785.1
Mesophyllum_expansum_MEX45
Lithothamnion_sp._MH281627_MH281627.1
Porolithon_onkodes_KY212106.1
Lithophyllum_sp._LIT44_LIT44
Neogoniolithon_spectabile_MH281628.1
Corallina_officinalis_MT211887.1
Storeatula_sp._CCMP1868_KY856940.1
Guillardia_theta_AF041468_AF041468.1
Cryptomonas_curvata_CCAP979_52_LC484192.1
Chroomonas_mesostigmatica_CCMP1168_KY860574.1
Pavlovulina_rananculiformis_LC564893.1
Phaeocystis_globosa_CNS00067_MT471328.1
Chrysochromulina_parva_MG520331.1
Chrysotila_dentata_MZ819921.1
Isochrysis_galbana_MT304829.1
Emiliania_huxleyi_CCMP_373_AY741371.1
Heterosigma_akashiwo_EU168190.1
Vaucheria_litorea_EU912438.1
Ishige_okamurae_MW762687.1
Egregia_menziesii_MZ156039.1
Akkesiphycus_lubricus_MZ156027.1
Trachydiscus_minutus_KJ624065.1
Monodopsis_sp._MarTras21_KX839260.1
Ochromonas_sp._CCMP1393_KJ877675.1
Neotessella_volvocina_MH795132.1
Mallomonas_splendens_MH795131.1
Synura_uvella_MH795130.1
Synura_petersenii_MH795128.1
Rhizochromulina_marina_MK561360.1
Aureoumbra_lagunensis_GQ231542.1
Triparma_laevis_AP014625.1
Leptocylindrus_danicus_KC509524.1
Proboscia_sp._MG755791_MG755791.1
Coscinodiscus_granii_MW561225.1
Rhizosolenia_setigera_MG755793.1
Rhizosolenia_imbricata_KJ958482.1
Guinardia_striata_MG755796.1
Lithodesmium_undulatum_KC509525.1
Thalassiosira_weissflogii_KJ958485.1
Eunotogramma_sp._MG755797_MG755797.1
Toxarium_undulatum_KX619437.1
Plagiogrammopsis_vanheurckii_MG755794.1
Chaetoceros_tenuissimus_MW845778.1
Chaetoceros_laevisporus_MW845779.1
Eucampia_zodiacus_MW412838.1
Cerataulina_daemon_KJ958484.1
Attheya_longicornis_MG755798.1
Psammoneis_obaidii_MG755803.1
Asterionellopsis_glacialis_KC509520.1
Nanofrustulum_shiloi_MN276191.1
Asterionella_formosa_KC509519.1
Seminavis_robusta_MH356727.1
Phaeodactylum_tricornutum_EF067920.1
Nitzschia_supralitorea_MT383638.1
Tryblionella_apiculata_MT383642.1
Cylindrotheca_closterium_KC509522.1
0.125

Fig. S24


Cyanophora_paradoxa_KM198929.1
Gloeochaete_wittrockiana_MF167426.1
Glaucocystis_incrassata_MF167425.1
Prasinoderma_coloniale_KJ746598.1
Micromonas_sp._RCC299_FJ858267.1
Pyramimonas_parkeae_KX013546.1
Strombomonas_acuminata_JN674637.1
Eutreptiella_gymnastica_NC_017754.2
Nephroselmis_pyriformis_MW077730.1
Chloropicon_sieburthii_MK085986.1
Chlamydomonas_reinhardtii_BK000554.2
Scherffelia_dubia_KU167098.1
Pedinomonas_minor_FJ968740.1
Planctonema_lauterbornii_KM462880.1
Gymnochlora_stellata_AP014947.1
Lotharella_sp._CCMP622_KF438023.1
Bigelowiella_natans_DQ851108.1
Ostreobium_quekettii_LT593849.1
Bryopsis_plumosa_LN810504.1
Avrainvillea_mazei_KY509313.1
Caulerpa_okamurae_KX809677.1
Halimeda_minima_MH591101.1
Rhipilia_penicilloides_KY819065.1
Pseudocodium_devriesii_MH591109.1
Ignatius_tetrasporus_KY407659.1
Ulva_compressa_MW548841.1
Trichosarcina_mucosa_KY407656.1
Neodangemannia_microcystis_KY407660.1
Bornetella_oligospora_XXXI
Acetabularia_acetabulum_XXXA
Halicoryne_sp._HV04044_XXXD
Parvocaulis_polyphysoides_XXXK
Neomeris_sp._HV02668_XXXB
Dasycladus_sp._HV04045_XXXC
Batophora_oerstedii_XXXF
Chlorokybus_atmophyticus_DQ422812.2
Entransia_fimbriata_KU646490.1
Chara_vulgaris_DQ229107.1
Chaetosphaeridium_globosum_AF494278.1
Mesotaenium_endlicherianum_KJ461682.1
Cosmarium_botrytis_KU646492.1
Leiosporoceros_dussii_MH577299.1
Anthoceros_agrestis_MK087646.1
Haplomitrium_blumei_MH064516.1
Schistochila_macrodonta_MH064506.1
Ptilidium_pulcherrimum_HM222519.1
Sphaerocarpos_texanus_MT023020.1
Lunularia_cruciata_MW429511.1
Wiesnerella_denudata_MT712073.1
Marchantia_polymorpha_MH635409.1
Sphagnum_wulfianum_KU725459.1
Eosphagnum_sp._BIS-2016_KU725456.1
Andreaea_rothii_MW561627.1
Tetraphis_pellucida_KJ817846.1
Polytrichum_commune_MW528408.1
Buxbaumia_aphylla_MN496310.1
Syntrichia_filaris_MK852705.1
Bartramia_pomiformis_MT024676.1
Lycopodium_clavatum_MH549642.1
Selaginella_lepidophylla_MK089531.1
Isoetes_malinverniana_MH549640.1
Equisetum_arvense_GU191334.1
Tmesipteris_elongata_KJ569699.1
Mankyua_chejuensis_KP205433.1
Hymenophyllum_coreanum_MN583335.1
Danaea_sellowiana_MN412588.1
Christensenia_aesculifolia_MN056350.1
Ginkgo_biloba_JN867585.1
Cycas_debaoensis_KU743927.1
Taxus_baccata_KR476375.1
Sciadopitys_verticillata_KT601210.1
Welwitschia_mirabilis_EU342371.1
Ephedra_distachya_MG594460.1
Amborella_trichopoda_AJ506156.2
Ceratophyllum_demersum_EF614270.1
Lilium_candidum_MK753244.1
Acorus_calamus_AJ879453.1
Corydalis_chrysosphaera_MZ983401.1
Vicia_faba_MT120813.1
Apium_graveolens_MK036045.1
Chloranthus_henryi_MK922064.1
Tasmannia_lanceolata_MT555078.1
Saruma_henryi_MG520100.1
Cinnamomum_aromaticum_MN698963.1
Chimonanthus_grammatus_MW166217.1
Calycanthus_chinensis_MG561304.1
Myristica_yunnanensis_MN495964.1
Magnolia_virginiana_MN990608.1
Annona_reticulata_MT742547.1
Galdieria_sulphuraria_KJ700459.1
Cyanidium_caldarium_AF022186.2
Cyanidioschyzon_merolae_AB002583.1
Porphyridium_purpureum_MF401423.1
Flintiella_sanguinaria_KY709211.1
Erythrolobus_coxiae_MZ853148.1
Madagascaria_erythrocladioides_MZ681963.1
Erythrotrichia_carnea_KX284721.2
Pulvinaster_venetus_MZ681965.1
Boldiaceae_sp._CCMP3255_MZ681962.1
Tsunamia_transpacifica_MZ680366.1
Rhodospora_sordida_MZ603805.1
Bangiopsis_subsimplex_KX284718.1
Rhodella_violacea_MZ853150.1
Corynoplastis_japonica_KY709210.1
Glaucosphaera_vacuolata_MZ895126.1
Bulboplastis_apyrenoidosa_KY709209.1
Porphyra_umbilicalis_MF385003.1
Neopyropia_fucicola_KJ776837.1
Neoporphyra_perforata_KC904971.1
Hildenbrandia_rivularis_KX284723.1
Apophlaea_sinclairii_KX284716.1
Sheathia_arcuata_NC_035231.1
Palmaria_palmata_KX284726.1
Ahnfeltia_plicata_KX284715.1
Asparagopsis_taxiformis_KX284717.1
Gelidium_vagum_KT266787.1
Chondrus_crispus_HF562234.1
Rhodogorgon_sp._MH281630_MH281630.1
Renouxia_sp._MH281629_MH281629.1
Sporolithon_durum_KT266785.1
Mesophyllum_expansum_MEX45
Lithothamnion_sp._MH281627_MH281627.1
Porolithon_onkodes_KY212106.1
Lithophyllum_sp._LIT44_LIT44
Neogoniolithon_spectabile_MH281628.1
Corallina_officinalis_MT211887.1
Storeatula_sp._CCMP1868_KY856940.1
Guillardia_theta_AF041468_AF041468.1
Cryptomonas_curvata_CCAP979_52_LC484192.1
Chroomonas_mesostigmatica_CCMP1168_KY860574.1
Pavlomulina_rannunculiformis_LC564893.1
Phaeocystis_globosa_CNS00067_MT471328.1
Chrysochromulina_parva_MG520331.1
Chrysotila_dentata_MZ819921.1
Isochrysis_galbana_MT304829.1
Emiliania_huxleyi_CCMP_373_AY741371.1
Heterosigma_akashiwo_EU168190.1
Vaucheria_litorea_EU912438.1
Ishige_okamurae_MW762687.1
Egregia_menziesii_MZ156039.1
Akkesiphycus_lubricus_MZ156027.1
Trachydiscus_minutus_KJ624065.1
Monodopsis_sp._MarTras21_KX839260.1
Ochromonas_sp._CCMP1393_KJ877675.1
Neotessella_volvocina_MH795132.1
Mallomonas_splendens_MH795131.1
Synura_uvella_MH795130.1
Synura_petersenii_MH795128.1
Rhizochromulina_marina_MK561360.1
Aureoumbra_lagunensis_GQ231542.1
Triparma_laevis_AP014625.1
Leptocylindrus_danicus_KC509524.1
Proboscia_sp._MG755791_MG755791.1
Coscinodiscus_granii_MW561225.1
Rhizosolenia_setigera_MG755793.1
Rhizosolenia_imbricata_KJ958482.1
Guinardia_striata_MG755796.1
Lithodesmium_undulatum_KC509525.1
Thalassiosira_weissflogii_KJ958485.1
Eunotogramma_sp._MG755797_MG755797.1
Toxarium_undulatum_KX619437.1
Plagiogrammopsis_vanheurckii_MG755794.1
Chaetoceros_tenuissimus_MW845778.1
Chaetoceros_laevisporus_MW845779.1
Eucampia_zodiacus_MW412838.1
Cerataulina_daemon_KJ958484.1
Attheya_longicornis_MG755798.1
Psammoneis_obaidii_MG755803.1
Asterionellopsis_glacialis_KC509520.1
Nanofrustulum_shiloi_MN276191.1
Asterionella_formosa_KC509519.1
Seminavis_robusta_MH356727.1
Phaeodactylum_tricornutum_EF067920.1
Nitzschia_supralitorea_MT383638.1
Tryblionella_apiculata_MT383642.1
Cylindrotheca_closterium_KC509522.1
0.75
0.99
0.63
0.49
0.95
0.93
0.125

References

Aguirre J, Baceta JI, Braga JC. 2007. Recovery of marine primary producers after the Cretaceous–Tertiary mass extinction: Paleocene calcareous red algae from the Iberian Peninsula. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* 249:393–411.

Anderson HM. 1976. A review of the Bryophyta from the Upper Triassic Molteno Formation, Karroo Basin, South Africa. Available from: http://wiredspace.wits.ac.za/handle/10539/16189

Barfod GH, Albarède F, Knoll AH, Xiao S, Télouk P, Frei R, Baker J. 2002. New Lu–Hf and Pb–Pb age constraints on the earliest animal fossils. *Earth Planet. Sci. Lett.* 201:203–212.

Bengtson S, Sallstedt T, Belivanova V, Whitehouse M. 2017. Three-dimensional preservation of cellular and subcellular structures suggests 1.6 billion-year-old crown-group red algae. *PLOS Biol.* 15:e2000735.

Bonamo PM. 1977. Rellimia Thomsonii (progymnospermopsida) from the Middle Devonian of New York State. *Am. J. Bot.* 64:1272–1285.

Bown P. 1998. Calcareous nannofossil biostratigraphy. Chapman and Hall; Kluwer Academic

Bucur II. 1994. Lower Cretaceous Halimedaceae and Gymnocodiaceae from southern Carpathians and Apuseni Mountains (Romania) and the systematic position of the Gymnocodiaceae. *Beitr. Zur Paläontol.* 19:13–37.

Butterfield NJ. 2000. Bangiomorpha pubescens n. gen., n. sp.: implications for the evolution of sex, multicellularity, and the Mesoproterozoic/Neoproterozoic radiation of eukaryotes. *Paleobiology* 26:386–404.

Butterfield NJ, Knoll AH, Swett K. 1990. A bangiophyte red alga from the Proterozoic of arctic Canada. *Science* 250:104–107.

Carlisle EM, Jobbins M, Pankhania V, Cunningham JA, Donoghue PCJ. 2021. Experimental taphonomy of organelles and the fossil record of early eukaryote evolution. *Sci. Adv.* 7:eabe9487.

Cavalier-Smith T. 2006. Cell evolution and Earth history: stasis and revolution. *Philos. Trans. R. Soc. B Biol. Sci.* 361:969–1006.

Chatalov A, Bonev N, Ivanova D. 2015. Depositional characteristics and constraints on the mid-Valanginian demise of a carbonate platform in the intra-Tethyan domain, Circum-Rhodope Belt, northern Greece. *Cretac. Res.* 55:84–115.

Clarke JT, Warnock RCM, Donoghue PCJ. 2011. Establishing a time-scale for plant evolution. *New Phytol.* 192:266–301.

Condon D, Zhu M, Bowring S, Wang W, Yang A, Jin Y. 2005. U-Pb Ages from the Neoproterozoic Doushantuo Formation, China. *Science* 308:95–98.

Creaser RA, Grütter H, Carlson J, Crawford B. 2004. Macrocrystal phlogopite Rb–Sr dates for the Ekati property kimberlites, Slave Province, Canada: evidence for multiple intrusive episodes in the Paleocene and Eocene. *Lithos* 76:399–414.

Crepet WL, Nixon KC, Gandolfo MA. 2005. An extinct calycanthoid taxon, Jerseyanthus calycanthoides , from the Late Cretaceous of New Jersey. *Am. J. Bot.* 92:1475–1485.

Damsté JSS, Muyzer G, Abbas B, Rampen SW, Massé G, Allard WG, Belt ST, Robert J-M, Rowland SJ, Moldowan JM, et al. 2004. The Rise of the Rhizosolenid Diatoms. *Science* 304:584–587.

De vargas C, Aubry M-P, Probert I, Young J. 2007. CHAPTER 12 - Origin and Evolution of Coccolithophores: From Coastal Hunters to Oceanic Farmers. In: Falkowski PG, Knoll AH, editors. Evolution of Primary Producers in the Sea. Burlington: Academic Press. p. 251–285. Available from: https://www.sciencedirect.com/science/article/pii/B9780123705181500138

Dilcher DL, Sun G, Ji Q, Li H. 2007. An early infructescence Hyrcantha decussata (comb. nov.) from the Yixian Formation in northeastern China. *Proc. Natl. Acad. Sci.* 104:9370–9374.

DiMichele WA, Phillips TL. 1977. Monocyclic Psaronius from the lower Pennsylvanian of the

Illinois Basin. *Can. J. Bot.* 55:2514–2524.
Doyle JA, Endress PK. 2010. Integrating Early Cretaceous fossils into the phylogeny of living angiosperms: Magnoliidae and eudicots. *J. Syst. Evol.* 48:1–35.
Edwards D, Feehan J. 1980. Records of Cooksonia-type sporangia from late Wenlock strata in Ireland. *Nature* 287:41–42.
Edwards D, Feehan J, Smith DG. 1983. A late Wenlock flora from Co. Tipperary, Ireland. *Bot. J. Linn. Soc.* 86:19–36.
Farinacci A, Howe RW. 2022. The Farinacci & Howe Catalog of Calcareous Nannofossils (1969-2022). Available from: mikrotax.org/Nannotax3/index.php?id=50001
Flügel E. 1988. Halimeda: paleontological record and palaeoenvironmental significance. *Coral Reefs* 6:123–130.
Gensel PG, Albright VM. 2006. Leclercqia complexa from the Early Devonian (Emsian) of northern New Brunswick, Canada. *Rev. Palaeobot. Palynol.* 142:103–121.
Ghosh AK, Maithy PK. 1996. On the present status of coralline red alga Archaeolithothamnium Roth. from India.
Gibson TM, Shih PM, Cumming VM, Fischer WW, Crockford PW, Hodgskiss MSW, Wörndle S, Creaser RA, Rainbird RH, Skulski TM, et al. 2018. Precise age of Bangiomorpha pubescens dates the origin of eukaryotic photosynthesis. *Geology* 46:135–138.
Girard V, Martin SS, Buffetaut E, Martin J-PS, Néraudeau D, Peyrot D, Roghi G, Ragazzi E, Suteethorn V. 2020. Thai amber: insights into early diatom history? *BSGF - Earth Sci. Bull.* 191:23.
Gradstein FM ed. 2012. The geologic time scale 2012. 1st ed. Amsterdam ; Boston: Elsevier
Gulbrandsen ØS, Andresen IJ, Krabberød AK, Bråte J, Shalchian-Tabrizi K. 2021. Phylogenomic analysis restructures the ulvophyceae. *J. Phycol.* 57:1223–1233.
Guo C-Q, Edwards D, Wu P-C, Duckett JG, Hueber FM, Li C-S. 2012. *Riccardiothallus devonicus* gen. et sp. nov., the earliest simple thalloid liverwort from the Lower Devonian of Yunnan, China. *Rev. Palaeobot. Palynol.* 176:35–40.
Gustavson TC, Delevoryas T. 1992. Caulerpa-like marine alga from Permian strata, Palo Duro Basin, West Texas. *J. Paleontol.* 66:160–161.
Hagino K, Young JR. 2015. Biology and Paleontology of Coccolithophores (Haptophytes). In: Ohtsuka S, Suzaki T, Horiguchi T, Suzuki N, Not F, editors. Marine Protists: Diversity and Dynamics. Tokyo: Springer Japan. p. 311–330. Available from: https://doi.org/10.1007/978-4-431-55130-0_12
Hajós M. 1975. Late Cretaceous Archaeomonadaceae, Diatomaceae, and Silicoflagellatae from the South Pacific Ocean, Deep Sea Drilling Project, Leg 29, Site 275. *Initial Rep. Deep Sea Drill. Proj.* 29:913–1009.
Hedges SB, Tao Q, Walker M, Kumar S. 2018. Accurate timetrees require accurate calibrations. *Proc. Natl. Acad. Sci.* 115:E9510–E9511.
Henderiks J, Sturm D, Šupraha L, Langer G. 2022. Evolutionary Rates in the Haptophyta: Exploring Molecular and Phenotypic Diversity. *J. Mar. Sci. Eng.* 10:798.
Hrabovskỳ M. 2020. Leaf evolution and classification. 1. Lycopodiopsida. *Acta Bot. Univ. Comen.* 55:19–37.
Hübers M, Kerp H. 2012. Oldest known mosses discovered in Mississippian (late Visean) strata of Germany. *Geology* 40:755–758.
Hughes NF, McDougall AB. 1990. Barremian-Aptian angiospermid pollen records from southern England. *Rev. Palaeobot. Palynol.* 65:145–151.
Jackson C, Knoll AH, Chan CX, Verbruggen H. 2018. Plastid phylogenomics with broad taxon sampling further elucidates the distinct evolutionary origins and timing of secondary green plastids. *Sci. Rep.* 8:1523.
Jones HM, Simpson GE, Stickle AJ, Mann DG. 2005. Life history and systematics of Petroneis (Bacillariophyta), with special reference to British waters. *Eur. J. Phycol.* 40:61–87.
Kooistra WHCF, Gersonde R, Medlin LK, Mann DG. 2007. CHAPTER 11 - The Origin and Evolution of the Diatoms: Their Adaptation to a Planktonic Existence. In: Falkowski PG, Knoll AH, editors. Evolution of Primary Producers in the Sea. Burlington:

Academic Press. p. 207–249. Available from: https://www.sciencedirect.com/science/article/pii/B9780123705181500126
Kotyk ME, Basinger JF, Gensel PG, de Freitas TA. 2002. Morphologically complex plant macrofossils from the Late Silurian of Arctic Canada. *Am. J. Bot.* 89:1004–1013.
Männik P, Tinn O, Loydell DK, Ainsaar L. 2016. Age of the Kalana Lagerstätte, early Silurian, Estonia. *Est. J. Earth Sci.* 65.
Massoni J, Doyle J, Sauquet H. 2015. Fossil calibration of Magnoliidae, an ancient lineage of angiosperms. *Palaeontol. Electron.* 18:1–25.
Mastik V, Tinn O. 2015. New dasycladalean algal species from the Kalana Lagerstätte (Silurian, Estonia). *J. Paleontol.* 89:262–268.
Medlin LK. 2015. A timescale for diatom evolution based on four molecular markers: reassessment of ghost lineages and major steps defining diatom evolution. *Vie Milieu-Life Environ.*
Medlin LK. 2016a. Evolution of the diatoms: major steps in their evolution and a review of the supporting molecular and morphological evidence. *Phycologia* 55:79–103.
Medlin LK. 2016b. Opinion: Can coalescent models explain deep divergences in the diatoms and argue for the acceptance of paraphyletic taxa at all taxonomic hierarchies? *Nova Hedwig.*:107–128.
Medlin LK, Sáez AG, Young JR. 2008. A molecular clock for coccolithophores and implications for selectivity of phytoplankton extinctions across the K/T boundary. *Mar. Micropaleontol.* 67:69–86.
Mohr BA, Coiffard C, Bernardes-de-Oliveira ME. 2013. Schenkeriphyllum glanduliferum, a new magnolialean angiosperm from the Early Cretaceous of Northern Gondwana and its relationships to fossil and modern Magnoliales. *Rev. Palaeobot. Palynol.* 189:57–72.
Morris JL, Puttick MN, Clark JW, Edwards D, Kenrick P, Pressel S, Wellman CH, Yang Z, Schneider H, Donoghue PCJ. 2018a. The timescale of early land plant evolution. *Proc. Natl. Acad. Sci.* 115:E2274–E2283.
Morris JL, Puttick MN, Clark JW, Edwards D, Kenrick P, Pressel S, Wellman CH, Yang Z, Schneider H, Donoghue PCJ. 2018b. Reply to Hedges et al.: Accurate timetrees do indeed require accurate calibrations. *Proc. Natl. Acad. Sci.* 115:E9512–E9513.
Peña V, Vieira C, Braga JC, Aguirre J, Rösler A, Baele G, De Clerck O, Le Gall L. 2020. Radiation of the coralline red algae (Corallinophycidae, Rhodophyta) crown group as inferred from a multilocus time-calibrated phylogeny. *Mol. Phylogenet. Evol.* 150:106845.
Rösler A, Perfectti F, Peña V, Aguirre J, Braga JC. 2017. Timing of the evolutionary history of Corallinaceae (Corallinales, Rhodophyta). *J. Phycol.* 53:567–576.
Ross R, Sims PA. 1985. Some genera of the Biddulphiaceae (diatoms) with interlocking linking spines. *Bull. Br. Mus. Nat. Hist. Bot.*
Rubinstein CV, Gerrienne P, de la Puente GS, Astini RA, Steemans P. 2010. Early Middle Ordovician evidence for land plants in Argentina (eastern Gondwana). *New Phytol.* 188:365–369.
Sánchez-Baracaldo P, Cardona T. 2020. On the origin of oxygenic photosynthesis and Cyanobacteria. *New Phytol.* 225:1440–1446.
Satkoski AM, Beukes NJ, Li W, Beard BL, Johnson CM. 2015. A redox-stratified ocean 3.2 billion years ago. *Earth Planet. Sci. Lett.* 430:43–53.
Schlagintweit F, Pavlik W. 2008. New findings of halimedacean algae from the Late Triassic Dachstein Limestone of the Northern Calcareous Alps (Hochschwab Area, Styria, Austria). *Geol. Croat.* 61:129–133.
Sims PA, Mann DG, Medlin LK. 2006. Evolution of the diatoms: insights from fossil, biological and molecular data. *Phycologia* 45:361–402.
Siver PA, Wolfe AP. 2005. Eocene scaled chrysophytes with pronounced modern affinities. *Int. J. Plant Sci.* 166:533–536.
Su D, Yang L, Shi X, Ma X, Zhou X, Hedges SB, Zhong B. 2021. Large-Scale Phylogenomic Analyses Reveal the Monophyly of Bryophytes and Neoproterozoic Origin of Land

Plants. *Mol. Biol. Evol.* 38:3332–3344.
Tang Q, Pang K, Yuan X, Xiao S. 2020. A one-billion-year-old multicellular chlorophyte. *Nat. Ecol. Evol.* 4:543–549.
Tomás S, Aguirre J, Braga JC, Martín-Closas C. 2007. Late Hauterivian coralline algae (Rhodophyta, Corallinales) from the Iberian Chain (E Spain). Taxonomy and the evolution of multisporangial reproductive structures. *Facies* 53:79–95.
Trivett ML. 2015. Growth Architecture, Structure, and Relationships of Cordaixylon iowensis Nov. comb. (Cordaitales). *Int. J. Plant Sci.* [Internet]. Available from: https://www-journals-uchicago-edu.inee.bib.cnrs.fr/doi/10.1086/297031
Vachard D, Hauser M, Matter A, Peters T, Martini R, Zaninetti L. 2001. New algae and problematica of algal affinity from the permian of the Aseelah Unit of the batain plain (East Oman). *Geobios* 34:375–404.
Verbruggen H, Ashworth M, LoDuca ST, Vlaeminck C, Cocquyt E, Sauvage T, Zechman FW, Littler DS, Littler MM, Leliaert F, et al. 2009. A multi-locus time-calibrated phylogeny of the siphonous green algae. *Mol. Phylogenet. Evol.* 50:642–653.
Witkowski J, Harwood DM, Chin K. 2011. Taxonomic composition, paleoecology and biostratigraphy of Late Cretaceous diatoms from Devon Island, Nunavut, Canadian High Arctic. *Cretac. Res.* 32:277–300.
Witkowski J, Sims PA, Harwood DM. 2011. Rutilariaceae redefined: a review of fossil bipolar diatom genera with centrally positioned linking structures, with implications for the origin of pennate diatoms. *Eur. J. Phycol.* 46:378–398.
Xiao S, Knoll AH, Yuan X, Pueschel CM. 2004. Phosphatized multicellular algae in the Neoproterozoic Doushantuo Formation, China, and the early evolution of florideophyte red algae. *Am. J. Bot.* 91:214–227.
Xiao S, Yuan X, Steiner M, Knoll AH. 2002. Macroscopic carbonaceous compressions in a terminal Proterozoic shale: a systematic reassessment of the Miaohe biota, South China. *J. Paleontol.* 76:347–376.
Zhang Y, Yin L, Xiao S, Knoll AH. 1998. Permineralized Fossils from the Terminal Proterozoic Doushantuo Formation, South China. *Mem. Paleontol. Soc.* 50:1–52.