


# Cryogenian Origins of Multicellularity in Archaeplastida

Alexander M.C. Bowles <sup>1,2,\*</sup>, Christopher J. Williamson <sup>1</sup>, Tom A. Williams <sup>2</sup>,  
and Philip C.J. Donoghue <sup>2</sup>

<sup>1</sup>School of Geographical Sciences, University of Bristol, Bristol BS8 1SS, UK

<sup>2</sup>Bristol Palaeobiology Group, School of Biological Sciences and School of Earth Sciences, Life Sciences Building, University of Bristol, Bristol BS8 1TQ, UK

\*Corresponding author: E-mail: [zl20161@bristol.ac.uk](mailto:zl20161@bristol.ac.uk).

Accepted: February 03, 2024

## Abstract

Earth was impacted by global glaciations during the Cryogenian (720 to 635 million years ago; Ma), events invoked to explain both the origins of multicellularity in Archaeplastida and radiation of the first land plants. However, the temporal relationship between these environmental and biological events is poorly established, due to a paucity of molecular and fossil data, precluding resolution of the phylogeny and timescale of archaeplastid evolution. We infer a time-calibrated phylogeny of early archaeplastid evolution based on a revised molecular dataset and reappraisal of the fossil record. Phylogenetic topology testing resolves deep archaeplastid relationships, identifying two clades of Viridiplantae and placing Bryopsidales as sister to the Chlorophyceae. Our molecular clock analysis infers an origin of Archaeplastida in the late-Paleoproterozoic to early-Mesoproterozoic (1712 to 1387 Ma). Ancestral state reconstruction of cytomorphological traits on this time-calibrated tree reveals many of the independent origins of multicellularity span the Cryogenian, consistent with the Cryogenian multicellularity hypothesis. Multicellular rhodophytes emerged 902 to 655 Ma while crown-Anydrophyta (Zygnematophyceae and Embryophyta) originated 796 to 671 Ma, broadly compatible with the Cryogenian plant terrestrialization hypothesis. Our analyses resolve the timetree of Archaeplastida with age estimates for ancestral multicellular archaeplastids coinciding with the Cryogenian, compatible with hypotheses that propose a role of Snowball Earth in plant evolution.

**Key words:** plant evolution, multicellularity, Cryogenian, Archaeplastida, Streptophyta, phylogenetics.

## Significance

Glaciation events of the distant past (720 to 635 million years ago) have been proposed as a driving force for the origin of multicellularity in plants (Archaeplastida). Here, evolutionary analyses produce divergence time estimates for multicellular plants (e.g. red algae and streptophyte algae) spanning this period (the Cryogenian). Our work is compatible with the hypothesis that ancient glaciations facilitated the origin of multicellularity in plants.

## Introduction

The early evolution of the superkingdom Archaeplastida, from the first photosynthetic eukaryotes to the earliest land plants, transformed the biosphere (Lenton et al. 2016; Bengtson et al. 2017; One Thousand Plant Transcriptomes Initiative 2019; Burki et al. 2020; Delaux and Schornack 2021; Schön et al. 2021; Bowles et al. 2022; Irisarri et al. 2022), paving the way for flora that

would dominate terrestrial environments. Archaeplastid evolution encompasses a vast period in Earth history, with divergence time estimates spanning the middle Paleoproterozoic to late Neoproterozoic (Sánchez-Baracaldo et al. 2017; Betts et al. 2018; Lutzoni et al. 2018; Morris et al. 2018; Strasser et al. 2021). During this formative period, early archaeplastids influenced the composition of the atmosphere as the dominant primary producers (Brocks et al. 2017;

© The Author(s) 2024. Published by Oxford University Press on behalf of Society for Molecular Biology and Evolution.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

Braakman 2019), provided new ecological niches promoting the diversity of lineages spanning the tree of life (Lutzoni et al. 2018; Del Cortona et al. 2020) and led to the development of more complex food webs (Brocks et al. 2017). Descendants of the archaeplastid crown-ancestor include the rhodophytes, glaucophytes, chlorophytes, streptophyte algae, and land plants (Guiry et al. 2014; Christenhusz and Byng 2016; Corlett 2016; Lughadha et al. 2016), all of which have responded to environmental challenges with novel adaptive mechanisms.

In Archaeplastida, macroscopic organization ranges from filamentous species to siphonous chlorophytes to 3D multicellular organisms, with multicellularity evolving independently on several occasions (Umen 2014). These independent transitions to multicellularity likely occurred during the Neoproterozoic (Knoll 2011; Del Cortona et al. 2020). The Cryogenian (720 to 635 million years ago; Ma) is characterized by two major glaciation events, the Sturtian (717 to 643 Ma) and the Marinoan (656 to 626 Ma) (Hoffman et al. 1998; Stern et al. 2006). These global glaciations, termed Snowball Earth events, encased the Earth in ice (Hoffman et al. 2017), and they have been hypothesized as an evolutionary driver of multicellularity in Archaeplastida, Metazoa, and fungi (Brocks et al. 2017; Del Cortona et al. 2020; Simpson 2021). In particular, Simpson (Simpson 2021) has hypothesized that the 70 million year-long Sturtian glaciation would have led to the near-complete loss of temperate and tropical water habitats, providing a geologically unique, unexplored marine ecological space for the evolution of multicellularity (the Cryogenian multicellularity hypothesis). Seawaters of the Cryogenian would have increased in viscosity, limiting the ability of unicellular life to be motile, feed, and acquire nutrients. Multicellularity may have provided a mechanism to increase motility and size, enabling optimization of nutrient acquisition and metabolic rate (Solari et al. 2006; DeLong et al. 2010).

Cryogenian glaciations have also been proposed as a driver of the colonization of land by early archaeplastids in the Cryogenian Plant Terrestrialization hypothesis (CPT; Becker 2013; Williamson et al. 2019; Žárský et al. 2022). The CPT hypothesis argues that the terrestrial component of snowball Earth environments was a crucial driver for land plant terrestrialization, serving as an intermediary between aquatic and terrestrial habitats. This is because ice environments are dry, exposed, and extremely cold, in contrast to the high viscosity cold sub-glacial seawater habitats. Glacial environments would have exposed ancestral streptophytes to the same environmental challenges that their descendants would have faced in adapting to life on land outside, viz. extreme temperature, irradiance, and desiccation (de Vries and Archibald 2018). Thus, the CPT hypothesis argues that ancestral land plants would have been preadapted to life on land, exapting adaptations evolved on the icy surfaces of Snowball Earth.

Both the Cryogenian multicellularity and CPT hypotheses make specific predictions about the time intervals during which the evolution of multicellularity and terrestrialization, respectively, occurred within Archaeplastida. The Cryogenian multicellularity hypothesis requires that archaeplastid multicellularity evolved coincident with one or more of the Snowball Earth events and the CPT hypothesis requires that land plants evolved in their aftermath. Here, we examine whether these hypotheses are consistent with the timescale of archaeplastid evolution using phylogenetics, divergence time estimation, and ancestral trait reconstruction. While previous studies have aimed to understand the relationships and timescale of green plant evolution, these have often focused specifically on chlorophyte (Del Cortona et al. 2020; Nie et al. 2020; Hou et al. 2022; Yang et al. 2023) and streptophyte algae (Wickett et al. 2014; Morris et al. 2018; Puttick et al. 2018; Hess et al. 2022). Additional work has investigated the phylogeny and divergence times of nongreen plant lineages but at higher taxonomic resolution (e.g. eukaryotes) and with shallower species sampling (Sánchez-Baracaldo et al. 2017; Betts et al. 2018; Lutzoni et al. 2018; Strasser et al. 2021). This work incorporates both a breadth and depth of taxonomic sampling, spanning rhodophytes, glaucophytes, chlorophytes, and streptophyte algae.

We first inferred a rooted phylogeny of Archaeplastida, using both concatenation- and coalescent-based approaches, combined with tree topology testing. We then estimated the timescale of archaeplastid evolution, incorporating an updated appraisal of the fossil record and a broad taxonomic sampling of genes. Finally, we reconstructed the evolution of multicellularity across Archaeplastida, demonstrating that their timing coincides with the Cryogenian. These results are compatible with the hypotheses that environmental conditions during Cryogenian Snowball Earth glaciations underpinned the evolution of archaeplastid multicellularity and the exaptive evolution of land plant innovations envisaged by the CPT hypothesis.

## Results

### Phylogenomics Resolves the Relationships of Archaeplastida

Recent genome and transcriptome sequencing of under-represented lineages can be utilized to address poorly resolved nodes in the archaeplastid phylogeny. We built a comprehensive dataset of 241 species, with improved representation for previously under-sampled lineages, most notably streptophytes, early diverging chlorophytes, and rhodophytes (supplementary data S1, Supplementary Material online, supplementary information S1, Supplementary Material online). Maximum likelihood analysis, based on a concatenation of 96 genes using the best-fitting model for each gene partition, resolves the fundamental split within Archaeplastida between rhodophytes and glaucophytes

plus Viridiplantae, recovers two clades of green plants (chlorophytes and streptophytes), and supports Bryopsidales as sister group to the chlorophyceae (Fig. 1). Our phylogenetic analyses also support the monophyly of Archaeplastida, the sister group relationship between Zygnematomyceae and land plants (Anydrophyta) and the monophyly of bryophytes and tracheophytes (Embryophyta; Fig. 1), in agreement with recent work (Wickett et al. 2014; Puttick et al. 2018; Gawryluk et al. 2019; One Thousand Plant Transcriptomes Initiative 2019; Price et al. 2019; Harris et al. 2020). These relationships were also supported in our coalescent-based analyses (supplementary information S2, Supplementary Material online).

Phylogenetic conflict associated with contested relationships was addressed using approximately unbiased (AU) tests (Fig. 2), allowing for evaluation of the support for the recovered and alternative phylogenetic relationships. The relationships recovered receive support based on this analysis, while previously proposed alternative topologies for Prasinodermophyceae (Li et al. 2020) and Bryopsidales (Fang et al. 2017) were not supported by the AU test. Several relationships remain enigmatic, with either support for multiple competing topologies or conflicting support across datasets (Fig. 2). Overall, these analyses provide a robust framework to date early archaeplastid evolution.

### The Timescale of Early Archaeplastid Evolution

We estimated divergence times along the resolved phylogeny of Archaeplastida (Fig. 3), aided by a dataset of 15 fossil calibrations (supplementary information S3, Supplementary Material online). The red alga, *Bangiomorpha pubescens* and the green alga, *Proterocladus antiquus* are considered as the oldest convincing records of Archaeplastida (See Materials and Methods for further discussion on archaeplastid fossils). Within the multicellular land plants, robustly supported fossil calibrations were identified for the major nodes in the embryophyte phylogeny.

Our results support an origin of Archaeplastida during the Mesoproterozoic (1,712 to 1,387 Ma), while crown-rhodophytes, crown-glaucophytes, and crown-Viridiplantae emerged 1,613 to 1,284 Ma, 1,624 to 1,317 Ma, and 1,369 to 1,147 Ma, respectively (Fig. 3). Among Viridiplantae, crown-chlorophytes diverged by 1,297 to 1,088 Ma while crown-streptophytes diverged by 1,199 to 998 Ma (late Mesoproterozoic—earliest Neoproterozoic). Within the streptophytes, Phragmoplastophyta emerged 901 to 760 Ma, crown-Anydrophyta (the clade of Zygnematomyceae and land plants) diverged 796 to 671 Ma, while crown-Embryophyta emerged 586 to 526 Ma in a late Ediacaran to Cambrian interval.

To explore the impact of different models of substitution and rate heritability on divergence time estimation, we undertook analyses using the CAT-GTR model in Phylobayes and the correlated rates model in MCMCTree, respectively

(supplementary informations S6 and S7, Supplementary Material online). These different approaches produced divergence time estimates that were largely congruent with our main analysis. For example, the correlated rates analysis recovered a divergence time for crown-Anydrophyta of 742 to 654 Ma whilst the initial analysis recovered an estimate of 796 to 671 Ma. These results suggest that the divergence timescale derived from our main analysis, using the uncorrelated rates model in MCMCTree, is robust of these methodological variables.

### Independent Emergence of Multicellular Archaeplastids

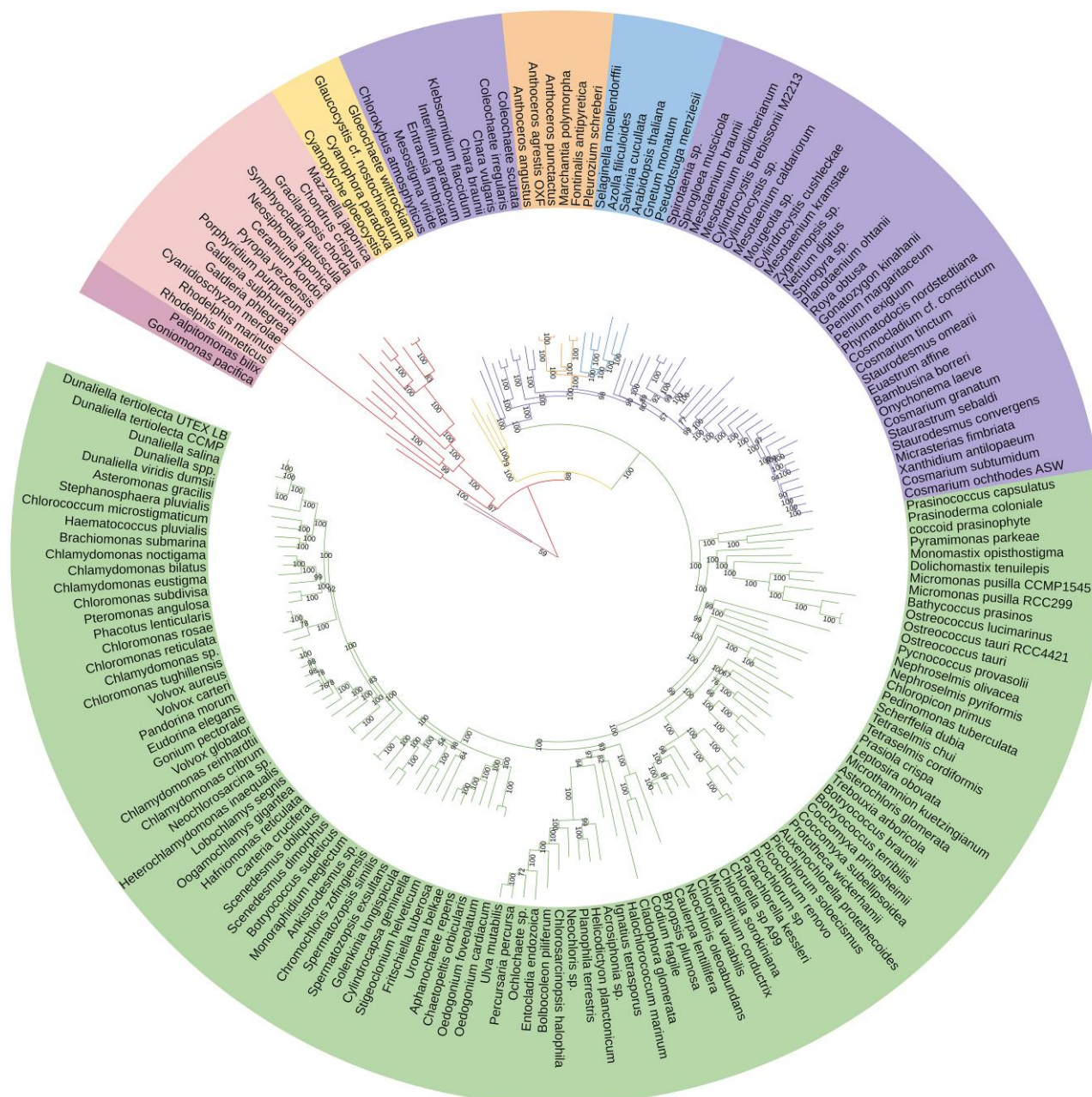
Stochastic mapping of ancestral states of multicellularity in early archaeplastids revealed that multicellularity in streptophytes and rhodophytes was present in the ancestors of Anydrophyta and Eurhodophytina, respectively (Fig. 4, supplementary informations S4 and S5, Supplementary Material online). Taking our timescale analysis into consideration, the radiation of these two multicellular lineages spans the Cryogenian (720 to 635 Ma). The diversification of chlorophytes was accompanied by multiple origins of multicellularity (e.g. Bryopsidales, Chaetophorales, and Oedogoniales). The diversification of multicellular Chaetaophorales (661 to 441 Ma) possibly occurred during the Cryogenian while the radiation of multicellular Bryopsidales (609 to 413 Ma) and Ulvales (364 to 237 Ma) occurred more recently.

## Discussion

### Comparison to Previous Studies

We present a time-calibrated phylogeny of archaeplastid evolution using a large-scale multigene dataset and an up to date suite of fossil calibrations. Modeling the evolution of cytomorphology within this framework reveals multiple, independent transitions from unicellular to multicellular growth in archaeplastids during the Neoproterozoic (Fig. 4). Recent studies have resulted in contradictory phylogenies. The phylogeny from the One Thousand Plant Transcriptome Project (1KP project) recovered Prasinococcales as sister to streptophytes (One Thousand Plant Transcriptomes Initiative 2019) while analysis of the *Prasinoderma coloniale* genome recovered Prasinococcales as the sister group to all other green plants (Li et al. 2020). The 1KP project also recovered support for two positions for Bryopsidales while recent phylogenomic analyses found support for contrasting positions for Bryopsidales between analytical approaches (One Thousand Plant Transcriptomes Initiative 2019; Del Cortona et al. 2020). Maximum likelihood analysis of the *Spirogloea muscicola* genome inferred Spirogloeaales as an early branching lineage within Zygnematomyceae, although Bayesian inference pointed to a potential sister group relationship between Spirogloeaales and land plants (Cheng et al. 2019). Our phylogenetic analyses, both concatenation- and coalescence-based, mark a



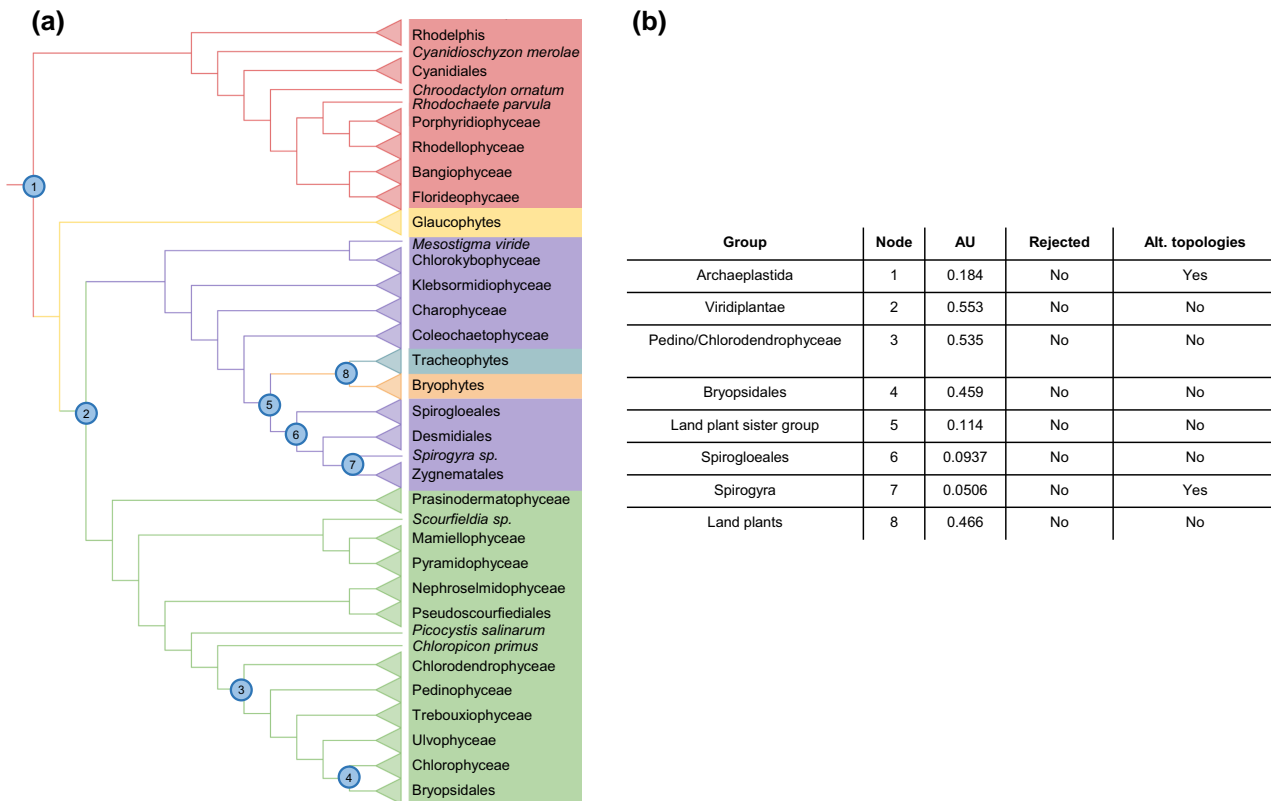


**Fig. 1.**—Phylogeny of Archaeplastida based on concatenation-based analysis of 96 genes representing 180 species. Branches and taxa are grouped into Outgroup (red group including *Palpitomonas bilix*), Rhodophytes (pink group including *Chondrus crispus*), Glaucophytes (yellow group including *Cyanophora paradoxa*), Chlorophytes (green group including *Volvox carteri*), Streptophyte algae (purple group including *Chara braunii* and *Spiroglaea muscicola*), Bryophytes (blue group including *Marchantia polymorpha*) and Tracheophytes (orange group including *Arabidopsis thaliana*).

significant consolidation of recent thinking on archaeplastid evolution, addressing some long-standing questions including the grouping of green plants and the nonmonophyly of Ulvophyceae. Although our study focuses on the early evolution of Archaeplastida, the relationships within land plants (e.g. monophyly of bryophytes) agree with the recent phylogenetic analysis (Puttick et al. 2018; One Thousand Plant Transcriptomes Initiative 2019; Harris et al. 2020, 2022),

providing additional support for these formerly uncertain regions of the archaeplastid tree.

Given the deep evolutionary history of early archaeplastid evolution, previous analyses have wide-ranging divergence time estimates. The origin of Archaeplastida has been estimated to the middle Palaeoproterozoic to early Mesoproterozoic interval (2137 to 1118 Ma; 5.7 to 9.11). Our analysis refines these estimates, recovering a divergence



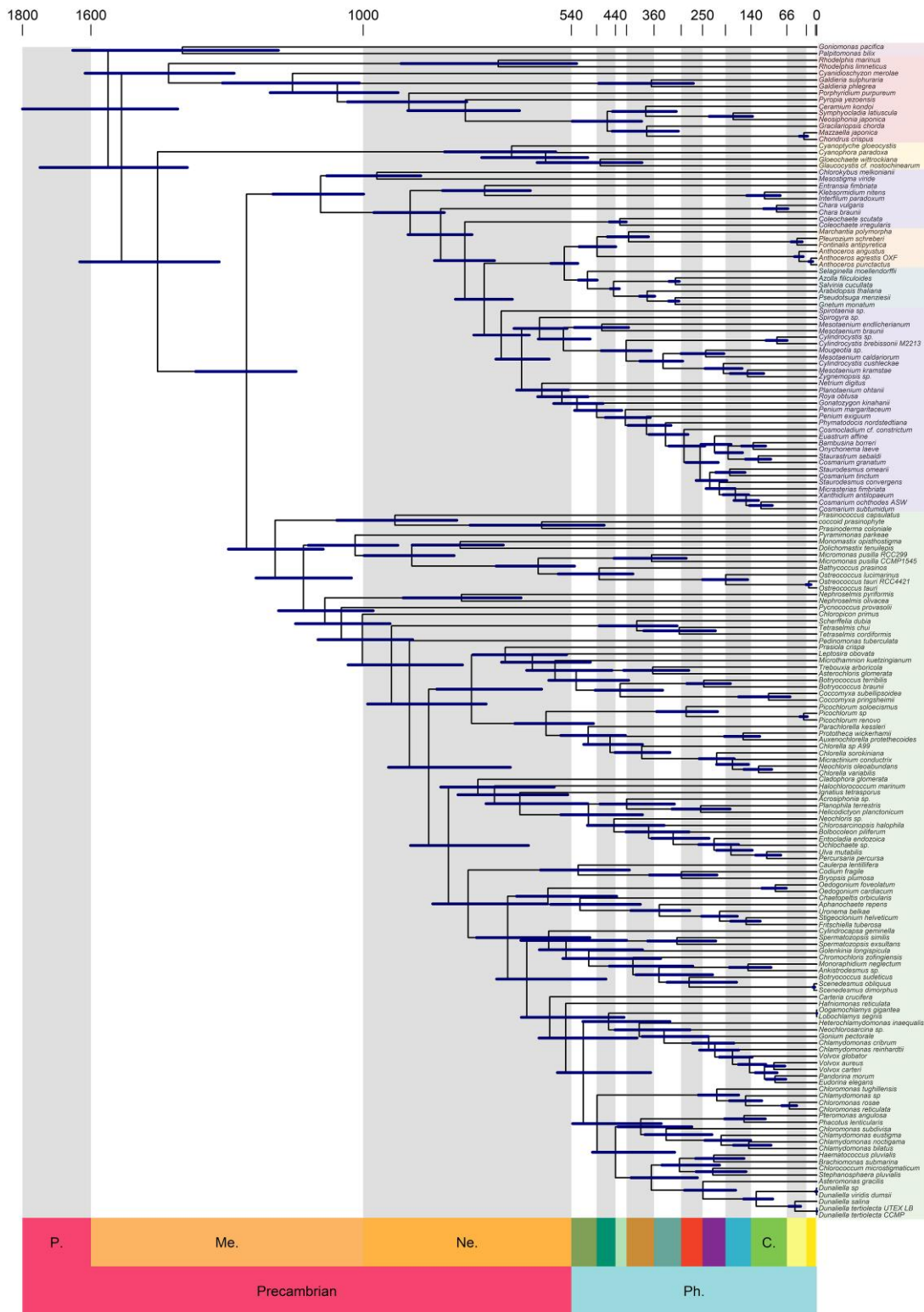
**FIG. 2.**—Support for alternative topologies of the Archaeplastida phylogeny. a) Phylogeny of Archaeplastida with key nodes numbered. b) Support for focal and alternative topologies for all major nodes in the Archaeplastida phylogeny.

time of 1712 to 1387 Ma, spanning the late-Paleoproterozoic to early-Mesoproterozoic. Viridiplantae has been estimated to emerge between the middle Mesoproterozoic to middle Neoproterozoic (1,400 to 669.9 Ma; 5,7,10,11) with our analysis recovering a divergence time of 1.369 to 1,147 Ma, during the mid-late Mesoproterozoic. Previous analyses have dated the origin of the first streptophytes during the middle Mesoproterozoic to late Neoproterozoic era (1,340 to 629.1 Ma; 5,10,11). Our analysis recovers a divergence time of 1,199 to 998 Ma (late Mesoproterozoic—earliest Neoproterozoic). These differences between our study and previously published work are likely due to taxon sampling as well as the interpretation of the fossil record, particularly for studies with taxa from the rhodophytes, glaucophytes, and green plants (Sánchez-Baracaldo et al. 2017; Betts et al. 2018; Lutzoni et al. 2018; Strassert et al. 2021).

Stochastic mapping, based on the time-calibrated phylogeny, is consistent with the Cryogenian multicellularity hypothesis, in which global glaciations promoted the origin of multicellularity across the eukaryotes (Fig. 4). Our timescale is compatible with this hypothesis, at least for red and streptophyte algae, though show contrasting patterns for the chlorophytes. For the latter, our data indicate that the Cryogenian multicellularity hypothesis is plausible for the ancestor of Chaetophorales and Oedogoniales. However, we

show here that multicellularity within Bryopsidales and the Ulvophyceae originated in the Phanerozoic.

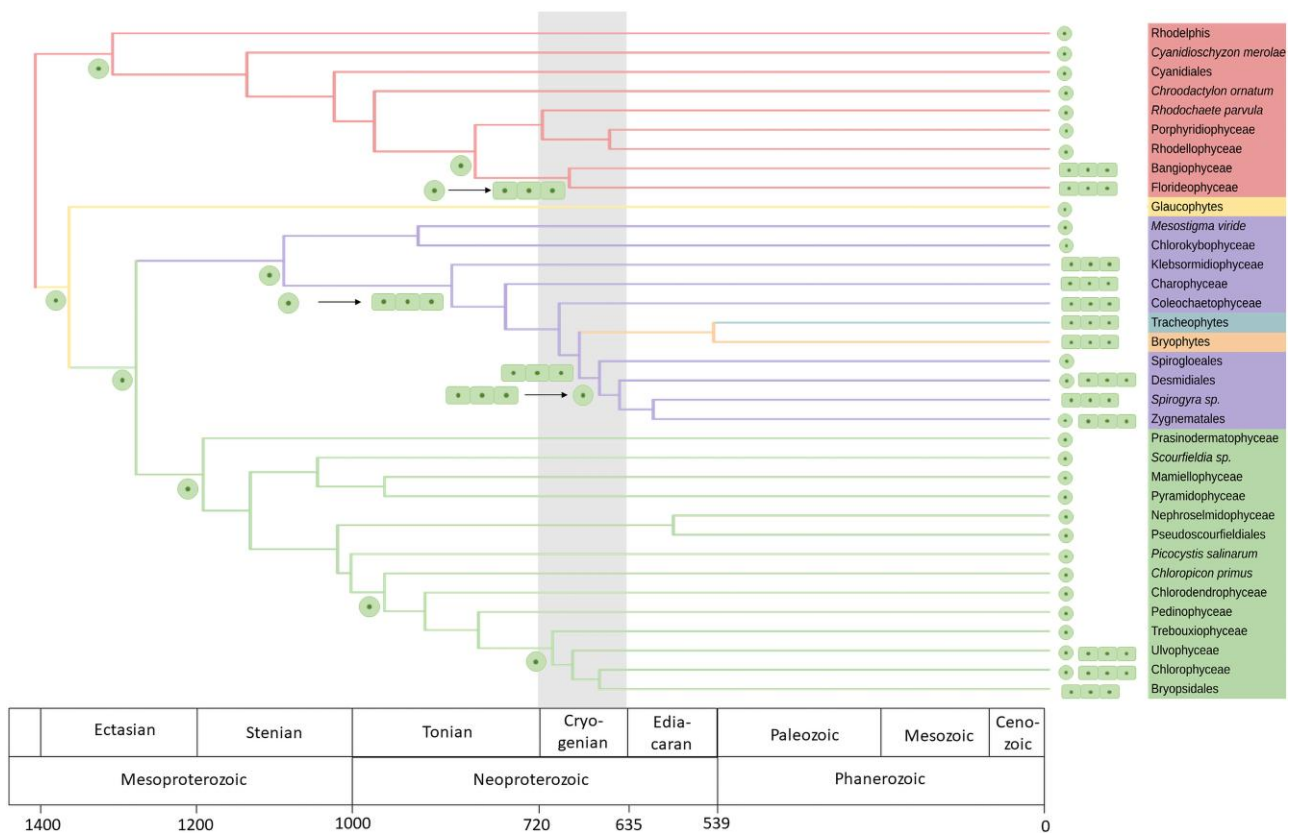
Stochastic mapping estimates ancestral states to a given node (crown group); however, the evolution of multicellularity may predate this ancestor (stem group). Therefore, while our analysis has localized the evolution of multicellularity to particular stem branches, the precise point of these state transitions is difficult to pinpoint. As such, while our analysis can provide a confident most recent divergence time of multicellular groups, there is a degree of uncertainty about the upper end of these estimates. This uncertainty may affect the potential temporal relationship between the origins of multicellularity and snowball Earth. The coding of characters also has important implications for ancestral character states. The coding of characters did not dramatically change the ancestral states in streptophyte and red algae. However, this did impact the ancestral state reconstruction in the chlorophyte algae, with the approaches coding more divisions of multicellularity (siphonous, sarcinoid, etc.) led to separate, distinct origins of multicellular groups (supplementary information S5a and b, Supplementary Material online). When these divisions of multicellularity (siphonous, sarcinoid, etc.) were placed within a single multicellular group, then multicellularity emerged once in the ancestor of Ulvophyceae and Chlorophyceae (supplementary information S5c, Supplementary Material



Downloaded from https://academic.oup.com/gbe/article/16/2/evae026/7604131 by University Library user on 12 March 2024

**Fig. 3.**—Time-calibrated phylogeny of Archaeplastida with 95% highest posterior density age uncertainties. Taxa are grouped into Outgroup (red group including *Palpitomonas bilix*), Rhodophytes (pink group including *Chondrus crispus*), Glaucophytes (yellow group including *Cyanophora paradoxa*), Chlorophytes (green group including *Volvox carteri*), Streptophyte algae (purple group including *Chara braunii* and *Spirogloea muscicola*), Bryophytes (blue group including *Marchantia polymorpha*) and Tracheophytes (orange group including *Arabidopsis thaliana*).





**FIG. 4.**—Ancestral state reconstruction of cytomorphological traits across the Archaeplastida. The topology of the tree is based on collapsed branches of the analysis presented in Fig. 3. Branch lengths are based on age estimates from the timescale analysis presented in Fig. 3. Ancestral state plotted onto the tree are summarized from [supplementary information S5b, Supplementary Material](#) online.

online). Analysis of the posterior probability density of this binary definition of multicellularity highlights the increasing likelihood of a multicellular ancestor along the spine of the streptophyte phylogeny ([supplementary information S5d, Supplementary Material](#) online).

### Multicellularity and the Cryogenian

With multiple origins across the tree of life as well as multiple definitions, the transition to multicellularity is not a simple story (Rose and Hammerschmidt 2021). Here, we use a broad definition of multicellularity to encompass any nonunicellular species. While these nonunicellular species are found across our phylogeny, complex multicellular organisms have traditionally been classified in only two clades of Archaeplastida, florideophyte red algae, and land plants (Knoll 2011). Our results suggest a stepwise evolution leading to more complex multicellularity in both these lineages. Our timescale is compatible with the interpretation of diverse taxa in the Ediacaran Weng’an biota (635 to 541 Ma), as stem-Corallinales (Xiao et al. 1998).

The paucity of the streptophyte algal fossil record prevents morphological comparisons, but phylogenetic analysis

provides mechanistic insights. The evolution of multicellularity required changes to gene networks regulating cell to cell adhesion, intercellular communication and cellular differentiation (Niklas 2014; Niklas and Newman 2020). The xyloglucan endotransglucosylase/hydrolases gene family, involved in plant cell wall expansion, were horizontally transferred from Alphaproteobacteria to the ancestor of Phragmoplastophyta and Klebsormidiophyceae and subsequently diversified in younger lineages (Shinohara and Nishitani 2021). Importantly, this gene was transferred during the Cryogenian. Several novel genes involved in cellulose synthesis, and therefore cell wall construction, emerge in the ancestor of Phragmoplastophyta and Klebsormidiophyceae (Lampugnani et al. 2019). Indeed, genome sequencing of zygmatophyte genomes indicated that most of the enzymes for cell wall innovations were present in this ancestor (Feng et al. 2023). Genome analysis has demonstrated how proliferations of genes involved in phytohormone signaling and cellular development lead to the emergence of the phragmoplast, a cytoplasmic structure involved in cell division, in the streptophyte lineage Phragmoplastophyta (Umen 2014; Nishiyama et al. 2018). Further genetic diversification and origin of new genes in Anydrophyta led to the

emergence of all major enzymes for the synthesis of cell wall components (Feng et al. 2023). The findings from the literature above and our analysis of the divergence time of Archaeplastida correlates the emergence of these genes to the Cryogenian, providing an initial mechanistic basis of multicellular adaptations.

Recent biomarker analysis has suggested that archaeplastid algal lineages became dominant over cyanobacteria after the Sturtian glaciation (Brocks et al. 2017; Hoshino et al. 2017; Brocks 2018). Some recent work has conflated the emergence of crown group eukaryotes with their early diversification (Porter et al. 2018; Budd and Mann 2020; Porter 2020), therefore proposing much later divergence time estimates (e.g. the Mesoproterozoic–Neoproterozoic boundary). However, our analysis identifies a distinction between the origin of photosynthetic eukaryotes in the late-Paleoproterozoic to early-Mesoproterozoic and their rise to ecological dominance (Figs. 3 and 4). Our analyses do not directly address when algae achieved ecological dominance though we note that there are alternative explanations of the biomarker record and that ecological modeling suggests they achieved significance early in the Proterozoic (Eckford-Soper et al. 2022).

Alongside environmental factors of the Snowball Earth conditions during the Cryogenian, several other drivers of multicellularity have also been proposed, including oxygen concentration, predation and nutrient concentration (Mills and Canfield 2014; Brocks et al. 2017; Herron et al. 2019). The ability to extract increased oxygen and nutrients, via multicellularity, may have provided ancestral organisms with competitive advantages in ancient environments (Knoll 2011; Brocks et al. 2017) while selection imposed by predators may have contributed to the emergence of multicellularity (Herron et al. 2019). These may have been additional important factors contributing to the colonization of snowball Earth by ancestral multicellular algae.

### Plant Terrestrialization and the Cryogenian

The estimate of 796 to 671 Ma for the divergence of Anydrophyta is entirely compatible with the Cryogenian plant terrestrialization hypothesis (Fig. 3). While temporal co-occurrence is distinct from causality, an origin for Anydrophyta during the Cryogenian is intriguing in the light of ancient cryospheric environments which are hypothesized as a major factor in the selection of key biological adaptations to terrestrial environments, leading ultimately to the origin of crown-embryophytes in the earliest Phanerozoic. This indicates that the time between the divergence of Embryophyta from Zygnematophyceae and the divergence of bryophytes from tracheophytes (i.e. the Embryophyte stem-lineage) was between 270 and 85 Myr in duration. In either instance, this is a long time span in which to assemble the body plan of crown-embryophytes, including the acquisition of embryonic development, 3D

growth, cuticle, stomata, meiospores and, likely, vascular tissues and axial branching (Donoghue et al. 2021). At the least, our ancestral state estimation indicates that the streptophyte lineage was already multicellular before the origin of Phragmoplastophyta, providing formal trait-based inference that Zygnematophyceae are secondarily simplified, corroborating recent inferences that these algae have lost many morphological traits (Buschmann and Zachgo 2016; Moody 2020; Hess et al. 2022) and genes since their common ancestor with land plants (Cheng et al. 2019; Bowles et al. 2020; Jiao et al. 2020; Gong and Han 2021; Harris et al. 2022). It remains unclear in which environment and in what order these novelties arose. Extant Zygnematophyceae algae (e.g. *Ancyronema nordenskiöldii* and *A. alaskana*) inhabit modern glaciers and are found in the sister lineage to land plants (Williamson et al. 2019; Procházková et al. 2021). Genome analysis of extant glacier algae will provide insight into these ancient physiological adaptations.

The origin of multicellular archaeplastids transformed the atmosphere, with larger organisms undertaking higher rates of photosynthesis, consuming more carbon dioxide and producing more oxygen (Simpson 2021). This evolutionary innovation led to ancestral archaeplastids becoming the dominant contributor to net primary production in ancient environments, surpassing cyanobacteria (Brocks et al. 2017). The consequent increase in atmospheric oxygen likely led to the development of new ecological niches and increased efficiency in the transfer of nutrients and energy between complex trophic structures. Elevated energy availability drove the rise of increasingly complex organisms and subsequently ecosystems. Ultimately, these drastic evolutionary changes would provide the backdrop to the origin of the first land plants, one of the most profound geobiological events in Earth history.

## Materials and Methods

### Dataset Selection

We used a latest sampling of transcriptome and genome data to infer the evolutionary history of Archaeplastida. Specifically, these included nonembryophyte transcriptomes from the one thousand plant transcriptomes project and genomes across the archaeplastid tree of life (supplementary information S1, Supplementary Material online, supplementary data S1, Supplementary Material online). Taxonomic sampling and genome completeness are key factors influencing the reconstruction of phylogeny. Considering this, distinct datasets were curated focusing on different branches of early archaeplastid evolution and the quality of the input data. These different datasets were used to assess the robustness of recovered topologies. On the whole, topologies across datasets were congruent. In total, eight datasets were assembled, focusing on



Streptophyta, Viridiplantae, and Archaeplastida with and without outgroup taxa. For each of these taxonomic groups, a data rich and taxon rich dataset were assembled. For the data rich datasets, BUSCO analysis, with the Eukaryota dataset, was used to assess the quality of transcriptome and genome data (Waterhouse et al. 2018) ([supplementary data S1, Supplementary Material](#) online). A benchmark of 30% missing BUSCO genes was used to filter high quality data. For the taxon rich dataset, all taxa were included. Therefore, the datasets consist of Streptophyta data rich (57 species), Streptophyta taxon rich (64 species), Viridiplantae data rich (166 species), Viridiplantae taxon rich (203 species), Archaeplastida without outgroup data rich (178 species), Archaeplastida without outgroup taxon rich (239 species), Archaeplastida data rich (180 species), and Archaeplastida taxon rich (241 species) ([supplementary data S1, Supplementary Material](#) online).

### Orthology Inference

OrthoFinder (Version 2.3.7) was used to cluster protein coding genes into orthogroups (Emms and Kelly 2019), using default settings (`orthofinder -f data_folder`). For all datasets, no universally present single copy orthologs were identified. Single copy genes are important for accurate phylogenetic inference. Therefore, near-single copy orthologs were identified using a previously described python script (Harris et al. 2020) which removes paralogous genes from orthogroups. The script enables the user to specify a minimum taxonomic occupancy of each orthogroup, set at 80%. Therefore, the datasets consist of Streptophyta data rich (222 genes), Streptophyta taxon rich (122 genes), Viridiplantae data rich (204 genes), Viridiplantae taxon rich (134 genes), Archaeplastida without outgroup data rich (133 genes), Archaeplastida without outgroup taxon rich (69 genes), Archaeplastida data rich (96 genes), and Archaeplastida taxon rich (48 genes).

### Species Tree Inference

#### Supermatrix Analysis

Single copy orthologs were aligned using MAFFT (Kato et al. 2002) using `-auto` parameter and trimmed with Trimal (Capella-Gutiérrez et al. 2009) using the `-automated1` parameter. Multiple sequence alignments were concatenated using Phyutility to create a supermatrix (Smith and Dunn 2008). A bootstrapped maximum likelihood phylogeny was inferred using IQ-Tree (Nguyen et al. 2015) using the Bayesian Information Criterion (BIC) to select best-fitting substitution model including empirical profile mixture models (C10–C60), which allow for compositional heterogeneity across sites and are much less prone to mutational saturation than simple single-matrix models (Williams et al. 2021). 1,000 ultrafast bootstrap replicates were used.

#### Supertree Analysis

Individual maximum likelihood gene trees were inferred for each single copy gene using IQ-Tree (Nguyen et al. 2015), using the BIC to select the best-fitting substitution model including empirical profile mixture models (C10–C60), which allow for compositional heterogeneity across sites. Supertree analysis was conducted using ASTRAL (Zhang et al. 2018) with default settings ([supplementary information S2, Supplementary Material](#) online).

#### Testing the Proposed Relationships of Early Archaeplastids

Uncertain and contested relationships across the early archaeplastid phylogeny were identified (Fig. 2b), including the position of the Zygnematophyceae, Bryopsidales, Prasinodermatophyta, Glaucophytes, and Rhodophytes (One Thousand Plant Transcriptomes Initiative 2019). Phylogenies were then constrained in IQ-Tree with these possible topologies and the support for these relationships was assessed using the Kishino and Hasegawa (Kishino and Hasegawa 1989; Goldman et al. 2000), Shimodaira and Hasegawa (Shimodaira and Hasegawa 1999), and AU (Shimodaira 2002) tests built within IQ-Tree. Based on these inferences, unlikely species tree hypotheses could then be rejected (Fig. 2).

#### Testing for the Impact of Across-branch Compositional Heterogeneity

Due to the deep evolutionary relationships of Archaeplastida, we tested the impact of compositional heterogeneity in our dataset. We used a previously described python script (Williams 2023), to filter out the 50% most compositional heterogeneous amino acid sites in our alignment. Phylogenetic analysis, using the methods from our supermatrix analysis, was repeated for this reduced dataset. This recovered the same topology suggesting our analysis is based a phylogenetically informative alignment ([supplementary information S8, Supplementary Material](#) online).

#### Divergence Time Estimation

We undertook a critical review of the fossil evidence used to calibrate archaeplastid evolution (see calibration descriptions in the [supplementary information S3, Supplementary Material](#) online). Interpreting the fossil record of early archaeplastids is challenging as the remains of unicellular algae are simple, precluding confident assignment to extant clades, while fossilized multicellular algae are often difficult to discriminate from filamentous cyanobacteria (e.g. *Archaeophycus yunnanensis* has been alternately interpreted as a bangiophyte red alga (Xiao et al. 1998), a trebouxioophyte green alga (Shang and Liu 2022) or a cyanobacterium (Pan et al. 2022)). With this in mind, 15 calibrations

were assigned based on stringent criteria (Parham et al. 2012), with the most influential being the multicellular green algae, *Proterocladus antiquus* (Gibson et al. 2018), used to calibrate the ancestor of Viridiplantae and the multicellular red algae, *Bangiomorpha pubescens* (Tang et al. 2020), used to calibrate the ancestor of Archaeplastida (supplementary Information S3, Supplementary Material online). While we are confident in the assignment of these fossils to Chlorophyta and Rhodophyta (respectively), their assignment to more derived lineages within these clades is based on their general gestalt, not on evidence of a nested hierarchy of characters that would support their assignment within the crown clades of red or green algae (Bowles et al. 2022). Similarly, a suite of fossil taxa from the Ediacaran Weng'an biota have been assigned to Bangiophyta and to derived clades within Florideophyta (Xiao et al. 2004, 1998) and they have been used widely to calibrate these clades (e.g. Yoon et al. 2004; Parfrey et al. 2011; Yang et al. 2016); unfortunately, they lack key phenotypic synapomorphies of crown-Rhodophyta (e.g. Gabrielson et al. 1985). We therefore interpret these taxa as records of the rhodophyte and chlorophyte total groups, reflecting uncertainty over whether they are member of their respective stem- or crown-groups. As we map the transition of multicellularity in rhodophytes to the common ancestor of Bangiophyceae and Florideophyceae (see results), this could be used to argue that these fossils should be used to calibrate a more derived position within red and green algae. However, it is difficult to exclude the loss of phenotypic features as well as extinction of evolutionary intermediates, hence our conservative use of the fossil record. Other possible early records of Archaeplastida (e.g. *Rafatazamia chitrakootia*, *Ramathallus lobatus*, *Tawuia*, and many unicellular algae) (Bengtson et al. 2017; Tang et al. 2021; Moczyłowska and Liu 2022) are too uncertain to be used as calibrations, at least at present (Carlisle et al. 2021; Bowles et al. 2022).

Node distributions using minimum and maximum constraints were specified, with full phylogenetic and age justifications listed in supplementary Information S3, Supplementary Material online. To specify the prior distributions on node ages, all calibrated land plant nodes were given a hard minimum age and a soft maximum age. For all algal calibrations, as their dates are harder to restrict, a 2.5% probability of exceeding both the minimum and maximum was used.

Initially, analyses were run without sequence data to obtain effective time priors, to ensure that the calibration densities and time priors were appropriate. A dataset of 164 protein coding genes were identified based on a taxonomic occupancy of 78% amongst the Archaeplastida data rich taxa. The single copy orthogroups were divided into 4 partitions according to their evolutionary rate, based on total tree length in IQTree (Nguyen et al. 2015) and grouped

using k-means clustering in R (R Core Team 2023). A log-normal independent relaxed clock model was used. Given the protein coding gene dataset, branch lengths were first estimated using codeml (Yang 2007). The gamma distribution for the mean substitution rate was assigned a diffuse shape parameter of 2 and a scale parameter of 10, based on pairwise distance between *Arabidopsis thaliana* and *Marchantia polymorpha*, assuming a divergence time of 469 Ma (Morris et al. 2018). The tree topology was fixed based on the focal maximum likelihood analysis above and was analyzed using the normal approximation method in MCMCTree (Yang 2007). The rate variation was assigned a shape parameter of 1 and a scale parameter of 10. The birth and death parameters were set to 1, specifying a uniform kernel. Burn-in was set at 40,000 with 400,000 samples and sampling frequency of 5. Two independent runs were completed. Trees were plotted using MCMCTreeR (Puttick 2019). A full list of posterior divergence time estimates are summarized in supplementary informations S9 and S10, Supplementary Material online.

### Comparison to Other Approaches

The impact of correlated rates on divergence times was assessed within MCMCTree, whilst all other parameters and input data remained the same as above. These were plotted in MCMCTreeR (supplementary Information S6, Supplementary Material online). Analysis was also conducted in Phylobayes (4.1: Lartillot et al. 2009), under log-normal autocorrelated relaxed clock models, using site-heterogenous CAT + GTR + G models. The same input data, tree topology and fossil calibration strategies were used as the main MCMCTree analysis. Two independent MCMC chains were run, the first 10% discarded as burnin. These were plotted in MCMCTreeR (supplementary information S7, Supplementary Material online).

### Stochastic Mapping of Multicellularity

Cytomorphology of species in this analysis was established from the published literature (supplementary information S4, Supplementary Material online). Previous analyses have coded characters for chlorophytes (Del Cortona et al. 2020; Li et al. 2021) which were transferred to this study. The tree topology was fixed based on the focal maximum likelihood analysis (Fig. 1) whilst branch lengths correspond to the timescale analysis presented in Fig. 3. Traits were discretely characterized as unicellular, filamentous, colony, siphonocladous, siphonous, sarcinoid, and multicellular. Ancestral states were estimated in R (R Core Team 2023) using the package phytools (Revell 2012). Briefly, likelihood ancestral states for discretely valued traits (e.g. multicellularity) were estimated using a continuous-time Markov chain model (Revell 2012). The MCMC approach is used to sample character histories from their posterior probability distribution

(Huelsenbeck et al. 2003). To sample a greater portion of the character history, 1,000 stochastic maps were produced and summarized (supplementary Information S4, Supplementary Material online).

As multicellularity can be classified in different ways, various classifications of multicellularity were tested (supplementary Informations S4 and S5, Supplementary Material online). These included coding filamentous organisms as a separate classification to multicellular organisms (supplementary information S5a, Supplementary Material online), incorporating filamentous within multicellularity (supplementary information S5b, Supplementary Material online) and coding multicellularity as a binary character e.g. unicellular or multicellular (supplementary information S5c, Supplementary Material online). Additionally for this binary definition of multicellularity, posterior probability density was calculated with the densityMap function from phytools (supplementary information S5d, Supplementary Material online). These analyses were largely congruent with each other. The ancestral states presented in Fig. 4 are summarized from supplementary information S5b, Supplementary Material online, based on the classification incorporating filamentous within multicellularity.

## Supplementary Material

Supplementary material is available at *Genome Biology and Evolution* online.

## Author Contributions

The work was devised by A.M.C.B., C.J.W., T.A.W., P.C.J.D. A.M.C.B. conducted the analyses and drafted the manuscript. All authors contributed to writing, reviewing, and editing the finished manuscript.

## Funding

We wish to acknowledge funding from the Leverhulme Trust (RPG-2020-199 “iDAPT” project to C.W., P.D., and T.W.; RF-2022-167 to P.D.); the Natural Environment Research Council (NE/P013678/1 to P.D.; part of the Biosphere Evolution, Transitions and Resilience (BETR) programme co-funded by the Natural Science Foundation of China (NSFC); the John Templeton Foundation (62220 to P.D. and T.W.; the opinions expressed here do not necessarily reflect the views of the John Templeton Foundation); the Gordon and Betty Moore Foundation (GBMF9741 to P.D. and T.W.) and the Royal Society for a University Research Fellowship to T.W. (URFR\201024).

## Conflict of Interest

The authors declare that they have no competing interests.

## Data Availability

All data are available in the main text, the supplementary materials or available on Figshare: <https://figshare.com/s/460c2a4704493e5bb3b9>.

## Literature Cited

- Becker B. Snow ball earth and the split of Streptophyta and Chlorophyta. *Trends Plant Sci.* 2013;18(4):180–183. <https://doi.org/10.1016/j.tplants.2012.09.010>.
- Bengtson S, Sallstedt T, Belivanova V, Whitehouse M. Three-dimensional preservation of cellular and subcellular structures suggests 1.6 billion-year-old crown-group red algae. *PLoS Biol.* 2017;15(3):e2000735. <https://doi.org/10.1371/journal.pbio.2000735>.
- Betts HC, Puttick MN, Clark JW, Williams TA, Donoghue PCJ, Pisani D. Integrated genomic and fossil evidence illuminates life's early evolution and eukaryote origin. *Nat Ecol Evol.* 2018;2(10):1556–1562. <https://doi.org/10.1038/s41559-018-0644-x>.
- Bowles AMC, Bechtold U, Paps J. The origin of land plants is rooted in two bursts of genomic novelty. *Curr Biol.* 2020;30(3):530–536.e2. <https://doi.org/10.1016/j.cub.2019.11.090>.
- Bowles AMC, Williamson CJ, Williams TA, Lenton TM, Donoghue PCJ. The origin and early evolution of plants. *Trends Plant Sci.* 2022;28(3):312–329. <https://doi.org/10.1016/j.tplants.2022.09.009>.
- Braakman R. Evolution of cellular metabolism and the rise of a globally productive biosphere. *Free Radic Biol Med.* 2019;140:172–187. <https://doi.org/10.1016/j.freeradbiomed.2019.05.004>.
- Brocks JJ. The transition from a cyanobacterial to algal world and the emergence of animals. *Emerg Top Life Sci.* 2018;2(2):181–190. <https://doi.org/10.1042/ETLS20180039>.
- Brocks JJ, Jarrett AJM, Sirantoine E, Hallmann C, Hoshino Y, Liyanage T. The rise of algae in Cryogenian oceans and the emergence of animals. *Nature.* 2017;548(7669):578–581. <https://doi.org/10.1038/nature23457>.
- Budd GE, Mann RP. The dynamics of stem and crown groups. *Sci Adv.* 2020;6(8):eaaz1626. <https://doi.org/10.1126/sciadv.aaz1626>.
- Burki F, Roger AJ, Brown MW, Simpson AGB. The new tree of eukaryotes. *Trends Ecol Evol.* 2020;35(1):43–55. <https://doi.org/10.1016/j.tree.2019.08.008>.
- Buschmann H, Zachgo S. The evolution of cell division: from streptophyte algae to land plants. *Trends Plant Sci.* 2016;21(10):872–883. <https://doi.org/10.1016/j.tplants.2016.07.004>.
- Capella-Gutiérrez S, Silla-Martínez JM, Gabaldón T. Trimal: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics.* 2009;25(15):1972–1973. <https://doi.org/10.1093/bioinformatics/btp348>.
- Carlisle EM, Jobbins M, Pankhania V, Cunningham JA, Donoghue PCJ. Experimental taphonomy of organelles and the fossil record of early eukaryote evolution. *Sci Adv.* 2021;7(5):eabe9487. <https://doi.org/10.1126/sciadv.abe9487>.
- Cheng S, Xian W, Fu Y, Marin B, Keller J, Wu T, Sun W, Li X, Xu Y, Zhang Y, et al. Genomes of subaerial zygnematophyceae provide insights into land plant evolution. *Cell.* 2019;179(5):1057–1067. <https://doi.org/10.1016/j.cell.2019.10.019>.
- Christenhusz MJM, Byng JW. The number of known plant species in the world and its annual increase. *Phytotaxa.* 2016;261(3):201–217. <https://doi.org/10.11646/phytotaxa.261.3.1>.
- Corlett RT. Plant diversity in a changing world: status, trends, and conservation needs. *Plant Divers.* 2016;38(1):10–16. <https://doi.org/10.1016/j.pld.2016.01.001>.

- Delaux PM, Schornack S. Plant evolution driven by interactions with symbiotic and pathogenic microbes. *Science*. 2021;371(6531): eaba6605. <https://doi.org/10.1126/science.aba6605>.
- Del Cortona A, Jackson CJ, Bucchini F, Van Bel M, D'hondt S, Škaloud P, Delwiche CF, Knoll AH, Raven JA, Verbruggen H, et al. Neoproterozoic origin and multiple transitions to macroscopic growth in green seaweeds. *Proc Natl Acad Sci U S A*. 2020; 117(5):2551–2559. <https://doi.org/10.1073/pnas.1910060117>.
- DeLong JP, Okie JG, Moses ME, Sibly RM, Brown JH. Shifts in metabolic scaling, production, and efficiency across major evolutionary transitions of life. *Proc Natl Acad Sci U S A*. 2010;107(29):12941–12945. <https://doi.org/10.1073/pnas.1007783107>.
- de Vries J, Archibald JM. Plant evolution: landmarks on the path to terrestrial life. *New Phytol*. 2018;217(4):1428–1434. <https://doi.org/10.1111/nph.14975>.
- Donoghue PCJ, Harrison CJ, Paps J, Schneider H. The evolutionary emergence of land plants. *Curr Biol*. 2021;31(19):1281–1298. <https://doi.org/10.1016/j.cub.2021.07.038>.
- Eckford-Soper LK, Andersen KH, Hansen TF, Canfield DE. A case for an active eukaryotic marine biosphere during the Proterozoic era. *Proc Natl Acad Sci U S A*. 2022;119(41):e2122042119. <https://doi.org/10.1073/pnas.2122042119>.
- Emms DM, Kelly S. OrthoFinder: phylogenetic orthology inference for comparative genomics. *Genome Biol*. 2019;20(1):238. <https://doi.org/10.1186/s13059-019-1832-y>.
- Fang L, Leliaert F, Zhang ZH, Penny D, Zhong BJ. Evolution of the Chlorophyta: insights from chloroplast phylogenomic analyses. *J Syst Evol*. 2017;55(4):322–332. <https://doi.org/10.1111/jse.12248>.
- Feng X, Zheng J, Irisarri I, Yu H, Zheng B, Ali Z, de Vries S, Keller J, Fürst-Jansen JMR, Dadrás A, et al. Chromosome-level genomes of multicellular algal sisters to land plants illuminate signaling network evolution. *bioRxiv* 526407. <https://doi.org/10.1101/2023.01.31.526407>, 01 February 2023, preprint: not peer reviewed.
- Gabrielson PW, Garbary DJ, Scagel RF. The nature of the ancestral red alga: inferences from a cladistic analysis. *Biosystems*. 1985; 18(3-4):335–346. [https://doi.org/10.1016/0303-2647\(85\)90033-4](https://doi.org/10.1016/0303-2647(85)90033-4).
- Gawryluk RMR, Tikhonenkov DV, Hehenberger E, Husnik F, Mylnikov AP, Keeling PJ. Non-photosynthetic predators are sister to red algae. *Nature*. 2019;572(7768):240–243. <https://doi.org/10.1038/s41586-019-1398-6>.
- Gibson TM, Shih PM, Cumming VM, Fischer WW, Crockford PW, Hodgskiss MS, Wöhrndle S, Creaser RA, Rainbird RH, Skulski TM, et al. Precise age of *Bangiomorpha pubescens* dates the origin of eukaryotic photosynthesis. *Geology*. 2018;46(2):135–138. <https://doi.org/10.1130/G39829.1>.
- Goldman N, Anderson JP, Rodrigo AG. Likelihood-based tests of topologies in phylogenetics. *Syst Biol*. 2000;49(4):652–670. <https://doi.org/10.1080/106351500750049752>.
- Gong Z, Han G-Z. Flourishing in water: the early evolution and diversification of plant receptor-like kinases. *Plant J*. 2021;106(1): 174–184. <https://doi.org/10.1111/tpj.15157>.
- Guiry MD, Guiry GM, Morrison L, Rindi F, Miranda SV, Mathieson AC, Parker BC, Langangen A, John DM, Bárbara I, et al. AlgaeBase: an on-line resource for algae. *Cryptogam Algal*. 2014;35(2): 105–115. <https://doi.org/10.7872/crya.v35.iss2.2014.105>.
- Harris BJ, Clark JW, Schrepf D, Szöllösi GJ, Donoghue PCJ, Hetherington AM, Williams TA. Divergent evolutionary trajectories of bryophytes and tracheophytes from a complex common ancestor of land plants. *Nat Ecol Evol*. 2022;6(11):1634–1643. <https://doi.org/10.1038/s41559-022-01885-x>.
- Harris BJ, Harrison CJ, Hetherington AM, Williams TA. Phylogenomic evidence for the monophyly of bryophytes and the reductive evolution of stomata. *Curr Biol*. 2020;30(11):2001–2012. <https://doi.org/10.1016/j.cub.2020.03.048>.
- Herron MD, Borin JM, Boswell JC, Walker J, Chen IK, Knox CA, Boyd M, Rosenzweig F, Ratcliff WC. De novo origins of multicellularity in response to predation. *Sci Rep*. 2019;9(1):2328. <https://doi.org/10.1038/s41598-019-39558-8>.
- Hess S, Williams SK, Busch A, Irisarri I, Delwiche CF, de Vries S, Darienko T, Roger AJ, Archibald JM, Buschmann H, et al. A phylogenetically informed five-order system for the closest relatives of land plants. *Curr Biol*. 2022;32(20):4473–4482.e7. <https://doi.org/10.1016/j.cub.2022.08.022>.
- Hoffman PF, Abbot DS, Ashkenazy Y, Benn DI, Brocks JJ, Cohen PA, Cox GM, Creveling JR, Donnadiou Y, Erwin DH, et al. Snowball Earth climate dynamics and Cryogenian geology-geobiology. *Sci Adv*. 2017;3(11):e1600983. <https://doi.org/10.1126/sciadv.1600983>.
- Hoffman PF, Kaufman AJ, Halverson GP, Schrag DP. A Neoproterozoic snowball Earth. *Science*. 1998;281(5381):1342–1346. <https://doi.org/10.1126/science.281.5381.1342>.
- Hoshino Y, Poshibaeva A, Meredith W, Snape C, Poshibaev V, Versteegh GJM, Kuznetsov N, Leider A, van Maldegem L, Neumann M, et al. Cryogenian evolution of stigmateroid biosynthesis. *Sci Adv*. 2017;3(9):e1700887. <https://doi.org/10.1126/sciadv.1700887>.
- Hou Z, Ma X, Shi X, Li X, Yang L, Xiao S, De Clerck O, Leliaert F, Zhong B. Phylotranscriptomic insights into a Mesoproterozoic–Neoproterozoic origin and early radiation of green seaweeds (Ulvoophyceae). *Nat Commun*. 2022;13(1):1610. <https://doi.org/10.1038/s41467-022-29282-9>.
- Huelsenbeck JP, Nielsen R, Bollback JP. Stochastic mapping of morphological characters. *Syst Biol*. 2003;52(2):131–158. <https://doi.org/10.1080/10635150390192780>.
- Irisarri I, Strasser JFH, Burki F. Phylogenomic insights into the origin of primary plastids. *Syst Biol*. 2022;71(1):105–120. <https://doi.org/10.1093/sysbio/syab036>.
- Jiao C, Sørensen I, Sun X, Sun H, Behar H, Alseikh S, Philippe G, Palacio Lopez K, Sun L, Reed R, et al. The *Penium margaritaceum* genome: hallmarks of the origins of land plants. *Cell*. 2020;181(5): 1097–1111. <https://doi.org/10.1016/j.cell.2020.04.019>.
- Katoh K, Misawa K, Kuma K, Miyata T. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res*. 2002;30(14):3059–3066. <http://www.ncbi.nlm.nih.gov/pubmed/12136088>. <https://doi.org/10.1093/nar/gkf436>.
- Kishino H, Hasegawa M. Evaluation of the maximum likelihood estimate of the evolutionary tree topologies from DNA sequence data, and the branching order in hominoidea. *J Mol Evol*. 1989;29(2):170–179. <https://doi.org/10.1007/BF02100115>.
- Knoll AH. The multiple origins of complex multicellularity. *Annu Rev Earth Planet Sci*. 2011;39(1):217–239. <https://doi.org/10.1146/annurev.earth.031208.100209>.
- Lampugnani ER, Flores-Sandoval E, Tan QW, Mutwil M, Bowman JL, Persson S. Cellulose synthesis—central components and their evolutionary relationships. *Trends Plant Sci*. 2019;24(5):402–412. <https://doi.org/10.1016/j.tplants.2019.02.011>.
- Lartillot N, Lepage T, Blanquart S. PhyloBayes 3: a Bayesian software package for phylogenetic reconstruction and molecular dating. *Bioinformatics*. 2009;25(17):2286–2288. <https://doi.org/10.1093/bioinformatics/btp368>.
- Lenton TM, Dahl TW, Daines SJ, Mills BJW, Ozaki K, Saltzman MR, Porada P. Earliest land plants created modern levels of atmospheric oxygen. *Proc Natl Acad Sci U S A*. 2016;113(35):9704–9709. <https://doi.org/10.1073/pnas.1604787113>.
- Li X, Hou Z, Xu C, Shi X, Yang L, Lewis LA, Zhong B. Large phylogenomic data sets reveal deep relationships and trait evolution in chlorophyte green algae. *Genome Biol Evol*. 2021;13(7): evab101. <https://doi.org/10.1093/gbe/evab101>.



- Li L, Wang S, Wang H, Sahu SK, Marin B, Li H, Xu Y, Liang H, Li Z, Cheng S, et al. The genome of *Prasinoderma coloniale* unveils the existence of a third phylum within green plants. *Nat Ecol Evol*. 2020;4(9):1220–1231. <https://doi.org/10.1038/s41559-020-1221-7>.
- Lughadha EN, Govaerts R, Belyaeva I, Black N, Lindon H, Allkin R, Magill RE, Nicolson N. Counting counts: revised estimates of numbers of accepted species of flowering plants, seed plants, vascular plants and land plants with a review of other recent estimates. *Phytotaxa*. 2016;272(1):82–88. <https://doi.org/10.11646/phytotaxa.272.1.5>.
- Lutizoni F, Nowak MD, Alfaro ME, Reeb V, Miadlikowska J, Krug M, Arnold AE, Lewis LA, Swofford DL, Hibbett D, et al. Contemporaneous radiations of fungi and plants linked to symbiosis. *Nat Commun*. 2018;9(1):5451. <https://doi.org/10.1038/s41467-018-07849-9>.
- Mills DB, Canfield DE. Oxygen and animal evolution: did a rise of atmospheric oxygen “trigger” the origin of animals? *BioEssays*. 2014;36(12):1145–1155. <https://doi.org/10.1002/bies.201400101>.
- Moczydłowska M, Liu P. Ediacaran algal cysts from the Doushantuo Formation, South China. *Geol Mag*. 2022;159(7):1050–1070. <https://doi.org/10.1017/S0016756820001405>.
- Moody LA. Three-dimensional growth: a developmental innovation that facilitated plant terrestrialization. *J Plant Res*. 2020;133(3):283–290. <https://doi.org/10.1007/s10265-020-01173-4>.
- Morris JL, Puttick MN, Clark JW, Edwards D, Kenrick P, Pressel S, Wellman CH, Yang Z, Schneider H, Donoghue PCJ. The timescale of early land plant evolution. *Proc Natl Acad Sci U S A*. 2018;115(10):2274–2283. <https://doi.org/10.1073/pnas.1719588115>.
- Nguyen L-T, Schmidt HA, von Haeseler A, Minh BQ. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol Biol Evol*. 2015;32(1):268–274. <https://doi.org/10.1093/molbev/msu300>.
- Nie Y, Foster CSP, Zhu T, Yao R, Duchêne DA, Ho SYW, Zhong B. Accounting for uncertainty in the evolutionary timescale of green plants through clock-partitioning and fossil calibration strategies. *Syst Biol*. 2020;69(1):1–16. <https://doi.org/10.1093/sysbio/syz032>.
- Niklas KJ. The evolutionary-developmental origins of multicellularity. *Am J Bot*. 2014;101(1):6–25. <https://doi.org/10.3732/ajb.1300314>.
- Niklas KJ, Newman SA. The many roads to and from multicellularity. *J Exp Bot*. 2020;71(11):3247–3253. <https://doi.org/10.1093/jxb/erz547>.
- Nishiyama T, Sakayama H, de Vries J, Buschmann H, Saint-Marcoux D, Ullrich KK, Haas FB, Vanderstraeten L, Becker D, Lang D, et al. The *Chara* genome: secondary complexity and implications for plant terrestrialization. *Cell*. 2018;174(2):448–464. <https://doi.org/10.1016/j.cell.2018.06.033>.
- One Thousand Plant Transcriptomes Initiative. One thousand plant transcriptomes and the phylogenomics of green plants. *Nature*. 2019;574(7780):679–685. <https://doi.org/10.1038/s41586-019-1693-2>.
- Pan X, Xiong L, Dai Q, Luo J, Liu Z, Wang T, Hua H. Phosphatized *Obruchevella* and other microfossils from the Ediacaran-Cambrian transition. *Precambrian Res*. 2022;380:106825. <https://doi.org/10.1016/j.precambres.2022.106825>.
- Parfrey LW, Lahr DJ, Knoll AH, Katz LA. Estimating the timing of early eukaryotic diversification with multigene molecular clocks. *Proc Natl Acad Sci U S A*. 2011;108(33):13624–13629. <https://doi.org/10.1073/pnas.1110633108>.
- Parham JF, Donoghue PC, Bell CJ, Calway TD, Head JJ, Holroyd PA, Inoue JG, Irmis RB, Joyce WG, Ksepka DT, et al. Best practices for justifying fossil calibrations. *Syst Biol*. 2012;61(2):346–359. <https://doi.org/10.1093/sysbio/syr107>.
- Porter SM. Insights into eukaryogenesis from the fossil record. *Interface Focus*. 2020;10(4):20190105. <https://doi.org/10.1098/rsfs.2019.0105>.
- Porter SM, Agić H, Riedman LA. Anoxic ecosystems and early eukaryotes. *Emerg Top Life Sci*. 2018;2(2):299–309. <https://doi.org/10.1042/ETLS20170162>.
- Price DC, Goodenough UW, Roth R, Lee J-H, Kariyawasam T, Mutwil M, Ferrari C, Facchinelli F, Ball SG, Cenci U, et al. Analysis of an improved *Cyanophora paradoxa* genome assembly. *DNA Res*. 2019;26(4):287–299. <https://doi.org/10.1093/dnares/dsz009>.
- Procházková L, Řezanka T, Nedbalová L, Remias D. Unicellular versus filamentous: the glacial alga *Ancylonema alaskana* comb. et stat. nov. and its ecophysiological relatedness to *Ancylonema nordenskiöldii* (Zygnematomyxaceae, Streptophyta). *Microorganisms*. 2021;9(5):1103. <https://doi.org/10.3390/microorganisms9051103>.
- Puttick MN. MCMCtree: functions to prepare MCMCtree analyses and visualize posterior ages on trees. *Bioinformatics*. 2019;35(24):5321–5322. <https://doi.org/10.1093/bioinformatics/btz554>.
- Puttick MN, Morris JL, Williams TA, Cox CJ, Edwards D, Kenrick P, Pressel S, Wellman CH, Schneider H, Pisani D, et al. The interrelationships of land plants and the nature of the ancestral embryophyte. *Curr Biol*. 2018;28(5):733–745. <https://doi.org/10.1016/j.cub.2018.01.063>.
- R Core Team. R: a language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2023. <https://www.R-project.org/>.
- Revell LJ. Phytools: an R package for phylogenetic comparative biology (and other things). *Methods Ecol Evol*. 2012;3(2):217–223. <https://doi.org/10.1111/j.2041-210X.2011.00169.x>.
- Rose CJ, Hammerschmidt K. What do we mean by multicellularity? The evolutionary transitions framework provides answers. *Front Ecol Evol*. 2021;9:730714. <https://doi.org/10.3389/fevo.2021.730714>.
- Sánchez-Baracaldo P, Raven JA, Pisani D, Knoll AH. Early photosynthetic eukaryotes inhabited low-salinity habitats. *Proc Natl Acad Sci U S A*. 2017;114(37):7737–7745. <https://doi.org/10.1073/pnas.1620089114>.
- Schön ME, Zlatogursky VV, Singh RP, Poirier C, Wilken S, Mathur V, Strasser JFH, Pinhassi J, Worden AZ, Keeling PJ, et al. Single cell genomics reveals plastid-lacking picrozoa are close relatives of red algae. *Nat Commun*. 2021;12(1):6651. <https://doi.org/10.1038/s41467-021-26918-0>.
- Shang X, Liu P. Diverse multicellular algae from the early Ediacaran Doushantuo chert nodules and their palaeoecological implications. *Precambrian Res*. 2022;368:106508. <https://doi.org/10.1016/j.precambres.2021.106508>.
- Shimodaira H. An approximately unbiased test of phylogenetic tree selection. *Syst Biol*. 2002;51(3):492–508. <https://doi.org/10.1080/10635150290069913>.
- Shimodaira H, Hasegawa M. Multiple comparisons of log-likelihoods with applications to phylogenetic inference. *Mol Biol Evol*. 1999;16(8):1114–1116. <https://doi.org/10.1093/oxfordjournals.molbev.a026201>.
- Shinohara N, Nishitani K. Cryogenian origin and subsequent diversification of the plant cell-wall enzyme XTH family. *Plant Cell Physiol*. 2021;62(12):1874–1889. <https://doi.org/10.1093/pcp/pcab093>.
- Simpson C. Adaptation to a viscous snowball earth ocean as a path to complex multicellularity. *Am Nat*. 2021;198(5):590–609. <https://doi.org/10.1086/716634>.
- Smith SA, Dunn CW. Phyutility: a phyloinformatics tool for trees, alignments and molecular data. *Bioinformatics*. 2008;24(5):715–716. <https://doi.org/10.1093/bioinformatics/btm619>.
- Solari CA, Kessler JO, Michod RE. A hydrodynamics approach to the evolution of multicellularity: flagellar motility and germ-soma

- differentiation in volvoclean green algae. *Am Nat.* 2006;167(4):537–554. <https://doi.org/10.1086/501031>.
- Stern RJ, Avigad D, Miller NR, Beyth M. Evidence for the snowball earth hypothesis in the Arabian-Nubian shield and the east African Orogen. *J African Earth Sci.* 2006;44(1):1–20. <https://doi.org/10.1016/j.jafrearsci.2005.10.003>.
- Strassert JFH, Irisarri I, Williams TA, Burki F. A molecular timescale for eukaryote evolution with implications for the origin of red algal-derived plastids. *Nat Commun.* 2021;12(1):1879. <https://doi.org/10.1038/s41467-021-22044-z>.
- Tang Q, Pang K, Li G, Chen L, Yuan X, Sharma M, Xiao S. The Proterozoic macrofossil Tawuia as a coenocytic eukaryote and a possible macroalga. *Palaeogeogr Palaeoclimatol Palaeoecol.* 2021;576:110485. <https://doi.org/10.1016/j.palaeo.2021.110485>.
- Tang Q, Pang K, Yuan X, Xiao S. A one-billion-year-old multicellular chlorophyte. *Nat Ecol Evol.* 2020;4(4):543–549. <https://doi.org/10.1038/s41559-020-1122-9>.
- Umen JG. Green algae and the origins of multicellularity in the plant kingdom. *Cold Spring Harb Perspect Biol.* 2014;6(11):a016170. <https://doi.org/10.1101/cshperspect.a016170>.
- Waterhouse RM, Seppay M, Simão FA, Manni M, Ioannidis P, Klioutchnikov G, Kriventseva EV, Zdobnov EM. BUSCO applications from quality assessments to gene prediction and phylogenomics. *Mol Biol Evol.* 2018;35(3):543–548. <https://doi.org/10.1093/molbev/msx319>.
- Wickett NJ, Mirarab S, Nguyen N, Warnow T, Carpenter E, Matasci N, Ayyampalayam S, Barker MS, Burleigh JG, Gitzendanner MA, et al. Phylotranscriptomic analysis of the origin and early diversification of land plants. *Proc Natl Acad Sci U S A.* 2014;111(45):4859–4868. <https://doi.org/10.1073/pnas.1323926111>.
- Williams TA. phylo. 2023. [accessed 2024 Feb 16]. <https://github.com/Tancata/phylo>.
- Williams TA, Schrepf D, Szöllösi GJ, Cox CJ, Foster PG, Embley TM. Inferring the deep past from molecular data. *Genome Biol Evol.* 2021;13(5):evab067. <https://doi.org/10.1093/gbe/evab067>.
- Williamson CJ, Cameron KA, Cook JM, Zarsky JD, Stibal M, Edwards A. Glacier algae: a dark past and a darker future. *Front Microbiol.* 2019;10:524. <https://doi.org/10.3389/fmicb.2019.00524>.
- Xiao S, Knoll AH, Yuan X, Pueschel CM. Phosphatized multicellular algae in the Neoproterozoic Doushantuo Formation, China, and the early evolution of florideophyte red algae. *Am J Bot.* 2004;91(2):214–227. <https://doi.org/10.3732/ajb.91.2.214>.
- Xiao S, Zhang Y, Knoll AH. Three-dimensional preservation of algae and animal embryos in a Neoproterozoic phosphorite. *Nature.* 1998;391(6667):553–558. <https://doi.org/10.1038/35318>.
- Yang Z. PAML 4: phylogenetic analysis by maximum likelihood. *Mol Biol Evol.* 2007;24(8):1586–1591. <https://doi.org/10.1093/molbev/msm088>.
- Yang EC, Boo SM, Bhattacharya D, Saunders GW, Knoll AH, Fredericq S, Graf L, Yoon HS. Divergence time estimates and the evolution of major lineages in the florideophyte red algae. *Sci Rep.* 2016;6:21361. <https://doi.org/10.1038/srep21361>.
- Yang Z, Ma X, Wang Q, Tian X, Sun J, Zhang Z, Xiao S, De Clerck O, Leliaert F, Zhong B. Phylotranscriptomics unveil a Paleoproterozoic–Mesoproterozoic origin and deep relationships of the Viridiplantae. *Nat Commun.* 2023;14(1):5542. <https://doi.org/10.1038/s41467-023-41137-5>.
- Yoon HS, Hackett JD, Ciniglia C, Pinto G, Bhattacharya D. A molecular timeline for the origin of photosynthetic eukaryotes. *Mol Biol Evol.* 2004;21(5):809–818. <https://doi.org/10.1093/molbev/msh075>.
- Žárský J, Žárský V, Hanáček M, Žárský V. Cryogenian glacial habitats as a plant terrestrialisation cradle—the origin of the anydrophytes and zygnetophyceae split. *Front Plant Sci.* 2022;12:735020. <https://doi.org/10.3389/fpls.2021.735020>.
- Zhang C, Rabiee M, Sayyari E, Mirarab S. ASTRAL-III: polynomial time species tree reconstruction from partially resolved gene trees. *BMC Bioinformatics.* 2018;19(Suppl 6):153. <https://doi.org/10.1186/s12859-018-2129-y>.

Associate editor: Beatriz Mello